Lu et al. - Single-cell analyses identify distinct and intermediate states of zebrafish pancreatic islet development

Supplementary Materials and Methods

Whole mount in situ hybridization

The digoxygenin-labeled probe was synthesized by *in vitro* transcription of PCR products with T7 promoter. The whole mount *in situ* hybridization of zebrafish embryos was performed following the previous instructions (Tong et al., 2014). The primer sequences for probe synthesis were as the following:

Primers for PCR amplification of cDNA templates for making *in situ* probes	
Gene and primer name	Primer sequence
gcga-forward	5'-GCGTCCAGTATTTTGCCAGTC-3'
gcga-reverse	5'-TAATACGACTCACTATAGGGAGA TTAGCCAGGCCACAAAGCTC-3'
hoxb8a-forward	5'-AGCCGCCCAGTAATTGGAAA-3'
hoxb8a-reverse	5'-TAATACGACTCACTATAGGGAGA AGCTTGAGGTCGCATCCAC-3'
ins-forward	5'-TTGGTCGTGTCCAGTGTAAG-3'
ins-reverse	5'-TAATACGACTCACTATAGGGAGA TGCCTCTCTTCCTTATCAGC-3'
sst2-forward	5'-ATGGCCTCCTCGCAACTCCACC-3'
sst2-reverse	5'-TAATACGACTCACTATAGGGAGA AGCATGCCGTTCTGGATTTCCA-3'

RT-PCR detection for enriched genes in single-cell libraries

New single-cell RNA-seq libraries prepared from pancreatic islets isolated from 18 hpf, 22 hpf and 30 hpf *TgBAC(neurod1:EGFP)* zebrafish embryos were pooled together, and PCR was performed using DNA polymerase (MCLAB, Cat# I5HD-OEM) following the manufacturer's instructions. 10 μL PCR products were loaded for electrophoresis. Primers used for RT-PCR are listed in the following table:

Primers for RT-PCR analysis of enriched genes in single-cell libraries	
Gene and primer name	Primer sequence
cabz01070258.1-forward	5'-ACGTTTGAAGAAAGCGGCAC-3'
cabz01070258.1-reverse	5'-GAGCTGGTGTATCCGTCCTG-3'
coe2-forward	5'-CGCACAGTTTGAATGGGTCTC-3'
coe2-reverse	5'-TTTCTCCCGTCTTGTCCTGC-3'
dek-forward	5'-TCCAATCGGCGCAACAGTAA-3'
dek-reverse	5'-CTTTCCCGAGTTGGTTGGCT-3'
dnmt3bb.2-forward	5'-CAGAGTGCAGGAGTTCTTTG-3'
dnmt3bb.2-reverse	5'-ATGACAGGAACGCTCCAGGA-3'
ehmt1b-forward	5'-TGCGTAAATGCAGTGGACAG-3'
ehmt1b-reverse	5'-GATGAGCTGGTATCGGGTAG-3'

gabpa-forward	5'-CAACCGCACAGGTAACAACG-3'
gabpa-reverse	5'-AACTGTGGTGACGGGTTGAC-3'
hat1-forward	5'-TCTCTGTCCTGGAGAAAGAG-3'
hat1-reverse	5'-TTCTCCAGTTCCTCGTGCTG-3'
hoxb8a-forward	5'-AGCCGCCCAGTAATTGGAAA-3'
hoxb8a-reverse	5'-AGCTTGAGGTCGCATCCAC-3'
id1-forward	5'-CAAAGTTGGAGGAGAGGATG-3'
id1-reverse	5'-GGGAAATGCATGAACAGTAAG-3'
klf16-forward	5'-GGAAGCGCAATGGAGACAAC-3'
klf16-reverse	5'-TCGAACAGCGAAAAACGCAG-3'
meis3-forward	5'-TACCACAGCCCACTACCCTCAGC-3'
meis3-reverse	5'-TCAGCAGGATTTGGTGCAGTTGT-3'
noc3l-forward	5'-TCCGTCTTACTGTCCTCCGT-3'
noc3l-reverse	5'-GCCCTGCGTCTCATTATCCA-3'
onecutl-forward	5'-AAGTCTGGCCGTGAAACCTT-3'
onecutl-reverse	5'-CAGAGCAGAAAACGGCTCATT-3'
setdb1b-forward	5'-AAGCGGTGCCGCTGTAATAT-3'
setdb1b-reverse	5'-TCATACTGTGGGACTTCAAA-3'
si:ch211-288g17.3-forward	5'-GCTCCCTCTCCCAAACAGTC-3'
si:ch211-288g17. 3-reverse	5'-TCGACTAAAACCGTCAGGGC-3'
tox3-forward	5'-AGGTGTCCAAGATAGTAGCC-3'
tox3-reverse	5'-TCAGTAGATGCTGACTTGCG-3'

