The Company of
Biologists

RESEARCH ARTICLE

Dynamics of chromatin marks and the role of JMJD3 during pancreatic endocrine cell fate commitment

Xin-Xin Yu[1,2,*], Wei-Lin Qiu[1,3,*], Liu Yang[1,2], Lin-Chen Li[1], Yu-Wei Zhang[1] and Cheng-Ran Xu[1,‡]

ABSTRACT

Pancreatic endocrine lineages are derived from pancreatic progenitors that undergo a cell fate transition requiring a switch from low to high Ngn3 expression. However, the underlying chromatin regulatory mechanisms are unclear. Here, we performed epigenomic analysis of gene regulatory loci featuring histone marks in cells with low or high level of Ngn3 expression. In combination with transcriptomic analysis, we discovered that in Ngn3-high cells, the removal of H3K27me3 was associated with the activation of key transcription factors and the establishment of primed and active enhancers. Deletion of *Jmjd3*, a histone demethylase for H3K27me3, at the pancreatic progenitor stage impaired the efficiency of endocrine cell fate transition and thereafter islet formation. Curiously, single-cell RNA-seq revealed that the transcriptome and developmental pathway of Ngn3-high cells were not affected by the deletion of *Jmjd3*. Our study indicates sequential chromatin events and identifies a crucial role for Jmjd3 in regulating the efficiency of the transition from Ngn3-low to Ngn3-high cells.

KEY WORDS: Pancreatic endocrine lineages, Chromatin modifications, Jmjd3, Cell fate transition, Pioneer factor, Single-cell RNA-seq

INTRODUCTION

Induction of mature pancreatic β cells *in vitro* is an effective way to treat diabetes mellitus. Because the developmental mechanisms of pancreatic development, especially pancreatic endocrine development, are not fully understood, artificially generated β cells remain functionally immature (Johnson, 2016; McKnight et al., 2010; Pagliuca and Melton, 2013). Therefore, defining the regulatory mechanism that controls endocrine cell fate commitment is extremely important in the study of pancreatic development.

During pancreatic development, all endocrine lineages arise from multipotent pancreatic progenitor cells. The developmental potential of these progenitors is gradually constrained to the establishment of exocrine and islet endocrine lineages (Shih et al., 2013). The basic helix-loop-helix transcription factor (TF) neurogenin 3 (Neurog3 or Ngn3) plays a central role in endocrine lineage specification, and lineage-tracing experiments have shown that all endocrine lineages, and even some exocrine cells, are derived from Ngn3-expressing progenitors (Gu et al., 2002; Schonhoff et al., 2004). Levels of Ngn3 are also important for driving endocrine lineage specification, as demonstrated by the fact that genetically ablating *Ngn3* expression or modifying the Ngn3 dose can affect the induction of other key TF genes involved in endocrine cell specification, including *Neurod1*, *Isl1*, *Pax4* and *Pax6* (Gradwohl et al., 2000), or can cause Ngn3-induced cells to follow exocrine cell fate (Beucher et al., 2012; Wang et al., 2010). These findings indicate that early-stage Ngn3-expressing cells retain their plasticity to differentiate into multiple pancreatic lineages.

During mouse embryogenesis, *Ngn3* is expressed at a relatively low level in the dorsal endoderm domain beginning at embryonic day (E) 8.5, and this expression is maintained until E11.0. The second wave of Ngn3[+] cell specification in both the dorsal and ventral pancreas begins at ~E12.0, peaks at E15.5, and then abruptly decreases (Gradwohl et al., 2000; Schwitzgebel et al., 2000). However, *Ngn3* is transiently expressed in individual cells and undergoes a transition from low expression (Ngn3-low) to high expression (Ngn3-high) during the second wave of endocrine differentiation. In addition, Ngn3-low cells are cell-cycle-active progenitors, whereas most Ngn3-high cells are cell-cycle-silent endocrine-committed cells (Bankaitis et al., 2015; Bechard et al., 2016). A recent EdU pulse-chase study demonstrated that pancreatic endocrine progenitor cells take ~12 h to pass through the Ngn3[+] state (Bankaitis et al., 2015).

Chromatin modifications play important roles in regulating cell fate transitions. During early developmental stages, both activating (H3K4me3) and repressive (H3K27me3) epigenetic marks are frequently located in the promoter regions of developmentally poised regulatory genes (Bernstein et al., 2006; Mikkelsen et al., 2007). Upon differentiation, these regions, which are known as bivalent domains, must be resolved to facilitate either rapid transcriptional activation or stable gene silencing (Voigt et al., 2013). The histone modifications H3K4me1 and H3K27ac are both chromatin marks of enhancers and show dramatic dynamics during cell fate differentiation. However, it is unclear how these chromatin modifications coordinate and promote the quick cell fate transition from Ngn3-low to Ngn3-high.

In this study, we performed transcriptomic analysis of sorted Ngn3-low and Ngn3-high cells from E13.5 *Ngn3-GFP* mouse embryos and discovered dramatic changes in gene expression between these two populations. Based on low-cell-number ChIP-seq analysis, we identified that the removal of a repressive H3K27me3 histone mark was linked to the activation of key TFs for the specification of Ngn3-high cells. The activation of one TF, NeuroD1, is sequentially accompanied by the deposition of histone modifications associated with active and primed enhancers. Moreover, we demonstrate that the histone H3K27 demethylase JMJD3 regulates the efficiency of the transition from Ngn3-low to Ngn3-high.

[1]Ministry of Education Key Laboratory of Cell Proliferation and Differentiation, College of Life Sciences, Peking-Tsinghua Center for Life Sciences, Peking University, Beijing 100871, China. [2]Academy for Advanced Interdisciplinary Studies, Peking University, Beijing 100871, China. [3]PKU-Tsinghua-NIBS Graduate Program, Peking University, Beijing 100871, China.
*These authors contributed equally to this work

‡Author for correspondence (cxu@pku.edu.cn)

C.-R.X., 0000-0002-0583-4464

DEVELOPMENT

1

RESULTS

Dramatic transcriptomic differences between Ngn3-low and Ngn3-high cells

To profile the global changes in gene expression associated with dorsal pancreatic endocrine lineage specification in mice, we performed RNA-seq analyses on E13.5 Ngn3-low and Ngn3-high cells from *Ngn3-GFP* pancreata (Lee et al., 2002), and on E17.5 nascent β cells from *Ins1-RFP* pancreata (Piccand et al., 2014). All of these cells were purified via fluorescence-activated cell sorting (FACS) (Fig. S1A,B and Table S1). Although passing through the Ngn3$^+$ state of pancreatic endocrine progenitor cells only takes ~12 h (Bankaitis et al., 2015), hierarchical clustering revealed a striking difference in the gene expression profiles between Ngn3-low and Ngn3-high cell populations. The gene expression profile of E13.5 Ngn3-high cells more closely resembled those of differentiated E17.5 β cells than those of E13.5 Ngn3-low cells (Fig. S1C). These findings indicate that a significant cell fate transition towards pancreatic endocrine development occurred between Ngn3-low and Ngn3-high stages.

A comparative transcriptomic analysis identified 5197 genes that were differentially expressed between Ngn3-low and Ngn3-high cell populations (Fig. S1D and Table S2), among which 154 genes encode TFs (Fig. S1E). The expression levels of *Sox9*, *Onecut1* and *Hes1* (Fig. S1E) were enriched in Ngn3-low cells, which is consistent with the progenitor property of Ngn3-low cells (Shih et al., 2013), whereas the expression levels of many key TFs for endocrine development, such as *Ngn3*, *Neurod1*, *Insm1*, *Rfx6*, *Isl1* and *Pax6* were elevated in Ngn3-high cells (Ahlgren et al., 1997; Gierl et al., 2006; Mellitzer et al., 2006; Naya et al., 1997; Sander et al., 1997; Smith et al., 2010; St-Onge et al., 1997) (Fig. S1E). Cell cycle regulatory and DNA replication genes were enriched in Ngn3-low cells but displayed a drastic change during the transition from Ngn3-low to Ngn3-high (Fig. S1F,G and Table S2). KEGG pathway (Kanehisa and Goto, 2000) enrichment analysis revealed that Ngn3-high cells were enriched for pathways related to pancreatic functions (Fig. S1F and Table S2). Taken together, these transcriptomic analyses indicate that during the transition from Ngn3-low to Ngn3-high, endocrine progenitor cells lose their progenitor properties and begin specification.

The removal of H3K27me3 is associated with the transition from Ngn3-low to Ngn3-high

To obtain a genomic view of the dynamics of chromatin bivalency during the endocrine fate transition, we mapped the modifications of H3K4me3 and H3K27me3 in E13.5 Ngn3-low and Ngn3-high cells using low-cell-number ChIP-seq (Brind'Amour et al., 2015) (Fig. S2A,B and Table S1). In Ngn3-low cells, 2131 genes displayed bivalency at their promoter regions [transcription start sites (TSSs) ±2 kb] (Fig. 1A). During the transition from Ngn3-low to Ngn3-high, bivalency was resolved in 590 of these genes, whereas 226 genes gained bivalent marks (Fig. 1B, top). Of the genes showing resolved bivalency in Ngn3-high cells, 80% (471 genes) lost only the H3K27me3 mark (Fig. 1B, top). However, only 49 genes became bivalent via the sole gain of H3K27me3 (Fig. 1B, top). To determine whether endocrine cells maintain this unbalanced loss and gain of bivalency-associated H3K27me3 during later development, we compared the dynamics of bivalent marks between E13.5 Ngn3-high and E17.5 β cells (Fig. 1A). Curiously, the numbers of genes gaining H3K27me3, H3K4me3 or both marks were similar to the numbers of genes losing those marks (Fig. 1B, bottom). This finding indicates that

the removal of bivalency-associated H3K27me3 is a specific characteristic of the transition from Ngn3-low to Ngn3-high.

The composite plot shows that the removal of bivalency-associated H3K27me3 did not change the overall corresponding H3K4me3 level (Fig. 1C). However, according to the RNA-seq analyses of Ngn3-low and Ngn3-high cell populations, the genes associated with the loss of bivalency-associated H3K27me3 showed a net increase in their expression levels (Fig. 1D). Notably, this group of genes included key TFs for endocrine development, such as *Neurod1*, *Rfx6* and *Insm1* (Gierl et al., 2006; Mellitzer et al., 2006; Naya et al., 1997; Smith et al., 2010) (Fig. 1E,F and Table S3), but did not include *Ngn3*, which indicates that the regulation of *Ngn3* is independent of bivalency resolution (Fig. 1G). These studies indicate that the resolution of bivalent domains, mainly by the removal of H3K27me3, is related to pancreatic endocrine cell specification.

We also examined the changes of H3K27me3 levels at promoters (TSSs ±2 kb) marked by H3K27me3 alone. To our surprise, during the transition from Ngn3-low to Ngn3-high, 124 promoter loci showed decreases (>2-fold changes) in H3K27me3 levels, whereas only six regions displayed increases in H3K27me3 levels (Fig. 2A,B). However, during the transition from Ngn3-high cells to β cells, we observed that the number of genes adding H3K27me3 (137 genes) was higher than the number removing H3K27me3 (36 genes) (Fig. 2B), which indicates a substantial increase in the number of promoters marked by H3K27me3 alone during β cell differentiation. Furthermore, the decrease in the H3K27me3 level at promoters during the transition from Ngn3-low to Ngn3-high was associated with an increase in H3K4me3 level (Fig. 2C) and a net increase in the expression level of corresponding genes (Fig. 2D). This class of genes included key TFs for endocrine lineage differentiation, such as *Pax6*, *Arx*, *Fev* and *Neurod2* (Sander et al., 1997; St-Onge et al., 1997; Collombat et al., 2003; Gasa et al., 2008; Ohta et al., 2011) (Fig. 2E,F and Table S3). Taken together, these results indicate that the removal of bivalency-independent H3K27me3 is also associated with the regulation of endocrine cell specification.

