Organogenesis requires the complex interactions of multiple cell lineages that coordinate their expansion, differentiation, and maturation over time. Here, we profile the cell types within the epithelial and mesenchymal compartments of the murine pancreas across developmental time using a combination of single-cell RNA sequencing, immunofluorescence, in situ hybridization, and genetic lineage tracing. We identify previously underappreciated cellular heterogeneity of the developing mesenchyme and reconstruct potential lineage relationships among the pancreatic mesothelium and mesenchymal cell types. Within the epithelium, we find a previously undescribed endocrine progenitor population, as well as an analogous population in both human fetal tissue and human embryonic stem cells differentiating toward a pancreatic beta cell fate. Further, we identify candidate transcriptional regulators along the differentiation trajectory of this population toward the alpha or beta cell lineages. This work establishes a roadmap of pancreatic development and demonstrates the broad utility of this approach for understanding lineage dynamics in developing organs.
Pancreatic organogenesis is a complex and dynamic process that ultimately results in the generation of multiple cell lineages that perform the functions of the mature organ: the regulation of glucose homeostasis by the endocrine compartment and the production of digestive enzymes by the exocrine compartment. In the mouse, all known epithelial lineages of the pancreas derive from a small field of epithelial precursor cells within the foregut endoderm specified by the expression of Pancreatic duodenal transcription factor 1 (Pdx1) (Fig. 1a)\(^1\). These Pdx1+ cells evaginate into a cap of surrounding mesenchymal cells around embryonic day 9 (E9), proliferate, and begin the process of branching morphogenesis. Further epithelial lineage diversification continues with the specification of Pdx1+ cells into tip and trunk domains by E12 and progresses to the restriction of tip cells to a digestive enzyme-producing acinar fate and of trunk cells to either a ductal or endocrine cell fate\(^1\). Within the trunk domain, induction of Neurogenin 3 (Ngn3) expression defines the cells that will differentiate into one of five endocrine lineages: alpha, beta, delta, gamma, or epsilon cells, marked by expression of the hormones Glucagon (Gcg), Insulin (Ins), Somatostatin (Sst), Pancreatic polypeptide (PP), or Ghrelin (Ghrl), respectively\(^2\). Gastrin+ cells have also been recently described\(^3\). Despite previous work focused on the formation of cellular heterogeneity in the pancreatic mesenchyme during development and in the adult organ. Therefore, a deeper understanding of the full diversity of the mesenchymal cell types, as well as their global gene-expression profiles, will serve as a basis for understanding these key cellular interactions.

Recent studies of late embryonic, postnatal, and adult alpha and beta cells have demonstrated the power of single-cell transcriptomic profiling for unraveling endocrine lineage heterogeneity and revealing distinct transcriptional states of beta cell maturation\(^4–8\). Here, we perform droplet-based, single-cell RNA sequencing of entire murine embryonic pancreata at earlier developmental timepoints to describe the cellular diversity and dynamics of gene expression in both the epithelial and mesenchymal compartments. We further identify and validate these populations within mouse and human pancreatic tissue, as well as human embryonic stem cell (hESC)-derived endocrine progenitor cells. Finally, we predict lineage relationships, identify previously unappreciated intermediate progenitor cells, and validate our methodology using in vivo genetic lineage tracing.

**Results**

**Cellular heterogeneity in the murine pancreas.** We first characterized the major sources of cellular heterogeneity in the developing pancreas. Two batches of mouse pancreata at E14.5, a particularly active time of expansion, morphogenesis, and diversification\(^2\) (Fig. 1a), were dissected from individual litters, dissociated into single-cell suspensions, sorted for live cells, and sequenced using the 10x Chromium Single-Cell version 1 (v1) kits (Fig. 1b and Supplementary Fig. 1a–e). We performed filtering, normalization, variable gene identification, linear regression for batch, and principal component analysis (PCA) with the R package Seurat (Supplementary Fig. 1d, e and 2a, b). Graph-based clustering\(^6\) of batch-adjusted, merged data identified 19 distinct cell populations, classified as epithelial, mesenchymal, immune, or vascular populations based on the expression of known markers (Fig. 1c, d and Supplementary Data 1). We identified expected populations, including endocrine, exocrine (acinar and ductal), and endothelial cells (Fig. 1e). The proportions of endocrine, mesenchymal, immune, and vascular populations were similar between E14.5 batches (Supplementary Fig. 2b–d). Down-sampling analysis confirmed that sufficient sequencing depth had been reached for calling clusters (Supplementary Fig. 2c–g). These results reveal the power of single-cell RNA sequencing to identify a broad range of cell types during development.

**Characterization of mesenchymal heterogeneity.** While previous studies have identified numerous markers of pancreatic epithelial populations\(^2\), comparatively little is known about heterogeneity among pancreatic mesenchymal cells. We characterized the mesenchymal compartment by subclustering only mesenchymal cells (5069 cells) and reperforming the clustering analysis (Fig. 2a and Supplementary Fig. 3a). Despite being less divergent from one another than were cells in the epithelial compartment (Fig. 2b and Supplementary Fig. 3b), mesenchymal cells could still be subdivided into ten transcriptionally distinct mesenchymal clusters (Fig. 2a, c and Supplementary Data 2). We verified the differential gene-expression analysis with three tests: bimodal likelihood ratio test\(^8\), Wilcoxon rank sum, and MAST\(^9\) (Supplementary Fig. 3c and Supplementary Data 2). We annotated two clusters based on the expression of known marker genes: cluster 1 is pancreatic mesothelial cells (Wt1, Krt19, and Lpkapb)\(^10,11\) and cluster 3 represents vascular smooth muscle (VSM) cells (Acta2, Tagln, and Myf9) (Fig. 2c and Supplementary Data 2)\(^12\). Indeed, in E14.5 pancreas, Wt1 expression was restricted to the tissue edge, as expected for mesothelial cells, while Acta2 expression was localized to cells surrounding vessels, as expected for VSM cells (Supplementary Fig. 3d, e). Cells in the mesothelial cluster also expressed the secreted factors Fgf9, Pdgfc, Rspo1, and Igfbp5 (Supplementary Fig. 3f) and genes regulating prostaglandin hormone signaling and tight junctions (Fig. 2d and Supplementary Data 3).

The remaining mesenchymal clusters included proliferating cells (clusters 6–8), a large cluster (10) expressing pan-mesenchymal markers, and four clusters (2, 4, 5, and 9) each expressing a signature distinct from that of cluster 10 (Fig. 2a, c and Supplementary Data 2). Cluster 2 was defined by differential expression of Stathmin 2 (Stmn2), a gene involved in neurite outgrowth and osteogenesis\(^13,14\). We also found two populations, clusters 4 and 5, that differentially expressed multiple secreted factors. Cluster 4 expressed Ace2, the chemokines Cxcl12 and Cxcl13, and Vegfd, while cluster 5 expressed high levels of the Wnt antagonists Secreted frizzled-related protein 1 and 2 (Sfrp1 and Sfrp2) (Fig. 2c–e and Supplementary Data 2). Cluster 5 also expressed the transcription factor BarH-like homeobox 1 (Barx1) and members of the Id DNA-binding protein family (Fig. 2c–e and Supplementary Data 2). Cluster 9 expressed Nkx2 homeobox 5 (Nkx2-5) and Tlx1, transcription factors reported to contribute to splenic development during a window in which the embryonic pancreas and spleen share a mesenchymal compartment (Fig. 2e)\(^15\). Pathway analysis identifies multiple signaling pathways that may be functionally relevant in these populations (Fig. 2d and Supplementary Data 3). We validated a subset of these distinct clusters using dual in situ hybridization/immuno-fluorescence (ISH/IF) on E14.5 pancreas for differentially expressed markers of clusters 1 (Cav1 and Barx1), 2 (Stmn2), and 5 (Barx1) (Fig. 2e–h). These gene-expression profiles demonstrate a previously underappreciated level of heterogeneity in the mesenchymal compartment of the developing pancreas.
Mesothelial cells undergo changes across developmental time. During organogenesis, the dynamics of each lineage are defined by the expansion, differentiation, and maturation of its constituent cells. To address how these processes change across chronological time within the developing pancreas, we performed single-cell sequencing at two additional timepoints, E12.5 and E17.5 (Fig. 3a). We identified mesenchymal cells from E12.5, E14.5, and E17.5 timepoints, merged them into one dataset, and reperformed the clustering analysis. We identified the clusters detected in our E14.5 analysis (clusters 1–10) along with seven new clusters (11–17) (Fig. 3a, Supplementary Fig. 3g–i, and Supplementary Data 4). The addition of E12.5 and E17.5 cells revealed further sub-division of the mesothelium into timepoint-specific clusters (1, 11, and 17), each with unique transcriptomic signatures (Fig. 3a, b). Within the mesothelium, we verified Paired-like homeodomain transcription factor 2 (Pitx2) expression...
at E12.5 and its absence at E17.5 and Mesothelin (Msnl) expression at E17.5 and its absence at E12.5 (Fig. 3c), consistent with the single-cell data. These data provide evidence of transcriptional maturation over developmental time within the mesothelial compartment.