Morpholino

The *tbx2b* splice morpholino was obtained from GeneTools and designed to target the splice site between exon 3 and intron 3 (5'-ATCAAAATATGGGTACATACCTTGT-3'). Embryos at the one-cell stage were injected with approximately 20 ng morpholino and the knockdown effect was evaluated by RT-PCR using the following primers: 5'-CATCCGGATAGCCCAGCTAC-3' and 5'-TGCTCGCTATCAGTGCATGT-3'.

RNA extraction, cDNA synthesis and qPCR

Total RNA was extracted from embryos at 30 hpf and 52 hpf using TRIzol reagent (Invitrogen, 15596018). cDNAs were synthesized using the FastQuant RT Kit (with gDNase) (Tiangen, KR106). Quantitative PCR was performed with EvaGreen 2× qPCR MasterMix-Low ROX (Abm, MasterMix-LR) in 8-well PCR tubes using the Agilent Mx3000P qPCR System. Primers used for qPCR are listed in the following table:

Primers used for qPCR analysis	
Gene and primer name	Primer sequence
cdx4-forward	5'-GGTTCTTCCTCCACCACCTG-3'

cdx4-reverse	5'-GCTCCAGCCTTTGGTGATCT-3'
gapdh-forward	5'-GGGTGATGCAGGTGCTACTT-3'
gapdh-reverse	5'-AAAGGAGCCAGGCAGTTGG-3'
gcga-forward	5'-AAGGCGACAGCACAAGCACA-3'
gcga-reverse	5'-GCCCTCTGCATGACGTTTGACA-3'
ins-forward	5'-TAAGCACTAACCCAGGCACA-3'
ins-reverse	5'-TTTAGGAGGAAGGAAACCCA-3'
sst2-forward	5'-CACTACAGGCGTCTCGCGGC-3'
sst2-reverse	5'-TGAGCAGCCACTCCAGCTCAGA-3'

Immunofluorescent staining and imaging

The following antibodies were used for immunofluorescent staining: p38 MAPK antibody [N1C3-2] (Genetex, GTX110720, 1:200) (Yeh et al., 2017), goat anti-rabbit IgG (H+L) Alexa Fluor 488 (Invitrogen, A11008, 1:400).

Embryos were fixed overnight at 4°C in 4% PFA and washed with PBST (PBS containing 0.1% Tween-20) at room temperature three times (5 minutes each time). Thereafter, proteinase-K (10 µg/ml in PBST) was used to digest embryos for 8 minutes at room temperature. After the digestion, embryos were washed with PBDTX (PBS containing 1% BSA, 1% DMSO and 0.3% Triton X-100) at room temperature three times (5 minutes each time) and blocked in blocking buffer (PBDTX containing 2% new born calf serum) for 2 hours at room temperature. Primary antibodies were diluted in the blocking buffer and used overnight at 4°C. Embryos were washed in PBDTX at room temperature four times (15 minutes each time) and then incubated with secondary antibodies and DAPI (1 µg/mL in PBDTX) overnight at 4°C. After the incubation, embryos were washed in PBST at room temperature six times (5 minutes each time) to remove unbound antibodies. Finally, images were taken using a confocal laser scanning microscope (Zeiss, LSM710NLO) with 20× or 40× objectives.