Establishment of primed and active enhancers in Ngn3-high cells

Enhancers, which function as integrated TF-binding platforms, play central roles in controlling cell type- and cell state-specific gene expression patterns (Calo and Wysocka, 2013; Long et al., 2016). Active enhancers are characterized by the presence of both H3K4me1 and H3K27ac modifications, and are generally associated with active genes. However, primed enhancers are marked solely by H3K4me1 prior to the activation of gene transcription (Creyghton et al., 2010). To understand how enhancer-associated chromatin modifications change during the pancreatic endocrine cell fate transition, we mapped H3K4me1 and H3K27ac profiles in Ngn3-low and Ngn3-high cells from E13.5 dorsal pancreata (Fig. S2A and Table S1). We identified 9366 active and 30,600 primed putative enhancers in Ngn3-low cells, and 13,953 active and 54,514 primed putative enhancers in Ngn3-high cells (Fig. 3A). Genes associated with active enhancers had higher net expression levels than genes associated with primed enhancers in both Ngn3-low and Ngn3-high cell populations (Fig. 3B). Of the active enhancers in Ngn3-high cells, 20.7% were primed and 49.0% were unmarked in the Ngn3-low population (Fig. 3C). These 9718 newly formed active enhancers were associated with the upregulation of gene expression levels during the transition from Ngn3-low to Ngn3-high (Fig. 3D,E). Furthermore, MOTIF
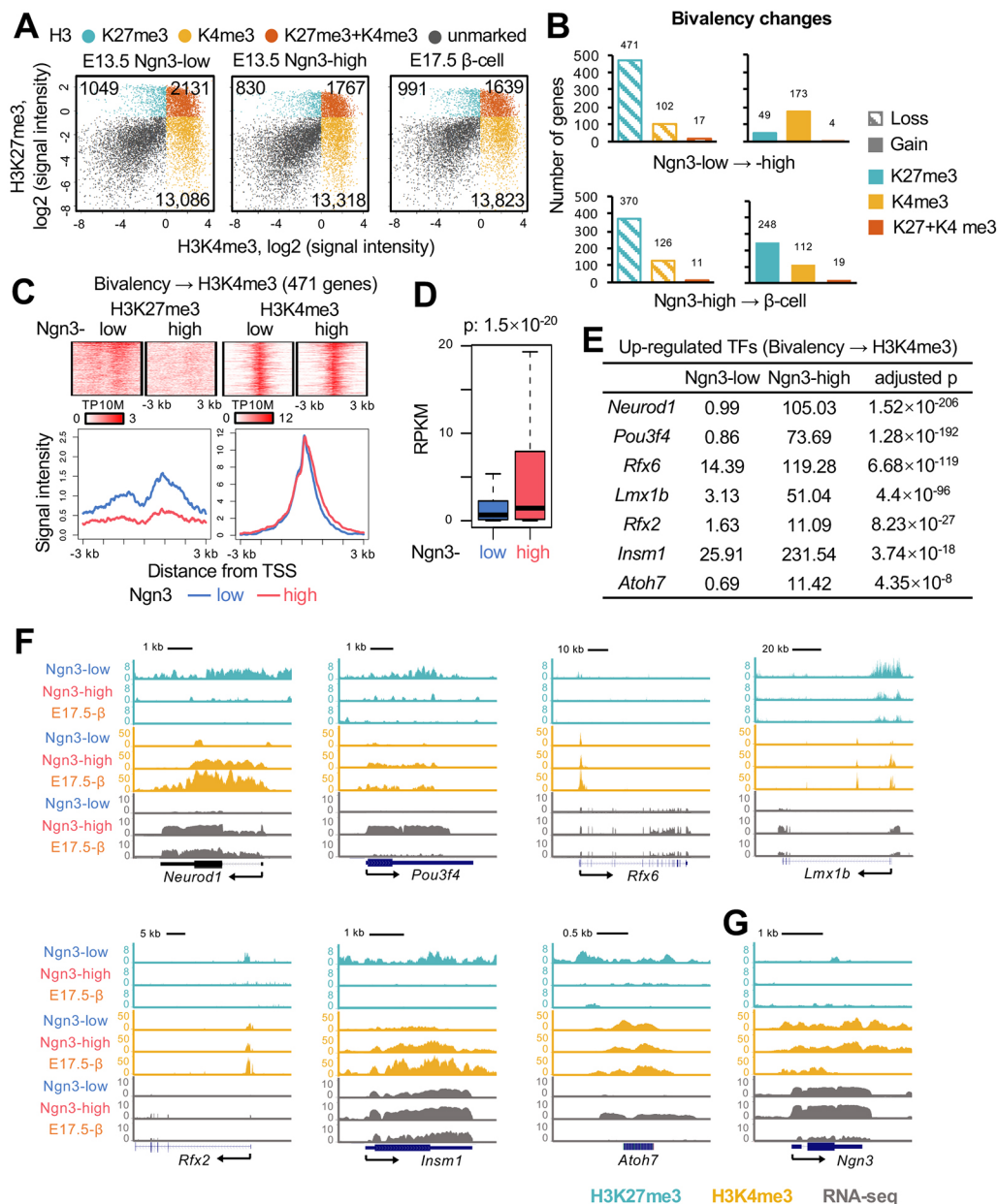
Fig. 1. Dynamics of bivalent promoter-associated histone marks during the transition from Ngn3-low to Ngn3-high. (A) Global patterns of H3K27me3 and H3K4me3 in E13.5 Ngn3-low and Ngn3-high cells, and in E17.5 β cells. The plots represent genes classified according to H3K4me3 and H3K27me3 levels. Different categories are color coded. (B) Quantification of the loss and gain of bivalent domains during fate transitions. (C) Density of H3K27me3 and H3K4me3 ChIP-seq tags at TSSs ±3 kb of genes losing bivalency through H3K27me3 resolution (top). Composite plots showing the distribution of H3K27me3 and H3K4me3 marks around TSSs (bottom). (D) Boxplot showing the expression levels of genes identified in C. The bottom and top of the box indicate the first and third quartiles, and the band inside the box indicates median; the bottom whisker indicates the minimum value, and the top whisker indicates 1.5× interquartile range. Wilcoxon rank-sum test. (E) List and expression levels (RPKM) of the key upregulated TFs (adjusted $P<0.05$, fold change >2, RPKM >10 in Ngn3-high) identified in C. (F,G) Genome browser images of histone H3K4me3 and H3K27me3 modifications and RNA-seq patches of *Ngn3* (G), and genes that lost bivalency through the resolution of H3K27me3 during the transition from Ngn3-low to Ngn3-high (F).

enrichment analysis revealed that the newly formed active enhancers in Ngn3-high cells were enriched for binding motifs, such as those for NEUROD1, FOXA2 and FOXO1 (Fig. 3F). KEGG pathway enrichment analysis showed that the genes associated with this class of enhancer were related to the regulation of pancreatic development and several cell signaling pathways (Fig. 3G and Table S4).

Of the 54,514 primed enhancers identified in Ngn3-high cells, 57.9% were unmarked in the Ngn3-low cells (Fig. 3C). The selection of new enhancer loci during cell fate transitions may be guided by 'pioneer' factors (Calo and Wysocka, 2013; Iwafuchi-Doi and Zaret, 2014), which are capable of directly binding to even compacted chromatin, thus making DNA-binding sites accessible to chromatin modifiers and other TFs. In Ngn3-high cells, the newly formed primed enhancers predominantly enriched the binding motif of NEUROD1 (Fig. 3H), which has been demonstrated to be a pioneer factor responsible for the establishment of a fraction of new enhancer sites during neuronal cell differentiation (Pataskar et al.,

2016). Curiously, the newly primed enhancers in Ngn3-high cells were associated with genes involved in processes specific to the islet lineage, such as pancreatic secretion and metabolic pathways (Fig. 3I and Table S4). To investigate whether these newly formed primed enhancers in Ngn3-high cells contribute to the proportion of active enhancers in β cells, we performed H3K4me1 and H3K27ac ChIP-seq analyses on E17.5 β cells, and detected 5031 active enhancers newly formed from primed enhancers during the transition from Ngn3-high cells to β cells (Fig. S3A). Of this class of active enhancers, 63.5% originated from the newly primed enhancers in Ngn3-high cells (Fig. S3A). The genes associated with this enhancer state change showed an increase in net expression level during the transition from Ngn3-high cells to β cells (Fig. S3B,C).

To verify whether the occupancy of NeuroD1 was accompanied by the establishment of primed and active enhancers during the transition from Ngn3-low to Ngn3-high status, we performed NeuroD1 ChIP-seq in cells from E14.5 whole pancreas (Fig. S4A).
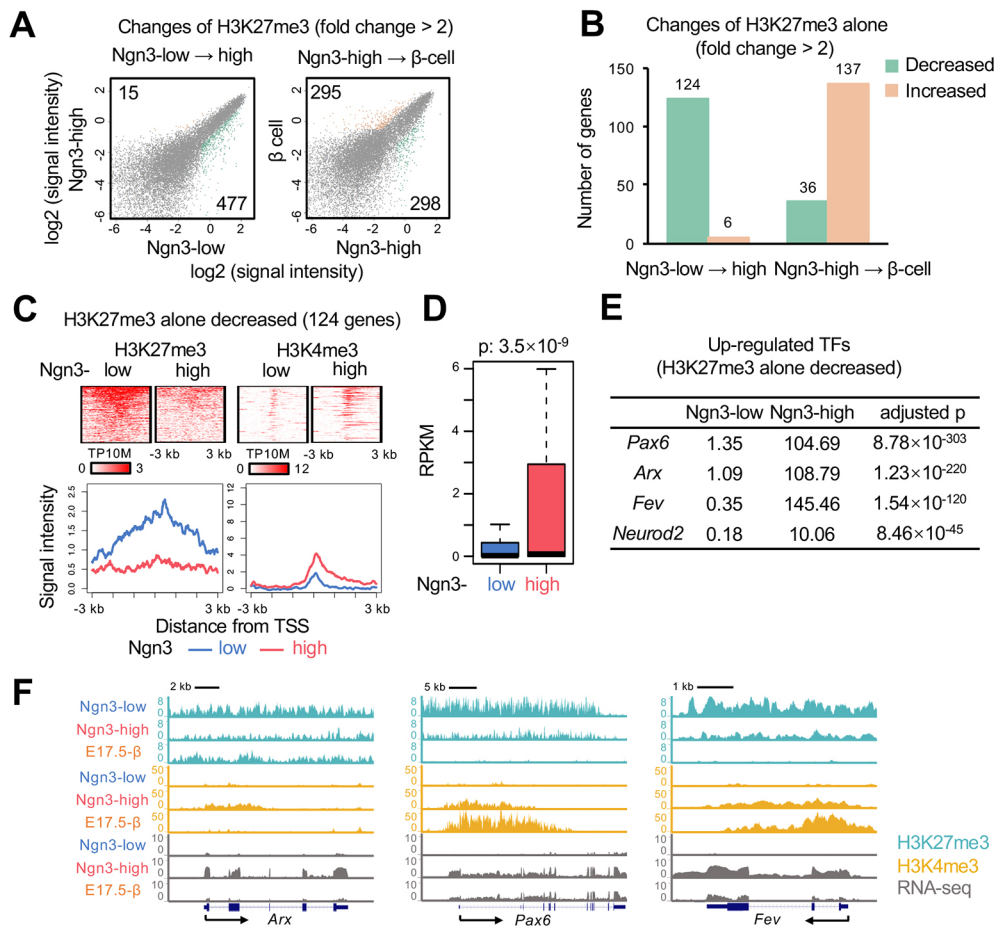
Fig. 2. Dynamics of H3K27me3-alone promoter-associated histone marks during the transition from Ngn3-low to Ngn3-high. (A) Comparison of H3K27me3 levels on promoters with >2-fold changes in H3K27me3 during the differentiation from E13.5 Ngn3-low to Ngn3-high cells (left) or the differentiation from E13.5 Ngn3-high cells to E17.5 β cells (right). The plots represent the genes in the whole genome. (B) Quantification of genes with >2-fold changes in H3K27me3 levels on promoters marked by H3K27me3 alone during the cell fate transition. (C) Density of H3K27me3 and H3K4me3 ChIP-seq tags at TSSs±3 kb of promoters for H3K27me3 alone with reduced H3K27me3 levels (>2-fold changes) during the transition from Ngn3-low to Ngn3-high (top). Composite plots showing the distributions of H3K27me3 and H3K4me3 marks around TSSs (bottom). (D) Boxplot showing the expression levels of genes identified in C. See Fig. 1D for explanation of box and whiskers. Wilcoxon rank-sum test. (E) List and expression levels (RPKM) of the key upregulated TFs (adjusted $P<0.05$, fold change >2, RPKM >10 in Ngn3-high) identified in C. (F) Genome browser images of histone H3K4me3 and H3K27me3 modifications, and RNA-seq patches of genes with reduced H3K27me3 levels (>2-fold changes) at H3K27me3-alone promoters during the transition from Ngn3-low to Ngn3-high.

We identified 1609 peaks, representing the NeuroD1-binding sites in endocrine lineages expressing *Neurod1*. Notably, 65.4% of the 1609 peaks contained the NEUROD1 motif (Fig. 4A), which indicated that the peaks identified in this study were reliable. After normalization for the small size of gene regulatory regions within the entire genome, it was clear that NeuroD1 binding-site peaks were preferentially enriched in enhancer and promoter regions (Fig. 4B,C). Further comprehensive analysis revealed that NeuroD1 bound to 569 Ngn3-high active putative enhancers, among which 79.4% were newly formed enhancers, and associated genes showed an increased net expression level during the transition from Ngn3-low to Ngn3-high (Fig. 4D-G). We also found that 274 Ngn3-high primed putative enhancers were occupied by NeuroD1, and 77.4% of those were newly formed primed enhancers (Fig. 4D). Gene ontology (GO) analysis revealed that the genes associated with these newly formed active and primed enhancers were involved in developmental processes and pancreatic endocrine functions (Fig. 4H and Table S5). Therefore, these data indicate that the pancreatic endocrine-specific TF NeuroD1 may function as a pioneer factor accompanied by the establishment of a subset of active or primed enhancers in Ngn3-high cells. However, NeuroD1 may not be the only pioneer factor during the transition, given that many other newly formed enhancers contain other motifs, such as FOXA2 and FOXO1, which suggests that the generation of Ngn3-high cells probably requires multiple pioneer factors (Fig. 3F,H). In addition to the enhancer loci, we identified NeuroD1-binding peaks at 67 promoter regions (Fig. S4B). The related genes, including *Hes6* and *Neurod1* (which encode TFs) showed increased net expression

levels during the transition from Ngn3-low to Ngn3-high (Fig. 4C and Fig. S4C).

Taken together, these data characterize the dramatic dynamics of primed and active enhancers, which are accompanied by the transition from Ngn3-low to Ngn3-high. Our study also indicates that the activation of *Neurod1* after the resolution of H3K27me3 from its bivalent promoter is accompanied by the establishment of active and primed enhancers in Ngn3-high cells. Given that the removal of H3K27me3 was associated with the activation of other key TFs related to endocrine development, in addition to the activation of NeuroD1 (Fig. 1E,F), we hypothesized that a disturbance in the removal of H3K27me3 would perturb the cell fate transition from Ngn3-low to Ngn3-high.