While the mesothelium is a well-established mesenchymal progenitor cell population for VSM and fibroblasts in multiple other organs, both the role of the mesothelium and the origin of the mesenchymal cell types within the pancreas remain uncharacterized16–19. We utilized our single-cell mesenchymal dataset to determine whether the pancreatic mesothelium may function as a mesenchymal progenitor cell population during development. We found six populations (clusters 2, 3, 4, 5, 12, and 13) that expressed VSM cell genes, such as Acta2 and Tagln, or genes known to regulate VSM development, such as Mgp20, Fhl121,22, Barx123, and Pitx224 (Fig. 3d). Based on these VSM-related gene-expression profiles, we hypothesized that these populations could represent VSM progenitors derived from the pancreatic mesothelium. To test the lineage relationships among these populations, we ordered cells in pseudotime based on their transcriptional similarity25. This analysis placed mesothelial cells on one side of the pseudotime trajectory (Fig. 3e). Mesothelial branches corresponded to either a maturation process, based on placement of E17.5 cells at the branch terminus, or proliferating mesothelium, based on expression of proliferation genes (Fig. 3e and Supplementary Fig. 3j). VSM-related populations were placed on the other side of the trajectory (Fig. 3e and Supplementary Fig. 3j). We calculated the proportion of each population over pseudotime and found a transition from the E12.5 mesothelial population (cluster 11) to cluster 12, both of which share expression of the gene Pitx2 (Fig. 3e–g). Cluster 12 then transitioned into the Stmn2-expressing cluster 2, which split into a branch composed of VSM populations, clusters 3 and 13 (Branch 1), and a branch composed of clusters 4 and 5 (Branch 2) (Fig. 3e–g). Thus, this analysis proposes clusters 2 and 12 as potential mesothelial-derived mesenchymal progenitor populations that can contribute to the VSM lineages (Fig. 3g). Our analysis has identified and validated multiple mesenchymal subtypes and possible lineage relationships among them.

A previously undescribed endocrine progenitor population.

After assessing the heterogeneity within the mesenchymal compartment, we next focused on the epithelial cells. We first subclustered the 2049 cells from our E14.5 dataset that comprised just the epithelial populations (Fig. 4a and Supplementary Fig. 4a). We identified 10 clusters, including acinar, ductal, beta, alpha, and Ngn3+ progenitor populations, as revealed by differential expression of known markers (Fig. 4a, b, Supplementary Fig. 4b, and Supplementary Data 5). Our analysis highlighted previously uncharacterized markers of acinar, Ngn3+, beta, and alpha cell populations, such as Reep5, Btbd17, Gng12, and Peg10, respectively (Fig. 4b and Supplementary Data 5). We also found Sst- and Pancreatic polypeptide (Ppy)-expressing cells, but they did not cluster into their own populations (Supplementary Fig. 4c).

After the ductal, acinar, Ngn3+, and hormone+ populations had been accounted for, there still remained one population that eluded classification based on known marker genes. This population was distinguished from all other epithelial populations by high-level expression of the E26 transformation-specific transcription factor Fev, previously shown to be expressed within the developing pancreas but not described as a marker of a distinct epithelial population26 (Fig. 4a, b and Supplementary Data 5). This Fev+ population expressed genes marking endocrine lineage cells, such as Paired box 4 (Pax4), chromogranins A/B (Chga/b) and Neurod113 (Supplementary Fig. 4d), but not mature endocrine markers, such as Insulin1 (Ins1) or Gcg, or the transitory early endocrine lineage marker, Ngn3 (Fig. 4b, c and Supplementary Data 5). Pairwise comparison between the Fev+ and Ngn3+ clusters identified 99 genes more highly expressed in Fev+ and 87 more highly expressed in Ngn3+ cells, suggesting that they are distinct populations (Fig. 4d). This Fev+, Ngn3−, hormone− population will henceforth be referred to as the FevHi population. Pathway analysis of the Ngn3+ and FevHi populations revealed enrichment of cell cycle and Notch signaling pathways in Ngn3+ cells (Fig. 4e and Supplementary Data 6), likely reflecting the exit of Ngn3+ progenitors from the cell cycle27 and the role of Ngn3 in Notch signaling28. The FevHi cluster expressed genes in pathways related to serotonin, insulin, Activating transcription factor 2 (Atf2), and sphingosine-1-phosphate signaling, which have been reported to regulate endocrine differentiation29,30. This relationship to serotonin is consistent with prior work establishing Fev as a critical transcription factor in serotonergic neurons26,31.

Further subclustering of all 661 cells within the endocrine lineage revealed additional sub-groups of Fev-expressing cells. The first was marked by high expression of Pax4 and Runx1 Translocation Partner 1 (Runx111) and lower levels of Ngn3. The second was marked by Chgb and Vimentin (Vim) (Fig. 4f, Supplementary Fig. 4e, f, and Supplementary Data 7). Therefore, our analysis proposed the existence of multiple intermediate states, marked by Fev, within the endocrine lineage. The Fev gene was also expressed at lower levels in a subset of the hormone-producing alpha, beta, and epsilon cell populations, which will collectively be referred to as hormone+//FevLo populations (Fig. 4b).

Given that the Fev+ populations expressed endocrine lineage genes, we utilized pseudotime ordering25 to test the hypothesis that both Fev+ populations were lineage-related to the Ngn3+ progenitors that give rise to the endocrine compartment of the pancreas32. This de novo reconstruction of the developmental trajectory placed both the Fev+/Pax4+ and FevHi/Chgb+ cells between Ngn3+ endocrine progenitors and alpha and beta cells (Fig. 4g and Supplementary Fig. 4g), suggesting that FevHi cells comprise a progenitor stage following Ngn3 expression and before hormone acquisition. The Fev+/Pax4+ population was placed
closer in pseudotime to the Ngn3+ population and was followed by the FevHi/Chgb+ population (Fig. 4g), indicating that the former represents an earlier cell state. Unlike alpha and beta cells, epsilon cells were found throughout the trajectory populated by the Fev+/Pax4+ and FevHi/Chgb+ populations (Fig. 4g), possibly reflecting their function as multipotent progenitor cells for alpha and gamma lineages during development33.

