Molecular and genetic regulation of pig pancreatic islet cell development

Seokho Kim[1,9], Robert L. Whitener[1,9], Heshan Peiris[1], Xueying Gu[1], Charles A. Chang[1], Jonathan Y. Lam[1], Joan Camunas-Soler[2], Insung Park[3], Romina J. Bevacqua[1], Krissie Tellez[1], Stephen R. Quake[2,4], Jonathan R. T. Lakey[4], Rita Bottino[5], Pablo J. Ross[3], Seung K. Kim[1,7,8]

[1]Department of Developmental Biology, Stanford University School of Medicine, Stanford, CA, 94305 USA

[2]Department of Bioengineering, Stanford University, Stanford, CA, 94305 USA

[3]Department of Animal Science, University of California Davis, Davis, CA, 95616 USA

[4]Chan Zuckerberg Biohub, San Francisco, CA 94518, USA.

[5]Department of Surgery, University of California Irvine, Irvine, CA, 92868 USA

[6]Institute of Cellular Therapeutics, Allegheny Health Network, Pittsburgh, PA, 15212 USA

[7]Department of Medicine, Stanford University School of Medicine, Stanford, CA, 94305 USA

[8]Stanford Diabetes Research Center, Stanford University School of Medicine, Stanford, CA, 94305 USA

[9]These authors contributed equally

Correspondence and requests for materials should be addressed to S.K.K (email: seungkim@stanford.edu)

Abstract

Reliance on rodents for understanding pancreatic genetics, developmental biology and islet function could limit progress in developing interventions for human diseases like diabetes mellitus. Similarities of pancreas morphology and function, and pancreatic disease modeling in pigs suggest that porcine and human pancreas developmental biology may have useful homologies. However, little is known about pig pancreas development. To fill this knowledge gap, we investigated fetal and neonatal pig pancreas at multiple, crucial developmental stages using modern experimental approaches. Purification of islet β-, α- and δ-cells via flow cytometry followed by high-throughput transcriptome analysis (RNA-Seq) provided comprehensive gene expression profiles. Morphometric analysis revealed dynamic developmental endocrine cell allocation and islet architectural similarities between pig and human islets. Gene expression analysis identified cell- and stage-specific expression in sorted pig β- and α-cells, and revealed that pig and human β-cells and α-cells shared characteristic molecular and developmental features not observed in mouse β-cells or α-cells. Over 150 causal or candidate 'diabetes risk' genes identified by human studies had dynamic developmental expression in pig β- and α-cells. Our analysis also unveiled scores of signaling pathways linked to native islet β-cell functional maturation. Thus, the findings, conceptual advances, and resources detailed here show how pig pancreatic islet studies complement or surpass other systems for understanding the developmental programs that generate functional β- and α-cells, and that are relevant to human diseases like diabetes mellitus.

Progress from studies of pancreas biology in humans[1,2] has advanced our understanding of post-natal islet regulation and function. The recognition that mouse pancreas biology – while essential – has limitations for understanding human pancreas formation or diseases[3,4], have intensified interest in experimental systems that more closely reflect human pancreas development[5,6]. There is a specific, crucial knowledge gap in our understanding of human islet and pancreas development from mid-gestation through neonatal and juvenile stages, a period when critical aspects of islet development are known to occur[7]. The use of human primary tissues to address such questions is limited by inter-individual heterogeneity, unpredictable and restricted access to tissue from key developmental stages, and challenges implementing high-throughput molecular approaches. Likewise, reconstitution of islet development from stem cell sources or fetal human cell lines remains imperfect[8], thereby limiting interpretation of developmental studies in those systems.

Humans and pigs are omnivorous mammals with similar physiology and metabolic diseases, including diabetes mellitus and its complications, from diet-induced obesity, insulinopathy and β-cell stress[9-11]. Recent experimental advances expand possible studies of pig pancreas development to include gain and loss-of-function genetics[12-15], fetal surgery and immunomodulation[16], and primary islet cell genetics[17]. Prior studies of pig pancreas development have largely relied on immunohistological survey of tissue cell types[18-21]. Moreover, studies linking hallmark islet functions, like regulated insulin secretion and dynamic changes in gene expression at advancing developmental stages, have not been previously reported. Thus, pancreas and metabolic research could benefit from systematic application of powerful methods like cell purification by flow cytometry, high-throughput transcriptome analysis, and islet physiological studies across a comprehensive range of pig developmental stages.

Here we apply these and other modern approaches to delineate pig pancreas development, with a focus on islet β-cells, the sole source of insulin, and α-cells, the principal systemic source of the hormone glucagon. Dysregulation of hormone secretion by these cells is thought to be a leading cause of both type 1 diabetes (T1D) and type 2 diabetes (T2D) in humans[22,23]. Our findings provide molecular evidence of greater similarity between pig and human β-cells and α-cells than to cognate mouse cells, and demonstrate how studies of pig pancreatic cells can advance our understanding of the genetic, molecular, signaling and physiological programs that generate functional β- and α-cells.

Results

Pancreas dissociation and FACS purification of islet cells from fetal and neonatal pigs. To measure gene expression and functional changes in the developing pig pancreas, we established a reliable infrastructure to procure pancreata from fetal or neonatal pigs using rigorous criteria that optimized tissue quality and gene expression analysis (Fig. 1a). To investigate the crucial period of ontogeny culminating in functional islet cells, we systematically isolated pancreata from gestational days (d) 40, 70, 90 and post-natal days (P) 8 and 22 (Fig. 1a). We surveyed sorting methods based on cell surface epitopes, like in prior work with mouse and human cells[7,24,25], but have not yet achieved reliable cell fractionation based with this strategy (R. Whitener and S.K. Kim, unpubl. results). Instead, we used intracellular sorting with specific antibodies against insulin-, glucagon- and somatostatin- to isolate β-, α- and δ-cells, respectively, from fetal or neonatal pancreata (Supplementary Fig. 1a), like in prior work on human islets[7,17,26]. Quantitative RT-PCR (qRT-PCR; Fig. 1b) confirmed enrichment or depletion of expected marker genes for β- (*INS*), α- (*GCG*), δ- (*SST*), ductal (*KRT19*) and acinar cells (*AMY2*).

To obtain comprehensive gene expression profiles, we performed RNA-Seq on 37 libraries generated from fetal and neonatal cells (Supplementary Data 1). An average of 7 million reads per sample library uniquely mapped to a current pig reference transcriptome (Sscrofa11.1). Unsupervised hierarchical clustering of transcriptome data with Pearson correlation analysis revealed distinct expression profiles that changed with advancing developmental stage within specific cell types (Supplementary Fig. 1b). Principal component analysis with these samples showed that transcriptomes clustered in the first principal component according to cell type (β-, α- or δ-cells: PC1, 35% of variance), and according to advancing developmental stage in the second principal component (PC2, 25% of variance: Supplementary Fig. 1c). Samples from 'late fetal' stages d70 and d90 clustered closely and separately from d40 samples, and were grouped together for subsequent analysis. In summary, our cell purification strategy generated high-quality gene expression profiles of developing pig β-, α- and δ-cells.

High-depth transcriptome maps in pig β-, α- and δ-cells. To evaluate stage-specific differential gene expression, we analyzed transcriptome data in early fetal (d40), late fetal (d70 and 90) and neonatal stages (P8-P22) (Figs. 1c-h). We used the DESeq2 algorithm[27] to find differentially expressed genes (DEGs) encoding transcripts whose level changed at least 1.5-fold (adjusted $P<0.05$) between these stages in β-, α- or δ-cells. Between early and late fetal stages, we identified 2696 DEGs in β-cells, with transcripts from 1270 genes increased and 1426 decreased. Between late fetal and neonatal stages, we identified 1463 DEGs: 785 genes increased, and 678

genes decreased (Fig. 1c and Supplementary Data 2). In α-cells between early and late fetal stages, we identified 2823 DEGs with transcripts from 1311 genes increased and 1512 genes decreased. Between late fetal and neonatal stages we identified 2291 DEGs: 1200 genes increased, and 1091 genes decreased (Fig. 1e and Supplementary Data 2). In δ-cells between late fetal and neonatal stages, we identified 1515 DEGs with transcripts from 866 genes increased, and 649 genes decreased (Fig. 1g and Supplementary Data 2). Among DEGs, we observed subsets of genes enriched in either fetal or neonatal stages in β-, α- or δ-cells, confirming the presence of regulated developmental gene expression programs in specific pig islet cell types. Transcript abundance data from β-, α- and δ-cells are also provided in a searchable dataset (Supplementary Data 3).

To evaluate the relevance of differential gene expression, and effects on signaling pathways at specific stages, we performed gene ontology (GO) term analysis for biological processes with Clusterprofiler[28]. In the transition from the early to the late fetal stages, we observed enrichment of specific terms linked to islet β- or α-cell development, like glucose homeostasis, endocrine pancreas development, store-operated calcium entry, response to glucose, and regulation of hormone secretion (Figs. 1d, f). In the transition from late fetal and neonatal stages in β-, α- or δ-cells, we observed enrichment of biological processes distinct from those at earlier developmental stages. These included MAP kinase signaling, response to metal ion, hormone secretion, cAMP metabolic process, and Rho protein signaling transduction (Figs. 1d, f, h).

Pig β-, α- and δ-cells had cell type-specific gene expression similar to that in human cells (Figs. 2a-g; ref. 7,29,30). For example, transcripts encoding the factors PDX1, MAFA and NKX6.1 were exclusively expressed or highly enriched in pig β-cells throughout development, while genes encoding ARX and IRX2 were exclusively expressed in pig α-cells (Figs. 2a, b and Supplementary Data 3). In addition, consistent with previous findings[31,32], mRNAs encoding HHEX, RBP4 and GCGR (Figs. 2a, c), and CRHR2 (Supplementary Data 3) were enriched in δ-cells, while mRNAs encoding GLP1R and GCGR were enriched both in β- and δ-cells (Fig. 2a). Like in humans[7,29,32], *MAFB* and *UCN3* were expressed in both β- and α-cells (Fig. 2d). Moreover, *MAFB* expression increased during late fetal stages, while *MAFA* expression increased continuously from fetal to post-natal stages, like in humans[7,29] (Figs. 2a, d). Thus, in addition to similarities of *cell type-specific* gene expression between pig and human islet β-, α- and δ-cells, our studies revealed conserved *stage-specific* gene expression dynamics.

Molecular regulators of β- and α-cell proliferation and functional maturation. Transcription factors like PDX1, MAFA, MAFB, NKX6.1 and ARX regulate expression of genes in mouse and human islet cells governing β- and α-cell differentiation[33-37]. These and other factors also govern expression of effector genes regulating islet cell proliferation, expansion and functional maturation. We observed dynamic expression of these in developing pig β- and α-cells (Fig. 2e-g). For example, *MKI67*, *BUB1B* and *NUSAP1*, which encode products known to regulate mitosis and cell cycle progression, declined as β- and α-cell development advanced (Fig. 2e), consistent with other measures of proliferation (see below).

Likewise, we found that expression of 'disallowed' genes like *COX5A, SLC16A1*, *MGST3* and *LDHA*, originally identified in β-cells and thought to restrain β-cell function and maturation[38-40], declined during development (Fig. 2f). We also detected dynamic or cell type-specific expression of genes encoding established effectors of islet hormone secretion, like the proprotein convertases PCSK2, the ion transporter SLC30A8, and the glucose transporter SLC2A2 (Fig. 2g). These data delineate gene expression changes at genome-scale that orchestrate age-dependent changes in proliferation and maturation of islet β-, α- and δ-cells.

We also identified unanticipated gene expression dynamics in developing β- and α-cells, not previously noted in human or mouse studies (Figs. 2h-j). For example, *PCSK1* encodes an endopeptidase normally expressed in intestinal L-cells to regulate GLP-1 production[41,42], but not in early fetal or adult α-cells (Fig. 2h). However, we observed a transient increase of *PCSK1* mRNA in late fetal and neonatal α-cells (Fig. 2h); immunostaining confirmed that a subset of neonatal α-cells produced PCSK1 protein (Fig. 2i). Likewise, *DPP4* encodes a protease that inactivates GLP-1 and whose expression is extinguished in postnatal human β-cells[7,29] (Fig. 2j). However, we observed a significant *increase* of *DPP4* mRNA levels in β-cells at late fetal stages, when β-cells also express *GLP1R* (Fig. 2j and Supplementary Data 3). Together, these findings raise the unexpected possibility that fetal α-cell GLP-1 production and regulated signaling to fetal β-cells could impact islet cell development.

Stage-specific expression of maturation factors during development. To assess gene expression in developing islets and validate our RNA-Seq findings, we performed immunostaining of pig pancreas at specific developmental stages. At fetal stages, islet β-, α- and δ-cells formed small clusters, which enlarged subsequently in neonatal and adult pancreas (Figs. 3a-c). Thus, like in other species[18,43,44], islet cells formed clusters at the inception of their ontogeny, then continued their morphogenesis in multicellular structures through post-natal stages. In fetal,

neonatal and adult islets, standard morphometry revealed that allocation of islets into specific lineages was dynamic (Supplementary Fig. 2a). In early fetal pancreas, the number of α-cells is slightly higher than that of β-cells, while the δ-cell fraction was very low (<4%). However, from late fetal stages to the neonatal stages, the number of β-cells is higher than that of α-cells, while the number of δ-cells is similar to that of α-cells. Consistent with our RNA-Seq data, antibodies that detect the transcription factors (TFs) PDX1, NKX6.1 and MAFB showed restricted TF production in specific islet cell subsets (Figs. 3d-f). In β-cells, we detected nuclear PDX1, NKX6.1 and MAFB. In contrast to mice, human β-cells express the transcription factor SIX3 in an age-dependent manner; moreover human α-cells do not express *SIX3*[7]. qPCR (Fig. 3g) of pig islets revealed adult-specific production of *SIX3*, and immunohistology (Fig. 3h) demonstrated SIX3 production specifically in a subset of β-cells. Thus, stage-specific and cell-specific expression of multiple TFs in pig islets resembled that observed previously in human islet cell subsets.