Jmjd3 regulates the transition from Ngn3-low to Ngn3-high

Only two histone demethylases, JMJD3 and UTX, have been proved to be responsible for the removal of H3K27me3 mark (Agger et al., 2007; Lan et al., 2007). Studies have indicated that JMJD3 regulates promoter-specific H3K27me3 levels in human ES cells (Akiyama et al., 2016). Hence, we examined whether the precocious loss of *Jmjd3* expression would affect cell fate commitment from Ngn3-low to Ngn3-high. We used a *Jmjd3* conditional allele (Fig. S5A) and a *Pdx1-Cre* transgene to delete *Jmjd3* from pancreatic progenitor cells on an *Ngn3-GFP* background (Hingorani et al., 2003). The expression level of *Jmjd3* was reduced in sorted Ngn3-low and Ngn3-high cells from E13.5 *Pdx1-Cre;Jmjd3^fl/fl;Ngn3-GFP* dorsal pancreata (Fig. S5B). Flow cytometric analysis also showed that in E13.5 *Pdx1-Cre; Jmjd3^fl/fl* pancreata, the number of Ngn3-low cells increased
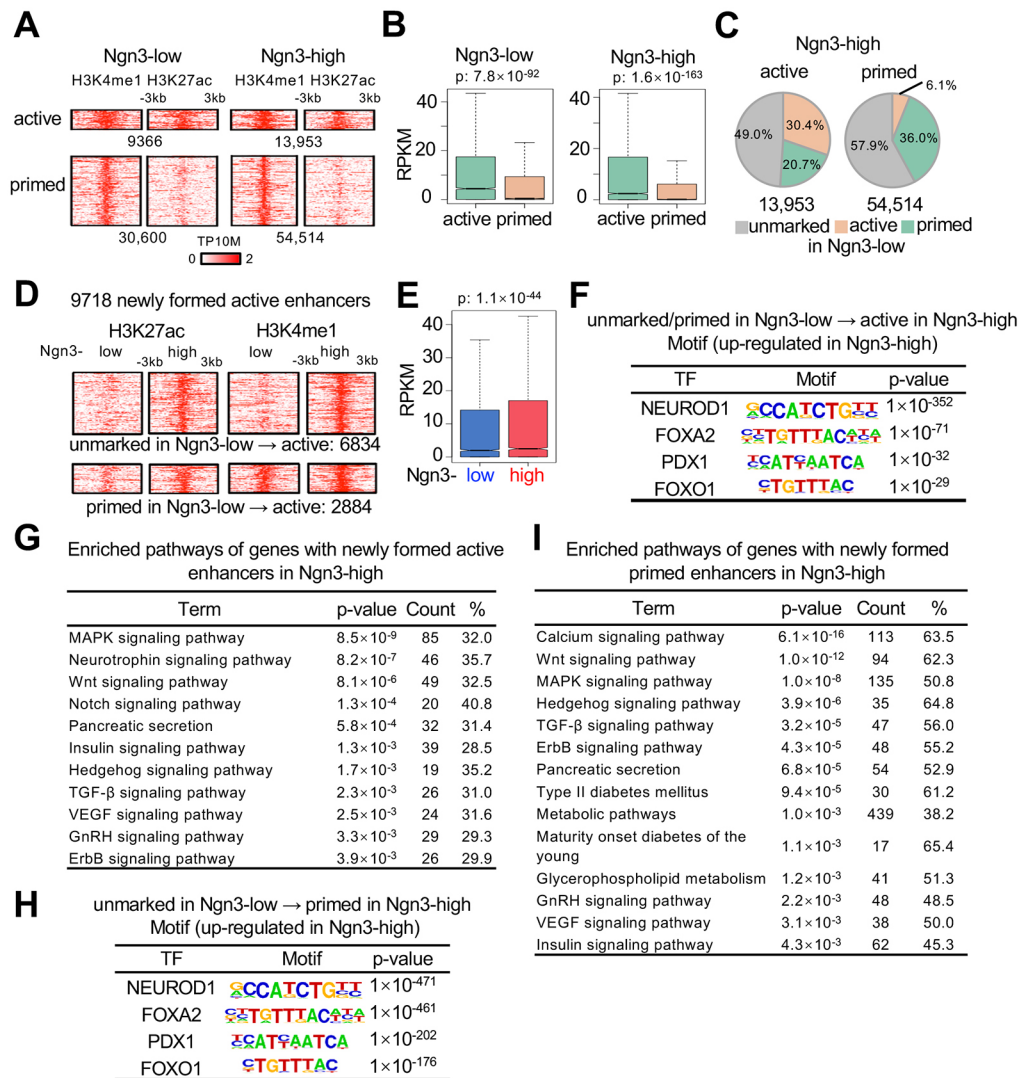
Fig. 3. Dynamics of primed and active enhancer-associated histone modifications during the transition from Ngn3-low to Ngn3-high.
(A) Density of H3K4me1 and H3K27ac ChIP-seq tags at active and primed putative enhancers in Ngn3-low and Ngn3-high cell populations. (B) Boxplots showing mRNA expression levels of genes associated with active and primed enhancers in Ngn3-low and Ngn3-high cells. (C) Global changes of active and primed enhancers during the cell fate transition. (D) Density of ChIP-seq tags on newly formed active enhancers in Ngn3-high cells. (E) Boxplot showing upregulated mRNA expression levels of genes associated with newly formed active enhancers in D. See Fig. 1D for explanation of box and whiskers. Wilcoxon rank-sum test. (F) Enriched TF-binding motifs for newly formed active enhancers in Ngn3-high cells. Expression levels of related TFs showing significant upregulation (adjusted $P<0.05$, fold change >2, RPKM >10) in Ngn3-high cells. (G) Selected enriched KEGG pathway terms for genes associated with newly formed active enhancers in Ngn3-high cells. (H) Enriched TF-binding motifs for newly formed primed enhancers in Ngn3-high cells. Expression levels of related TFs showing significantly upregulation (adjusted $P<0.05$, fold change >2, RPKM >10) in Ngn3-high cells. (I) Selected enriched KEGG pathway terms for genes associated with newly primed enhancers in Ngn3-high cells.

A

Ngn3-low — Ngn3-high
H3K4me1 H3K27ac — H3K4me1 H3K27ac
−3kb 3kb / −3kb 3kb

active — 9366 / 13,953
primed — 30,600 / 54,514
TP10M 0–2

B

Ngn3-low p: 7.8×10^{-92} / Ngn3-high p: 1.6×10^{-163}

RPKM — active primed / active primed

C

Ngn3-high
active / primed — 6.1%

active: 49.0%, 30.4%, 20.7%
primed: 57.9%, 36.0%

13,953 / 54,514
unmarked active primed in Ngn3-low

D

9718 newly formed active enhancers
H3K27ac / H3K4me1
Ngn3- low high / low high
−3kb 3kb / −3kb 3kb

unmarked in Ngn3-low → active: 6834
primed in Ngn3-low → active: 2884

E

p: 1.1×10^{-44}

RPKM — Ngn3- low high

F

unmarked/primed in Ngn3-low → active in Ngn3-high
Motif (up-regulated in Ngn3-high)

TF	Motif	p-value
NEUROD1		1×10^{-352}
FOXA2		1×10^{-71}
PDX1		1×10^{-32}
FOXO1		1×10^{-29}

G

Enriched pathways of genes with newly formed active enhancers in Ngn3-high

Term	p-value	Count	%
MAPK signaling pathway	8.5×10^{-9}	85	32.0
Neurotrophin signaling pathway	8.2×10^{-7}	46	35.7
Wnt signaling pathway	8.1×10^{-6}	49	32.5
Notch signaling pathway	1.3×10^{-4}	20	40.8
Pancreatic secretion	5.8×10^{-4}	32	31.4
Insulin signaling pathway	1.3×10^{-3}	39	28.5
Hedgehog signaling pathway	1.7×10^{-3}	19	35.2
TGF-β signaling pathway	2.3×10^{-3}	26	31.0
VEGF signaling pathway	2.5×10^{-3}	24	31.6
GnRH signaling pathway	3.3×10^{-3}	29	29.3
ErbB signaling pathway	3.9×10^{-3}	26	29.9

H

unmarked in Ngn3-low → primed in Ngn3-high
Motif (up-regulated in Ngn3-high)

TF	Motif	p-value
NEUROD1		1×10^{-471}
FOXA2		1×10^{-461}
PDX1		1×10^{-202}
FOXO1		1×10^{-176}

I

Enriched pathways of genes with newly formed primed enhancers in Ngn3-high

Term	p-value	Count	%
Calcium signaling pathway	6.1×10^{-16}	113	63.5
Wnt signaling pathway	1.0×10^{-12}	94	62.3
MAPK signaling pathway	1.0×10^{-8}	135	50.8
Hedgehog signaling pathway	3.9×10^{-6}	35	64.8
TGF-β signaling pathway	3.2×10^{-5}	47	56.0
ErbB signaling pathway	4.3×10^{-5}	48	55.2
Pancreatic secretion	6.8×10^{-5}	54	52.9
Type II diabetes mellitus	9.4×10^{-5}	30	61.2
Metabolic pathways	1.0×10^{-3}	439	38.2
Maturity onset diabetes of the young	1.1×10^{-3}	17	65.4
Glycerophospholipid metabolism	1.2×10^{-3}	41	51.3
GnRH signaling pathway	2.2×10^{-3}	48	48.5
VEGF signaling pathway	3.1×10^{-3}	38	50.0
Insulin signaling pathway	4.3×10^{-3}	62	45.3

1.8-fold, whereas the number of Ngn3-high cells decreased 1.5-fold, with the ratio of Ngn3-high to Ngn3-low cells decreasing 2.6-fold compared with the ratio in *Pdx1-Cre; Jmjd3*[+/+] pancreata (Fig. 5A,B). Later, at E14.5, similar analysis also revealed that the ratio of Ngn3-high to Ngn3-low cells in knock-out (KO) pancreata decreased 1.7-fold compared with the ratio in wild-type pancreata (Fig. S5C). To determine whether cell death occurred in *Jmjd3* KO cells, we performed TUNEL detection of E14.5 FACS-sorted Ngn3-low and Ngn3-high cells from wild-type or KO pancreata. We observed few apoptotic cells in *Jmjd3* wild-type and KO Ngn3-low and Ngn3-high cells (Fig. S5D). Therefore, the decreased ratio of Ngn3-high to Ngn3-low cells in *Jmjd3* KO pancreata was not induced by cell death. These findings indicate that the cell fate transition from Ngn3-low to Ngn3-high is constrained in *Jmjd3*-ablated mice.

We then examined whether the deletion of *Jmjd3* by *Pdx1-Cre* in pancreatic progenitor cells affected the properties of individual Ngn3-low and Ngn3-high cells. We performed single-cell transcriptomic analyses of 49 Ngn3-low and 59 Ngn3-high cells purified from wild-type pancreata and of 69 Ngn3-low and 44 Ngn3-high cells from *Pdx1-Cre;Jmjd3*[fl/fl] pancreata at E13.5 (Fig. S6A-D). Principal component analysis (PCA) (Pearson, 1901; Hotelling, 1933) revealed a developmental pathway from Ngn3-low to Ngn3-high cells (Fig. 5C). However, the sources of Ngn3-low and Ngn3-high cells, i.e. wild-type or *Pdx1-Cre;Jmjd3*[fl/fl] pancreata, could not be distinguished based on the PCA plot (Fig. 5C). Hierarchical clustering analyses on Ngn3-low and Ngn3-high populations showed that the single cells from wild-type and *Jmjd3* KO pancreata were intermingled in the heat maps (Fig. S5F). Furthermore, a comparative transcriptomic analysis barely identified differentially expressed genes between wild-type and *Jmjd3*-deleted Ngn3-low or Ngn3-high cells (Fig. 5D). This result indicates that *Jmjd3* KO decreased the rate of transition from Ngn3-low to Ngn3-high but did not affect the cell properties before or after cell fate transition.

Considering the dramatic change in the ratio of Ngn3-low to Ngn3-high cells, we next assessed the associated developmental consequence by quantifying the number of nascent β cells from E17.5 *Pdx1-Cre;Jmjd3*[fl/fl] pancreata on an *Ins1-RFP* transgenic background. Using a different genetic background can eliminate the effects of the Ngn3-GFP knock-in mouse line, which has been shown to have an increased capacity for exocrine differentiation (Wang et al., 2010). We observed a 62.3% decrease in the proportion of β cells in *Pdx1-Cre;Jmjd3*[fl/fl] dorsal pancreata compared with that in wild-type mice (Fig. 6A,B). In addition, pancreatic islet areas were reduced and the shape of islet was
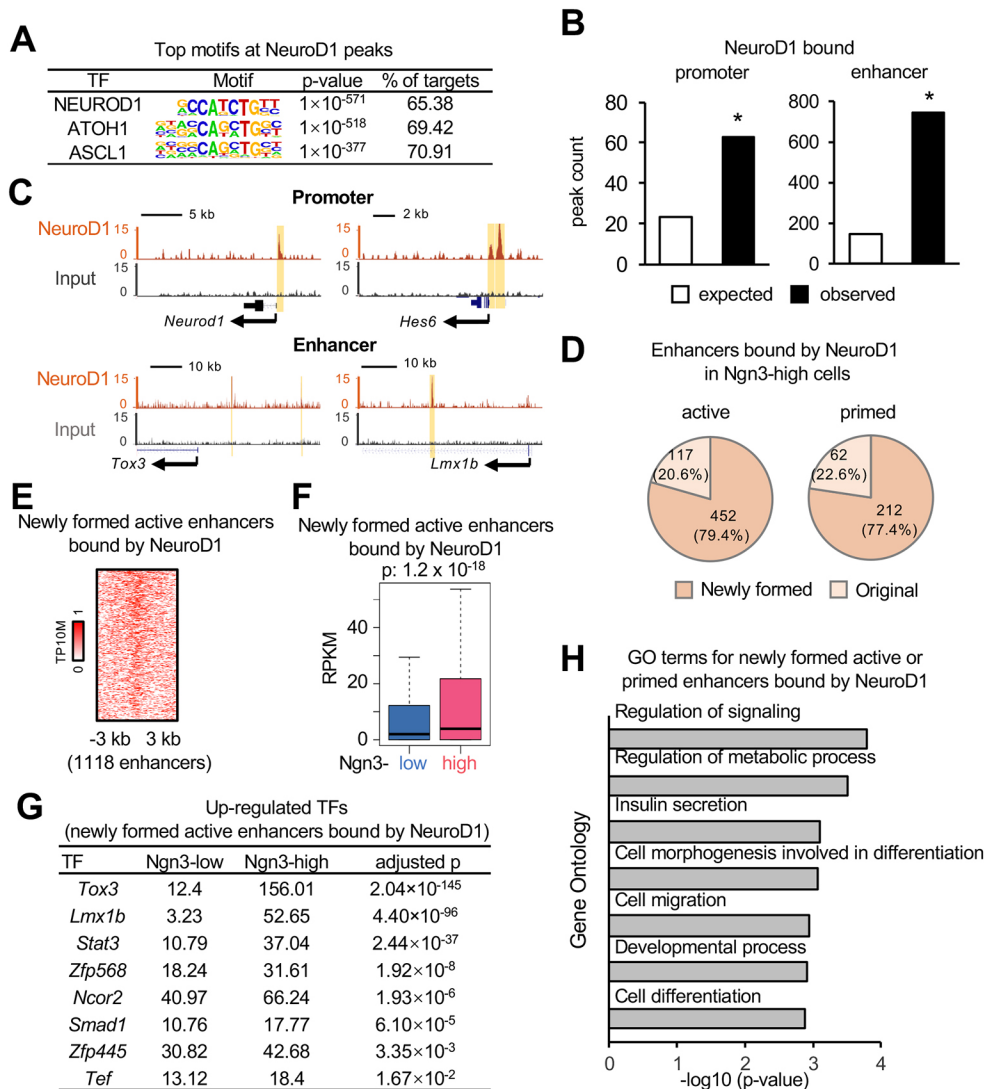
A