To validate these lineage relationships, we performed an in vivo lineage trace of Ngn3+ cells. In E14.5 Ngn3-Cre; ROSA26mTmG mouse pancreata, where lineage-traced cells are membrane-green fluorescent protein (GFP+)34, approximately 20% of all Ngn3-lineage-traced cells were identified as the Fev population by the presence of Fev and the absence of both Ngn3 and the pan-differentiated endocrine cell marker Islet1 (Isl1) (Fig. 5a, e, yellow
and VSM-related mesenchymal populations based on pseudotime ordering in where we captured 584, 1267, and 1837 endocrine cells at E12.5, Data 8 mesenchymal clusters. Color intensity indicates level of expression. development36 in vivo, it is more synchronized in the hESC endocrine differentiation progresses as a wave throughout (blue arrows), consistent with a model in which Fev Hi cells 1 and 2) to increase our capture of epithelial cells and then epithelial cells (Supplementary Fig. 5b, c). Given the high numbers depletion of CD140a cell RNA-sequencing experiments at E12.5, E14.5, and E17.5 after (v1) runs, we reperformed an entirely new (version 2) set of single-cell data at E14.5 and E17.5 version 1 points (Supplementary Fig. 5a). To capture more epithelial cells and then captured comparatively fewer epithelial cells at E12.5 and E17.5 version 1, we still identify the FevHi cells at both time- points (Supplementary Fig. 5a). To capture more epithelial cells and account for those that were missing from E12.5 and E17.5 version 1 (v1) runs, we reperformed an entirely new (version 2) set of single-cell RNA-sequencing experiments at E12.5, E14.5, and E17.5 after depletion of CD140a+ mesenchymal cells in order to enrich for epithelial cells (Supplementary Fig. 5b, c). Given the high numbers of red blood cells at E17.5, we ran two wells of E17.5 cells (replicates 1 and 2) to increase our capture of epithelial cells and then aggregated the datasets. We first analyzed the exocrine compartment and identified acinar, ductal, and proliferating populations of both at all timepoints (Supplementary Fig. 5d and Supplementary Data 8–10). We then focused on the endocrine compartment, where we captured 584, 1267, and 1837 endocrine cells at E12.5, E14.5, and E17.5, respectively. We found similar gene-expression topologies as in our v1 dataset but gained additional resolution with increased cell numbers and transcriptomic coverage (Supplementary Fig. 5e and Supplementary Data 11–13).

To analyze how endocrine populations change over time, we merged all three v2 timepoints into one dataset using canonical correlation analysis39. We correlated the v2 dataset to the v1 dataset and could identify all populations present in the v1 dataset (Supplementary Fig. 5f and Supplementary Data 14). We also found additional populations, including a cluster characterized by decreased expression of Fev and increased expression of Pdx1 and Mafb, genes with known roles in endocrine lineage decisions (Fig. 5h and Supplementary Fig. 5g). This Pdx1+/Mafb+ population correlates most strongly with the FevHi/Chgb+ population, as well as both the alpha and beta cell populations in the v1 dataset (Supplementary Fig. 5f). We also found a second beta cell population characterized by increased expression of Ins1 and Ins2 and lower expression of Pdx1, perhaps representing more mature beta cells (Supplementary Fig. 5g). Indeed, this second beta cell group is almost entirely comprised of cells from the E17.5 timepoint (Fig. 5i). To examine how these populations shift over developmental time, we calculated the proportion of these populations at each timepoint (Fig. 5j). We found shifts in cell proportions that match those reported in literature, such as a high proportion of alpha cells early in development at E12.5 and increasing proportions of beta and delta cells at later time-points40. The Ngn3− population decreased over time, while the Fev+/Pax4+, Fev+Chgb+, and Pdx1+/Mafb+ populations peaked at E14.5, consistent with previous studies that reported peak Ngn3 expression at approximately E14.5 and its subsequent downregulation as differentiation into endocrine lineage ensues40. At E17.5, we also found an increasing proportion of proliferating endocrine cells, presumably those responsible for the expansion of endocrine cell mass in later embryonic development41. These results from the larger v2 dataset confirm our initial findings from the v1 dataset and add additional resolution to the endocrine populations during pancreatic development.

**Endocrine dynamics over developmental time.** Although we had captured comparatively fewer epithelial cells at E12.5 and E17.5 than at E14.5, we could still identify the FevHi cells at both time-points (Supplementary Fig. 5a). To capture more epithelial cells and account for those that were missing from E12.5 and E17.5 version 1 (v1) runs, we reperformed an entirely new (version 2) set of single-cell RNA-sequencing experiments at E12.5, E14.5, and E17.5 after depletion of CD140a+ mesenchymal cells in order to enrich for epithelial cells (Supplementary Fig. 5b, c). Given the high numbers of red blood cells at E17.5, we ran two wells of E17.5 cells (replicates 1 and 2) to increase our capture of epithelial cells and then aggregated the datasets. We first analyzed the exocrine compartment and identified acinar, ductal, and proliferating populations of both at all timepoints (Supplementary Fig. 5d and Supplementary Data 8–10). We then focused on the endocrine compartment, where we captured 584, 1267, and 1837 endocrine cells at E12.5, E14.5, and E17.5, respectively. We found similar gene-expression results from the larger v2 dataset confirm our initial findings from the v1 dataset and add additional resolution to the endocrine populations during pancreatic development.

**Lineage decisions within the endocrine compartment.** As the in vivo lineage tracing data had revealed that the FevHi population is derived from the Ngn3+ population, we hypothesized that the FevHi population could then function as a progenitor for the endocrine populations of the developing pancreas. We utilized a Fev-Cre; ROSA26mTmG lineage tracing strategy to label Fev-expressing cells and their progeny. We found that 100% of alpha, beta, and delta cells, 90.1% of gamma cells, and 23.2% of epsilon cells were lineage-traced in E14.5 pancreas (Fig. 6a–e). These proportions of lineage labeling held true later in development (E17.5) and in adulthood (6 weeks) (Supplementary Fig. 6 and 7). Epsilon cells are rare in the adult pancreas33 and still exhibited only partial lineage tracing in E17.5 pancreas (47.8% traced) (Supplementary Fig. 6e). These results demonstrate that the
majority of endocrine cells pass through a Fev-expressing stage during development.

We next combined this lineage tracing approach with single-cell RNA sequencing to identify transcriptional regulators of endocrine differentiation. FACS sorting was used to enrich for Fev-expressing cells and their progeny (membrane GFP+) from Fev-Cre; ROSA26<sup>mTmG</sup> pancreata at E14.5 (Fig. 6f, g). All expected endocrine populations were identified in the resulting single-cell dataset (Fig. 6h, i). In addition, we found that eGFP reads mapped to all endocrine populations except the Ngn3+ population (Fig. 6i), further confirming that Fev expression turns on after Ngn3.
We next set out to model the lineage relationships among the endocrine cells and identify transcriptional regulators of differentiation. Pseudotime ordering identified a trajectory that began with Ngn3+ cells, transitioned into Fev+ cells, and then split into two main branches (Fig. 7a; see similar branching pattern in analysis of our first v1 dataset, Supplementary Fig. 8a). The termini of the branches were populated by differentiated beta and alpha cells, suggesting that the branches represent a transition from a progenitor to fully differentiated hormone+ cell (Fig. 7a).