Conservation analysis

Conservation analysis followed the previous studies (Athanasiadis et al., 2017; Carmona et al., 2017). Orthologous genes were obtained from Ensembl (Ensembl Genes 91), and differentially expressed genes were calculated by Seurat without further filtration. The percentage of genes with orthologs in each of the differentially expressed gene clusters was calculated.

Gene duplication analysis

Gene duplication analysis followed previous studies (Athanasiadis et al., 2017; Carmona et al., 2017). Paralog genes were obtained from Ensembl (Ensembl Genes 91), and classified into pre-ray-finned fish duplicated genes and post-ray-finned fish

duplicated genes. Differentially expressed genes were calculated by Seurat without further filtration. For the analysis of expression pattern divergence, genes were defined as conserved if all paralog genes were expressed in the same cell type and considered as diverged if paralog genes were expressed in more than one cell type based on our data.

Supplementary References

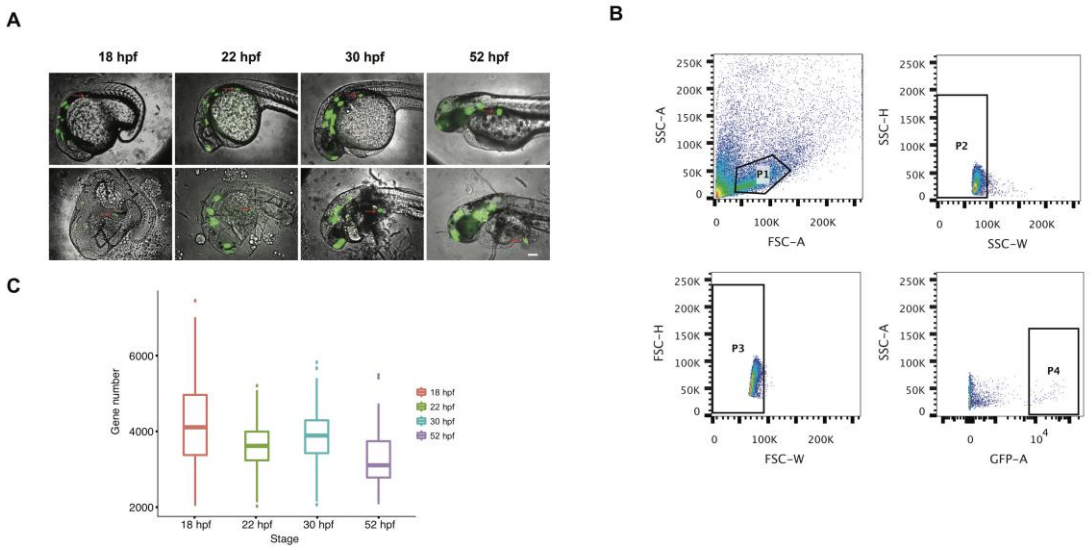
Athanasiadis, E.I., Botthof, J.G., Andres, H., et al. (2017). Single-cell RNA-sequencing uncovers transcriptional states and fate decisions in haematopoiesis. Nat. Commun. *8*, 2045.

Carmona, S.J., Teichmann, S.A., Ferreira, L., et al. (2017). Single-cell transcriptome analysis of fish immune cells provides insight into the evolution of vertebrate immune cell types. Genome Res. *27*, 451-461.