Top motifs at NeuroD1 peaks

TF	Motif	p-value	% of targets
NEUROD1		1×10^{-571}	65.38
ATOH1		1×10^{-518}	69.42
ASCL1		1×10^{-377}	70.91

C

Promoter

Enhancer

E

Newly formed active enhancers bound by NeuroD1

-3 kb   3 kb
(1118 enhancers)

F

Newly formed active enhancers bound by NeuroD1

p: 1.2×10^{-18}

Ngn3- low high

G

Up-regulated TFs
(newly formed active enhancers bound by NeuroD1)

TF	Ngn3-low	Ngn3-high	adjusted p
Tox3	12.4	156.01	2.04×10^{-145}
Lmx1b	3.23	52.65	4.40×10^{-96}
Stat3	10.79	37.04	2.44×10^{-37}
Zfp568	18.24	31.61	1.92×10^{-8}
Ncor2	40.97	66.24	1.93×10^{-6}
Smad1	10.76	17.77	6.10×10^{-5}
Zfp445	30.82	42.68	3.35×10^{-3}
Tef	13.12	18.4	1.67×10^{-2}

B

NeuroD1 bound

promoter   enhancer

☐ expected ■ observed

D

Enhancers bound by NeuroD1
in Ngn3-high cells

active   primed

☐ Newly formed ☐ Original

H

GO terms for newly formed active or
primed enhancers bound by NeuroD1

Fig. 4. NeuroD1 is associated with the establishment of primed and active enhancers. (A) The top three enriched motifs at NeuroD1 peaks. (B) Expected counts by random chance versus observed counts of NeuroD1-binding loci on promoters and enhancers. *$P<2.2\times10^{-16}$, chi-square test. (C) Genome browser images of NeuroD1 patches (shaded regions) at promoters of the TFs *Neurod1* and *Hes6* (top), and putative enhancers of the TFs *Tox3* and *Lmx1b* (bottom). (D) The numbers of active and primed enhancers bound by NeuroD1 in E14.5 Ngn3-high cells. (E) Density of NeuroD1 ChIP-seq tags on newly formed active enhancers in E14.5 Ngn3-high cells. (F) Boxplot showing upregulated mRNA expression levels of genes associated with the newly formed active enhancers in D. See Fig. 1D for explanation of box and whiskers. Wilcoxon rank-sum test. (G) List and expression levels (RPKM) of the upregulated TFs (adjusted *P*<0.05, RPKM >10 in Ngn3-high) whose newly formed enhancers are bound by NeuroD1. (H) Selected GO terms for genes with newly formed active or primed enhancers targeted by NeuroD1.

deficient in 6-week-old *Pdx1-Cre;Jmjd3^{fl/fl}* male mice (Fig. 6C). In *Jmjd3* KO islets, the level of H3K27me3 modifications were higher than those in wild-type islets (Fig. S5G). Moreover, *Pdx1-Cre; Jmjd3^{fl/fl}* animals were smaller than wild-type mice (Fig. 6D and Fig. S5H) and had diabetes, demonstrating defects in glucose metabolism and serum insulin secretion after overnight fasting and glucose stimulation (Fig. 6E,F). Our data demonstrate that Jmjd3 promotes the cell fate transition from Ngn3-low to Ngn3-high cells, whereas the deletion of *Jmjd3* by *Pdx1-Cre* at an earlier stage impairs islet development.

DISCUSSION

In this study, we compared the transcriptomic profiles of cell populations at the key stages of pancreatic endocrine cell development. To our surprise, the dramatic changes in gene expression occurred during the transition from Ngn3-low to Ngn3-high, which was completed within ~12 h (Bankaitis et al., 2015). During this process, the cells release the suppression of key TFs, mainly by removing the H3K27me3 mark. Some of these TFs might in turn help establish the primed and active enhancers, which facilitate the fate commitment to Ngn3-high cells (Fig. 7). Therefore, our studies suggest the existence of sequential chromatin modification events during lineage specification.

Cell fate decisions are usually linked to changes in cell cycle progression, and loss of pluripotency is associated with an extension of the G1 phase of the cell cycle (Dalton, 2015). However, whereas pancreatic endocrine progenitor cells lose multipotency and become specified in the islet lineage, cell cycle regulatory genes are broadly repressed, culminating in G1 arrest in Ngn3-high cells. During the G1 cell phase, the key TF genes for endocrine differentiation, such as *Neurod1*, *Pax4*, *Insm1* and *Rfx6*, start to be expressed. Considering the putative role of NeuroD1 in establishing enhancer-related chromatin marks, we presume that a significant number of dynamic changes in H3K4me1 and H3K27ac occur during the G1 phase. However, based on current findings, we cannot speculate whether the resolution of H3K27me3 and H3K4me3 is related to the cell cycle phase. Therefore, this system is suitable for investigating the mechanisms related to the regulation of the cell cycle and chromatin modifications during the loss of multipotency and lineage specification of progenitor cells.

The histone H3K27me3 balance is regulated by the histone methyltransferases EZH2 and EZH1, and their histone demethylase counterparts JMJD3 and UTX (Voigt et al., 2013). Our previous work demonstrated that EZH2 restricts the specification of pancreatic progenitors to an Ngn3-low population (Xu et al., 2014), but the current study, using the same *Pdx1-Cre* system,
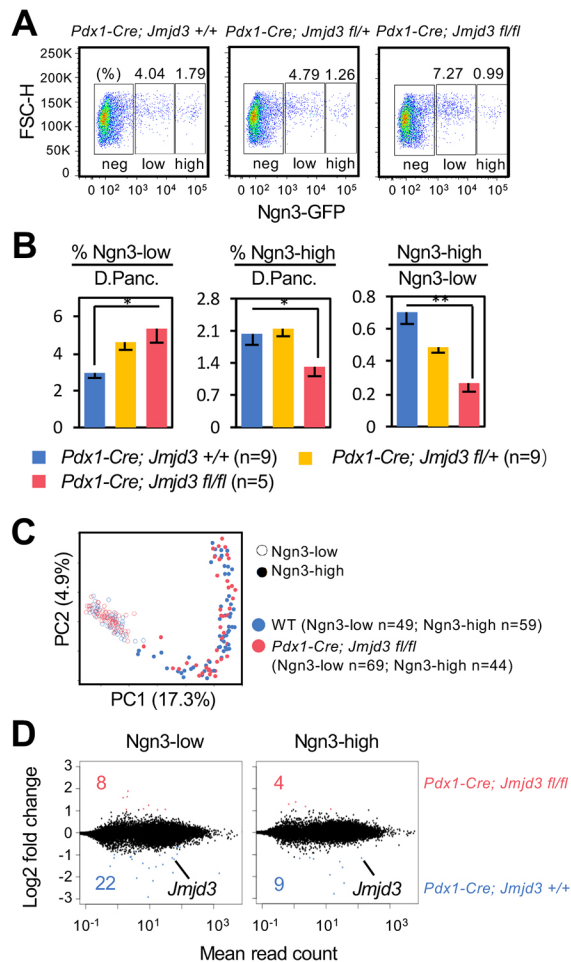
Fig. 5. The histone demethylase Jmjd3 regulates endocrine development.
(A) Flow cytometric analysis shows E13.5 Ngn3-GFP$^-$, Ngn3-low and Ngn3-high cells in dorsal pancreata from *Ngn3-GFP* wild-type embryos and from *Jmjd3* heterozygous and homozygous KO embryos. (B) Statistics for the percentages of E13.5 Ngn3-low and Ngn3-high cells among all dorsal pancreatic cells and the proportions of Ngn3-high/low cells in wild type, and *Jmjd3* heterozygous and homozygous KO pancreata. *n*, number of embryos. Data are mean±s.e.m. *P<0.05, **P<0.01 (Student's *t*-test). (C) PCA plot of single Ngn3-low and Ngn3-high cells from E13.5 wild-type and *Jmjd3* homozygous KO pancreata. *n*, number of cells. (D) Identification of genes differentially expressed in Ngn3-low (left) and Ngn3-high (right) cells from E13.5 wild-type or *Pdx1-Cre;Jmjd3$^{fl/fl}$* pancreata (fold change >1.5, adjusted *P*<0.1). The numbers indicate significantly highly expressed genes in wild-type (blue) or *Pdx1-Cre;Jmjd3$^{fl/fl}$* (red) pancreata.

revealed that Jmjd3 promotes the transition from Ngn3-low to Ngn3-high. Therefore, rather than playing opposing roles during the same step of cell lineage commitment, Ezh2 and Jmjd3 function at different stages of endocrine cell specification.

In our study, we performed single-cell RNA transcriptomic analyses in Ngn3-low and Ngn3-high cells from *Jmjd3* wild-type or KO embryos. Surprisingly, we found no obvious differences between wild-type and KO cells in terms of the transcriptome and developmental pathways (Fig. 5C,D). This result indicated that knocking out *Jmjd3* blocked the transition from Ngn3-low to Ngn3-high. We discussed that in Ngn3-low cells, the histone modification H3K27me3 produces a chromatin barrier to pancreatic endocrine specification by repressing the expression of key TFs. During the transition from Ngn3-low to Ngn3-high, Jmjd3 reduced this chromatin barrier, thereby facilitating the cell fate transition. In

Jmjd3 KO pancreas, the proportion of cells that could overcome this chromatin barrier was significantly decreased; however, once these cells overcame the barrier, they developed normally, which was demonstrated at the single-cell level. In addition, in E17.5 *Jmjd3* KO pancreata, we observed differentiated β cells (Fig. 6A,B). This finding is consistent with our earlier studies of foregut endoderm differentiation, which indicated that the histone modifications, the histone modifier itself, or both play a modulatory role rather than an absolute governance role in cell fate induction (Xu et al., 2011). This finding also suggests that transiently regulating the activities of chromatin modifiers using small molecules at specific times *in vitro* might be a safe way to regulate cell fate transition.

The development of endocrine cells *in vivo* is controlled in a coordinated manner by cell signaling pathways, cell cycle regulators, key TFs and histone modifiers. However, how the cells integrate all of these different levels of regulation into the determination of cell fate at a precise developmental stage needs further investigation. Understanding the mechanisms underlying endocrine cell specification will undoubtedly facilitate the *in vitro* generation of functional β cells.

MATERIALS AND METHODS
Mouse strains
Pdx1-Cre (Hingorani et al., 2003), *Ngn3-GFP* (Lee et al., 2002), *Ins1*-RFP (Piccand et al., 2014) and *Jmjd3$^{fl/fl}$* mice were used. *Jmjd3$^{fl/fl}$* mice were generated at the Nanjing Biomedical Research Institute of Nanjing University. All protocols were approved by the Institutional Animal Care and Use Committees of Peking University. All mice were housed and maintained under specific pathogen-free conditions on a 12 h/12 h day/night cycle at 23±2°C and received an autoclaved standard diet and water *ad libitum*.

Cell preparation and sorting
Dissected E13.5 and E14.5 dorsal pancreata were dissociated into single-cell suspensions by digestion with 0.25% trypsin for 5 min at 37°C. The digestion was terminated by adding 0.4 volumes of fetal bovine serum. E17.5 pancreata were digested with 0.5 mg/ml of collagenase P (Roche, 11213873001) for 2 min at 37°C, followed by trypsin treatment. Cells were sorted using a BD FACSAria II cell sorter.

Bulk-cell RNA-seq and library construction
Total RNA was isolated from ~4-6×10^4 sorted cells using an RNeasy Micro Kit (Qiagen, 74004). mRNA was purified from ~200 ng of total RNA using an NEBNext Poly(A) mRNA Magnetic Isolation Module (NEB, E7490). Libraries were prepared using an NEBNext Ultra Directional RNA Library Prep Kit for Illumina (NEB, E7420), according to the manufacturer's instructions. The quality of the cDNA libraries was assessed using an Advanced Analytical Fragment Analyzer. RNA-seq analyses were performed on two independent biological replicates. Correlations between the replicates are shown in Fig. S1B.