We next used Monocle’s branched expression analysis modeling (BEAM) to identify the genes that distinguish the paths along the two branches to either alpha or beta cells. We found gene clusters that were upregulated along different segments of the pseudotime trajectory (Fig. 7b and Supplementary Data 15) and performed pathway analysis to identify pathways enriched at each stage of pseudotime (Supplementary Fig. 8c and Supplementary Data 16). Genes upregulated at the beginning of pseudotime in gene cluster 2, included early markers of endocrine differentiation, such as Sox4 and Ngn3 (Fig. 7b). Fev was in gene cluster 6 and increased in both branches before ultimately decreasing in expression at the branch termini (Fig. 7b, c). Gene cluster 6 also included other genes expressed within the FevHI population, including Chdh, Vim, and Chgb (Fig. 7b, c and Supplementary Fig. 8b). We found branch-specific clusters that included known markers of beta (Ins1) and alpha (Gcg) cells and known differentiation regulators of alpha (Ark, Pou3f4, Irx1, Sclk38a5, and Tmem27) and beta (Pdx1, Pak3, and Nkx6-1) cells (Fig. 7c and Supplementary Fig. 8b). These clusters also contained genes that were enriched in either the alpha or beta branch but were expressed before acquisition of hormone expression (Supplementary Fig. 8b). Within the alpha cell branch, Peg10, Smarca1, Auts2, and Wnk3 increased in expression before upregulation of Gcg occurred (Supplementary Fig. 8b). Peg10 and Auts2 have roles in differentiation46,47 and migration48, but a role in endocrine differentiation has not been described. As a regulator of chromatin states and an adult human alpha cell marker49, Smarca1 may be involved in the epigenetic regulation of alpha cell differentiation. Within the beta cell branch, Gng12, Tssc4, Ecc1, Tmem108, Wip1, and Papss2 increased in expression before upregulation of Ins1 (Supplementary Fig. 8b). To our knowledge, a role in endocrine lineage decisions have not been described for these beta branch-specific genes. We found a similar endocrine differentiation trajectory by an orthogonal method that uses force-directed layouts to visualize gene topologies and infer lineage relationships within single-cell data50,51 (Supplementary Fig. 8d). We hypothesize that the genes identified by the analysis above may represent regulators of the differentiation of an endocrine progenitor to a fully differentiated hormone-expressing cell.

To validate our pseudotime results, we performed ISH for markers that defined each branch of the trajectory. First, we confirmed the expression of Peg10 and Gng12 within the FevHI population (Fig. 7d, e, indigo and teal gradient arrows), validating the expression of these genes in a stage before hormone acquisition. We also validated the enrichment of Peg10 and Gng12 in alpha and beta cells, respectively (Fig. 7f, g, solid indigo and teal arrows). First, 95.8% of beta cells expressed Gng12 (n = 46 cells, 6 pancreata), while 30.5% expressed Peg10 (n = 71 cells, 7 pancreata) (Fig. 7f and Supplementary Fig. 9a). Additionally, 100% of alpha cells expressed Peg10 (n = 31 cells, 6 pancreata), while only 5.4% expressed Gng12 (n = 32 cells, 4 pancreata) (Fig. 7g and Supplementary Fig. 9b). The lineage relationships generated by pseudotime ordering, combined with the validation in vivo, lead us to hypothesize that the Fev+/Peg10+ cells are fated toward an alpha cell identity and Fev+/Gng12+ cells toward a beta cell identity (Fig. 7h). These results suggest that lineage allocation of endocrine progenitors toward alpha or beta cell fates may occur after the onset of Fev expression.

Discussion

The mesenchyme is critical for epithelial specification and proliferation throughout pancreatic development52–54, yet the individual cell types responsible for these processes remain unidentified. Our single-cell dataset has enabled the identification of multiple mesenchymal subpopulations and gene candidates for regulating epithelial–mesenchymal interactions. Secreted factors, such as mesothelial-derived Fgf9, may play a similar role in the pancreas as in the lung, where it regulates mesenchymal cell proliferation and vascular formation55. Additionally, secretion of Wnt antagonists by cluster 5 may regulate Wnt signaling in the developing pancreas, influencing processes such as epithelial specification, expansion, and exocrine development56. Future work can focus on uncovering the functions of these individual mesenchymal populations in the development, physiology, and pathology of the pancreas.

With the cell types of the mesenchyme now enumerated and their markers identified, we can begin to elucidate the maturation and lineage relationships of the pancreatic mesenchymal compartment. Our time-course data have provided evidence of maturation within the mesothelial population. Genes such as Pitx2, kalikrein 13 (Klk13) and 8 (Klk8), were differentially expressed in younger, E12.5 mesothelial cells. Pitx2 regulates differentiation in multiple systems24,57,58, and the kalikrein family are serine proteases involved in the degradation of extracellular matrix and adhesion molecules59. Expression of these genes suggests that the E12.5 mesothelial population may be primed for migration and differentiation. In contrast, the E17.5 mesothelial population expressed genes related to barrier or immune function, such as dermokine (Dmkn)60,61, bone marrow stromal antigen 2 (Bst2), and retinoic acid receptor responder 2 (Rarres2)62. These results suggest stage-dependent roles for the mesothelium throughout development.

The different roles for the mesothelium across time are also evident from our pseudotime analysis, which proposes that the mesothelium serves as a progenitor for other mesenchymal cell types during development. The mesothelium is a critical mesenchymal progenitor population in other organs, such as the heart, intestine, lung, and liver16–19. Our data suggest that mesothelial progenitor activity occurs at E12.5 or earlier during pancreatic development, consistent with other organ systems11,16,18. Indeed, a recent study identified that parietal mesothelial cells can function as progenitor cells prior to pancreatic specification63. The transcriptomic information obtained
by this study will allow the development of tools to target individual populations within the mesenchyme and perform lineage tracing, ablation, and expression studies in vivo. Furthermore, this developmental dataset can be compared to mesenchymal population dynamics during adult disease progression, where aberrant recapitulation of developmental pathways can lead to disease states in the pancreas\textsuperscript{64,65}. Thus, this dataset is a broad resource for the implementation of future studies in pancreatic mesenchymal biology.

Within the epithelial compartment, our identification of a Fev\textsuperscript{Hi} endocrine progenitor population provides increased resolution of endocrine differentiation. The relative timing of expression of canonical endocrine lineage genes can now be mapped onto these additional differentiation stages. Several lines of evidence suggest that the gene Fev may be a direct target of Ngn3: Fev is the transcription factor most strongly expressed in Ngn3\textsuperscript{+} endocrine progenitors\textsuperscript{66}, and Ngn3 knockout embryos do not express Fev in the developing pancreas\textsuperscript{26}. Known target genes
of Ngn3, such as Pax6\(^+\) and Runx1\(^+\), are expressed by the early-stage Fev\(^+\)/Pax4\(^+\) population. Additionally, Pax6 is upregulated within the Fev\(^+\) population. Although Chga and Chgb are often utilized as markers of differentiated endocrine lineages, we found that both are expressed in the Fev\(^+\) population prior to hormone acquisition. This result is consistent with previous work that identified Chga\(^+\), hormone- \(^+\) cells in rodent pancreatic development\(^{69}\). The Fev\(^+\) cell stage likely represents the cell stage during endocrine differentiation preceding specialized hormone production and may now serve as a cellular landmark for understanding endocrine lineage gene-expression dynamics.

The gene Fev has been previously studied mainly in serotonergic neurons, where it is a master transcriptional regulator required for cellular differentiation, maturation, and serotonin synthesis\(^{31}\). Fev switches transcriptional targets from differentiation genes during development to maturation genes post-natally in serotonergic neurons\(^{70}\). In an insulinoma cell line, Fev directly binds to the regulatory regions of serotonergic genes, such as Tph1, Tph2, Ddc, Slc18a2, and Slc6a4, as well as the Ins1 promoter itself\(^{26}\). Future ChIP-seq studies of embryonic pancreas will globally identify direct targets of Fev and Fev-regulated transcriptional networks in developing endocrine cells.