We also quantified the nuclear factor Ki67, a marker of cell proliferation encoded by *MKI67* in multiple species, including pigs, humans and mice (Supplementary Fig. 2b; ref. 7,45-47). In pigs, pancreas immunostaining revealed a high percentage (10%) of Ki67^{+} early fetal islet cells (d40), that declined to 2% by late fetal and neonatal stages, then further declined to <0.1% by adulthood (Supplementary Fig. 2b). These changes were mirrored by mRNA levels encoding *MKI67* (Fig. 2e). Thus, we observed evidence of vigorous fetal-stage proliferation of islet cells that, like in mice and humans, declined further after birth.

Similarities of developmental gene expression in pig and human β-cells and α-cells. To compare gene expression between pig, human and mouse β-cells, we performed unsupervised hierarchical clustering using distance matrix analysis with our RNA-Seq data, and extant human and mouse data[7,29,48]. This analysis revealed that neonatal β-cell transcriptomes in pig and human were globally more similar than to the β-cell transcriptome of mice (Fig. 4a). The absence of sufficient fetal or neonatal mouse α- or δ-cell transcriptome data precluded analogous global comparisons of human, pig and mouse α-cells or δ-cells. Next, we performed pairwise comparison of pig and human[29] cells from fetal and postnatal stages (Figs. 4b-g and Supplementary Data 4). This analysis revealed conserved cell type-specific and stage-specific regulation of multiple factors. For example, *GLP1R* encodes a G protein-coupled receptor for the incretin hormone GLP-1. Like in humans[49], *GLP1R* is expressed in pig β-cells but not α-cells and this expression increases from fetal to neonatal stages (Fig. 4b). By contrast the G protein coupled receptor encoded by *GPR87* is expressed in pig and human α-cells but not β-cells and has increasing

expression during postnatal islet maturation (Fig. 4c). Similarities in cell type-specificity or stage-specific expression for other regulators, including *DNMT1*, *PPID*, *KLF9*, and *DCHS1,* provide further evidence of conserved genetic programs underlying pig and human β- and α-cell development (Figs. 4d-g).

In human diabetes, genetic risk may reflect monogenic or polygenic mechanisms[3,50-52]. Genome wide association studies (GWAS) in humans have identified candidate 'risk genes' associated with diabetes and related metabolic traits, such as fasting glucose impairment, altered fasting insulin levels. However, little is known about these candidate genes, including their regulation by risk-associated single nucleotide polymorphisms (SNPs), by cell-specific factors, or by physiological roles. We compared causal or candidate risk genes linked by prior studies to T1D, T2D, neonatal diabetes (NDM), maturity onset diabetes of the young (MODY), or relevant metabolic traits to genes whose developmental expression changed in pig β- or α-cell development. We identified 153 causal or diabetes risk genes whose expression changed at least 1.5-fold in β- or α-cell development (Supplementary Data 5: 58 of 153 genes shown in Fig. 4h). Overall, this linkage was significant (P=0.001) compared to genes without dynamic developmental transcript changes. These genes included those whose developmental expression changed (Fig. 4i) exclusively in β-cells (*G6PC2*, *EFR3B, RAPGEF4*), in α-cells (*ADAMTS9*), or in both cell types (*IGF2BP2*).

Dynamic gene regulation during pig β- and α-cell development. Our studies of pig islet development provided an opportunity to assess gene expression dynamics throughout fetal and postnatal stages. Between early and late fetal stages, we identified 3111 and 3668 genes with altered expression from early fetal stage to neonatal day 22 in β- and α-cells, respectively (Figs. 5a, b and Supplementary Data 6). Thus, our analysis permitted grouping of genes in 'clusters', based on the sequence of changes observed between these three discrete developmental periods. For example, we observed genes in β- and α-cells that increased through late fetal stages, then did not change thereafter (cluster 2, 'Up_NC') and other genes whose transcript levels declined through late fetal stages then did not change thereafter (cluster 7, 'Down_NC'). We grouped genes in 8 clusters (Up_Down, Down_Up, etc.), and GO term analysis of each cluster revealed dynamic biological processes (Figs. 5a, b and Supplementary Figs. 3a, b). For example, we observed that the β- and α-cell "Down_NC" clusters were significantly enriched for genes that regulate cell proliferation and cell cycle progression (like *CDK2, CDK14* and *CDC20*; Figs. 5a, b and Supplementary Data 6). The "Up_NC" cluster in both cell types was enriched for genes

associated with multiple signaling pathways involved in glucose and steroid hormone signaling (Supplementary Figs. 3a, b). Thus, the data identified multiple biological processes involved in islet differentiation and maturation that are dynamically regulated in development.

The frequency of transcript measurement in our workflow also identified unanticipated gene expression dynamics in developing β- and α-cells, particularly in gene clusters with "Up_Down" or "Down_Up" trajectories (clusters 3 and 6: Figs. 5c, d). For example, from fetal to neonatal stages in islet β-cells, we observed increased levels of transcript encoding pyruvate dehydrogenase kinase 4 (PDK4), a factor previously characterized[40] as 'disallowed' in β-cells (Fig. 5c). Consistent with this, we observed that expression of *PPARGC1A*, a known positive regulator of *PDK4* transcription[53] also increased at this stage in β-cell development ("NC_Up" cluster; Supplementary Data 6). Thus, unlike other disallowed genes whose expression continuously declined in developing β- and α-cells, *PDK4* expression appeared to increase after late fetal stages, raising the possibility of as-yet unidentified functions for PDK4 in pig β-cell development.

Genetic regulation of β-cell functional maturation. Age-dependent increases of insulin secretion have been described in mice and humans[7,33,54-56], but have not been reported in pigs. We compared glucose-stimulated insulin secretion in islets isolated from fetal (d90), newborn (P2), and weaning-age donors (P22). In d90 islet clusters after 1 hr in basal (2.8mM) or elevated (20mM) glucose, only 0.02 to 0.04% of total insulin content was secreted. By contrast, in similar conditions we observed that 0.5-2% of total insulin content was secreted from P2 and P22 islets (Fig. 6a). Moreover, insulin released from P22 islets exposed to 20mM glucose + IBMX (a cAMP-increasing agent), was significantly greater than insulin levels from islets at 2.8 mM and 20 mM glucose, while this IBMX effect was not observed in P2 islets (Fig. 6a). Thus, like in humans [7] and other species, pig islet insulin secretion matured from fetal to post-natal stages.

To assess the genetic basis of this β-cell functional maturation, we compared transcriptomes from d90 and P22 β-cells (Fig. 6b). Unsupervised hierarchical clustering and principal component analysis revealed that d90 and P22 pig β-cells had distinct gene expression (Supplementary Figs. 4a, b). DESeq2 analysis revealed significantly increased expression of 1337 genes and decreased expression of 734 genes between d90 and P22 (Supplementary Data 7). For example, transcripts of genes encoding critical regulators of human insulin secretion like *HCN4*, *CACNA1A* and *GLIS3* were up-regulated whereas disallowed or fetal β-cells enriched genes[2,40,57] like *RND3*, *GSTA4*, and *RCOR2* were down-regulated in P22 pig β-cells (Figs 6b, d). GO term analysis revealed that up-regulated transcripts were encoded by genes associated with pathways like

regulation of peptide secretion, and other signaling pathways corresponding to β-cell maturation (Figs. 6e, f and Supplementary Data 8). Thus, combined developmental, functional and molecular investigations here revealed molecular and genetic factors governing maturation of neonatal β-cells and unveiled conserved mechanisms underlying islet cell development and function in pigs and humans. We also observed that, compared to d90, 677 genes were up- or down-regulated in P22 α-cells (Supplementary Figs. 4c-e and Supplementary Data 7). These included genes and pathways known to be enriched in α-cells, or related to α-cell differentiation (Supplementary Figs, 4f, g; ref. 31,58). Further studies of glucagon secretion at specific stages are required to assess the functional significance of these changes in α-cell development.

Discussion

Based on the crucial role of pancreatic islets in human diseases, and the expanding catalog of differences between human and mouse islet developmental biology or regulation[1,7,59,60], there is intense interest in developing complementary experimental systems for investigating islet development and maturation. To address this need, we used an integrated approach to delineate pancreas development in pigs. This multimodal assessment of genetic and developmental phenotypes across a comprehensive range of fetal and post-natal stages of pig pancreas ontogeny and maturation revealed multiple similarities in formation and regulation of pig and human islet β-, α- and δ-cells, including features not observed in mice. Systematic phenotyping of pig pancreas development – particularly gene expression and identification of developmental signaling pathways – revealed unexpected dynamic gene regulation, and links to human diabetes genetics. Moreover, studies in pigs are cost-efficient (see Methods). Our experimental approaches and findings provide a unique roadmap for diabetes and metabolic research, including attempts to direct differentiation and maturation of replacement islet cell types from renewable human cell sources.

Here, we established cell purification strategies based on flow cytometry to investigate developmental genetics in fetal and neonatal pig β-, α- and δ-cells; an approach we and others have successfully used in pancreas from mice and humans[7,26,29,48,61]. Like in prior studies[7,26], we isolated primary β-, α- and δ-cells throughout fetal and postnatal pancreas development using antibodies against hormones specific to these cell types, an approach permitting comprehensive, high-quality stage-specific transcriptome analysis. Similar approaches for primary ductal and acinar cells should provide opportunities in the future to delineate the transcriptome of these important components of the exocrine pancreas. Stage-specific gene expression patterns

revealed that pig and human β-cells and α-cells shared characteristic molecular and developmental features not observed in mouse β-cells or α-cells. For example, clustering analysis revealed that global human β-cell gene expression was more similar to that of β-cells from pigs than mice. Like in human islets[58], the transcription factor MAFB and the hormone UCN3 are expressed in both pig β- and α-cells (Fig 2). Similarly, we found that the transcription factor SIX3 is expressed in pig β-cells in an age-dependent manner, like in human islets[7] (Fig 3). Our studies also provide a reference transcriptome analysis of developing islet δ-cells, whose recognized role in health and diabetes is growing[62]. Together, the resources and methods described here provide evidence that developmental genetic studies of pancreas in pigs could complement similar studies in humans and mice. Discovery of flow cytometry-based methods for purifying live primary endocrine or exocrine cells, with antibodies that recognize cell-surface epitopes, should expand our capacity to investigate other important aspects of pancreatic gene regulation and function, using approaches like ATAC-Seq, promoter capture Hi-C, and Patch-Seq[24,63,64].

Transcriptome analysis described here provides evidence of genetic regulatory mechanisms and cell development in pigs that are also conserved in other mammals, including (1) cell type-specific expression of transcription factors (*PDX1, MAFA, ARX, IRX, HHEX*) and functional regulators (*PCSK1, PCSK2, SLC2A2, SLC30A8, GLP1R, GCGR*), (2) transient β-, α- and δ-cell proliferation, peaking in fetal/perinatal stages and regulated by conserved factors like *MKI67, BUB1B* and *NUSAP1*[45-47], (3) dynamic expression of histone or chromatin regulators like *EZH2* and *DNMT1*[49,65] and (4) reduction or elimination of transcripts encoding 'disallowed' genes[40,57,66,67]. These findings from transcriptome analysis also correspond well with observations from our other stage-specific tissue phenotyping (see below). To facilitate assessment of transcriptome data on these or other genes of interest, we will provide online browser tools.

In addition to the resemblance of pancreatic gene regulation in pigs and humans, we observed similarities of islet morphogenesis, cell specification, regulation and islet cell allocation. Scattered islet β-, α- and δ-cells coalesced into small clusters at late fetal stages, which contained proliferating cells (Fig. 3). At birth, the proportion of α- and δ-cells in pig islets was each ≥20%, like that of neonatal human islets[59] (M. Brissova, A. Powers, S.K.K., unpubl. results), and unlike in mice. After birth, these clusters enlarged, and the geometry of β-, α- and δ-cells in islets appeared intermingled, like in human islets[59]. Studies of insulin secretion further demonstrated that β-cell function matures from fetal to post-natal stages like in humans[7] and other mammals. Thus, our study reveals and quantifies morphological, cellular and functional features in pig islet development that are conserved in human pancreas, suggesting that pigs could be a reliable

surrogate animal model to study the islet function of human gene orthologs. This includes studies of monogenic, dominant forms of diabetes like maturity onset of diabetes in the young (MODY) and neonatal diabetes mellitus (NDM), which have not been reconstituted in mice[3,4]. By contrast, dominant monogenic diabetes resulting from mutations in *INS* and the MODY3 gene *HNF1α* has been reconstituted in pigs[10,68]. Thus, studies of gene function and pathogenesis in diseases like MODY and NDM could be advanced through investigations in pigs.