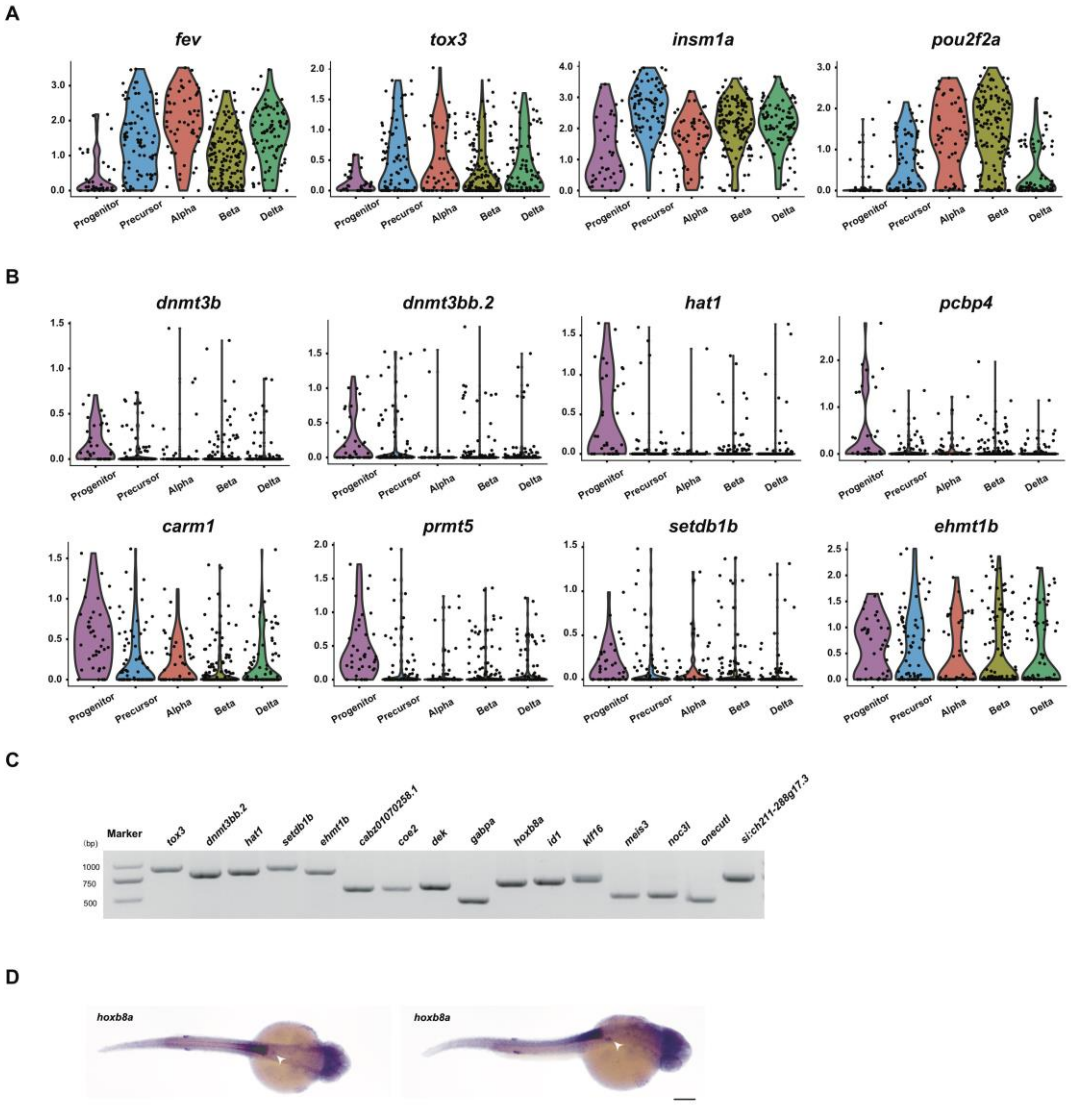
Tong, X., Zu, Y., Li, Z., et al. (2014). Kctd10 regulates heart morphogenesis by repressing the transcriptional activity of Tbx5a in zebrafish. Nat. Commun. *5*, 3153.

Yeh, C.J., Chen, C.C., Leu, Y.L., et al. (2017). The effects of artocarpin on wound healing: in vitro and in vivo studies. Sci. Rep. *7*, 15599.