Single-cell RNA-seq and library construction
Single cells were manually picked up from FACS-purified bulk cells by using a mouth pipette. cDNA synthesis was performed following the Smart-seq2 method described by Picelli et al. (2014). Briefly, 0.1 µl of ERCC spike-in RNA (1:1,000,000, Life Technologies, 4456740) was added to 4 µl of lysis buffer. The libraries were prepared with 2.5 ng of cDNA using a TruePrep DNA Library Prep Kit (Vazyme, TD502) (Qiu et al., 2017).

ChIP-seq and library construction
Low-cell-number ChIP-seq was performed following the method of Brind'Amour et al. (2015). We used ~1.5×10^4 cells per ChIP with primary antibodies (1 ng each) recognizing H3K4me3 (Abcam, ab8580, GK9535-2), H3K27me3 (Millipore, 07-449, 2318778), H3K4me1 (Millipore, 07-436, 2343581) and H3K27ac (Active motif, 39133,
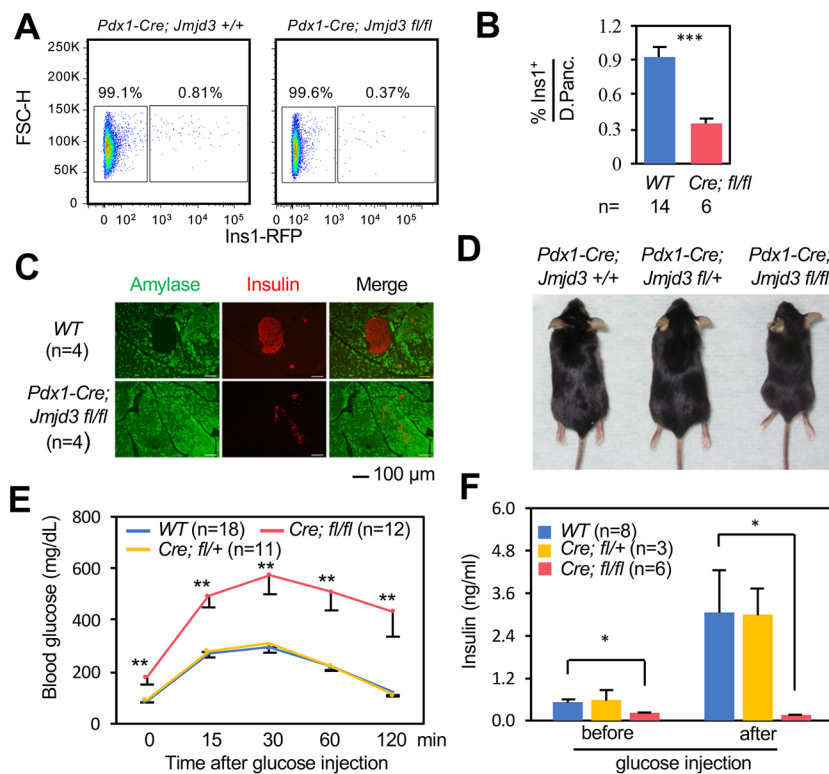
Fig. 6. Islet development is impaired in *Jmjd3* KO mice.
(A) Flow cytometric analysis shows E17.5 Ins1-RFP⁺ cells in dorsal pancreata from wild-type and *Jmjd3* homozygous KO embryos. (B) Percentage of Ins1-RFP⁺ cells among all E17.5 dorsal pancreatic cells. *n*, number of embryos. (C) Insulin and amylase immunofluorescence of pancreata from 6-week-old male mice. *n*, number of mice. (D) Body sizes of 6-week-old animals with homozygous *Jmjd3* deletions were smaller than those of wild type and animals with heterozygous *Jmjd3* deletions. (E,F) Blood glucose levels (E) and insulin secretion levels (F) after glucose injections in fasted 6-week-old male mice. *n*, number of mice. Data are mean±s.e.m. *P<0.05, **P<0.01, ***P<0.001 (Student's *t*-test).

01613007). NeuroD1 ChIP-seq was performed following the method of Johnson et al. (2007). We used ~8×10⁶ cells per ChIP with an antibody against NeuroD1 (Santa Cruz Biotechnology, sc-1084, A0616), which has now been discontinued at Santa Cruz Biotechnology. The library was prepared using the Ultra DNA Library Prep Kit for Illumina (NEB, E7370). Correlations between the replicates are shown in Fig. S2A.

Immunofluorescence staining and microscopy
FACS-sorted E14.5 Ngn3-low and Ngn3-high cells were attached on poly-prep slides (Sigma, P0425-72EA), fixed with 4% paraformaldehyde and stained with a TUNEL apoptosis detection kit (Alexa Fluor 647) (Yesen, 40308), following the manufacturer's instructions. Pancreata from 6-week-old male mice were fixed with 4% paraformaldehyde and

embedded in paraffin wax. Thereafter, 5 μm paraffin sections were stained with guinea pig anti-insulin (1:500, Abcam, ab7842, GR119778-1) and mouse anti-amylase (1:500, Santa Cruz Biotechnology, sc-46657, B0911) or goat anti-Pdx1 (1:2000, Abcam, ab47383, GR16180-1) and rabbit anti-H3K27me3 (1:800, Millipore, 07-449, 2318778) primary antibodies. The staining of Pdx1 and H3K27me3 was with antigen recovery using Vector Antigen Unmasking Solutions (Vector Labs, H-3300) at high pressure for 8 min. The secondary antibodies used were Alexa Fluor 488 donkey anti-mouse (Invitrogen, A21202), Alexa Fluor 594 goat anti-guinea pig (Invitrogen, A11076), Alexa Fluor 488 donkey anti-goat (Invitrogen, A11055) and Alexa Fluor 594 donkey anti-rabbit (Invitrogen, A21207). Fluorescent images were obtained using a Zeiss Axioimager M2 or a Zeiss Axio Scan Z1.

RNA isolation and RT-qPCR
Total RNA was isolated from FACS-sorted cells using the RNeasy Micro Kit (Qiagen) and reverse-transcribed using HiScript II Q RT SuperMix (Vazyme, R223). The RT-qPCR primers are: *Gapdh* forward, 5′-ATGGT-GAAGGTCGGTGTGAAC-3′; *Gapdh* reverse, 5′-GCCTTGACTGTGC-CGTTGAAT-3′; *Jmjd3* forward, 5′-CCTCACCGCCTATCAGTA-3′; *Jmjd3* reverse, 5′-GCCATTCTCACTTGTAACG-3′.

Glucose tolerance test and insulin secretion assay
Six-week-old male *Jmjd3* KO mice and their littermates were fasted for 16 h and then intraperitoneally injected with glucose at 2 mg/g body weight. Blood glucose levels were monitored before injection and at 15, 30, 60 and 120 min after injection. For insulin secretion assays, blood was drawn both before and 30 min after glucose injection, and the serum insulin levels were assayed using a Rat/Mouse Insulin ELISA Kit (Millipore, EZRMI-13K), following the manufacturer's instructions. *Jmjd3*-deleted mice with or without the *Ngn3-GFP* background showed similar phenotypes; thus, we pooled their results for statistical analyses.

Western blot
Adult islet cells were lysed in RIPA buffer with protease inhibitor cocktail (Sigma, P8340) and subjected to standard immunoblotting analysis. Goat anti-NeuroD1 (1:100, Santa Cruz Biotechnology, sc-1084, Lot# A0616)
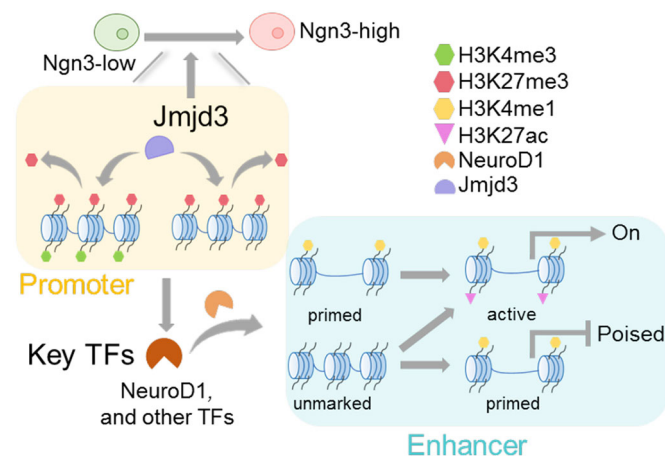

Fig. 7. Model of sequential chromatin regulatory steps during pancreatic endocrine lineage specification. During the transition from Ngn3-low to Ngn3-high cells, the histone demethylase Jmjd3 releases the suppression of key TF genes by removing the H3K27me3 mark from their promoter loci. Some of the activated TFs might in turn facilitate the establishment of primed and active enhancers to promote cell fate commitment.

and donkey anti-goat IgG-HRP (1:500, Santa Cruz Biotechnology, sc-2020) antibodies were used.

Processing of bulk-cell RNA-seq data

Bulk-cell RNA-seq libraries were sequenced as 101 bp paired-end or 51 bp single-end clusters using an Illumina HiSeq 2000 System to a depth of 2-4×10^7 reads. The sequencing reads were aligned to the *Mus musculus* genome (mm10) using TopHat (v2.1.0) (Kim et al., 2013) with the parameters '-o out_dir -G gtf –transcriptome-index trans_index bowtie2_index input_fastq1 input_fastq2'. Reads aligned to genes were counted using HTSeq (v0.6.0) (Anders et al., 2015) with the parameters 'htseq-count -f bam -r pos -s reverse -a 30 input_bam gtf'. Gene expression levels were quantified as reads per kilobase of the longest transcript per million mapped reads (RPKM). The normalization of RPKM values was performed by size factor (Anders and Huber, 2010). Ward's hierarchical clustering was performed with Spearman's correlation distance metric (Fig. S1C). Genes shown in Fig. S1C were identified based on the following criteria: coefficient of variation >0.3 and expressed in at least two samples (normalized RPKM >5). Differential expression analysis was performed using the R package DESeq2 (v1.12.4) (Love et al., 2014) with adjusted *P*-value (padj) <0.05 and fold change >2. Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa and Goto, 2000) pathway and Gene Ontology (GO) (Ashburner et al., 2000) enrichment analyses were performed using the R package GOstats (v2.38.1) (Falcon and Gentleman, 2007).

Processing of ChIP-seq data

ChIP-seq libraries were sequenced as 51 bp single-end reads on an Illumina HiSeq 2500 system to a depth of $1-4 \times 10^7$ reads. The sequencing reads were aligned to the *Mus musculus* genome (mm10) using bowtie2 (v2.2.5) (Langmead and Salzberg, 2012) with default parameters, and only reads with a mapping quality ≥30 were retained for subsequent analyses. PCR duplicate reads were removed using samtools (v1.3.1) (Li et al., 2009) with the commands 'samtools sort in_bam|samtools rmdup -s - out_bam'. Reads from biological replicates were pooled. TP10M (tags per 10 million) values were calculated and stored in the bedGraph format using Homer (Heinz et al., 2010) with the parameters 'makeUCSCfile out_dir –o out.bdg –name sample_name -color track_color –fragLength 150 –avg -fsize 1e20'. To calculate signal intensity, we subtracted the input TP10M from the ChIP TP10M. Regions around the transcription start sites (TSSs) ±2 kb were considered as promoters. Promoter regions that overlapped with ENCODE blacklisted genomic regions (Encode Project Consortium, 2012) were excluded from subsequent analyses. The average signal intensity of H3K27me3 and H3K4me3 at the promoters was determined using the ScoreMatrixBin function of the R genomation package (v1.4.2) (Akalin et al., 2015). For genes with multiple transcript isoforms, the start sites of the isoform with the highest average H3K4me3 signal intensity was selected as the TSSs of these genes. Promoters for which the average H3K4me3 (or H3K27me3) signal intensity was greater than 1 (or 0.7) were considered to be marked by H3K4me3 (or H3K27me3), and promoters marked by both H3K4me3 and H3K27me3 were considered to be bivalent (Fig. 1A). A comparison of the averaged H3K27me3 signal intensities at the promoters is shown in Fig. 2A, in which genes with an H3K27me3 signal intensity >0.7 and a fold change >2 were identified.

Highly enriched regions (peaks) of H3K4me1 were identified using MACS (v1.4.2) (Zhang et al., 2008) with default settings. We merged H3K4me1 peaks that did not overlap with ENCODE blacklisted genomic regions and regions ±3 kb in length around known TSSs as potential enhancers. The average signal intensities of H3K4me1 and H3K27ac at the potential enhancers were calculated using the ScoreMatrixBin function of the R genomation package (v1.4.2) (Akalin et al., 2015). Potential enhancers with an H3K4me1 signal intensity >1 and an H3K27ac signal intensity ≤1 were defined as primed enhancers, whereas potential enhancers with an H3K4me1 signal intensity >1 and an H3K27ac signal intensity >1 were defined as active enhancers. The closest gene to a given enhancer was regarded as its putative target. Motif enrichment analysis was performed using the Homer (Heinz et al., 2010) 'findMotifsGenome.pl' script with default parameters.

NeuroD1-binding sites were identified using MACS (v1.4.2) (Zhang et al., 2008) with parameters '-f BAM -g mm -n sample_name -t ip_bam -c

input_bam -p 1e-5'. The regions of enhancers (center ±2 kb) overlapping with NeuroD1 peaks were considered as NeuroD1-bound enhancers. Promoter regions (−800 to +200 bp flanking the TSSs) overlapping with NeuroD1 peaks were considered as NeuroD1-targeted promoters. The percentage of NEUROD1 motifs occurring at NeuroD1 peaks was calculated using Homer (Heinz et al., 2010) with commands 'findMotifsGenome.pl Neruod1_peaks.bed mm10 motif_dir -size given'.