Using genetic lineage tracing in vivo, we have demonstrated that the majority of endocrine cells in the developing pancreas transit through a Fev-expressing stage, and that Fev-lineage cells contribute not only to embryonic, but also to adult pancreatic endocrine cells. The fraction of epsilon cells not derived from a Fev-lineage may represent the subset of Ghrl\(^+\) cells previously reported to give rise to the ducal and exocrine lineages\(^{53}\). As all adult gamma cells are Fev-lineage labeled, the small subset of gamma cells not lineage traced during pancreatic development may represent those that do not persist in the adult pancreas.

Further highlighting the relevance of Fev\(^+\) progenitors during pancreatic development, our pseudotime analysis revealed that Fev-expressing cells may be pre-specified towards an alpha or beta cell fate. As expected, we found expression of Ins1 and Gcg at the termini of the beta and alpha branches, and upregulation of Pdx1 and Arx, which are known regulators of endocrine cell fate decisions, earlier in pseudotime. In addition, our pseudotime analysis identified genes enriched along the alpha or beta branch and expressed prior to upregulation of hormones. These genes warrant further study as potential regulators of the acquisition of alpha or beta cell identity.

For the eventual application of this knowledge to human therapeutics, the findings in the murine model must be validated in human pancreatic development. Our staining of human fetal pancreas identified the analogous FEV\(^+\) population, consistent with our findings in murine pancreata. Directed differentiation of hESCs toward endocrine cell fates will provide a platform for modeling and manipulating the putative lineage regulators found in this study. Indeed, we have identified a Fev\(^+\) population within hESC-derived endocrine progenitor cells. Deeper knowledge of these lineage decisions may substantially improve directed differentiation efforts to efficiently generate functional beta cells for cellular replacement therapy for people with diabetes. This study highlights the power of combining single-cell transcriptomic information with in vivo lineage tracing to reconstruct developmental trajectories within cellular compartments. Discovery of populations and their lineage relationships will promote identification of the mechanisms that drive lineage decisions and commitment.

**Methods**

**Mice.** All mouse procedures were approved by the University of California, San Francisco (UCSF) Institutional Animal Care and Use Committee. Mice were housed in a 12-h light-dark cycle in a controlled temperature climate. Noon of the day of vaginal plug was considered embryonic day 0.5.

Timed-pregnant Swiss Webster mice were obtained from Charles River Laboratories. Ngn3-Cre\(^{21}\) (a gift from Dr. Matthias Hebrok), Fev-Cre\(^{22}\) (The Jackson Laboratory 012712), and ROSA26mTmG\(^{34}\) (the Jackson Laboratory 007676) mice were maintained on a C57BL/6 background. The Cre transgene was genotyped using the following primers: GGGGGCCATAGTGCAAGTT and CGGTCAGCAGGATGTTC.

**Human tissue procurement and isolation.** Human fetal pancreata were harvested from post-mortem fetuses at 23 weeks of gestation with permission from the ethical committee of UCSF. Tissue was fixed in 4% paraformaldehyde overnight at 4°C. After three washes in 1X phosphate-buffered saline (PBS), tissue was either frozen in 4% paraformaldehyde overnight at 4°C. After three washes in 1X phosphate-buffered saline (PBS), tissue was either fixed in 4% paraformaldehyde overnight at 4°C or placed in 40% ethanol then 70% ethanol at cutting temperature (OCT) compound, or placed in 40% ethanol then 70% ethanol before paraffin embedding. Sections measuring 8 μm were cut on the cryostat or microtome. In situ hybridization and immunofluorescence were then performed as described below.

Adult human islets were isolated from cadaveric donor tissue by the UCSF Islet Production Core with permission from the ethical committee of UCSF. Mice were used for the generation of hESC-derived UCIs. Pluripotent HUES8 cells were maintained as spherical clusters in suspension in mTeSR-1 (StemCell Technologies) in 500 mL spinner flasks (Corning, VWR) on a magnetic stir plate (Dura-Mag) within a 37°C incubator at 5% CO\(_2\), 100% humidity, and a rotation rate of 70 rpm. Cells were screened for mycoplasma contamination using the MycoProbe Mycoplasma Detection Kit (R&D Systems), according to the manufacturer’s instructions.

**Embryonic stem cell culture and differentiation.** The hESC line HUES8 was obtained from Harvard University and used for the generation of hESC-derived UCIs. Pluripotent HUES8 cells were maintained as spherical clusters in suspension in mTeSR-1 (StemCell Technologies) in 500 mL spinner flasks (Corning, VWR) on a magnetic stir plate (Dura-Mag) within a 37°C incubator at 5% CO\(_2\), 100% humidity, and a rotation rate of 70 rpm. Cells were screened for mycoplasma contamination using the MycoProbe Mycoplasma Detection Kit (R&D Systems), according to the manufacturer’s instructions.

hESC-derived endocrine progenitor cells were generated as previously described\(^{35}\). In brief, HUES8 cells were seeded into a spinner flask at a
concentration of $8 \times 10^5$ cells/mL in mTeSR-1 media with 10 μM Rock inhibitor Y27632 (StemCell Technologies) to allow formation of spherical clusters. Differentiation was initiated 72 h later. Differentiation was achieved in a step-wise fashion using the following growth factors and/or small molecules: definitive endoderm (Stage 1) (1 day of 100 ng/mL Activin A (R&D Systems) and 14 μg/mL of CHIR99021 (Stemgent); 2 days of 100 ng/mL Activin A); gut tube endoderm (Stage 2) (3 days of 50 ng/mL KGF (Peprotech)); early pancreatic progenitors (Stage 3) (1 day of 200 nM LDN193189 (Fisher Scientific), 50 ng/mL KGF, 0.25 μM Sant-1 (Sigma), 0.25 μM Retinoic Acid (Sigma), 500 nM PdbU (EMD Biosciences); 1 day of 50 ng/mL KGF, 0.25 μM Sant-1, 2 μM Retinoic Acid, 500 nM PdbU); later pancreatic progenitors (Stage 4) (5 days of 50 ng/mL KGF, 0.25 μM Sant-1, 0.1 μM Retinoic Acid); endocrine progenitors (Stage 5) (4 days of 0.25 μM Sant-1, 0.1 μM Retinoic Acid, 1 μM XXI (EMD Millipore), 10 μM Alk5i (Axon), 1 μM T3 (EMD Biosciences), 20 ng/mL Betacellulin (Fisher Scientific); 3 days of 25 nM Retinoic
Acid, 1 μM XXI, 10 μM Alk5i, 1 μM T3, 20 ng/ml Betacellulin); BLCs (Stage 6) (6 days of 10 μM Alk5i; 1 μM T3). Successful differentiation was assessed at Stages 1, 3, 4, 5, and 6 via immunofluorescence or FACS for stage-specific marker genes. To measure the expression of PEV at various stages of human endocrine differentiation, all pools of clusters were removed from the flask and analyzed at several timepoints: after 5 days in Stage 5 ("mid-stage endocrine progenitors"), and after 7 days in Stage 5 ("late-stage endocrine progenitors"), and after 5 days at the BLC stage. As a comparator, pluripotent, undifferentiated hESCs were harvested, cultured in mTeSR-1, and then stained with primary and secondary antibodies.

Immunofluorescence. Embryonic mouse pancreata were dissected and fixed in cold 1X PBS and then mounted in Fluoromount-G mounting medium (SouthernBiotech). Slides were stored at 4°C.