Here, we evaluated pancreas phenotypes through a full range of fetal and peri-natal stages. In β-, α- and δ-cells, this 'granularity' of phenotyping revealed changes in >6000 genes during development, and identified developmental dynamics not previously noted in studies of human islet gene expression with comparatively limited sampling during developmental stages[7,29,69], or in one study, a single stage[70]. For example, in β-cells, genes involved in glucose processing, including *G6PC2* and *PDK3*, are in the "Up_NC" cluster whereas genes regulating insulin secretion, such as *SLC2A2*, are in the "NC_Up" cluster (Supplementary Data 6), indicating that mechanisms governing glucose sensing and processing or insulin secretion are established at different developmental stages in pig β-cells. Moreover, we noted 366 β-cell or α-cell transcripts whose levels changed with a 'discontinuous' trajectory in development (Fig. 5 and Supplementary Data 6). That is, we observed heterogeneous patterns of expression, including early rise then fall (like *DPP4, FEV* in β-cells, and *FOXO1* or *KIT* in α-cells), or initial decline followed by a later rise in transcript levels, including those encoding a subset of disallowed gene products like PDK4 in β-cells (Fig. 5 and Supplementary Data 6). Thus, our ability to detect dynamic gene expression trajectories in β-cell or α-cells was enhanced by the comprehensiveness of developmental phenotyping afforded in pigs.

Cell- and stage-specific phenotyping here also revealed links between our data to human diabetes risk and islet replacement. Prior studies have noted a significant enrichment of candidate diabetes risk genes among those with dynamic developmental expression[7,71]. In our recent work, prioritization of candidate diabetes risk genes with dynamic fetal or post-natal expression led to identification of *SIX3* and *BCL11A* as unrecognized regulators of human β-cell development and function[7,17]. Here, we detected a significant enrichment of candidate diabetes risk genes (both T1D and T2D) or risk genes for diabetes-related traits in the subset of genes with changing expression in β- or α-cells (Fig. 4 and Supplementary Data 5). This included a set of genes with known or postulated functions in monogenic and polygenic forms of diabetes, and in immune cell signaling[63]. Thus, our datasets provide a foundation for systematic investigations of *cis*- and *trans*-regulatory elements governing the dynamic expression of diabetes risk genes in the native,

physiological setting of pig islet development. Additionally, we found significant enrichment of β- and α-cell gene sets involved in specific developmental signaling pathways. However only a subset of these pathways, like mTOR, retinoic acid receptor, glutamate receptor and TGF-β signaling, have been previously linked to β- and α-cell biology[72-75]. Thus, further studies of our data could reveal unrecognized conserved regulators of islet development, function and fate.

Identifying signals controlling islet development, including detection of signaling modulation, should enhance efforts to control and direct differentiation of functional replacement islet cells from stem cells and other renewable sources. Recent studies have explored use of neonatal pig islet cultures for identifying signals that promote β-cell functional maturation[21]. In their work with stepwise directed differentiation of human stem cells into distinct lineages, Loh and colleagues have demonstrated the value of inhibiting signaling pathways – not merely through signal withdrawal, but by exposure of cells to potent signaling inhibitors – to prevent development along undesired fates or lineages[76,77]. The identification of dynamically regulated genes and pathways, including those with discontinuous trajectories (Up then Down, etc.), in β-cell and α-cell development provides a new roadmap for modulating islet cell replacement efforts. Fetal and neonatal β-cells and α-cells develop in close proximity, and our work likely identifies intercellular signaling pathways (like those regulated by *PCSK1*, *GLP1R* and *DPP4*) that could regulate differentiation of these cells. Thus, our findings identify multiple pathways associated with fetal and post-natal β-cell maturation (Supplementary Data 8), including many not directly modulated in prior studies[78-81]. In summary, data, conceptual advances and resources detailed here provide an unprecedented molecular, cellular and developmental framework for investigating pancreas development that could inform our understanding of diabetes and obesity risk, and advance efforts to produce replacement human islets for diabetes.

Methods

Animals and pancreas harvest. Pigs used in this study were outbred and used in accordance to protocols approved by the Institutional Animal Care and Use Committees at University of California Davis and Stanford University. For collection of fetal pig pancreata, pregnancies were confirmed via ultrasound 30 days after mating and pregnant gilts were humanely euthanized at 40, 70, or 90 days. The reproductive tracts were removed after mid-ventral laparotomy, cooled on ice, and fetuses were collected. Fetuses remained on ice until dissection. Pancreas was identified according to anatomical location and harvested under a stereoscopic microscope. Neonatal pancreata were harvested from neonatal piglets at 2, 8, and 22 days of age, while adult pancreata were collected from pigs 6 months of age or older. Briefly, after euthanasia, a midline incision was performed to provide access to the peritoneum and thoracic cavity. The pancreas was rapidly chilled via either perfusion of sterile ice-cold saline solution into the descending aorta or by pouring the same solution directly into the peritoneal cavity and then surgically removed. Harvested pancreatic tissue was maintained at a temperature of 2-10°C in saline solution until further downstream processing.

Generation of pancreatic single cell suspensions. Morphological differences of the pancreas between developmental stages necessitated customized protocols to generate single cell suspensions suitable for downstream assays. Embryonic day 40 pancreata were diced finely with a razor blade and transferred into a 4°C solution of Collagenase D (Millipore-Sigma, Cat# C9263) at a concentration of 1 mg/ml in HBSS, and digested to single cells overnight at 4°C. Embryonic day 70 and 90 pancreata were finely diced with a razor blade and digested for 5-8 minutes at 37°C using the same Collagenase D solution above. The reaction was stopped with the addition of equal volume HBSS containing 10% porcine serum (Millipore-Sigma, Cat# P9783). If necessary, an additional digestion using TrypLE (Gibco, Cat#12605-010) for 3 minutes at 37°C was used to fully disperse any remaining cell clusters. Post-natal pancreata were digested following published methods[82]. Briefly, pancreata were trimmed to remove lymphatic, connective, and fat tissues, and chopped into 1-5 mm pieces. The tissue was digested using 2 mg/ml Collagenase D in HBSS at 37°C for 12-18 minutes. 1.5 volumes of HBSS with 10% porcine serum and 0.5% Antibiotic/Antimycotic solution was added to stop digestion. Neonatal Islet Cell clusters (NICCs) of the desired size were obtained after straining digest through a 500 μm mesh (PluriSelect, Cat# 43-50500-03) to remove large undigested tissue, washed, and allowed to gravity settle two times. To generate single cell suspensions, some NICCs were further dispersed with TrypLE for 3 minutes at 37°C. Remaining NICCs were spun down and resuspended in RPMI

1640 (Gibco, Cat# 11879-020) containing 5 mM glucose, 10% porcine serum, 10 mM HEPES (Caisson Labs, HOL06-6X100ML) and 1% Pen/Strep (Life Technologies, Cat# 15140-122) and placed in a humidified, 5% CO_2 tissue culture incubator.

The cost basis of pancreas procurement and islet isolation from pigs is currently lower than for humans. In 2019, reimbursement for a single human cadaveric donor pancreas is between $4,500 and $7,000 (US dollars). For human islets, investigators are charged $0.12 per islet equivalent using the NIH-supported Integrated Islet Distribution Program (https://iidp.coh.org). By contrast, the cost of a P22 pig pancreas is approximately $500, and after processing, a P22 pancreas yields approximately 90- to 100,000 NICCs. The total cost per pig NICC is <$0.008.

Cell flow cytometry and purification. Single cell suspensions were washed twice with cold PBS and filtered with a 70 µm filter (BD Falcon, Cat# 352350). Prior to fixation, cells were stained with LIVE/DEAD fixable Aqua (ThermoFisher, Cat# L34976) or LIVE/DEAD fixable Near-IR (ThermoFisher, Cat#L10119) dead cell stains according to manufacturer's protocol. Cells were fixed with 4% PFA in PBS (Electron Microscopy Sciences) for 10 minutes at room temperature. After fixation, cells were permeabilized with Permeabilization Buffer: 1X PBS containing permeabilization solution (BioLegend, Cat# 421002), 0.2% BSA (Sigma-Aldrich, Cat#A9576), mouse and rat IgG (Jackson Labs, Cat# 015-000-003 and 012-000-003), and 1:40 Ribolock RNase Inhibitor (Fisher Scientific, Cat# FEREO0384), for 30 minutes at 4°C. Cells were pelleted and resuspended with Wash Buffer: 1X PBS containing permeabilization solution, 0.2% BSA, and 1:100. Cells were pelleted again, supernatant removed, and stained with Alexa Fluor 488 labeled anti-insulin (R&D Systems, IC1417G-100UG), Alexa Fluor 647 labeled anti-glucagon (Santa Cruz, sc-57171 AF647) and PE labeled anti-somatostatin (Santa Cruz, sc-55565 PE) antibodies in Staining Buffer: 1X PBS containing permeabilization solution, 1% BSA, and 1:40 Ribolock, for 30 minutes at 4°C. Cells were washed three times in Wash Buffer and resuspended in Sort Buffer: 1X PBS containing 0.5% BSA, and 1:20 Ribolock. Labeled cells were sorted on a special-order, 5-laser FACS Aria II using FACSDiva software (BD Biosciences). After excluding doublets and dead cells, desired cell populations were identified using fluorescence minus one (FMO) controls. Cells were sorted using the purity mask in chilled low retention tubes containing 50 µl of Sorting Buffer. Post-hoc visualization of recorded data was performed using FlowJo vX software (FlowJo, LLC).

Total RNA extraction and quantitative Reverse Transcription PCR (qRT-PCR). Total RNA was extracted from the sorted cells as described[26] and concentrated using the RNA Clean &

Concentrator-5 kit (Zymo Research, Cat# R1013). Complementary DNA (cDNA) was generated using the Maxima First Strand cDNA Synthesis kit (Thermo Scientific, Cat# K1642) according to the manufacturer's instructions. Purity of sorted populations was confirmed via Quantitative PCR using an Applied Biosystems 7500 Real-Time PCR System. The following TaqMan probes were used: INS (Ss03386682_u1), GCG (Ss03384069_u1), SST (Ss03391856_m1), AMY2 (Ss03394345_m1) ACTB (Ss03376081_u1) and KRT19 (Forward: 5'-GAAGAGCTGGCCTACCTG-3', Reverse: 5'-ATA CTGGCTTCTCATGTCGC-3'). Relative mRNA abundance was calculated by the Comparative Ct (ΔΔCt) relative quantitation method.

Bulk RNA-Seq Library preparation. Multiplexed libraries were prepared using the SMARTer Stranded Total RNA-Seq Kit v2 - Pico Input Mammalian (TakaraBio, Cat# 634413) following manufacturer's instructions. Library fragments of approximately 350 bp were obtained, and the quality was assessed using a 2100 Bioanalyzer (Agilent). Barcoded libraries were multiplexed and sequenced on an Illumina NextSeq sequencer.

Bioinformatic and statistical analysis of RNA-Seq datasets. RNA-Seq libraries were sequenced as 75 bp pair-end reads to a depth of 30 to 60 million reads. To estimate transcriptome abundance, Salmon algorithm (v0.11.3) was used[83]. An index was built on the Sus Scrofa reference transcriptome (Sscrofa 11.1, version 96)[84] using Salmon with parameters "salmon index --keepDuplicates -t transcripts -i transcripts_index --type quasi -k 31". The reads were aligned to the index and quantified using Salmon with parameters "salmon quant -i $salmon_index -l A \-1 $Input1 \-2 $Input2 \ p 8 -o quants/${Input1}_quant \ --gcBias –seqBias". After getting the transcript counts, differentially expressed gene analysis was performed using the DEseq2 analysis pipeline[27,85]. Differentially expressed genes (DEGs) were obtained by comparisons of any two sequential stages (i.e. Early fetal versus Late fetal and Late fetal versus Neonatal) in a cell type specific manner. Genes were considered differentially expressed if the fold change between samples was greater than or equal to 1.5 with setting of Benjamini-Hochberg (BH) false discovery rate (FDR) adjusted P value of <0.05 in the package. Hierarchical clustering of samples were done with the cor() functions of R. Principal component analysis (PCA) was performed with the function *plotPCA* that comes with DEseq2 package. Visualization of RNA-seq data examined in this study were produced with pheatmap 1.0.12[86] and ggplot2[87]. GO term analysis was carried out using the R package clusterProfiler as described[28] with a Benjamini-Hochberg adjusted P value of <0.1.

GWAS, MODY and NDM genes data Comparison with pig RNA-seq data. Lists of genes were obtained from the NHGRI-EBI Catalog (https://www.ebi.ac.uk/gwas; All associations v1.0.2, data release date of 2019-05-03)[88] for Type 1 diabetes, Type 2 diabetes, fasting glucose levels, glycated hemoglobin levels glycemic traits, and proinsulin levels categories. A List of MODY and NDM genes were generated from [89]. Orthologs between human and pig were identified. Enrichment of genes associated with listed traits in among DEGs was tested using the Chi-square test.