Lu et al. - Single-cell analyses identify distinct and intermediate states of zebrafish pancreatic islet development

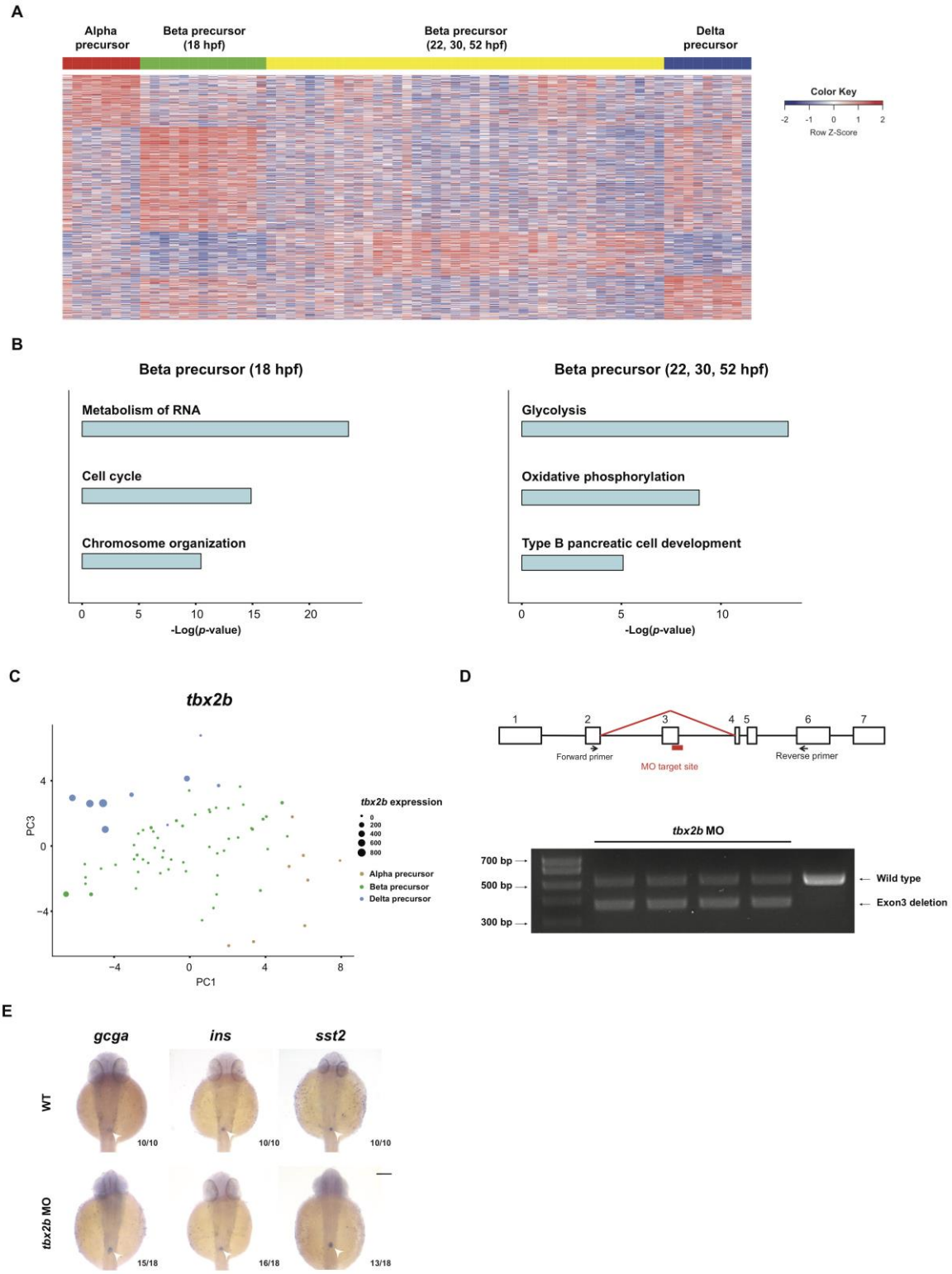


Supplementary Figure S1 Sample preparation and data processing. A) *TgBAC(Neurod1:EGFP)* embryos before (upper panel) and after (lower panel) dissection at different developmental stages. The red arrowheads indicate the dissected endocrine islet. The bar represents 100 μm. **B)** FACS sorting process of the dissected endocrine pancreas from 30 hpf zebrafish embryos (Sample sorting from other stages followed the same strategy). Cellular debris was removed by P1 gating, followed by P2 and P3 gating to collect single cells. Finally, GFP-positive cells were sorted by P4 gating. **C)** Boxplot results of detected genes (TPM > 0) in each cell. Related to Figure 1 and Table S1.