For immunofluorescence on cryosections, slides were removed from –80°C storage and allowed to reach RT. Sections were stained in 1X PBS three times and permeabilized in 0.3% Triton X-100 in PBS (PBT) for 1 h and then stained overnight at 4°C using the following primary antibodies: Acta2 (1:200, Abcam ab22107), Cav1 (1:200, Abcam ab2910), Chromogranin A (1:100, Abcam ab15160), E-cadherin (1:200, BD Transduction Lab 610182), Glucagon (1:100, Abcam ab82270), Insulin (1:50, DAKO A0564), Vimentin (1:200, Abcam ab2054), and WT1 (1:100, Abcam ab9901). All antibodies were validated by the manufacturer. The next day, sections were washed three times in 0.1% Tween 20 in 1X PBS and then incubated with species-specific Alexa Fluor 488, 594-, or 647-conjugated secondary antibodies (1:500, Jackson Immunoresearch) and DAPI in 5% NDS in 0.2% PBT for 1 h at RT. Sections were washed three times in 0.1% Tween 20 in 1X PBS, rinsed in 1X PBS, and then mounted in Fluoromount-G mounting medium (SouthernBiotech). Slides were stored at 4°C.

For immunofluorescence on cryosections, slides were removed from –80°C storage and allowed to reach RT. Sections were stained in 1X PBS three times and permeabilized in 0.5% PBT for 10 min at RT. Tissue sections were blocked in 5% NDS and, if needed, Mouse-on-Mouse IgG blocking reagent in 0.1% PBT for 1 h at RT. Tissue sections were then stained with primary and secondary antibodies.

Quantification of cell proportions. Quantification of pancreata was performed by manual counting using ImageJ software. Cell populations present at less than 1% in manual counting were deleted or excluded from further analysis.

Dissociation and FACS of embryonic pancreas. Embryonic mouse pancreata were dissected and placed in 1X PBS on ice, then dissociated into single cells using TrypLE Express dissociation reagent (Thermo Fisher) at 37°C with pipet triturations. After 5-min intervals during incubation, for v1 datasets, E12.5 pancreata were dissociated for 10 min, E14.5 pancreata for 15 min, and E17.5 pancreata for 30 min. For batch 1, we pooled 14 E14.5 pancreata from one litter. For batch 2, which was collected on a different day, we pooled tissue from each timepoint separately: 18 E12.5 pancreata from two litters, 11 E14.5 pancreata from one litter, and 8 E17.5 pancreata from one litter. Dissociations were neutralized with FACS buffer (10% FBS + 2 mM EDTA in phenol-red free HBSS). Dissociated cells were passed through a 30 μm cell strainer and stained with Sytox live/dead stain (Thermo Fisher). Stained cells were washed twice in FACS buffer and then sorted using a BD FACS Aria II. After size selection to remove doublets, all live cells were collected.

For v2 10X datasets, we pooled tissue from each time point separately, each performed on a different day: 14 E12.5 pancreata from one litter, 13 E14.5 pancreata from one litter, and 13 E17.5 pancreata from one litter. For the E14.5 Fev-Cre; ROSA26mTmG sample, we pooled three pancreata from one litter. Dissociations were performed as described above. Doublets were removed using an additional 21,000 TdTomato cells (8 x E14.5, 1 x E12.5, and 1 x E17.5), and after 5 days at the BLC stage. As a comparator, pluripotent, undifferentiated hESCs in mTeSR-1, as described above, were also collected.

Single-cell capture and sequencing. To capture individual cells, we utilized the Chromium Single Cell 3' Reagent Version 1 Kit (10X Genomics)44. For batch 1, 12,800 cells from E14.5 pancreata were loaded into one well of the 10X chip, while for batch 2, 18,000 cells per time point were each loaded into their respective wells to produce Gel Bead-in-Eмуulsions (GEMs). GEMs underwent reverse transcription to barcode RNA before cleanup and CDNA amplification. Libraries were prepared with the Chromium Single Cell 3' Reagent Version 1 Kit. Each sample was sequenced on 2 (Batch 1) or 1 (Batch 2) lanes of the HiSeq2500 (llumina) in Rapid Run Mode with paired-end sequencing parameters: Read1, 98 cycles; Index1, 14 cycles; Index2, 8 cycles; and Read2, 10 cycles.

The CD140a-depleted E12.5, E14.5, and E17.5 datasets were sequenced on the Chromium Single Cell 3' Reagent Version 2 kits (10X Genomics). In total, 27,000 cells were loaded onto their respective wells and underwent the same processing as the Version 1 kits, according to manufacturer instructions for Version 2 kits. Libraries were sequenced on the NovaSeq (llumina) with the same sequencing parameters as above.

Single-cell analysis for the v1 datasets, we utilized CellRanger v1.1.0 software for v1 datasets and v2.1.0 for v2 datasets with default settings for de-multiplexing, aligning reads to the mouse genome (10X Genomics pre-built mm10 reference genome) with STAR and counting unique molecular identifiers (UMIs) to build transcriptomic profiles of individual cells. For the v1 datasets, gene-barcode matrices were analyzed with the R package Seurat v4.1, using the
online tutorial as a guide. We first performed a filtering step, retaining only the cells that expressed a minimum of 200 genes and only the genes that were expressed in at least 3 cells. A large number of cells did not meet this threshold in the E17.5 time point and were determined to be red blood cells by the high expression of hemoglobin genes. Variable genes were determined by mean-variance relationship to identify highly expressed and variable genes with the Seurat function MeanVarPlot with default settings. UMI counts were log-normalized, and linear regression was performed with RegressOut to account for differences in the number of UMIs between cells. PCA was then utilized to determine sources of variability in the dataset with PCAfast. Significant PCs were determined based on the Scree plot and utilized for Seurat’s graph-based clustering algorithm (function FindClusters) with default parameters, except for the resolution parameter. To vary cluster numbers, the resolution parameter in FindClusters was adjusted from 0.6 to 3.0, and resulting clusters analyzed as
follows. Clusters were visualized by t-SNE with Seurat’s RunTSNE function using default settings27. Differentially expressed genes were determined with the FindAllMarkers function, which uses a bimodal likelihood ratio test6. We confirmed differential gene-expression analysis with the Wilcoxon rank sum test and Model-based Analysis of Single-cell Transcriptomics (MAST)9 utilizing Seurat v2’s FindMarkers function with default settings. These tests calculate adjusted p values for multiple comparisons. To determine the final number of clusters, clusters were required to have at least nine significantly (p < 0.05) differentially expressed genes with a twofold change in expression in comparison to all other clusters. Clusters were manually curated for differential gene expression, and those that did not meet this threshold were manually merged with the nearest cluster based on the phylogenetic tree from Seurat’s Build-ClusterTree. In some cases, clusters met the nine-gene threshold but appeared to have very similar differentially expressed genes to another cluster. It is likely a result of the comparison of individual clusters against all other clusters in determining differentially expressed genes. In these cases, a pairwise comparison between the two clusters was performed and the same nine-gene threshold applied. An exception to the nine-gene threshold was made to annotate the proliferating populations in the early stages of the cell cycle within the E14.5 mesenchymal analysis (Fig. 4, cluster 8). Additionally, cluster 10 in the E14.5 mesenchymal dataset did not meet the nine-gene threshold. Rather, clusters 1–9 had distinct transcriptomic signatures (with at least nine differentially expressed genes) that distinguished them from cluster 10. Lists of at least two-fold differentially expressed genes for individual analyses are provided in Supplementary Data 2.

For v2 datasets, Seurat v2.2 and v2.3 was utilizing to perform the analysis. Cells with fewer than 200 genes and genes expressed in fewer than 3 cells were removed, as above. UMI counts were normalized with NormalizeData using default settings. Variable genes were determined with FindVariableGenes, using the following cutoffs suggested by the online tutorial (klow = 10, khigh = 30, the number of top genes from each cluster, and v. pseudotime trajectory).

Custom genome build. The custom genome for alignment of reads to eGFP and TdTome sequences from the mTmG mouse line was created according to instructions provided by 10X Genomics reference support (https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines/latest/advanced/references). eGFP and TdTome sequences were concatenated to the mm10.21.0 reference genome (FASTA file) provided by 10X Genomics. eGFP and TdTome annotations were then concatenated to the mm10 annotations (GTF file) provided by 10X Genomics. The cellranger mkref command was then utilized with the genome and annotations with eGFP and TdTome, as described in the above link.