Comparison to extant human and mouse beta cell transcriptomes. Human[7,29] (GEO accession numbers GSE67543 and GSE79469) and mouse[48] (GSE87375) datasets were obtained from the NCBI Gene Expression Omnibus (GEO). Indices for human and mouse libraries were built from reference transcriptomes GRCh38 cDNA and GRCm38 cDNA, respectively (Homo sapiens, version 96; Mus musculus, version 96; ref.[84]). Pig and mouse Ensembl IDs were converted to their corresponding human Ensembl ID using predicted ortholog tables across these species obtained from Ensembl BioMart[84]. Sample-wise hierarchical clustering analysis (Fig. 3a) was performed using dist() function of R.

Generation of dynamic gene lists. To create lists of dynamic genes, DEGs from comparisons between early fetal and late fetal stages and between late fetal and neonatal stage were intersected, which created lists of genes showing patterns of gene expression with up, down or no change directions as development proceeds in islets development. Gene TPM values shifted by +1, log based 2 transformed were averaged across biological replicates and subsequently standardized to their Z-scores.

Immunohistochemistry. Biopsies of harvested pancreata were fixed in 4% paraformaldehyde (w/v) in PBS at 4°C overnight, washed twice in PBS, and cryoprotected via sucrose saturation. The tissues were imbedded in Tissue-Tek OCT compound (VWR, Cat# 25608-930) and stored at -80°C until sectioning. 10μm thick sections were collected on ThermoFisher Superfrost Plus slides. For staining, frozen slides were thawed at room temperature for 30 min and washed with PBS. Antigen retrieval was performed by boiling slides immersed in Target Retrieval Solution (DAKO, S169984-2) for 30 minutes, as needed. Subsequently, sections were incubated with streptavidin/biotin blocking solution (Vector Labs, SP-2002) according to manufacturer's instruction. Then, sections were incubated in permeabilization/blocking buffer (5% normal donkey serum, 0.3% Triton X-100 in PBS) at room temperature for 1 hour followed by overnight incubation

at 4°C with primary antibodies diluted in Antibody dilution buffer (1% bovine serum albumin, 0.3% Triton X-100 in PBS). Slides were washed and stained with appropriate secondary antibodies for 1 hour at room temperature. Primary antibodies and dilutions used are as follows: guinea pig anti-insulin (DAKO, A0564, 1:1000), mouse anti-glucagon (Sigma-Aldrich, G2654, 1:1000), rabbit anti-SIX3 (LifeSpan Biosciences, LS-B9336, 1:100), goat anti-PDX1 (R&D systems, AF2419, 1:200), mouse anti-human NKX6.1 (R&D systems, AF5857, 1:200,) rabbit anti-MAFB (Bethyl, IHC-00351, 1:200), rabbit anti-MAFA (AVIVA systems biology, ARP47760_P050, 1:200), goat anti-somatostatin (Santa Cruz biotechnology, 1:500), rabbit anti-Chromogranin A (Immunostar, 20085, 1:500), rat anti-E-cadherin (Life Technologies, 13-1900 1:500). Secondary antibodies used are donkey anti-isotype conjugated with Alexa Fluor 488/555/647 (1:500; Molecular Probes). After a final wash with PBS, slides were preserved with a mounting medium containing DAPI (Vector labs, H-1200) and coverslipped. Image acquisition was performed on a Leica SP8 confocal microscope. Islet cells were quantified by counting insulin positive β-cells, glucagon positive α-cells, and somatostatin positive δ-cells using pancreas sections from different stages.

In vitro insulin secretion assay. Functional maturity of NICCs isolated from 2-day or 22-day old piglets was assessed by in vitro glucose stimulated insulin secretion (GSIS) assay performed after 5 days of culture. In brief, 100 NICCs were washed and incubated at 2.8 mM glucose for 1 hour. After the initial preincubation, clusters were serially incubated in media above with containing 2.8 mM, 20 mM, and 20 mM plus IMBX glucose concentrations for 1 hour each. Finally, clusters were lysed in acid-ethanol solution to extract the total insulin content. Insulin from supernatants and the NICC lysates was quantified using the pig insulin ELISA kit (Mercodia, 10-12000-01). Secreted insulin is presented as a percent of total insulin detected in lysate. All steps were performed in triplicate at 37°C using RPMI 1640 supplemented with 2% porcine serum adjusted to the respective glucose concentrations.

Statistical analysis. For qPCR, GSIS and quantification of immunohistology experiments, the number of biological or technical replicates (n), measure of central tendency (e.g. average), error bars, and statistical analysis is detailed in the Fig. legend. Graphs and statistical analysis were produced and performed by using GraphPad Prism (version 8) software.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

References

1. Hart, N. J. & Powers, A. C. Use of human islets to understand islet biology and diabetes: progress, challenges and suggestions. *Diabetologia* **62**, 212-222, (2019).
2. Hrvatin, S. *et al.* Differentiated human stem cells resemble fetal, not adult, β cells. *Proceedings of the National Academy of Sciences* **111**, 3038, (2014).
3. Hattersley, A. T. & Patel, K. A. Precision diabetes: learning from monogenic diabetes. *Diabetologia* **60**, 769-777, (2017).
4. Maestro, M. A. *et al.* Distinct roles of HNF1beta, HNF1alpha, and HNF4alpha in regulating pancreas development, beta-cell function and growth. *Endocr Dev* **12**, 33-45, (2007).
5. McKnight, K. D., Wang, P. & Kim, S. K. Deconstructing pancreas development to reconstruct human islets from pluripotent stem cells. *Cell Stem Cell* **6**, 300-308, (2010).
6. Pan, F. C. & Brissova, M. Pancreas development in humans. *Current opinion in endocrinology, diabetes, and obesity* **21**, 77-82, (2014).
7. Arda, H. E. *et al.* Age-Dependent Pancreatic Gene Regulation Reveals Mechanisms Governing Human beta Cell Function. *Cell metabolism* **23**, 909-920, (2016).
8. Sneddon, J. B. *et al.* Stem Cell Therapies for Treating Diabetes: Progress and Remaining Challenges. *Cell Stem Cell* **22**, 810-823, (2018).
9. Dyson, M. C., Alloosh, M., Vuchetich, J. P., Mokelke, E. A. & Sturek, M. Components of metabolic syndrome and coronary artery disease in female Ossabaw swine fed excess atherogenic diet. *Comp Med* **56**, 35-45, (2006).
10. Renner, S. *et al.* Permanent neonatal diabetes in INS(C94Y) transgenic pigs. *Diabetes* **62**, 1505-1511, (2013).
11. Lim, H. *et al.* A novel autophagy enhancer as a therapeutic agent against metabolic syndrome and diabetes. *Nat Commun* **9**, 1438, (2018).
12. Matsunari, H. *et al.* Blastocyst complementation generates exogenic pancreas in vivo in apancreatic cloned pigs. *Proc Natl Acad Sci U S A* **110**, 4557-4562, (2013).
13. Wu, J. *et al.* Interspecies Chimerism with Mammalian Pluripotent Stem Cells. *Cell* **168**, 473-486 e415, (2017).
14. Kemter, E. *et al.* INS-eGFP transgenic pigs: a novel reporter system for studying maturation, growth and vascularisation of neonatal islet-like cell clusters. *Diabetologia* **60**, 1152-1156, (2017).
15. Sheets, T. P. *et al.* Targeted Mutation of NGN3 Gene Disrupts Pancreatic Endocrine Cell Development in Pigs. *Sci Rep* **8**, 3582, (2018).
16. Fisher, J. E. *et al.* In utero transplanted human hepatocytes allow postnatal engraftment of human hepatocytes in pigs. *Liver transplantation : official publication of the American Association for the Study of Liver Diseases and the International Liver Transplantation Society* **19**, 328-335, (2013).
17. Peiris, H. *et al.* Discovering human diabetes-risk gene function with genetics and physiological assays. *Nat Commun* **9**, 3855, (2018).
18. Carlsson, G. L., Heller, R. S., Serup, P. & Hyttel, P. Immunohistochemistry of pancreatic development in cattle and pig. *Anat Histol Embryol* **39**, 107-119, (2010).
19. Nagaya, M. *et al.* Distributions of endocrine cell clusters during porcine pancreatic development. *PLoS One* **14**, e0216254, (2019).
20. Ferrer, J. *et al.* Pig pancreas anatomy: implications for pancreas procurement, preservation, and islet isolation. *Transplantation* **86**, 1503-1510, (2008).
21. Hassouna, T., Seeberger, K. L., Salama, B. & Korbutt, G. S. Functional Maturation and In Vitro Differentiation of Neonatal Porcine Islet Grafts. *Transplantation* **102**, e413-e423, (2018).
22. Holst, J. J. *et al.* Insulin and Glucagon: Partners for Life. *Endocrinology* **158**, 696-701, (2017).
23. Lee, Y. H., Wang, M. Y., Yu, X. X. & Unger, R. H. Glucagon is the key factor in the development of diabetes. *Diabetologia* **59**, 1372-1375, (2016).

24. Arda, H. E. *et al.* A Chromatin Basis for Cell Lineage and Disease Risk in the Human Pancreas. *Cell Syst* **7**, 310-322 e314, (2018).
25. Sugiyama, T. *et al.* Reconstituting pancreas development from purified progenitor cells reveals genes essential for islet differentiation. *Proc Natl Acad Sci U S A* **110**, 12691-12696, (2013).
26. Hrvatin, S., Deng, F., O'Donnell, C. W., Gifford, D. K. & Melton, D. A. MARIS: method for analyzing RNA following intracellular sorting. *PLoS One* **9**, e89459, (2014).
27. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550, (2014).
28. Yu, G., Wang, L. G., Han, Y. & He, Q. Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284-287, (2012).
29. Blodgett, D. M. *et al.* Novel Observations From Next-Generation RNA Sequencing of Highly Purified Human Adult and Fetal Islet Cell Subsets. *Diabetes* **64**, 3172-3181, (2015).
30. Segerstolpe, A. *et al.* Single-Cell Transcriptome Profiling of Human Pancreatic Islets in Health and Type 2 Diabetes. *Cell metabolism* **24**, 593-607, (2016).
31. Muraro, M. J. *et al.* A Single-Cell Transcriptome Atlas of the Human Pancreas. *Cell Syst* **3**, 385-394 e383, (2016).
32. van der Meulen, T. *et al.* Urocortin3 mediates somatostatin-dependent negative feedback control of insulin secretion. *Nat Med* **21**, 769-776, (2015).
33. Aguayo-Mazzucato, C. *et al.* Mafa expression enhances glucose-responsive insulin secretion in neonatal rat beta cells. *Diabetologia* **54**, 583-593, (2011).
34. Artner, I. *et al.* MafB is required for islet beta cell maturation. *Proc Natl Acad Sci U S A* **104**, 3853-3858, (2007).
35. Artner, I. *et al.* MafA and MafB regulate genes critical to beta-cells in a unique temporal manner. *Diabetes* **59**, 2530-2539, (2010).
36. Schaffer, A. E. *et al.* Nkx6.1 controls a gene regulatory network required for establishing and maintaining pancreatic Beta cell identity. *PLoS Genet* **9**, e1003274, (2013).
37. Collombat, P. *et al.* Opposing actions of Arx and Pax4 in endocrine pancreas development. *Genes Dev* **17**, 2591-2603, (2003).
38. Lemaire, K., Thorrez, L. & Schuit, F. Disallowed and Allowed Gene Expression: Two Faces of Mature Islet Beta Cells. *Annu Rev Nutr* **36**, 45-71, (2016).
39. Sekine, N. *et al.* Low lactate dehydrogenase and high mitochondrial glycerol phosphate dehydrogenase in pancreatic beta-cells. Potential role in nutrient sensing. *J Biol Chem* **269**, 4895-4902, (1994).
40. Pullen, T. J., Huising, M. O. & Rutter, G. A. Analysis of Purified Pancreatic Islet Beta and Alpha Cell Transcriptomes Reveals 11beta-Hydroxysteroid Dehydrogenase (Hsd11b1) as a Novel Disallowed Gene. *Front Genet* **8**, 41, (2017).
41. Rouille, Y., Martin, S. & Steiner, D. F. Differential processing of proglucagon by the subtilisin-like prohormone convertases PC2 and PC3 to generate either glucagon or glucagon-like peptide. *J Biol Chem* **270**, 26488-26496, (1995).
42. Steiner, D. F. The proprotein convertases. *Curr Opin Chem Biol* **2**, 31-39, (1998).
43. Jeon, J., Correa-Medina, M., Ricordi, C., Edlund, H. & Diez, J. A. Endocrine cell clustering during human pancreas development. *J Histochem Cytochem* **57**, 811-824, (2009).
44. Jorgensen, M. C. *et al.* An illustrated review of early pancreas development in the mouse. *Endocr Rev* **28**, 685-705, (2007).
45. Teta, M., Rankin, M. M., Long, S. Y., Stein, G. M. & Kushner, J. A. Growth and regeneration of adult beta cells does not involve specialized progenitors. *Dev Cell* **12**, 817-826, (2007).
46. Meier, J. J. *et al.* Beta-cell replication is the primary mechanism subserving the postnatal expansion of beta-cell mass in humans. *Diabetes* **57**, 1584-1594, (2008).
47. Georgia, S. & Bhushan, A. Beta cell replication is the primary mechanism for maintaining postnatal beta cell mass. *J Clin Invest* **114**, 963-968, (2004).