Supplementary Figure S2 Newly identified endocrine markers and enriched epigenetic factors in progenitors. A) Violin plot of newly identified endocrine markers *fev*, *tox3*, *insm1a*, and *pou2f2a*. **B)** Violin plot showing expression patterns of progenitor cell enriched epigenetic factors *dnmt3b*, *dnmt3bb.2*, *hat1*, *pcbp4, carm1*, *prmt5*, *setdb1b* and *ehmt1b*. **C)** RT-PCR analysis of the expression of progenitor cell enriched transcription factors in pooled single-cell RNA-seq libraries from stages of 18 hpf, 22 hpf and 30 hpf zebrafiah embryos. **D)** mRNA expression of *hoxb8a* in a 30 hpf zebrafiah embryo revealed by whole mount *in situ* hybridization. The arrowheads indicate the pancreas region, and the bar represents 200 μM. Related to Figure 2 and Table S2.

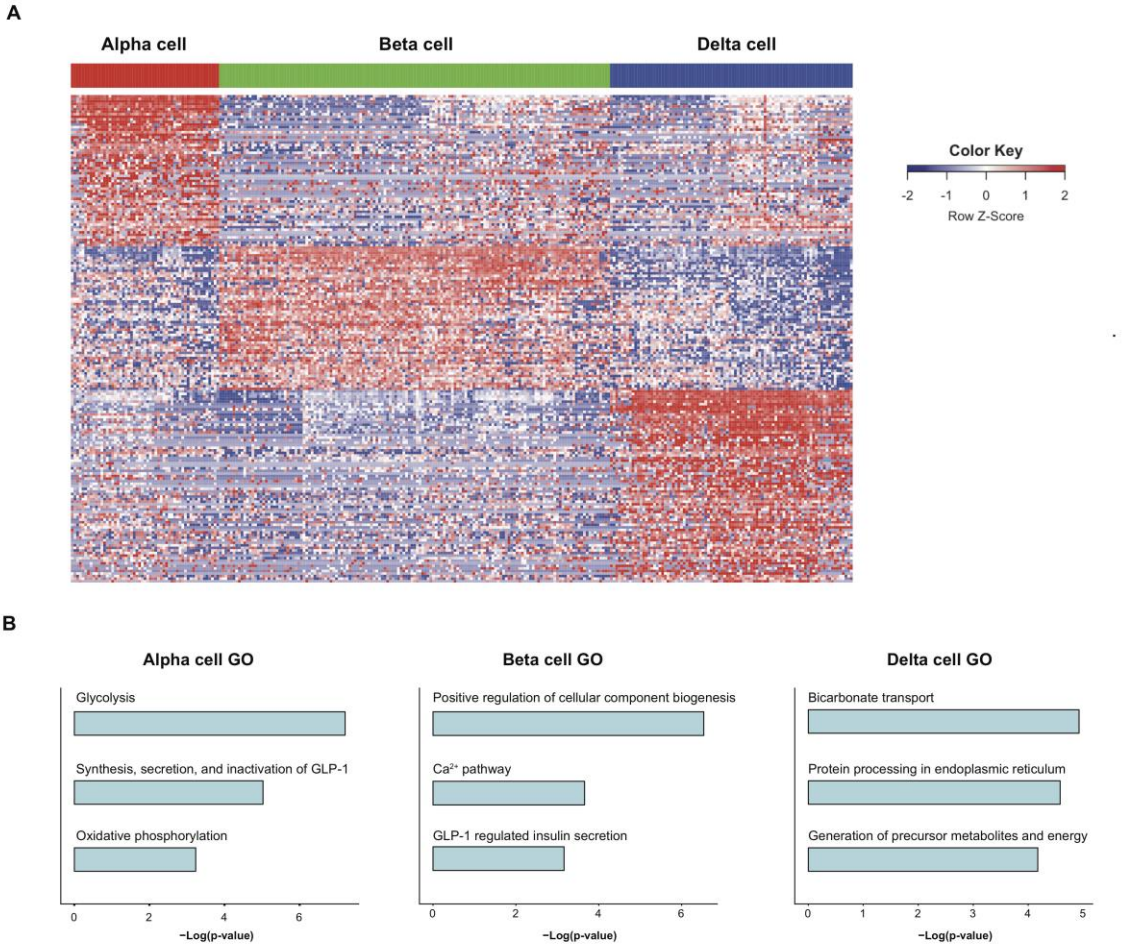
Lu et al. - Single-cell analyses identify distinct and intermediate states of zebrafish pancreatic islet development