Pathway analysis. Pathway analysis and calculation of associated p values were performed using the ConsensusPathDB overrepresentation analysis for pathway-based sets category (http://cpdb.molgen.mpg.de/)29.

Aggregating E17.5 v2 datasets. E17.5 technical replicates from the v2 dataset were aggregated with Cellranger v2.1, utilizing the agg function with default settings. The aggregated dataset was used for analysis and merging with the E12.5 and E14.5 v2 datasets.

Subclustering and merging datasets. Subclustering was performed by isolating clusters of interest with the Seurat function SubsetData and reanalyzing as described above (identification of variable gene, regression, and determination of significant PCs). Cells were classified as epithelial based on the expression of E-cadherin (Cdh1) and other known epithelial population markers. Cells that were Cdh1+, Vim−, and Collagen3a1 (Col3a1)+ were classified as mesenchymal. Multiple batches were merged with the MergeSeurat function. The v2 E14.5 batch 1 and batch 2 clusters were robust to the sampling differences between batches as evidenced by the contribution of cells from both batches to each cluster (Supplementary Fig. 2b). We found high correlation of cell type proportion between batches in all populations except the exocrine compartment (acinar and ductal) (Supplementary Fig. 2c), possibly due to technical challenges of pancreatic dissociation. Within each cluster, batch 1 cells correlated most highly with those of batch 2 contained in the same cluster, indicating proper cluster calling with the merged datasets (Supplementary Fig. 2d).

For v2 datasets (E12.5, E14.5, and E17.5), multiple canonical correlation analysis (multCCA) from Seurat v2.3 was utilized to merge the epithelial datasets30. The top 1000 most highly variable genes that were variable in at least 2 datasets were used for the alignment, as recommended in the Seurat tutorial. The shared correlation strength of each CC was measured with Seurat’s MetageneBicorPlot, and those before the drop-off were used for alignment, analogous to the Scree plot in choosing significant PCs. We then aligned the datasets with AlignSubspace and ran an integrated t-SNE and clustering analysis, as outlined in the Seurat tutorial. Clusters were required to have 9 significantly differentially expressed genes as described above. Clusters with similar differentially expressed genes were verified with pairwise comparisons to the most related clusters (based on BuildClusterTree) and merged if they did not meet the pairwise nine-gene threshold. The Beta 2 cluster in the v2 endocrine merged timecourse analysis. In the v2 datasets, rare cells (4 cells in E12.5 and 13 cells in E14.5 endocrine datasets) with high levels of hemoglobin gene expression were removed from the analysis.

Downsampling analysis. To determine if the sequencing depth was sufficient for calling clusters, downsampling analysis was performed for the v1 E14.5 batch 1 dataset. Reads were randomly downsampled from the 10x Cellranger bam file output to a specified percentage, then grouped based on UMI to generate a count profile for each cell. The number of genes with greater than 0 counts was then calculated. UMI downsampling was performed with the SampleUMI function. A new Seurat object was created with the downsampled matrix and reanalyzed as above.

The number of UMIs/cell was downsampled from an average of 4600 UMIs/cell in the full dataset to 200 UMIs/cell, and the median number of genes/cell and clustering robustness was then calculated. Clustering robustness was determined as the percentage of cells within the same cluster, with clusters required to maintain at least nine genes with a twofold change in expression in comparison to all other clusters. Within this dataset, robust clustering was maintained as long as down to 500 UMIs/cell, when the percentage of cells in the same cluster began to climb, indicating collapsing of individual clusters. Both of these downsampling analyses indicate that sufficient sequencing depth was reached.

Pseudotemporal ordering. We utilized Monocle 2.6.12 to order cells in pseudotime based on their transcriptomic similarity. For v1 timecourse datasets, batch-corrected values and variable genes from Seurat analysis were used as input, utilizing the gaussianf expressionFamily, and clusters were projected onto the minimum spanning tree after ordering.
For the Ftv-lineage-traced dataset, UMI counts and variable genes from the Seurat analysis were used as input, utilizing the negBinom expressionFamily. To find genes differentially expressed across the branch point in the trajectory, we used monocyte's internal BEAM analysis and selected genes with an FDR cutoff of 0.001. Gene expression patterns were plotted with plot_genes_branch Heatmap and plot_multiple_branches_pseudotime.

Data availability

RNA sequences for the single-cell RNA-sequencing analyses reported in this paper have been deposited in the GEO database under accession code GSE101099. The authors declare that all data supporting the findings of this study are available within the article and its supplementary information files or from the corresponding author upon reasonable request.