48. Qiu, W.-L. *et al.* Deciphering Pancreatic Islet β Cell and α Cell Maturation Pathways and Characteristic Features at the Single-Cell Level. *Cell metabolism* **25**, 1194-1205.e1194, (2017).
49. Dai, C. *et al.* Age-dependent human beta cell proliferation induced by glucagon-like peptide 1 and calcineurin signaling. *J Clin Invest* **127**, 3835-3844, (2017).
50. Sanyoura, M., Philipson, L. H. & Naylor, R. Monogenic Diabetes in Children and Adolescents: Recognition and Treatment Options. *Curr Diab Rep* **18**, 58, (2018).
51. Cerolsaletti, K., Hao, W. & Greenbaum, C. J. Genetics Coming of Age in Type 1 Diabetes. *Diabetes Care* **42**, 189-191, (2019).
52. Pearson, E. R. Type 2 diabetes: a multifaceted disease. *Diabetologia* **62**, 1107-1112, (2019).
53. Wende, A. R., Huss, J. M., Schaeffer, P. J., Giguere, V. & Kelly, D. P. PGC-1alpha coactivates PDK4 gene expression via the orphan nuclear receptor ERRalpha: a mechanism for transcriptional control of muscle glucose metabolism. *Mol Cell Biol* **25**, 10684-10694, (2005).
54. Rorsman, P. *et al.* Glucose-inhibition of glucagon secretion involves activation of GABAA-receptor chloride channels. *Nature* **341**, 233-236, (1989).
55. Blum, B. *et al.* Functional beta-cell maturation is marked by an increased glucose threshold and by expression of urocortin 3. *Nat Biotechnol* **30**, 261-264, (2012).
56. Avrahami, D. *et al.* Aging-Dependent Demethylation of Regulatory Elements Correlates with Chromatin State and Improved beta Cell Function. *Cell metabolism* **22**, 619-632, (2015).
57. Pullen, T. J. *et al.* Identification of genes selectively disallowed in the pancreatic islet. *Islets* **2**, 89-95, (2010).
58. Benner, C. *et al.* The transcriptional landscape of mouse beta cells compared to human beta cells reveals notable species differences in long non-coding RNA and protein-coding gene expression. *BMC Genomics* **15**, 620, (2014).
59. Brissova, M. *et al.* Assessment of human pancreatic islet architecture and composition by laser scanning confocal microscopy. *J Histochem Cytochem* **53**, 1087-1097, (2005).
60. Arda, H. E., Benitez, C. M. & Kim, S. K. Gene regulatory networks governing pancreas development. *Dev Cell* **25**, 5-13, (2013).
61. DiGruccio, M. R. *et al.* Comprehensive alpha, beta and delta cell transcriptomes reveal that ghrelin selectively activates delta cells and promotes somatostatin release from pancreatic islets. *Mol Metab* **5**, 449-458, (2016).
62. Rorsman, P. & Huising, M. O. The somatostatin-secreting pancreatic delta-cell in health and disease. *Nat Rev Endocrinol* **14**, 404-414, (2018).
63. Camunas-Soler, J. *et al.* Pancreas patch-seq links physiologic dysfunction in diabetes to single-cell transcriptomic phenotypes. *bioRxiv*, 555110, (2019).
64. Miguel-Escalada, I. *et al.* Human pancreatic islet 3D chromatin architecture provides insights into the genetics of type 2 diabetes. *bioRxiv*, 400291, (2018).
65. Chakravarthy, H. *et al.* Converting Adult Pancreatic Islet alpha Cells into beta Cells by Targeting Both Dnmt1 and Arx. *Cell metabolism* **25**, 622-634, (2017).
66. Lemaire, K. *et al.* How stable is repression of disallowed genes in pancreatic islets in response to metabolic stress? *PLoS One* **12**, e0181651, (2017).
67. Thorrez, L. *et al.* Tissue-specific disallowance of housekeeping genes: the other face of cell differentiation. *Genome Res* **21**, 95-105, (2011).
68. Umeyama, K. *et al.* Dominant-negative mutant hepatocyte nuclear factor 1alpha induces diabetes in transgenic-cloned pigs. *Transgenic Res* **18**, 697-706, (2009).
69. Bramswig, N. C. *et al.* Epigenomic plasticity enables human pancreatic alpha to beta cell reprogramming. *J Clin Invest* **123**, 1275-1284, (2013).
70. Ramond, C. *et al.* Understanding human fetal pancreas development using subpopulation sorting, RNA sequencing and single-cell profiling. *Development* **145**, (2018).

71. Perez-Alcantara, M. *et al.* Patterns of differential gene expression in a cellular model of human islet development, and relationship to type 2 diabetes predisposition. *Diabetologia* **61**, 1614-1622, (2018).
72. Blandino-Rosano, M. *et al.* Loss of mTORC1 signalling impairs beta-cell homeostasis and insulin processing. *Nat Commun* **8**, 16014, (2017).
73. Brun, P. J. *et al.* Retinoic acid receptor signaling is required to maintain glucose-stimulated insulin secretion and beta-cell mass. *FASEB J* **29**, 671-683, (2015).
74. Lin, H. M. *et al.* Transforming growth factor-beta/Smad3 signaling regulates insulin gene transcription and pancreatic islet beta-cell function. *J Biol Chem* **284**, 12246-12257, (2009).
75. Yokoi, N., Gheni, G., Takahashi, H. & Seino, S. beta-Cell glutamate signaling: Its role in incretin-induced insulin secretion. *J Diabetes Investig* **7 Suppl 1**, 38-43, (2016).
76. Ang, L. T. *et al.* A Roadmap for Human Liver Differentiation from Pluripotent Stem Cells. *Cell Rep* **22**, 2190-2205, (2018).
77. Loh, K. M. *et al.* Efficient endoderm induction from human pluripotent stem cells by logically directing signals controlling lineage bifurcations. *Cell Stem Cell* **14**, 237-252, (2014).
78. Nair, G. G. *et al.* Recapitulating endocrine cell clustering in culture promotes maturation of human stem-cell-derived beta cells. *Nat Cell Biol* **21**, 263-274, (2019).
79. Pagliuca, F. W. *et al.* Generation of functional human pancreatic beta cells in vitro. *Cell* **159**, 428-439, (2014).
80. Rezania, A. *et al.* Reversal of diabetes with insulin-producing cells derived in vitro from human pluripotent stem cells. *Nat Biotechnol* **32**, 1121-1133, (2014).
81. Russ, H. A. *et al.* Controlled induction of human pancreatic progenitors produces functional beta-like cells in vitro. *The EMBO journal* **34**, 1759-1772, (2015).
82. Lamb, M. *et al.* In vitro maturation of viable islets from partially digested young pig pancreas. *Cell Transplant* **23**, 263-272, (2014).
83. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods* **14**, 417-419, (2017).
84. Zerbino, D. R. *et al.* Ensembl 2018. *Nucleic Acids Res* **46**, D754-D761, (2018).
85. Soneson, C., Love, M. I. & Robinson, M. D. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Res* **4**, 1521, (2015).
86. Kolde, R. Pheatmap: pretty heatmaps. *R package version* **61**, 617, (2012).
87. Wickham, H. ggplot2: Elegant Graphics for Data Analysis. *Use R*, 1-212, (2009).
88. Buniello, A. *et al.* The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res* **47**, D1005-D1012, (2019).
89. Yang, Y. & Chan, L. Monogenic Diabetes: What It Teaches Us on the Common Forms of Type 1 and Type 2 Diabetes. *Endocr Rev* **37**, 190-222, (2016).

Acknowledgements

We thank past and current members of the Kim group for advice and encouragement, Dr. Jing Wang and Michael Alexander for assistance with pig pancreas work, Serena Whitener for help developing the pig pancreas gene expression browser, Dr. Kyle Loh for stimulating discussions, and Drs. Dan Lu and Ramesh Nair for help on bioinformatics. This work was supported by a Deans Innovation Fund from the Dept. of Developmental Biology (Stanford), and by fellowship awards from the L.L. Hillblom Foundation (2017-D-008-FEL) to S.K., from the Division of Endocrinology T32 training grant (DK007217-41, A. Hoffman and F. Kraemer) to R.L.W., and from the Stanford Child Health Research Institute (UL1 TR001085) and the American Diabetes Association (1-16-PDF-086) to H.P. Work in the Kim lab was also supported by grants from the U.S. National Institutes of Health P30 DK116074, and by the Stanford Islet Research Core and Diabetes Genomics and Analysis Core of Stanford Diabetes Research Center.

Author contributions

S.K., R.L.W. and S.K.K. designed the study; S.K, R.L.W, X.G., C.A.C., J.Y.L., I.P., R.J.B., and K.T. performed the experiments; J.C., S.R.Q., J.R.T.L., R.B., and P.J.R. shared materials; S.K., R.L.W., H.P., and S.K.K analyzed the data and wrote the manuscript.

Competing interests: Nothing to declare

Materials & Correspondence should be addressed to S.K.K

Figure legends

Fig. 1 Pancreas dissociation and FACS purification of islet cells from fetal and neonatal pigs. **a** Schematic representation of the study design. **b** qRT-PCR analysis of *INS*, *GCG*, *SST*, *KRT*, and *AMY2* of FACS sorted β-, α- and δ-cell populations. **c** Heatmap of differentially expressed genes (DEGs) in pig β-cells between early fetal, late fetal and neonatal stages. **d** GO Term enrichment analysis of genes increasing with age in pig β-cells. **e** Heatmap of DEGs in pig α-cells between early fetal, late fetal and neonatal stages. **f** GO Term enrichment analysis of genes increasing with age in pig α-cells. **g** Heatmap of DEGs in pig δ-cells between late fetal and neonatal stages. **h** GO Term enrichment analysis of genes increasing with age in pig δ-cells. Adjusted p value threshold for GO term analysis was 0.1.

Fig. 2 Analysis of gene expression in developing pig islet β-, α- and δ-cells**. a-i** Boxplots displaying normalized TPM (transcripts per million) counts of β-, α-, and/or δ-cell specific genes (**a-h**), markers associated with proliferation (**e**), disallowed genes (**f**) and β cell function regulators (**g-i**). Adjusted *P* value * *P≤0.05; ** P≤0.01; *** P≤0.001.* **j** Immunostaining of PCSK1 (white) with insulin (INS; green) and glucagon (GCG; red) in early fetal, neonatal and adult pig pancreas. White arrows indicate PCSK1 and GCG double positive cells. White arrow head indicate GCG positive cells Scale bars, 50 µm.

Fig. 3 Stage-specific expression of islet factors during development**. a-c** Immunostaining of INS (green) and GCG (white) paired with Somatostatin (SST; red) (**a**), E-Cadherin (ECad; white) (**b**), and Chromogranin-A (CHGA; white) (**c**) in the early fetal, neonatal, and adult pig pancreas. **d-f** Immunostaining of the transcription factors PDX1 (**d**), NKX6.1 (**e**), and MAFB (**f**) in the early fetal, neonatal, and adult pig pancreas. **g** qRT-PCR data probing for expression of the maturation factor *SIX3* in isolated neonatal and adult pig islets (n=3, t-test; Error bars indicate SD). **h** Immunostaining of SIX3 with INS and GCG in adult pig pancreas. Insets indicated by dashed yellow box. Scale bars, 50 µm.

Fig. 4 Comparison of gene expression in pig and human β- and α-cell development. **a** Unsupervised hierarchical clustering using distance matrix of neonatal pig, human, and mouse β cells. **b-g** Boxplots displaying normalized TPM (transcripts per million) counts of select genes whose expression changes in a conserved manner between human and pig development. **h** Matrix indicating overlap of genes whose expression is age-dependent in pig β- or α- cells and genes linked to diabetes or related metabolic traits in human genome wide association studies.

NDM, neonatal diabetes mellitus; MODY, maturity-onset diabetes of the young, (*Chi-square test). Green box indicates the specific linkage(s) for each gene. **i** Boxplots of genes identified in (**h**) show similar developmentally regulated expression patterns in both pigs and humans. Adjusted *P* value * *P≤0.05; ** P≤0.01; *** P≤0.001.*

Fig. 5 Dynamic gene regulation during pig β- and α-cell development. **a-b** Heat maps of genes with altered expression from early fetal to neonatal stages in β-cells (**a**) and α-cells (**b**), showing segregation into 8 clusters. The z-score scale represents $\log_2(X+1)$ transformed TPM counts. Cluster 2 includes genes that increased up to late fetal stages then did not change thereafter. Cluster 7 includes genes that declined then did not change thereafter. Enriched GO terms in each cluster are shown to the right. Adjusted p value threshold was 0.1. **c-d** The z-score scale represents $\log_2(X+1)$ transformed TPM counts with total genes in β-cells (**c**) and α-cells (**d**). Cluster 3 (Up-Down) or Cluster 6 (Down_Up). Boxplots displaying normalized TPM (transcripts per million) counts of genes shown. Adjusted *P* value * *P≤0.05; ** P≤0.01; *** P≤0.001.*

Fig. 6 Maturation of β-cell function in postnatal pigs. **a** Glucose-stimulated insulin secretion (GSIS) assay measured insulin secretion in response to 2.8mM (low), 20mM (high) glucose and 20mM glucose + IBMX. Data show average secreted insulin as a percentage of insulin content in the isolated islets from P2 (n=7) and P22 (n=8) piglets (t-test, **P ≤0.05*; error bars indicate SD). **b** Heat map of DEGs between d90 and P22 β-cells. **c-d** Boxplots displaying normalized TPM (transcripts per million) counts of up-regulated (**c**) or down-regulated (**d**) genes between d90 and P22. Adjusted *P* value * *P≤0.05; ** P≤0.01; *** P≤0.001.* **e-f** GO term analysis for biological processes with up-regulated (**e**) and down-regulated (**f**) genes between d90 and P22 β-cells. Adjusted *P* value threshold was 0.1.