Supplementary Figure S3 Comparison of separate lineage primed precursors. **A)** Heat map of differentially expressed genes in alpha precursors (red), beta

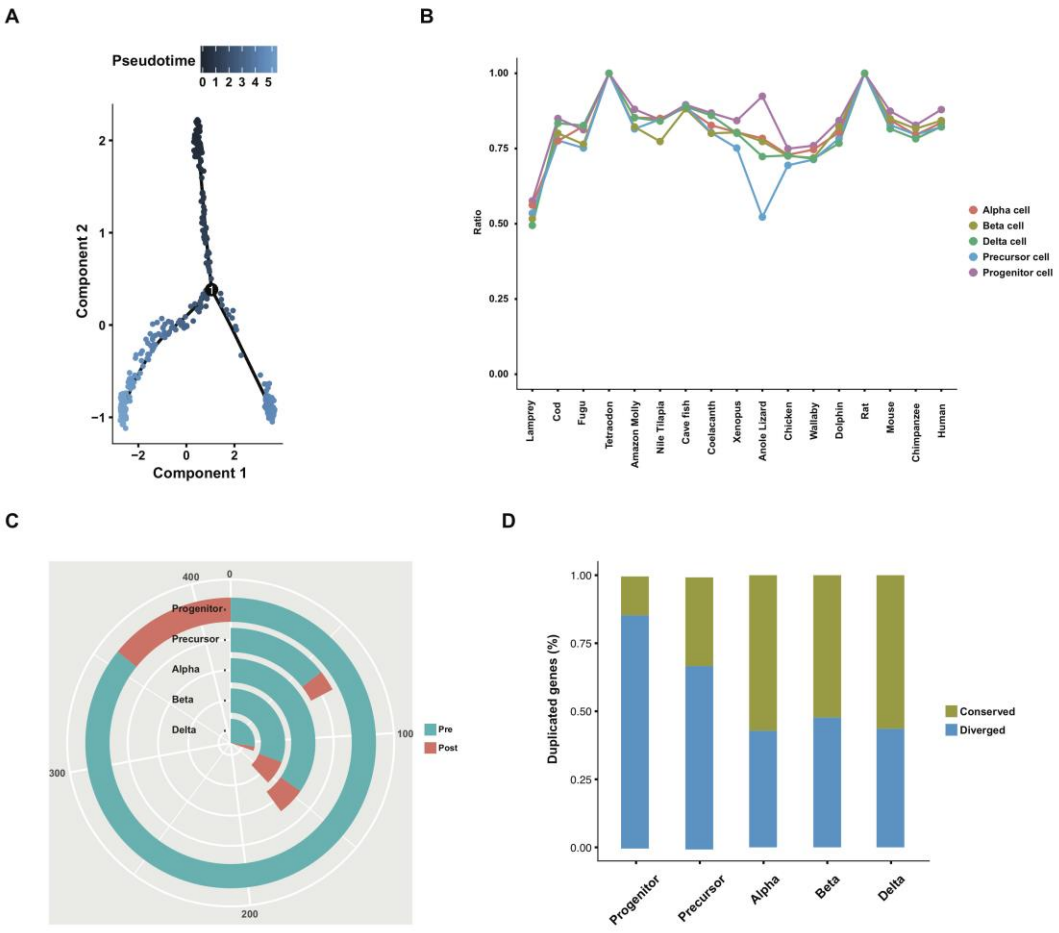
precursors at 18 hpf (green), beta precursors at 22, 30 and 52 hpf (yellow), and delta precursors (blue). **B)** GO terms of beta precursors at 18 hpf (left) and beta precursors at 22, 30 and 52 hpf (right) in comparison with the other precursor cells. **C)** PCA plot showing the expression of *tbx2b* across alpha, beta and delta precursors. **D)** The target design and evaluation of knowndown efficiency of the *tbx2b* morpholino by RT-PCR. **E)** Expression of alpha, beta and delta cell hormone markers *gcga*, *ins* and *sst2* in 30 hpf *tbx2b* morphants detected by whole mount *in situ* hybridization. The arrowheads indicate the pancreas region, and the bar represents 200 μM. Related to Figure 3 and Table S3.