Processing of single-cell RNA-seq data

Single-cell RNA-seq libraries were sequenced as 51 bp single-end reads on an Illumina HiSeq 2500 system to a depth of $0.5-2 \times 10^6$ reads. The sequencing reads were aligned to the *Mus musculus* genome (mm10) using TopHat (v2.1.0) (Kim et al., 2013) with the parameters '-o out_dir -G gtf –transcriptome-index trans_index bowtie2_index input_fastq'. Reads aligned to genes were counted using HTSeq (v0.6.0) (Anders et al., 2015) with the parameters 'htseq-count -f bam -r pos -s no -a 30 input_bam gtf'. Gene expression levels were quantified as RPKM, which were correlated with the number of mRNA molecules (Fig. S6A). Saturation analysis show that 0.5 million mapped reads were sufficient to detect most genes in the single-cell libraries (Fig. S6B). Therefore, the cells having more than 0.5 million mapped reads were retained for further analyses (Fig. S6C and Table S1). Cells expressing the leukocyte marker gene *Spi1* (*PU.1*) (Hromas et al., 1993) were also excluded from subsequent analyses. In general, 6000-9000 genes were detected in each cell (Fig. S6D). We accounted for technical noise and identified highly variable genes based on the ERCC spike-in as described previously (Brennecke et al., 2013). PCA was performed using the 'PCA' function of the FactoMineR package (v1.34) (Le et al., 2008) with the log2(RPKM+0.1) values of highly variable genes. The graphic displays for the PCA results were generated using the R package ggplot2 (v2.2.1) (Hadley, 2009). Ward's hierarchical clustering was performed with Spearman's correlation distance metric on highly variable genes (Fig. S5F).

Statistics

The standard error of the mean (s.e.m.) and *P*-values of unpaired two tails *t*-tests were calculated for analyses of FACS, blood glucose levels, insulin secretion levels, imaging, RT-qPCR and body weights. The numbers of independent experiments/biological replicates are indicated in the figures. *P*<0.05 was considered significant. Gene expression data are shown in box plots; unpaired or paired two-tailed Wilcoxon rank-sum tests were performed, and the *P*-values are presented above the box plots. The chi-square test was performed for Fig. 4B, and the *P*-value is provided in the figure legend.

Data availability

RNA-seq and ChIP-seq data have been deposited at Gene Expression Omnibus under accession number GSE84324.

## Supplementary information

Supplementary information available online at
http://dev.biologists.org/lookup/doi/10.1242/dev.163162.supplemental

References

Agger, K., Cloos, P. A. C., Christensen, J., Pasini, D., Rose, S., Rappsilber, J., Issaeva, I., Canaani, E., Salcini, A. E. and Helin, K. (2007). UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. *Nature* **449**, 731-734.

Ahlgren, U., Pfaff, S. L., Jessell, T. M., Edlund, T. and Edlund, H. (1997). Independent requirement for ISL1 in formation of pancreatic mesenchyme and islet cells. *Nature* **385**, 257-260.

Akalin, A., Franke, V., Vlahovicek, K., Mason, C. E. and Schubeler, D. (2015). Genomation: a toolkit to summarize, annotate and visualize genomic intervals. *Bioinformatics* **31**, 1127-1129.

Akiyama, T., Wakabayashi, S., Soma, A., Sato, S., Nakatake, Y., Oda, M., Murakami, M., Sakota, M., Chikazawa-Nohtomi, N., Ko, S. B. H. et al. (2016). Transient ectopic expression of the histone demethylase JMJD3 accelerates the differentiation of human pluripotent stem cells. *Development* **143**, 3674-3685.

Anders, S. and Huber, W. (2010). Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106.

Anders, S., Pyl, P. T. and Huber, W. (2015). HTSeq-a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166-169.

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T. et al. (2000). Gene ontology: tool for the unification of biology. *Nat. Genet.* **25**, 25-29.

Bankaitis, E. D., Bechard, M. E. and Wright, C. V. (2015). Feedback control of growth, differentiation, and morphogenesis of pancreatic endocrine progenitors in an epithelial plexus niche. *Genes Dev.* **29**, 2203-2216.

Bechard, M. E., Bankaitis, E. D., Hipkens, S. B., Ustione, A., Piston, D. W., Yang, Y.-P., Magnuson, M. A. and Wright, C. V. E. (2016). Precommitment low-level Neurog3 expression defines a long-lived mitotic endocrine-biased progenitor pool that drives production of endocrine-committed cells. *Genes Dev.* **30**, 1852-1865.

Bernstein, B. E., Mikkelsen, T. S., Xie, X., Kamal, M., Huebert, D. J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K. et al. (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125**, 315-326.

Beucher, A., Martín, M., Spenle, C., Poulet, M., Collin, C. and Gradwohl, G. (2012). Competence of failed endocrine progenitors to give rise to acinar but not ductal cells is restricted to early pancreas development. *Dev. Biol.* **361**, 277-285.

Brennecke, P., Anders, S., Kim, J. K., Kołodziejczyk, A. A., Zhang, X., Proserpio, V., Baying, B., Benes, V., Teichmann, S. A., Marioni, J. C. et al. (2013). Accounting for technical noise in single-cell RNA-seq experiments. *Nat. Methods* **10**, 1093-1095.

Brind'Amour, J., Liu, S., Hudson, M., Chen, C., Karimi, M. M. and Lorincz, M. C. (2015). An ultra-low-input native ChIP-seq protocol for genome-wide profiling of rare cell populations. *Nat. Commun.* **6**, 6033.

Calo, E. and Wysocka, J. (2013). Modification of enhancer chromatin: what, how, and why? *Mol. Cell* **49**, 825-837.

Collombat, P., Mansouri, A., Hecksher-Sorensen, J., Serup, P., Krull, J., Gradwohl, G. and Gruss, P. (2003). Opposing actions of Arx and Pax4 in endocrine pancreas development. *Genes Dev.* **17**, 2591-2603.

Creyghton, M. P., Cheng, A. W., Welstead, G. G., Kooistra, T., Carey, B. W., Steine, E. J., Hanna, J., Lodato, M. A., Frampton, G. M., Sharp, P. A. P. et al. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl. Acad. Sci. USA* **107**, 21931-21936.

Dalton, S. (2015). Linking the cell cycle to cell fate decisions. *Trends Cell Biol.* **25**, 592-600.

ENCODE Project Consortium (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57-74.

Falcon, S. and Gentleman, R. (2007). Using GOstats to test gene lists for GO term association. *Bioinformatics* **23**, 257-258.

Gasa, R., Mrejen, C., Lynn, F. C., Skewes-Cox, P., Sanchez, L., Yang, K. Y., Lin, C.-H., Gomis, R. and German, M. S. R. (2008). Induction of pancreatic islet cell differentiation by the neurogenin-neuroD cascade. *Differentiation* **76**, 381-391.

Gierl, M. S., Karoulias, N., Wende, H., Strehle, M. and Birchmeier, C. (2006). The zinc-finger factor Insm1 (IA-1) is essential for the development of pancreatic beta cells and intestinal endocrine cells. *Genes Dev.* **20**, 2465-2478.

Gradwohl, G., Dierich, A., Lemeur, M. and Guillemot, F. (2000). neurogenin3 is required for the development of the four endocrine cell lineages of the pancreas. *Proc. Natl. Acad. Sci. USA* **97**, 1607-1611.

Gu, G., Dubauskaite, J. and Melton, D. A. (2002). Direct evidence for the pancreatic lineage: NGN3+ cells are islet progenitors and are distinct from duct progenitors. *Development* **129**, 2447-2457.

Hadley, W. (2009). *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer Science & Business Media.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y. C., Laslo, P., Cheng, J. X., Murre, C., Singh, H. and Glass, C. K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576-589.

Hingorani, S., Petricoin, E. F., Maitra, A., Rajapakse, V., King, C., Jacobetz, M. A., Ross, S., Conrads, T. P., Veenstra, T. D., Hitt, B. A. R. et al. (2003). Preinvasive and invasive ductal pancreatic cancer and its early detection in the mouse. *Cancer Cell* **4**, 437-450.

Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *J. Educ. Psychol.* **24**, 417.

Hromas, R., Orazi, A., Neiman, R. S., Maki, R., Van Beveran, C., Moore, J. and Klemsz, M. (1993). Hematopoietic lineage- and stage-restricted expression of the ETS oncogene family member PU.1. *Blood* **82**, 2998-3004.

Iwafuchi-Doi, M. and Zaret, K. S. (2014). Pioneer transcription factors in cell reprogramming. *Genes Dev.* **28**, 2679-2692.

Johnson, J. D. (2016). The quest to make fully functional human pancreatic beta cells from embryonic stem cells: climbing a mountain in the clouds. *Diabetologia* **59**, 2047-2057.

Johnson, D. S., Mortazavi, A., Myers, R. M. and Wold, B. (2007). Genome-wide mapping of in vivo protein-DNA interactions. *Science* **316**, 1497-1502.

Kanehisa, M. and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27-30.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R. and Salzberg, S. L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14**, R36.

Lan, F., Bayliss, P. E., Rinn, J. L., Whetstine, J. R., Wang, J. K., Chen, S., Iwase, S., Alpatov, R., Issaeva, I., Canaani, E. F. et al. (2007). A histone H3 lysine 27 demethylase regulates animal posterior development. *Nature* **449**, 689-694.

Langmead, B. and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357-359.

Le, S., Josse, J. and Husson, F. (2008). FactoMineR: an R package for multivariate analysis. *J. Stat. Softw.* **25**, 1-18.

Lee, C. S., Perreault, N., Brestelli, J. E. and Kaestner, K. H. (2002). Neurogenin 3 is essential for the proper specification of gastric enteroendocrine cells and the maintenance of gastric epithelial cell identity. *Genes Dev.* **12**, 47-56.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G. and Durbin, R. H. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078-2079.

Long, H. K., Prescott, S. L. and Wysocka, J. (2016). Ever-changing landscapes: transcriptional enhancers in development and evolution. *Cell* **167**, 1170-1187.

Love, M. I., Huber, W. and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550.

McKnight, K. D., Wang, P. and Kim, S. K. (2010). Deconstructing pancreas development to reconstruct human islets from pluripotent stem cells. *Cell Stem Cell* **6**, 300-308.

Mellitzer, G., Bonné, S., Luco, R. F., Van De Casteele, M., Lenne-Samuel, N., Collombat, P., Mansouri, A., Lee, J., Lan, M., Pipeleers, D. G. et al. (2006). IA1 is NGN3-dependent and essential for differentiation of the endocrine pancreas. *EMBO J.* **25**, 1344-1352.

Mikkelsen, T., Ku, M., Jaffe, D. B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.-K., Koche, R. P. S. et al. (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**, 553-560.

Naya, F., Huang, H.-P., Qiu, Y., Mutoh, H., DeMayo, F. J., Leiter, A. B. and Tsai, M.-J. J. (1997). Diabetes, defective pancreatic morphogenesis, and abnormal enteroendocrine differentiation in BETA2/neuroD-deficient mice. *Genes Dev.* **11**, 2323-2334.

Ohta, Y., Kosaka, Y., Kishimoto, N., Wang, J., Smith, S. B., Honig, G., Kim, H., Gasa, R. M., Neubauer, N., Liou, A. Y. et al. (2011). Convergence of the insulin and serotonin programs in the pancreatic beta-cell. *Diabetes* **60**, 3208-3216.

Pagliuca, F. W. and Melton, D. A. (2013). How to make a functional beta-cell. *Development* **140**, 2472-2483.

Pataskar, A., Jung, J., Smialowski, P., Noack, F., Calegari, F., Straub, T. and Tiwari, V. K. A. (2016). NeuroD1 reprograms chromatin and transcription factor landscapes to induce the neuronal program. *EMBO J.* **35**, 24-45.

Pearson, K. (1901). LIII. On lines and planes of closest fit to systems of points in space. *Lond. Edinb. Dubl. Phil. Mag. J. Sci.* **2**, 559-572.

Piccand, J., Meunier, A., Merle, C., Jia, Z., Barnier, J. V. and Gradwohl, G. (2014). Pak3 promotes cell cycle exit and differentiation of beta-cells in the embryonic pancreas and is necessary to maintain glucose homeostasis in adult mice. *Diabetes* **63**, 203-215.

Picelli, S., Faridani, O. R., Björklund, A. K., Winberg, G., Sagasser, S. and Sandberg, R. (2014). Full-length RNA-seq from single cells using Smart-seq2. *Nat. Protoc.* **9**, 171-181.

Qiu, W. L., Zhang, Y.-W., Feng, Y., Li, L.-C., Yang, L. and Xu, C.-R. (2017). Deciphering pancreatic islet beta cell and alpha cell maturation pathways and characteristic features at the single-cell level. *Cell Metab.* **25**, 1194-1205 e1194.

Sander, M., Neubuser, A., Kalamaras, J., Ee, H. C., Martin, G. R. and German, M. S. (1997). Genetic analysis reveals that PAX6 is required for normal transcription of pancreatic hormone genes and islet development. *Genes Dev.* **11**, 1662-1673.

Schonhoff, S. E., Giel-Moloney, M. and Leiter, A. B. (2004). Neurogenin 3-expressing progenitor cells in the gastrointestinal tract differentiate into both endocrine and non-endocrine cell types. *Dev. Biol.* 270, 443-454.

Schwitzgebel, V. M., Scheel, D. W., Conners, J. R., Kalamaras, J., Lee, J. E., Anderson, D. J., Sussel, L., Johnson, J. D. and German, M. S. (2000). Expression of neurogenin3 reveals an islet cell precursor population in the pancreas. *Development* 127, 3533-3542.

Shih, H. P., Wang, A. and Sander, M. (2013). Pancreas organogenesis: from lineage determination to morphogenesis. *Annu. Rev. Cell Dev. Biol.* 29, 81-105.