Fig. 1 Pancreas dissociation and FACS purification of islet cells from fetal and neonatal pigs

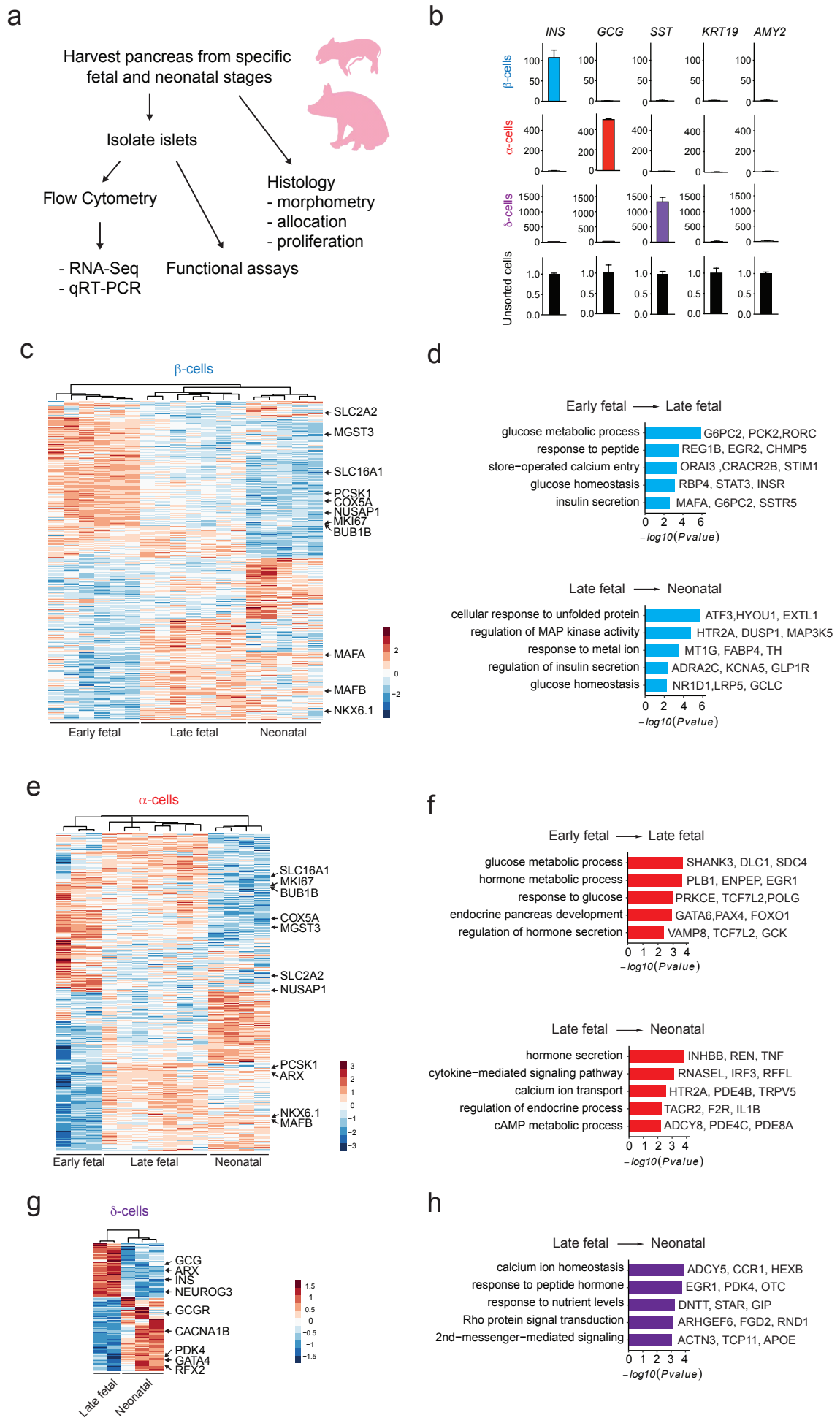


Fig. 2 Analysis of gene expression in developing pig islet β-, α- and δ-cells

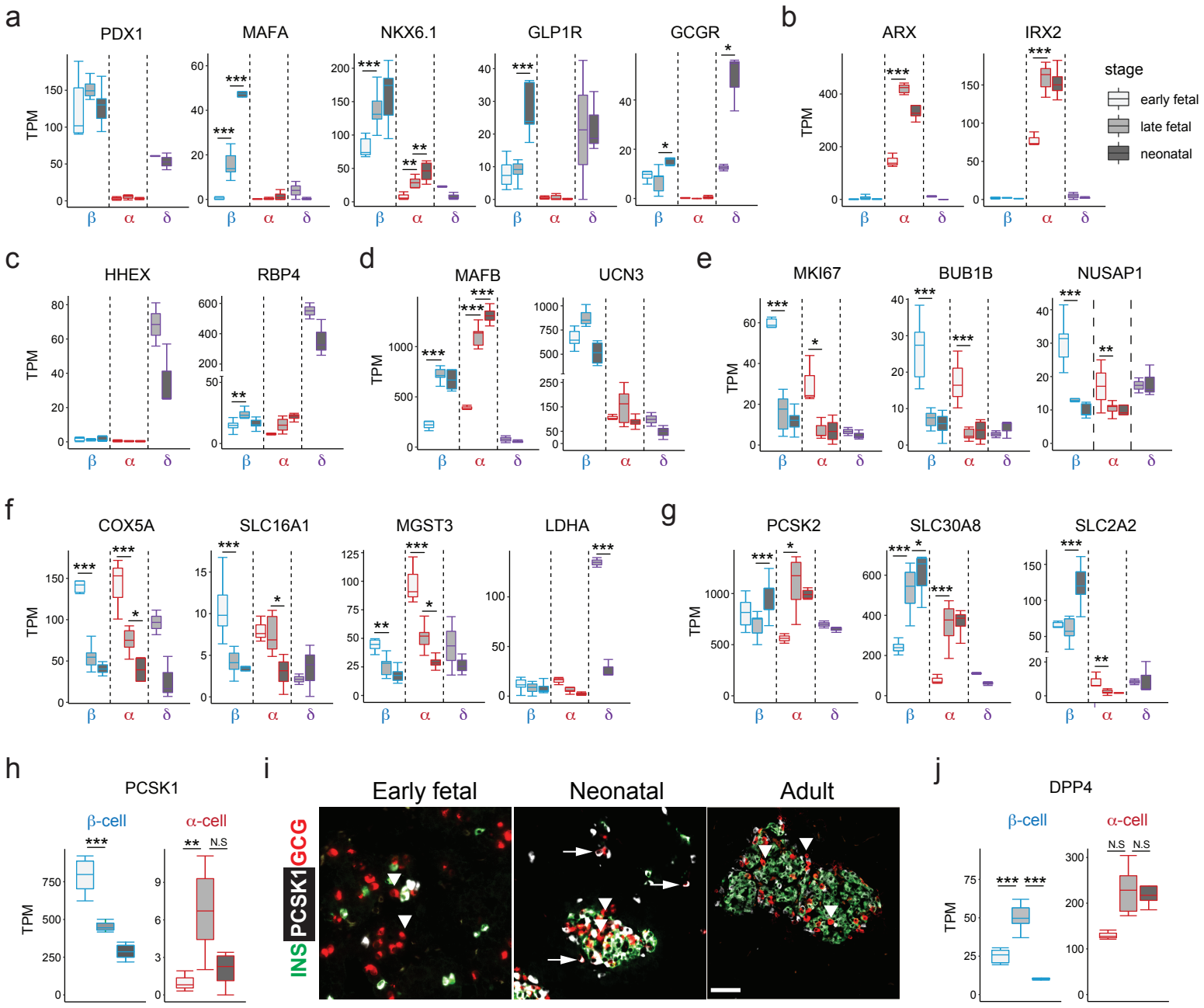


Fig. 3 Stage-specific expression of islet factors during development