Lu et al. - Single-cell analyses identify distinct and intermediate states of zebrafish pancreatic islet development



Supplementary Figure S4 Characterization of terminally differentiated islet cells. **A)** Heat map showing differentially expressed genes of alpha, beta and delta cells. **B)** Functional gene ontology terms of alpha, beta and delta cells. Related to Figure 4 and Table S4.

Lu et al. - Single-cell analyses identify distinct and intermediate states of zebrafish pancreatic islet development



Supplementary Figure S5 Cell types and genes identified by pseudo-time ordering. A) The trajectory of the islet lineage specification, and the color key is according to the pseudo-time value of each single cell. **B)** The ratio of orthologs of protein-coding genes for alpha (red), beta (brown), delta (green), precursor (blue) and progenitor (purple) cells. **C)** The total number of paralog genes duplicated pre- (green) or post- (red) ray-finned fish speciation enriched in each of the alpha, beta, delta, precursor or progenitor cell population. **D)** The ratio of diverged or conserved paralog genes of the post-ray-finned species. Related to Figure 5 and Table S5.

Lu et al. - Single-cell analyses identify distinct and intermediate states of zebrafish pancreatic islet development

Supplementary Tables

Supplementary Table S1 Sample information and statistics of single-cell RNA-seq analysis. Related to Figure 1.
(Supplied as Excel file: Supplementary Table S1.xlsx)

Supplementary Table S2 List of genes related to progenitor features. Related to Figure 2.
Sheet 1: Genes enriched in the progenitor cell.
Sheet 2: Progenitor-related ribosome biogenesis factors.
Sheet 3: Cell cycle phase-related genes.
Sheet 4: Progenitor-enriched transcription factors.
(Supplied as Excel file: Supplementary Table S2.xlsx)

Supplementary Table S3 List of genes related to the features of the lineage-primed precursors. Related to Figure 3.
Sheet 1: Alpha cell enriched genes compared with the alpha precursor.
Sheet 2: Beta cell enriched genes compared with the beta precursor.
Sheet 3: Delta cell enriched genes compared with the delta precursor.
Sheet 4: Enriched genes of the separate precursors.
(Supplied as Excel file: Supplementary Table S3.xlsx)

Supplementary Table S4 List of genes related to the features of terminally differentiated alpha, beta and delta cells. Related to Figure 4.
Sheet 1: Enriched genes of terminally differentiated alpha, beta and delta cells.
Sheet 2: Stage-specific genes for alpha cells.
Sheet 3: Stage-specific genes for beta cells.
Sheet 4: Stage-specific genes for delta cells.
(Supplied as Excel file: Supplementary Table S4.xlsx)

Supplementary Table S5 List of dynamically changed genes across pseudo-time. Related to Figure 5.
(Supplied as Excel file: Supplementary Table S5.xlsx)