Smith, S., Qu, H.-Q., Taleb, N., Kishimoto, N. Y., Scheel, D. W., Lu, Y., Patch, A.-M., Grabs, R., Wang, J., Lynn, F. C. et al. (2010). Rfx6 directs islet formation and insulin production in mice and humans. *Nature* 463, 775-780.

St-Onge, L., Sosa-Pineda, B., Chowdhury, K., Mansouri, A. and Gruss, P. (1997). Pax6 is required for differentiation of glucagon-producing alpha-cells in mouse pancreas. *Nature* 387, 406-409.

Voigt, P., Tee, W.-W. and Reinberg, D. A. (2013). A double take on bivalent promoters. *Genes Dev.* 27, 1318-1338.

Wang, S., Yan, J., Anderson, D. A., Xu, Y., Kanal, M. C., Cao, Z., Wright, C. V. and Gu, G. (2010). Neurog3 gene dosage regulates allocation of endocrine and exocrine cell fates in the developing mouse pancreas. *Dev. Biol.* 339, 26-37.

Xu, C.-R., Cole, P. A., Meyers, D. J., Kormish, J., Dent, S. and Zaret, K. S. (2011). Chromatin "prepattern" and histone modifiers in a fate choice for liver and pancreas. *Science* 332, 963-966.

Xu, C.-R., Li, L.-C., Donahue, G., Ying, L., Zhang, Y.-W., Gadue, P. and Zaret, K. S. (2014). Dynamics of genomic H3K27me3 domains and role of EZH2 during pancreatic endocrine specification. *EMBO J.* 33, 2157-2170.

Zhang, Y., Liu, T., Meyer, C. A., Eeckhoute, J., Johnson, D. S., Bernstein, B. E., Nusbaum, C., Myers, R. M., Brown, M., Li, W. et al. (2008). Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* 9, R137.

DEVELOPMENT
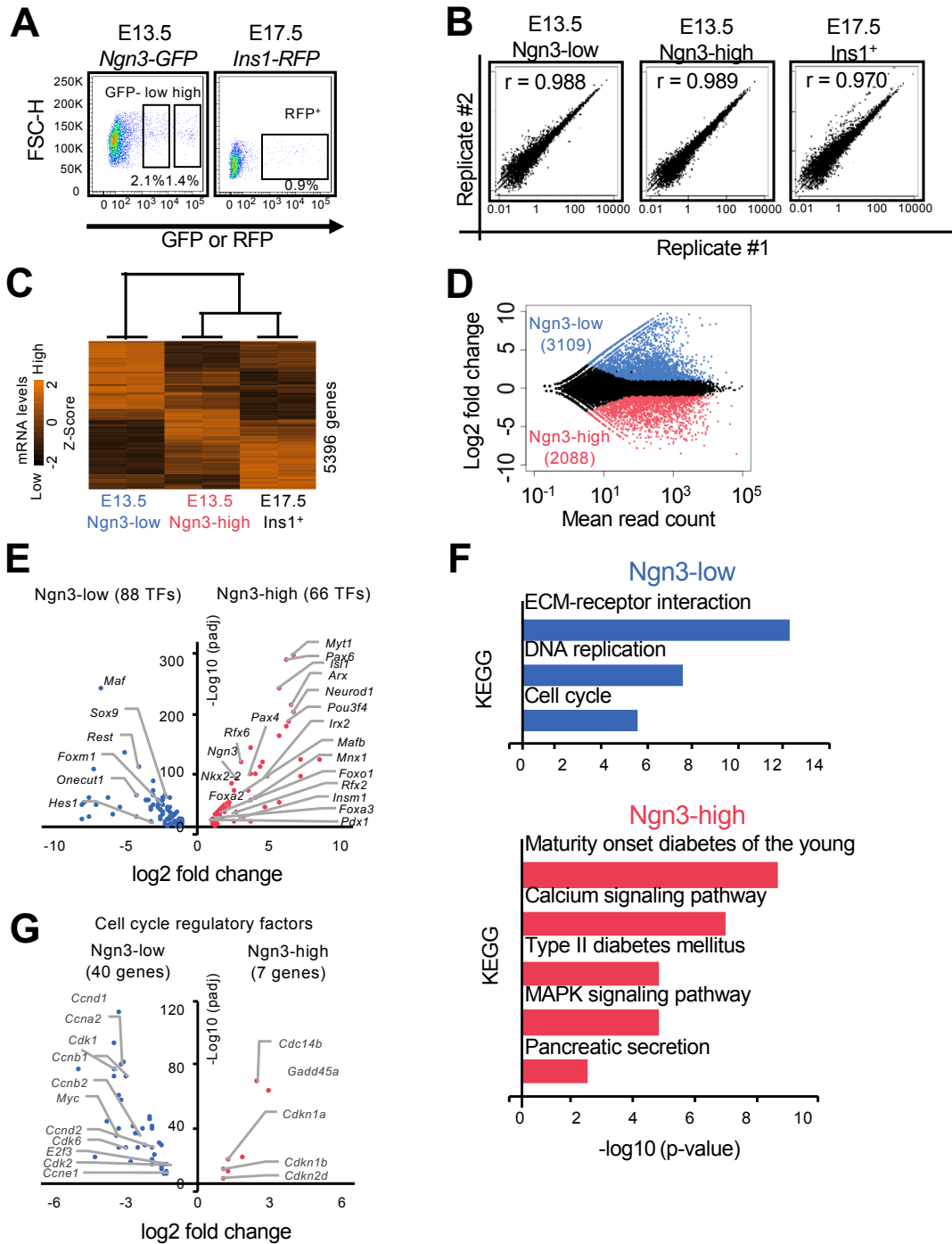
Supplementary Figure 1. Gene expression profiles of bulk Ngn3-low and Ngn3-high cells.

(A) FACS gating for sorting Ngn3-GFP-low and Ngn3-GFP-high cells from E13.5 dorsal pancreatic buds and RFP$^+$ cells from E17.5 dorsal pancreata. The y-axis shows forward scatter high (FSC-H) values. (B) Correlations between bulk-cell RNA-seq replicates. The x- and y-axes represent RPKM for each protein-coding gene. r: Pearson's correlation. (C) Heat map of variably expressed genes in the three cell populations in (A), analyzed by RNA-seq (two replicates). (D) Identification of genes highly expressed in Ngn3-low (blue) and Ngn3-high (red) cells. Fold change > 2, adjusted p < 0.05. (E) Volcano plots listing TFs highly expressed in Ngn3-low (blue) and Ngn3-high (red) cells with adjusted p < 0.05, fold change > 2 and RPKM > 5 in at least one population. The listed TFs have been reported to regulate pancreatic development. (F) Selected enriched KEGG pathway terms for Ngn3-low and Ngn3-high cells. (G) Volcano plots listing cell cycle regulatory genes highly expressed in Ngn3-low (blue) and Ngn3-high (red) cells with adjusted p < 0.05, fold change > 2 and RPKM > 5 in at least one population.

A
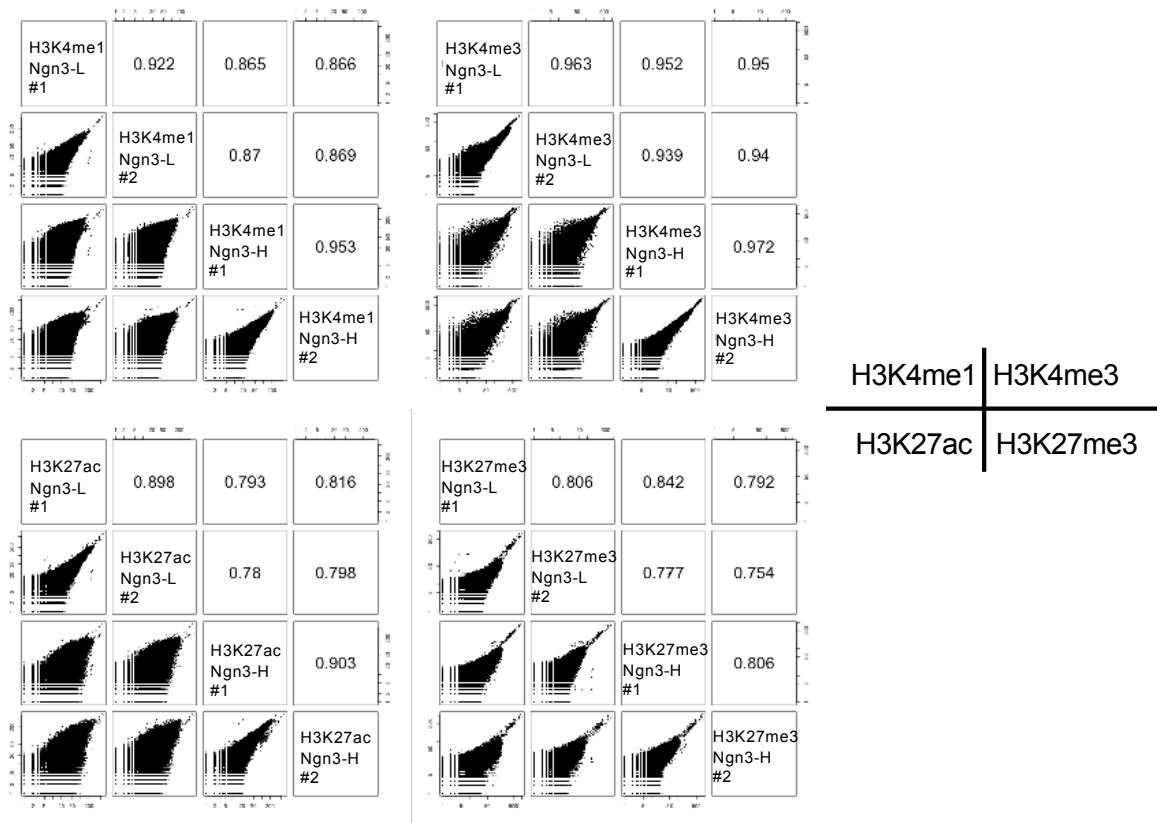
B
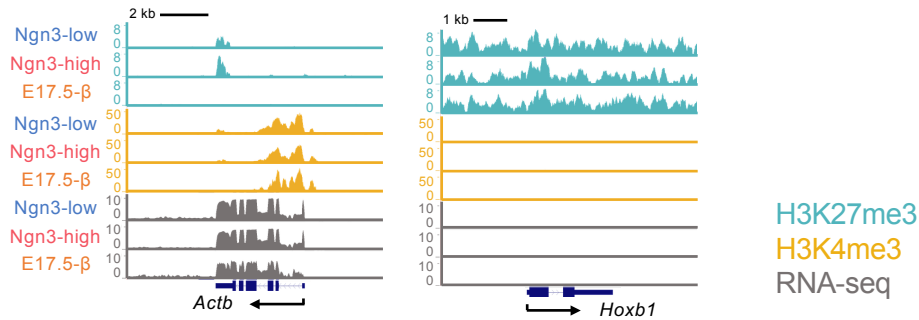
Supplementary Figure 2. Quality control of ChIP-seq data.

(A) Correlations between ChIP-seq replicates. The x- and y-axes of the small figures on the lower left represent read counts for each 1000 bp genomic window. The numbers on the upper right are Pearson's correlations. (B) Genome browser images of histone H3K4me3 and H3K27me3 modifications and RNA-seq patches of control genes during the transition from Ngn3-low to Ngn3-high.
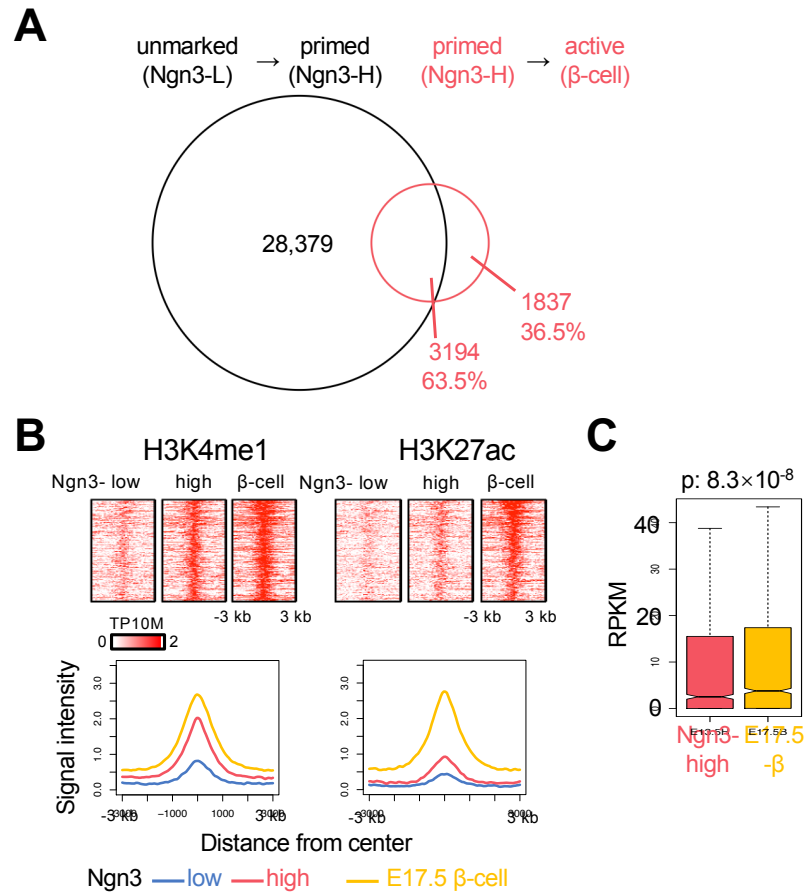
Supplementary Figure 3. A fraction of newly formed primed enhancers in Ngn3-high cells become active enhancers in β cells.