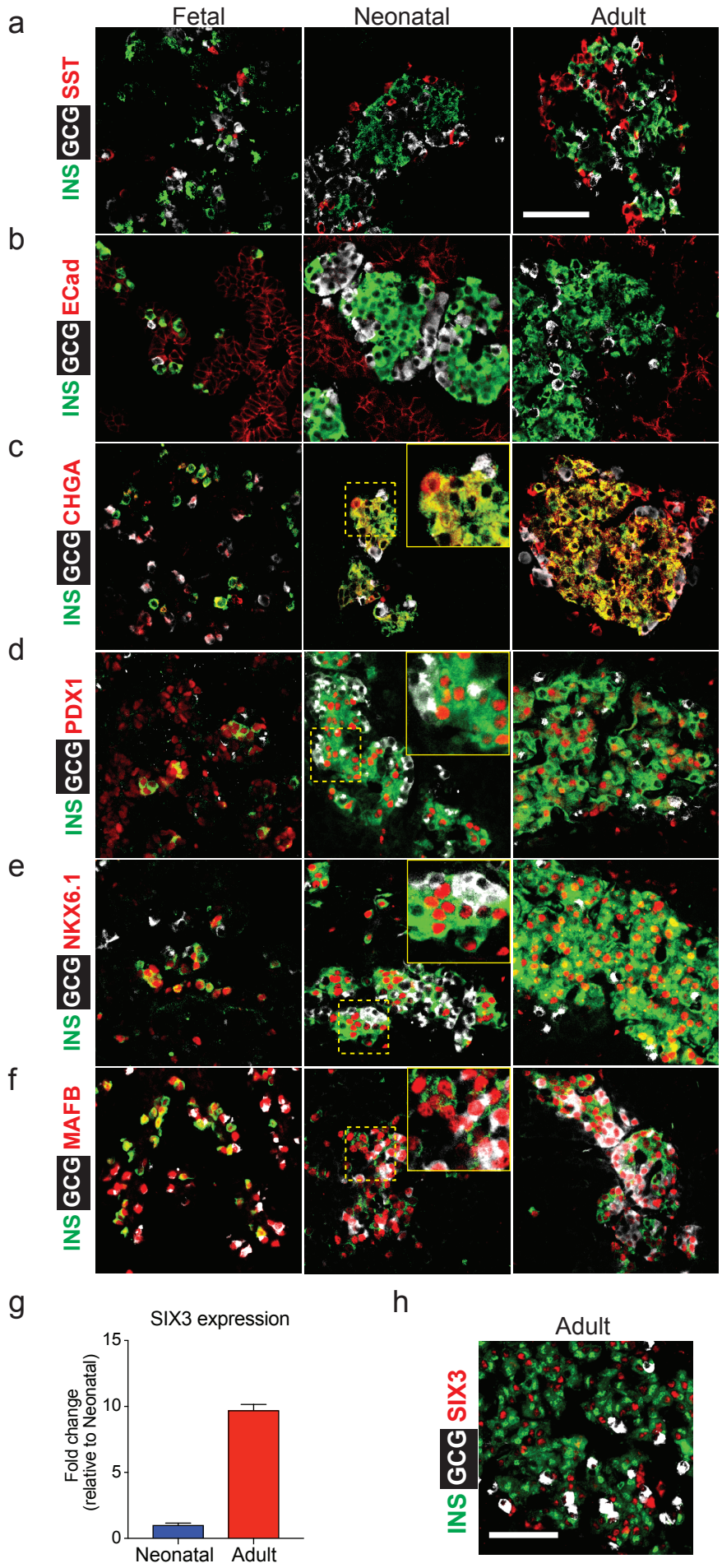


Fig. 4 Comparison of gene expression in pig and human β- and α-cell development

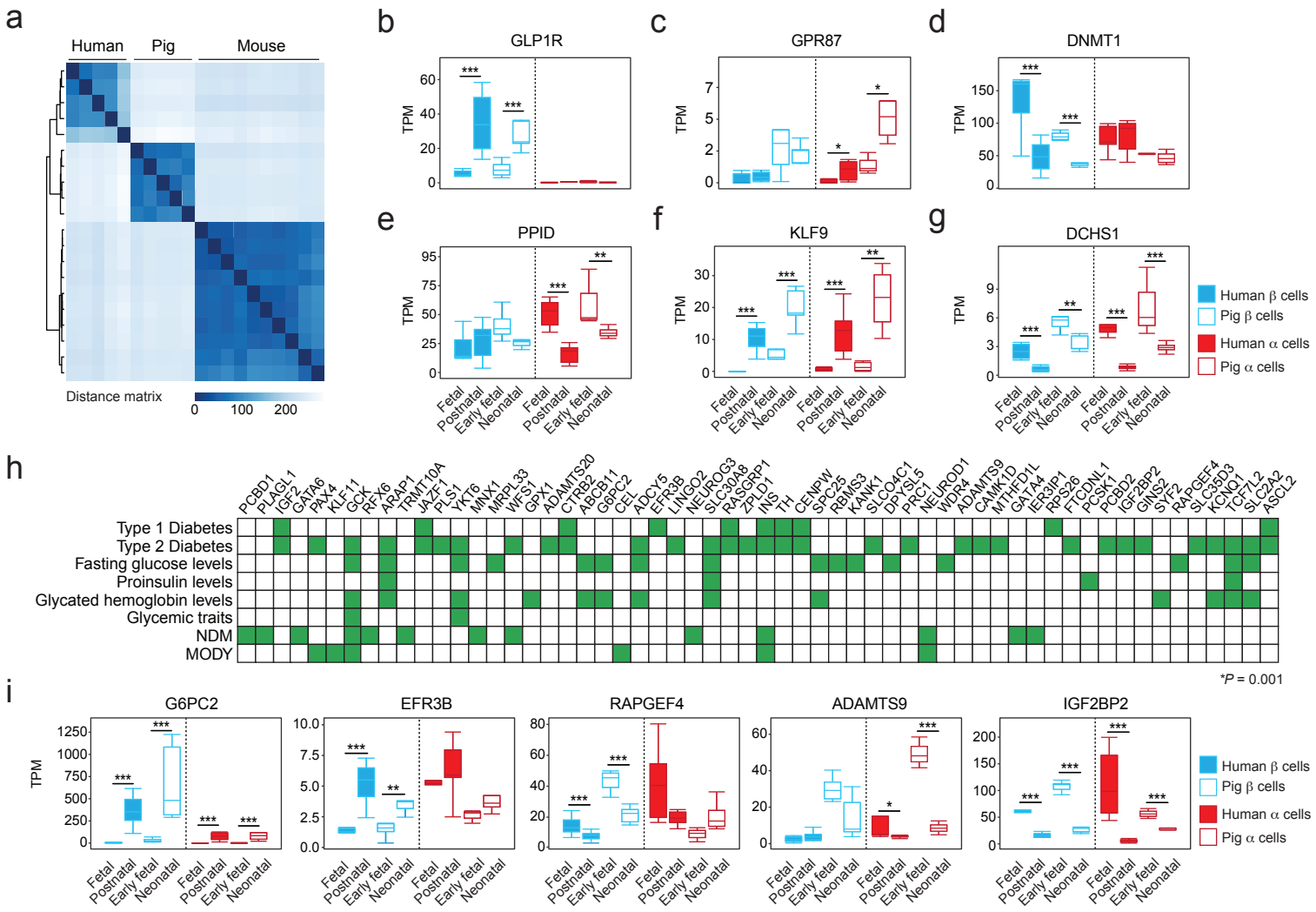


Fig. 5 Dynamic gene regulation during pig β- and α-cell development

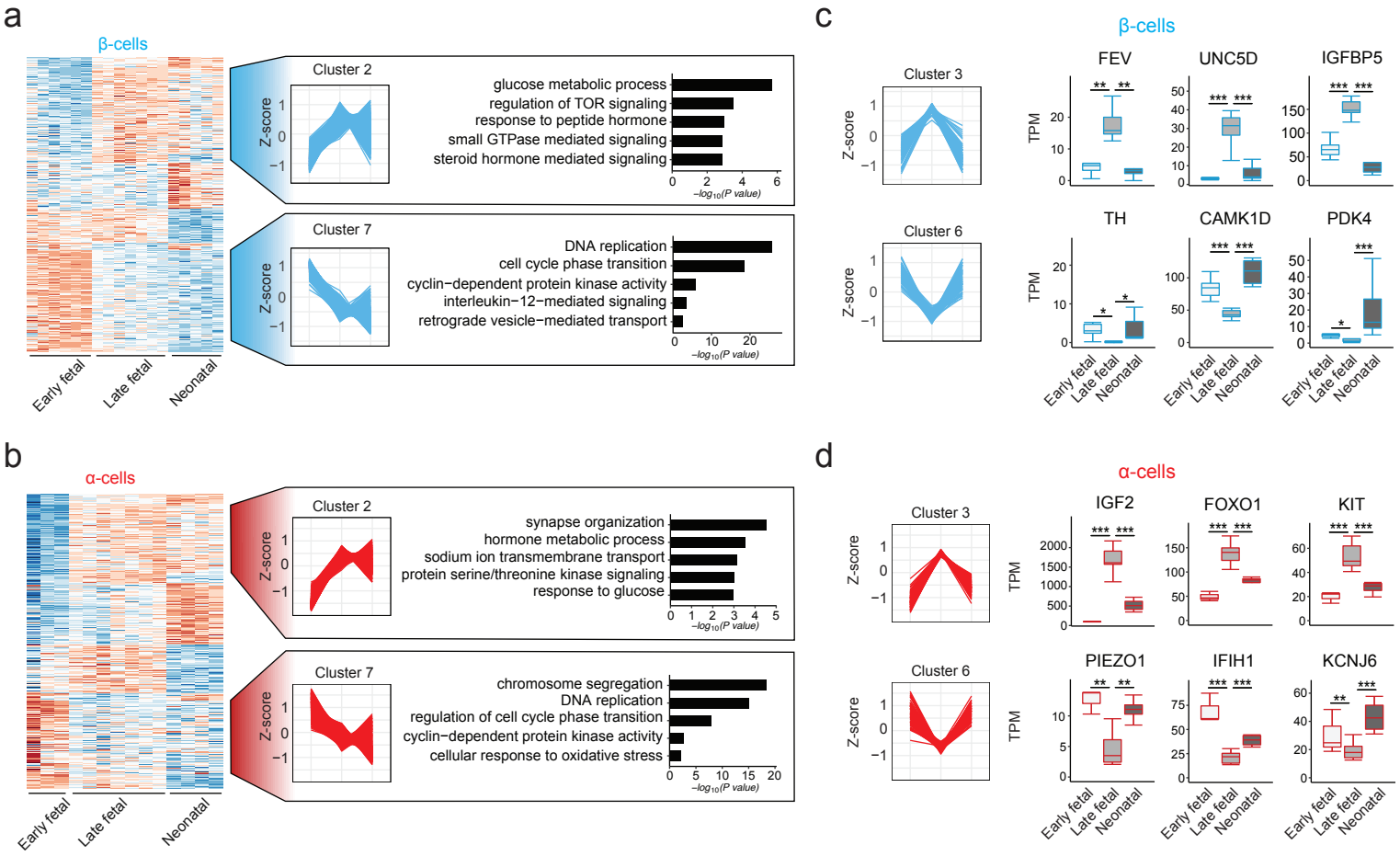
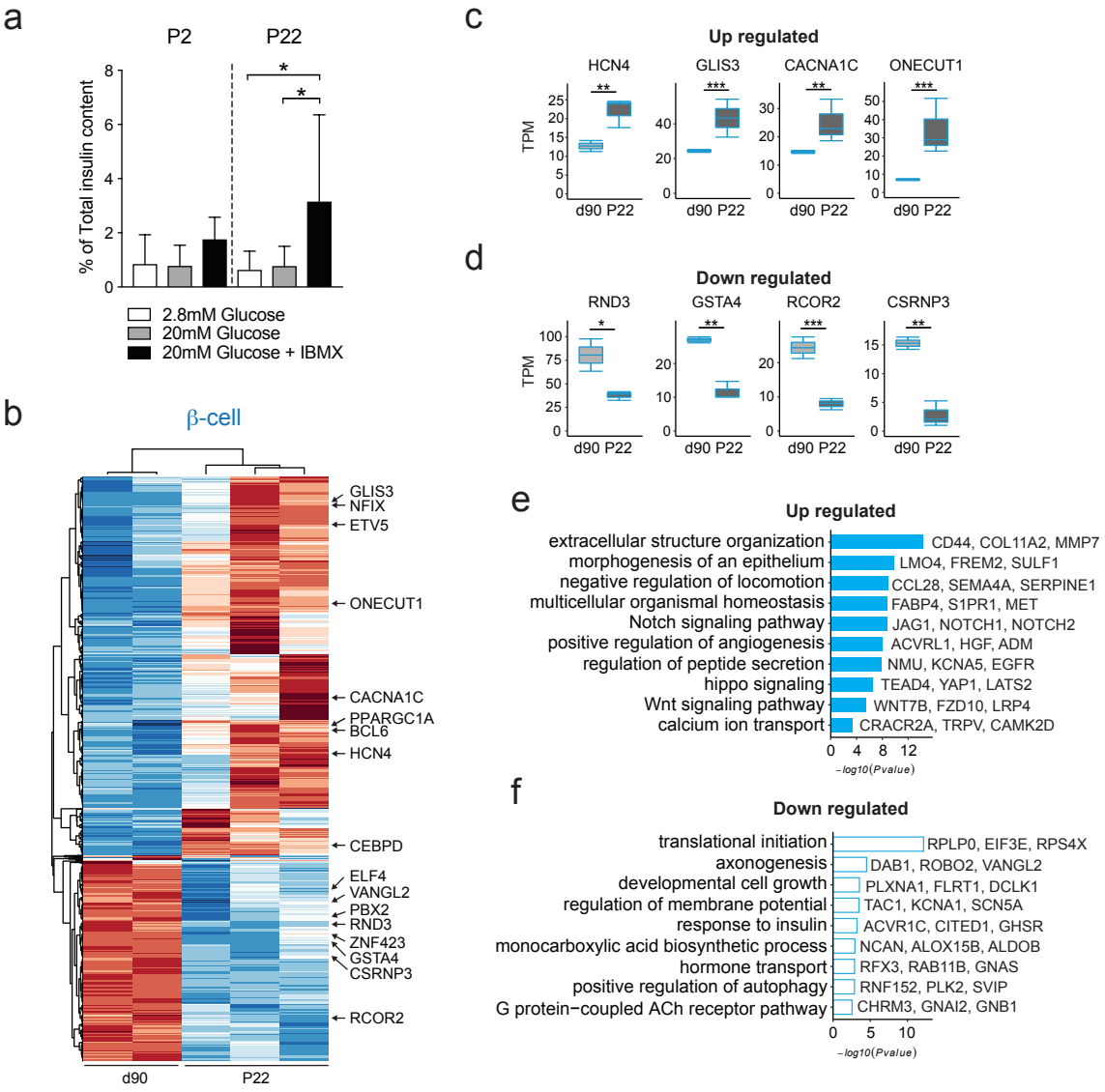
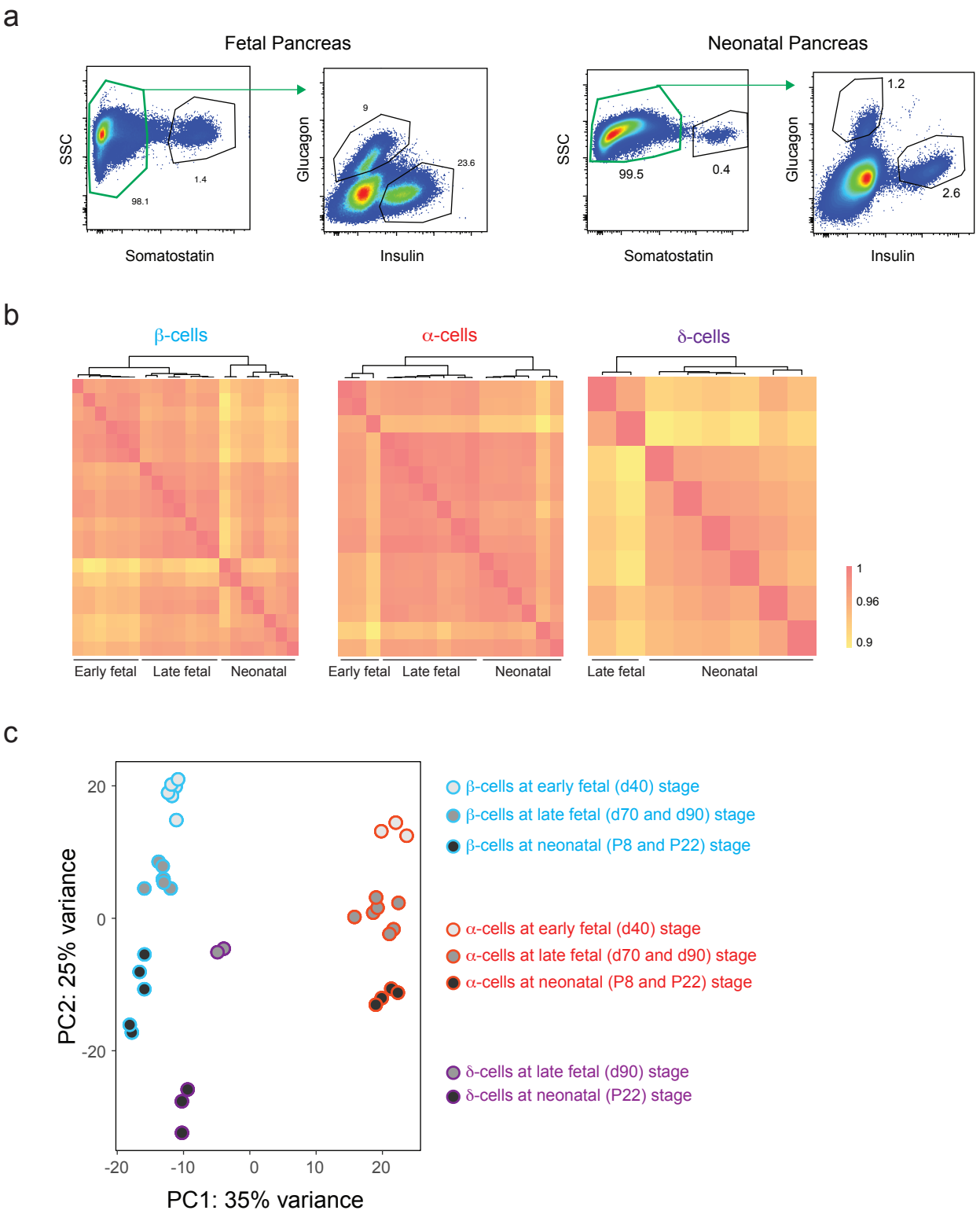