References

1. Shih, H. P., Wang, A. & Sander, M. Pancreas organogenesis: from lineage determination to morphogenesis. Annu. Rev. Cell Dev. Biol. 29, 81–105 (2013).
2. Pan, F. C. & Wright, C. Pancreas organogenesis: from bud to plexus to gland. Dev. Dyn. 240, 530–565 (2011).
3. Sussa, Y. et al. Gastrin: a distinct fate of neurogenin 3 positive progenitor cells in the embryonic pancreas. PLoS One 8, e70397 (2013).
4. Qiu, W.-L. et al. Deciphering pancreatic islet β cell and α cell maturation pathways and characteristic features at the single-cell level. Cell Metab. 25, 1194–1205.e4 (2017).
5. Zeng, C. et al. Pseudotemporal ordering of single cells reveals metabolic control of postnatal β-cell proliferation. Cell Metab. 25, 1160–1175.e11 (2017).
6. Dorrell, C. et al. Human islets contain four distinct subtypes of β cells. Nat. Commun. 7, 11756 (2016).
7. Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. & Regev, A. Spatial reconstruction of single-cell gene-expression data. Nat. Biotechnol. 33, 495–502 (2015).
8. McDavid, A. et al. Data exploration, quality control and testing in single-cell qPCR-based gene-expression experiments. Bioinformatics 29, 461–467 (2012).
9. Finak, G. et al. MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA-sequencing data. Genome Biol. 16, 278 (2015).
10. Kanamori-Katayama, M. et al. LRRK4 and UKB3 are markers of primary mesothelial cells. PLoS One 6, e25391 (2011).
11. Winters, N. & Bader, D. Development of the serosal mesothelium. JDB 1, 64–81 (2013).
12. Majesky, M. W., Dong, X. R., Regan, J. N., Hoglund, V. J. & Schneider, M. Vascular smooth muscle progenitor cells. Circ. Res. 108, 365–377 (2011).
13. Chiellini, C. et al. Stathmin-like 2, a developmentally-associated neuronal marker, is expressed and modulated during osteogenesis of human mesenchymal stem cells. Biochem. Biophys. Res. Commun. 374, 64–68 (2008).
14. Grenningloh, G., Soehrman, S., Bondallaz, P., Ruchti, E. & Casades, H. Role of the microtubule destabilizing proteins SCG10 and stathmin in neuronal growth. J. Neurobiol. 58, 60–69 (2003).
15. Heitkamp-Sørensen, J. et al. The splanchic mesodermal plate directs spleen and pancreatic laterality, and is regulated by Bapx1/NKxs2.2. Development 131, 4665–4675 (2004).
16. Bin Zhou et al. Epidermal progenitors contribute to the cardiomyocyte lineage in the developing heart. Nature 454, 109–113 (2008).
17. Ashahina, K., Zhou, B., Wu, P. T. & Tsukamoto, H. Septum transversum-derived mesothelium gives rise to hepatic stellate cells and perivascular mesenchymal cells in developing mouse liver. Hepatology 53, 983–995 (2011).
18. Que, J. et al. Mesothelium contributes to vascular smooth muscle and mesenchyme during lung development. Proc. Natl. Acad. Sci. USA 105, 16626–16630 (2008).
19. Wilm, B., Ipenberg, A., Hastie, N. D., Burch, J. B. E. & Bader, D. M. The serosal mesothelium is a major source of smooth muscle cells of the gut vasculature. Development 132, 5317–5328 (2005).
20. Speer, M. Y. et al. Smooth muscle cells give rise to osteo-/chondrogenic precursors and chondrocytes in calcifying arteries. Circ. Res. 104, 733–741 (2009).
52. Golosow, N. & Grobstein, C. Epitheliomesenchymal interaction in pancreatic morphogenesis. Dev. Biol. 4, 242–253 (1962).
53. Landsman, L. et al. Pancreatic mesenchyme regulates epithelial organogenesis throughout development. PLoS Biol. 9, e1001141 (2011).
54. Bhushan, A. et al. Fgf10 is essential for maintaining the proliferative capacity of epithelial progenitor cells during early pancreatic organogenesis. Development 128, 5109–5117 (2001).
55. Yenumula, D. M. Mesothelial- and epithelial-derived Fgf9 have distinct functions in the regulation of lung development. Development 138, 3169–3177 (2011).
56. Murtaugh, L. C. The what, where, when and how of Wnt/β-catenin signaling in pancreas development. Organogenesis 4, 81–86 (2008).
57. Hernandez-Torres, F., Rodriguez-Outeiriño, L., Franco, D. & Aranega, A. E. Pitx2 in embryonic and adult myogenesis. Front. Cell Dev. Biol. 5, 211 (2017).
58. Cao, H. et al. The Pitx2:miR-200c/141:noggin pathway regulates Bmp signaling and ameloblast differentiation. Development 140, 3348–3359 (2013).
59. Kapadia, C., Ghosh, M. C., Grass, L. & Diamandis, E. P. Human kallikrein 13 involvement in extracellular matrix degradation. Biochem. Biophys. Res. Commun. 323, 1084–1090 (2004).
60. Huang, C. et al. Dermokine contributes to epithelial-mesenchymal transition through increased activation of signal transducer and activator of transcription 3 in pancreatic cancer. Cancer Sci. 108, 2130–2141 (2017).
61. Hasegawa, M. et al. Dermokine inhibits ELR + CXC chemokine expression and delays early skin wound healing. J. Dermatol. Sci. 70, 34–41 (2013).
62. Ernst, M. C. & Sinai, C. J. Chemerin: at the crossroads of inflammation and obesity. Trends Endocrinol. Metab. 21, 660–667 (2010).
63. Angelo, J. R. & Tremblay, K. D. Identification and fate mapping of the pancreatic mesenchyme. Dev. Biol. 435, 15–25 (2018).
64. Jansen, J. N. et al. Recapitulation of elements of embryonic development in adult mouse pancreatic regeneration. Gastroenterology 128, 728–741 (2005).
65. Rhim, A. D. & Stanger, B. Z. Molecular biology of pancreatic ductal adenocarcinoma progression: aberrant activation of developmental pathways. Prog. Mol. Biol. Transl. Sci. 97, 41–78 (2010).
66. Miyatsuka, T., Li, Z. & German, M. S. Chronology of islet differentiation revealed by temporal cell labeling. Diabetes 58, 1863–1868 (2009).
67. Collombat, P. et al. Opposing actions of Arx and Pax4 in endocrine pancreas development. Genes Dev. 17, 2591–2603 (2003).
68. Benitez, C. M. et al. An integrated cell purification and genomics strategy reveals multiple regulators of pancreas development. PLoS Genet. 10, e1004645 (2014).
69. Butler, A. E. et al. β-Cell deficit in obese type 2 diabetes, a minor role of β-cell dedifferentiation and degranulation. J. Clin. Endocrinol. Metab. 101, 525–532 (2016).
70. Wyler, S. C. et al. Pet-1 switches transcriptional targets postnatally to regulate maturation of serotonin neuron excitability. J. Neurosci. 36, 1758–1774 (2016).
71. Schonhoff, S. E., Giel-Moloney, M. & Leiter, A. B. Neurogenin 3-expressing progenitor cells in the gastrointestinal tract differentiate into both endocrine and non-endocrine cell types. Dev. Biol. 270, 443–454 (2004).
72. Scott, M. M. et al. A genetic approach to access serotonin neurons for in vivo and in vitro studies. PNAS 102, 16472–16477 (2005).
73. Wang, F. et al. A novel in situ rna analysis platform for formalin-fixed, paraffin-embedded tissues. J. Mol. Diagn. 14, 22–29 (2012).
74. Zheng, G. X. Y. et al. Massively parallel digital transcriptional profiling of single cells. Nat. Commun. 8, 41049 (2017).
75. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21 (2012).
76. R core team. R: A language and environment for statistical computing. R Foundation for Statistical Computing (2016).
77. Maaten, L. V. D. & Hinton, G. Visualizing data using t-SNE. J. Mach. Learn. Res. 9, 2579–2605 (2008).
78. Kamburov, A., Wierling, C., Lehrach, H. & Herwig, R. ConsensusPathDB—a database for integrating human functional interaction networks. Nucleic Acids Res. 37, D623–D628 (2008).
79. Frank, P. G. Caveolin, caveolae, and endothelial cell function. Arterioscler. Thromb. Vasc. Biol. 23, 1161–1168 (2003).

Acknowledgments

This work was supported by grants from the UCSF Resource Allocation Program and the UCSF Program for Breakthrough Biomedical Research, which is partially funded by the Sandler Foundation. L.E.B. was supported by an Achievement Rewards for College Scientists (ARCS) Foundation Scholar Award, and D.M.W. was supported by a National Science Foundation (NSF) Graduate Research Fellowship and the Diabetes, Endocrinology, and Metabolism Training Program (T32 DK007418). V2 Chromium reagent kits were generously provided by 10X Genomics. The authors would like to acknowledge expert technical assistance from Eunice Wan and Simon Wong at the UCSF Institute for Human Genetics Core, the UCSF Parnassus Flow Cytometry Core, William Walantus at the UCSF Broad Imaging Core, Adam Fries at the Diabetes Center Imaging Core, and Gregory Stotz of the UCSF Islet Production Core. We thank Dr. Matthias Hebrok’s laboratory for providing the Ngn3-Cre mouse line, and members of the Sneddon Lab for thoughtful discussions. Lastly, we thank Kevin Yu, Daniel Nguyen, and members of the Ye lab for assistance with analysis, as well as Drs. Steven Garcia, Michael German, and Evan Deneris for helpful discussions.

Author contributions

Conceptualization: L.E.B., D.M.W. and J.B.S.; Methodology: L.E.B., D.M.W., M.S., A.D. T., C.I.Y. and J.B.S.; Software: L.E.B., M.S., A.D.T., C.I.Y. and J.B.S.; Validation: L.E.B., D. M.W., M.S., N.P.M., A.D.T., C.I.Y. and J.B.S.; Formal analysis: L.E.B., D.M.W., M.S., A.D. T., C.I.Y. and J.B.S.; Investigation: L.E.B., D.M.W., N.P.M., C.I.G. and J.B.S.; Resources: S. M.K., A.D.T., C.I.Y. and J.B.S.; Data curation: L.E.B. and J.B.S.; Writing—original draft: L.E.B., D.M.W. and J.B.S.; Writing—review and editing: L.E.B., D.M.W., A.D.T., C.I.Y. and J.B.S.; Visualization: L.E.B., D.M.W., A.D.T., C.I.Y. and J.B.S.; Supervision: J.B.S.; Project administration: J.B.S.; Funding acquisition: J.B.S.

Additional information

Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-018-06176-3.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.