(A) Venn diagram showing the overlap between newly formed primed enhancers in E13.5 Ngn3-high cells and newly formed active enhancers in E17.5 β cells that were primed enhancers in Ngn3-high cells. (B) Density of H3K4me1 and H3K27ac ChIP-seq tags (midpoint ± 3 kb) for the overlapping enhancers shown in (A) (top). Composite plot showing the distributions and signal intensities of H3K4me1 and H3K27ac marks (bottom). (C) Boxplot showing mRNA expression levels of the overlapping enhancers in (A) associated genes in Ngn3-high and β cells. Wilcoxon rank-sum test.
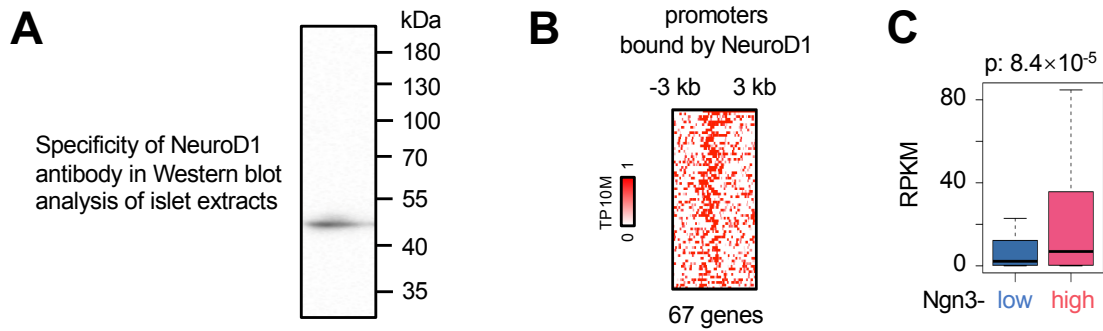
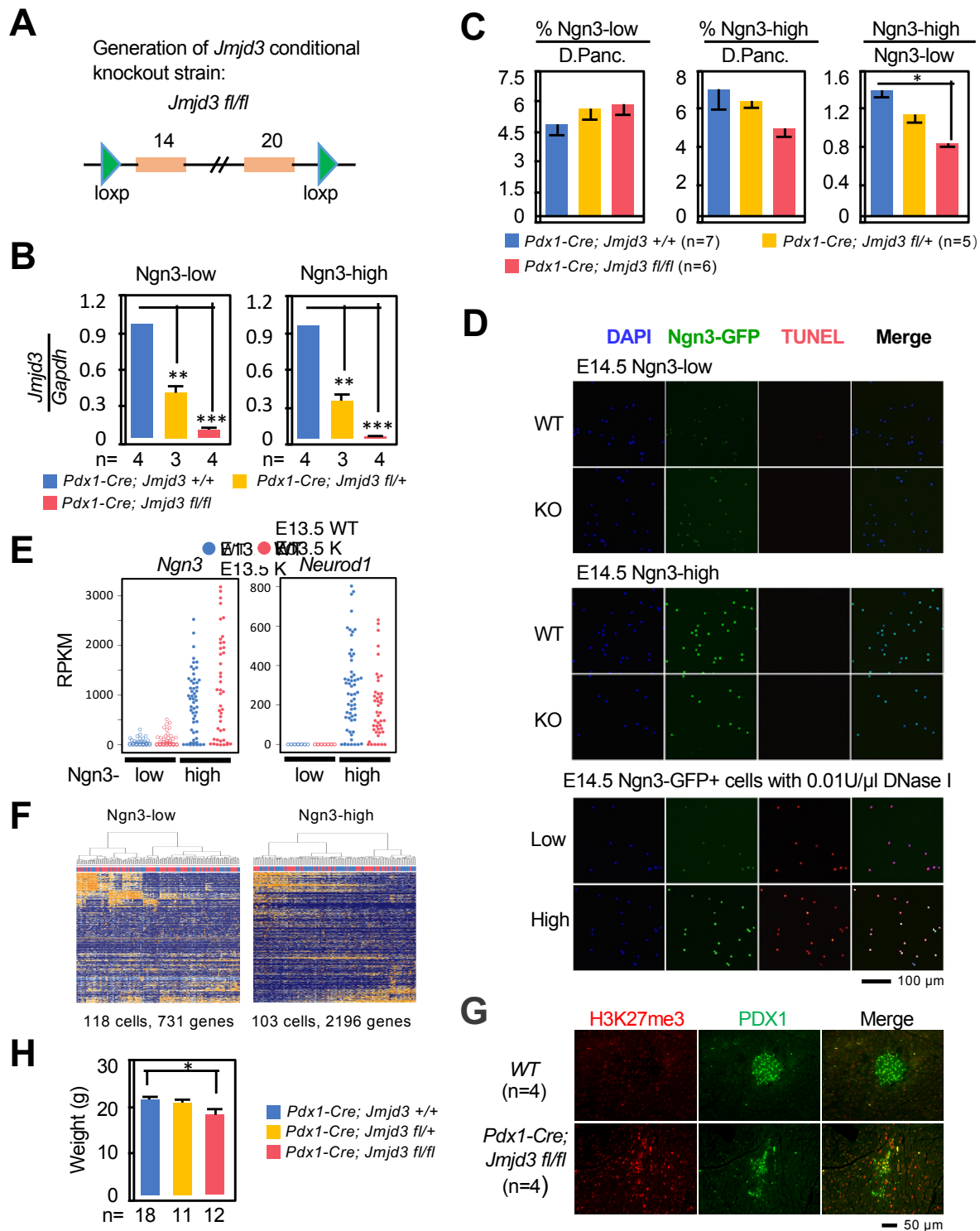**Supplementary Figure 4. NeuroD1 targets promoters in pancreatic endocrine cells.**

(A) Validation of the anti-NeuroD1 antibody by Western blot using adult pancreatic islet extracts.

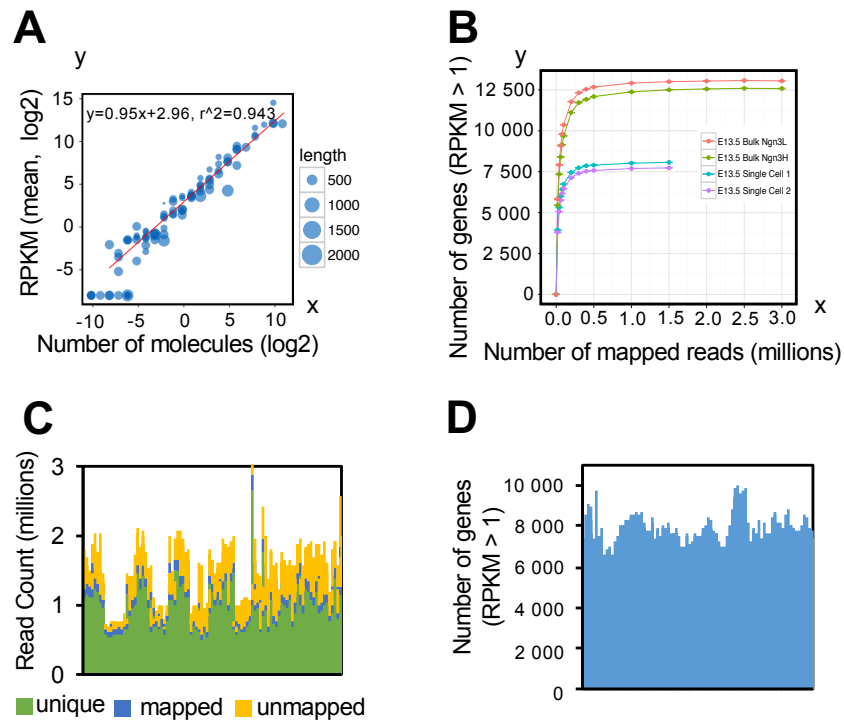(B) Density of NeuroD1 ChIP-seq tags at the promoter regions. (C) Boxplot showing up-regulated mRNA expression levels of genes bound by NeuroD1 at promoters during the transition from Ngn3-low to Ngn3-high. Wilcoxon rank-sum test.

A

Generation of *Jmjd3* conditional
knockout strain:

Jmjd3 fl/fl

B

C

■ *Pdx1-Cre; Jmjd3 +/+* (n=7)   ■ *Pdx1-Cre; Jmjd3 fl/+* (n=5)
■ *Pdx1-Cre; Jmjd3 fl/fl* (n=6)

D

DAPI   Ngn3-GFP   TUNEL   Merge

E14.5 Ngn3-low

E14.5 Ngn3-high

E14.5 Ngn3-GFP+ cells with 0.01U/µl DNase I

100 µm

E

F

Ngn3-low   Ngn3-high

118 cells, 731 genes   103 cells, 2196 genes

G

H3K27me3   PDX1   Merge

WT
(n=4)

Pdx1-Cre;
Jmjd3 fl/fl
(n=4)

50 µm

H

■ *Pdx1-Cre; Jmjd3 +/+*
■ *Pdx1-Cre; Jmjd3 fl/+*
■ *Pdx1-Cre; Jmjd3 fl/fl*

Supplementary Figure 5. Deletion of *Jmjd3* from Ngn3-low and Ngn3-high cells.

(A) Generation of the *Jmjd3* conditional-deletion strain by inserting LoxP sites flanking exons 14-20. (B) RT-qPCR detection of *Jmjd3* expression level in sorted Ngn3-low and Ngn3-high cells from E13.5 WT, *Jmjd3* heterozygous and *Jmjd3* homozygous KO embryos. n: number of replicates. (C) Statistics for the percentages of E14.5 Ngn3-low and Ngn3-high cells among all dorsal pancreatic cells and the ratios of Ngn3-high/low cells in WT, *Jmjd3* heterozygous and homozygous KO pancreata. n: number of embryos. (D) TUNEL staining of sorted Ngn3-low and Ngn3-high cells from E14.5 WT or *Pdx1-Cre;Jmjd3*$^{fl/fl}$ pancreata. DNase I-treated WT Ngn3-GFP^{+} cells were used as positive controls. Scale bar: 100 μm. (E) Expression levels (RPKM) of the key TFs *Ngn3* and *NeuroD1* in Ngn3-low (empty circle) and Ngn3-high (solid circle) cells from E13.5 WT (blue) or *Pdx1-Cre;Jmjd3*$^{fl/fl}$ (red) pancreata. (F) Hierarchical clustering of variously expressed genes in Ngn3-low or Ngn3-high cells from WT and KO pancreata. Each column represents a single cell and each row represents a gene. (G) Histone H3K27me3 and PDX1 immunofluorescence of pancreata from 6-week-old male mice. Scale bar: 50 μm. n: number of male mice. (H) Body weights of *Jmjd3* homozygous KO animals were lower than those of 6-week-old WT and heterozygous KO mice. n: number of male mice. Values represent the mean ± SEM. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (Student's t-test).

A

y

y=0.95x+2.96, r^2=0.943

RPKM (mean, log2)

length
- 500
- 1000
- 1500
- 2000

Number of molecules (log2)

x

B

y

Number of genes (RPKM > 1)

- E13.5 Bulk Ngn3L
- E13.5 Bulk Ngn3H
- E13.5 Single Cell 1
- E13.5 Single Cell 2

Number of mapped reads (millions)

x

C

Read Count (millions)

■ unique ■ mapped ■ unmapped

D

Number of genes
(RPKM > 1)

Supplementary Figure 6. Quality control of single-cell RNA-seq data.

(A) RPKM values from single-cell RNA-seq are correlated with the number of mRNA molecules. Each point represents one of the 92 ERCC RNA spike-ins; the point size reflects the transcript length of the ERCC RNA spike-ins. The x-axis represents the molecular numbers (log2) of the ERCC RNA spike-ins, and the y-axis represents the mean RPKM (log2) values of all cells. Linear regression analysis shows a strong positive correlation between the RPKM value from the single-cell RNA-seq and the molecular number. (B) The sequencing depth required for single-cell RNA-seq was determined by saturation analysis. The x-axis represents the number of reads that were randomly sampled from the mapped reads, and the y-axis represents the number of detected genes (RPKM > 1). Error bars represent the SEM of five independent sub-samplings. The saturation curve shows that approximately 0.5 million mapped reads were sufficient to detect most genes contained within single-cell RNA-seq libraries. (C) Alignment statistics for all single-cell RNA-seq data. Each bar represents a single-cell RNA-seq sample. The mapped read counts for most cells are > 0.5 million. (D) The number of detected genes (RPKM > 1) for all single-cell RNA-seq data. Each bar represents a single-cell RNA-seq sample.

Table S1. Alignment statistics of RNA-seq and ChIP-seq.

Sheet 1: Alignment statistics of bulk-cell RNA-seq.

Sheet 2: Alignment statistics of single-cell RNA-seq. Sheet 3: Alignment statistics of ChIP-seq.

Click here to Download Table S1

Table S2. Differentially expressed genes and enriched KEGG terms of Ngn3-low and Ngn3-high cells.

Sheet 1: Genes highly expressed in Ngn3-low cells.

Sheet 2: Genes highly expressed in Ngn3-high cells.

Sheet 3: Enriched KEGG terms for genes highly expressed in Ngn3-low cells.

Sheet 4: Enriched KEGG terms for genes highly expressed in Ngn3-high cells.

Click here to Download Table S2

Table S3. Genes with removal or decrease of H3K27me3 during the transition from Ngn3-low to Ngn3-high.

Sheet 1: Genes with removal of H3K27me3 at the bivalent promoters.

Sheet 2: Genes with decreased H3K27me3 level at the H3K27me3 alone promoters.

Click here to Download Table S3

Table S4. Enriched KEGG terms for genes associated with newly formed active or primed enhancers in Ngn3-high cells.

Sheet 1: Enriched KEGG terms for genes associated with newly formed active enhancers in Ngn3-high cells.

Sheet 2: Enriched KEGG terms for genes associated with newly formed primed enhancers in Ngn3-high cells.

Click here to Download Table S4

Table S5. GO terms for genes associated with newly formed active or primed enhancers bound by NeuroD1.

Click here to Download Table S5