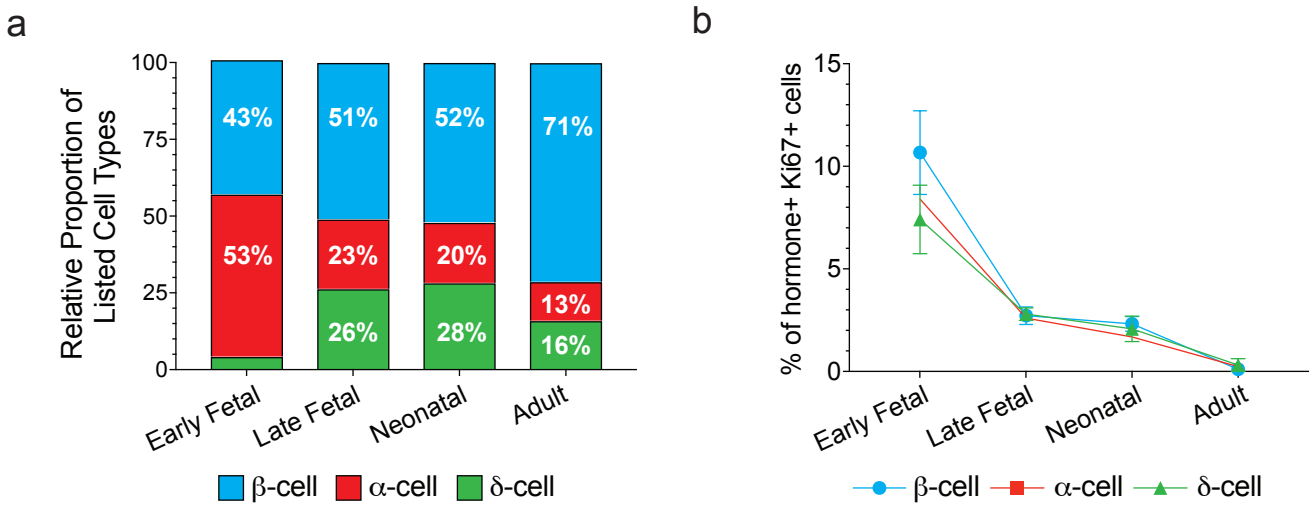
Fig. 6 Maturation of β-cell function in postnatal pigs



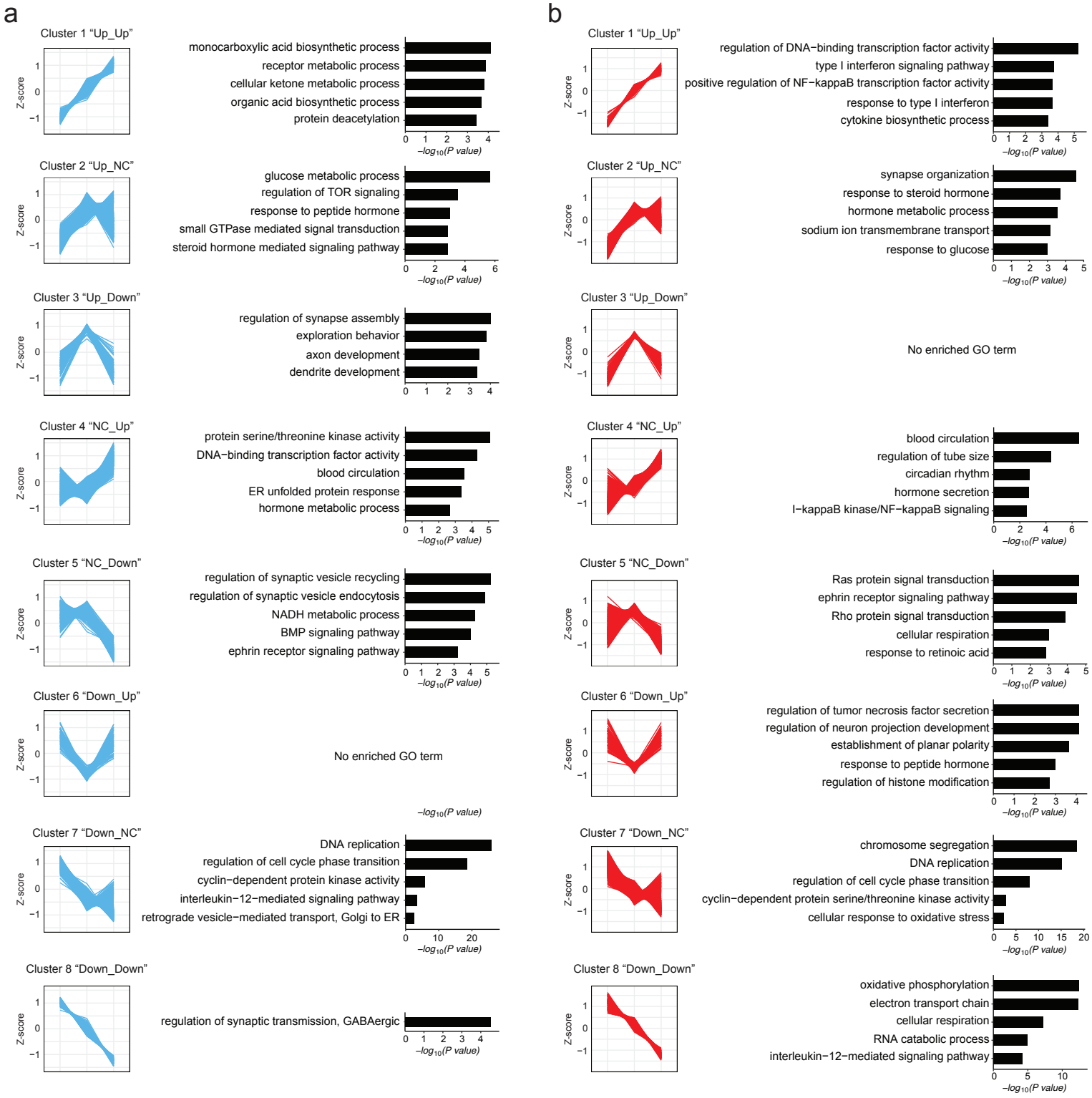
Supplementary Figures



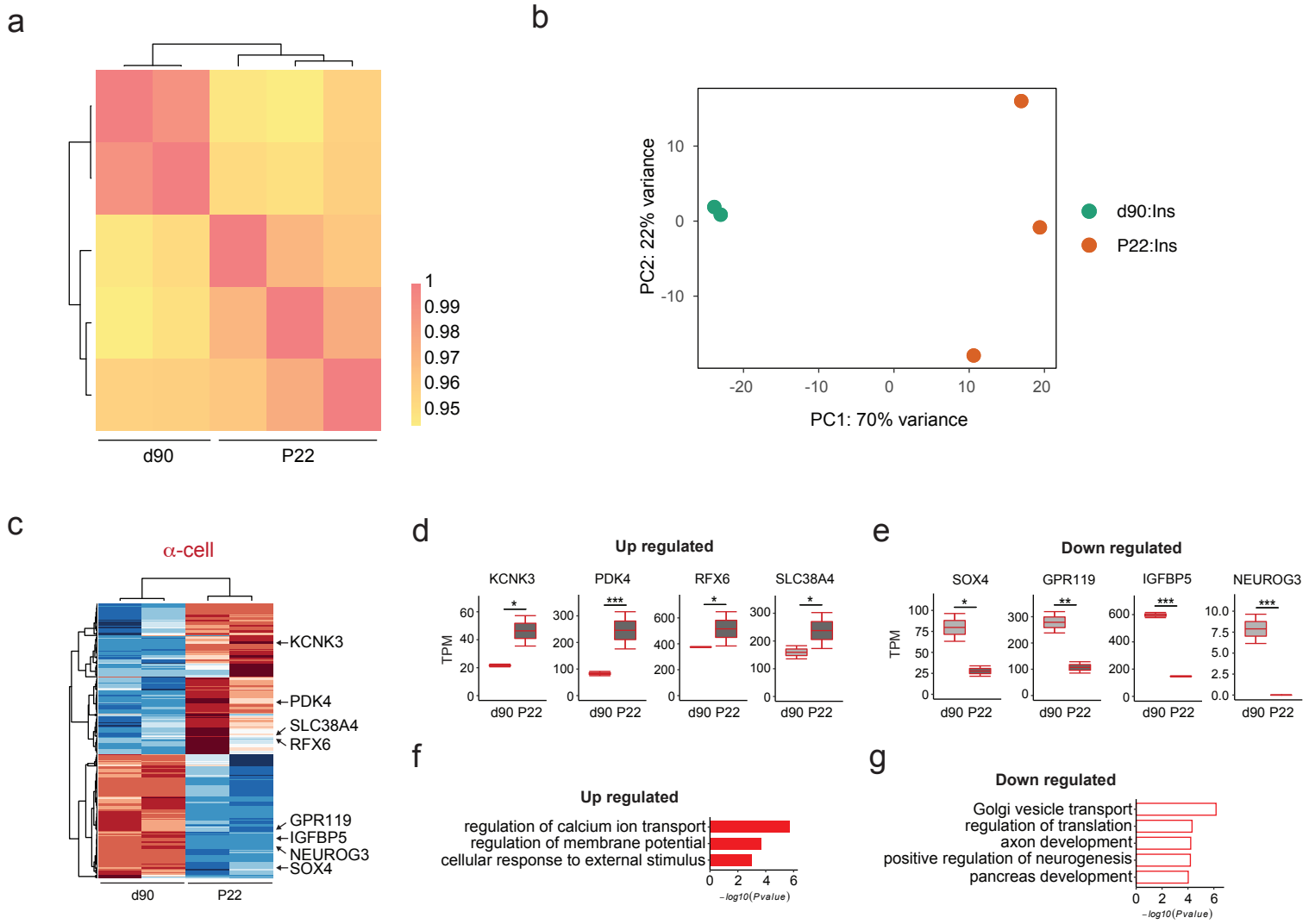
Supplementary Fig. 1 (a) FACS plots demonstrating distinct populations of Insulin (INS) positive β-, glucagon (GCG) positive α- and somatostatin (SST) positive δ-cells among dispersed fetal and neonatal pancreatic cells. (b) Pearson correlation matrix of all RNA-seq samples used in this study demonstrating distinct stage specific gene expression patterns in β-, α-, and δ-cells (c) Principal Component Analysis indicating the RNA-Seq libraries cluster according to cell type in principle component 1, followed by developmental stage in principle component 2



Supplementary Fig. 2 (a) Quantification graph showing relative proportion of INS+, GCG+ and SST+ cells in each developmental stage (n = 40 islets per group, from 3 pigs per group) (b) Quantification graph showing average percentage of INS+, GCG+ and SST+ cells with Ki67+ in each developmental stage (n = 40 islets per group, from 3 pigs per group; error bars are SD).



Supplementary Fig. 3 (a,b) Visualization of gene expression change with z score representing log2(X+1) transformed TPM counts of genes in 8 clusters. Genes were clustered based on expression changes from early fetal stage to neonatal day 22 in α- (a) and β- (b) cells. Enriched GO terms with genes in each cluster are shown to the right. Adjusted p value threshold was 0.1.



Supplementary Fig. 4 Identification of differentially expressed genes between P2 and P22 pig β- and α-cells. (a) Unsupervised hierarchical clustering of d90 and P22 pig β-cells RNA-seq samples (b) Principal Component Analysis of d90 and P22 pig β-cells RNA-seq samples (c) Heatmap of differentially expressed genes (DEGs) in pig β-cells between d90 and P22 stages. (d-e) Boxplots displaying normalized TPM (transcripts per million) counts of up- regulated (d) or down-regulated (e) genes between d90 and P22 in α-cells. Adjusted P-value * P≤0.05; ** P≤0.01; *** P≤0.001. (f) GO term analysis for biological processes with up-regulated (e) and down-regulated (f) genes between d90 and P22 β-cells. Adjusted p value threshold was 0.1.

Supplementary Data

Supplementary Data 1. List of RNA-Seq libraries sequenced in this study

Supplementary Data 2. List of differentially expressed genes in pig β-, α-, and δ-cells between two developmental stages (Early fetal vs Late fetal, Late fetal vs Neonatal)

Supplementary Data 3. TPM counts of genes in each RNA-seq library

Supplementary Data 4. List of differentially expressed genes in human β- and α-cells between early fetal and post-natal stages and in pig β- and α-cells between early fetal and neonatal stages.

Supplementary Data 5. Pig homologs of GWAS risk genes with differential expression from early fetal to neonatal stages

Supplementary Data 6. List of genes with dynamic expression in pig β- and δ-cells from early fetal to neonatal stages.

Supplementary Data 7. List of up- and down-regulated genes in pig β- and δ-cells between d90 and P22.

Supplementary Data 8. List of GO term for biological process with up- and down-regulated genes in pig β- and δ-cells between d90 and P22. Adjusted p value threshold was 0.1.