Methods

Assessment of Sample Purity. Sample purity of the sorted alpha and beta cell populations was calculated as percentage of contamination by the opposite cell type using qPCR expression values for cell-specific markers as follows:

$$\%\ \text{alpha-cell contamination (in beta cell population)} = \text{purity(alpha)} * \frac{\text{Gcg mRNA (b cells)}}{\text{Gcg mRNA (a cells)}}$$

$$\%\ \text{beta-cell contamination (in alpha cell population)} = \text{purity(beta)} * \frac{\text{Ins mRNA (a cells)}}{\text{Ins mRNA (b cells)}}$$

The maximum purity is 100%, and therefore we can compute an upper bound on the contamination without knowing the purity of the other cell population:

$$\%\ \text{alpha-cell contamination (in beta cell population)} \leq \frac{\text{Gcg mRNA (b cells)}}{\text{Gcg mRNA (a cells)}}$$

$$\%\ \text{beta-cell contamination (in alpha cell population)} \leq \frac{\text{Ins mRNA (a cells)}}{\text{Ins mRNA (b cells)}}$$

ChIP-Seq Analysis. Libraries for H3K4me3, H3K27me3, and input were sequenced on an Illumina GA-IIx to 36bp. Reads were aligned to hg18 using ELAND. Up to two mismatches were allowed and only reads with a best alignment to a single location were used for further analysis.

Enrichment of H3K4me3 was performed on individual donors and cell types using GLITR (FDR=5%) (Tuteja et al., 2009) and a pool of human input from various tissues. GLITR was also run on islet cell input samples and any peaks identified in the inputs were filtered from the H3K4me3 enriched regions. Enriched regions in the H3K27me3 data were identified using the program STAR (Lefterova et al., 2010) with a sliding

window of 5,000bp, step size of 1,000bp, and a FDR=0.5%. To ensure fair comparison, H3K4me3 and H3K27me3 profiles from CD4-positive T-cells (Barski et al., 2007) were determined using the same algorithms and settings.

Histone Modification Classification. To classify genes as H3K4me3-only, H3K27me3-only, bivalent, or unmarked we considered pools of all regions identified as enriched for either H3K4me3 or H3K27me3 for the individual cell types. Overlapping regions for the same mark were merged to produce a set of regions enriched for that mark in at least one donor. A gene was considered to have the H3K4me3 mark when a merged enriched region overlapped the region 2kb downstream of the transcription start site, by at least 500bp. A gene was considered to have the H3K27me3 mark when a merged enriched region overlapped the whole gene by at least 500bp. A gene was considered bivalently marked if it was marked by both H3K4me3 and H3K27me3 using these definitions.

Summarizing multiple transcript annotations for each gene was done using these rules:

1. Genes with any transcript marked bivalent are considered bivalent.
2. Genes with a transcript marked H3K4me3, and another transcript marked H3K27me3 are considered "Ambiguous".
3. Genes with at least one transcript marked H3K4me3, and no transcripts marked H3K27me3, are considered “H3K4me3-only”.
4. Genes with at least one transcript marked H3K27me3, and no transcripts marked H3K4me3, are considered “H3K27me3-only”.
5. All remaining genes are considered "Unmarked".

A very small number (alpha cells: 25 genes, beta cells: 30 genes, exocrine cells: 59 genes out of 21,457 genes) of genes were considered "Ambiguous" by these criteria, and were excluded from subsequent analysis.

Quantitative ChIP-Seq Analysis. Heatmap analysis was performed to assess whether the individual samples are correlated and equally contributing to the calls. H3K4me3 heatmaps were generated as follows. For each TSS in RefGene, we counted the number of reads from each sample in the region 2kb downstream of the TSS (same region used for assigning H3K4me3 GLITR peaks to genes), and normalized to reads per million per kb (RPKM). Alternate TSSs for the same gene were combined by taking the maximum normalized signal within each sample. H3K4me3 RPKM signals for each gene were then transformed to the log2 scale, and median normalized across samples. For each sorted cell population, normalized signals were plotted in a heatmap, with columns corresponding to individual samples, and rows corresponding to genes. The row order was determined by first grouping the genes called as H3K4me3 for the corresponding cell population (GLITR analysis, indicated by the solid blue bar), and then ordering genes by the average signal across all individual samples. H3K27me3 heatmaps were generated by a similar method, except that reads were counted and normalized across the entire gene body, and the H3K27me3 gene calls used for row grouping were based on STAR calls (indicated by the solid red bar). P-values were determined for the correlation between each pair of samples using the cor.test function in R. Normalized RPKM values were then averaged together for samples of the same cell type and used for comparisons of histone mark enrichments between cell-specific expression groups (defined below).

RNA-Seq. After total RNA extraction (see above), RNA-Seq libraries were prepared from sorted alpha, beta, and exocrine cells. The detailed protocol can be found on http://ngsc.med.upenn.edu (Lab protocol, Library preparation for RNA-Seq from total RNA). Libraries were single-end sequenced to 100bp on an Illumina hiSeq2000. Reads from ribosomal RNA and genomic repeats were identified by aligning the 5' 50bp of each read to ribosomal sequences and the human repeats in RepBase (version 14.10) using Bowtie (Langmead et al., 2010) and allowing for up to three mismatches. The remaining reads were processed with RUM (Grant et al., 2011) and aligned to the set of known transcripts included in RefSeq, UCSC known genes, and ENSEMBL transcripts, and the human genome (hg18, NCBI build 36.1). Transcript-, exon-, and intron-level quantification was done using only the uniquely aligning reads. This process was repeated separately for each cell type and donor.

Gene Expression Analysis. To analyze global gene expression profiles from each sample, we extracted the RNA-seq read counts aligning to exons of mRNA transcripts in RefSeq, and normalized these values to total uniquely aligning reads and transcript length (RPKM). We then summarized these data for individual genes by selecting a "representative transcript" with the highest RPKM value for each gene symbol (total=18,822 genes). Gene-level data was then quantile-normalized across all samples to remove bias from the variable expression levels of highly expressed genes, e.g. Insulin, Glucagon, and highly expressed ncRNAs, which dampen the relative signal of all other genes. Finally, quantile-normalized gene expression values were averaged together for samples of the same cell type to produce a single average gene profile for each cell type.

These values were used for all subsequent analyses, including expression comparisons for individual genes of interest.

For the heatmap of cell-type specific expression, we removed all genes below 0.5 RPKM normalized average expression in all three cell types. The remaining 14,003 genes were each normalized to percentage of their maximum expression across the three cell types, to focus our analysis on the cell-type specificity of each gene, rather than the absolute expression level. These values were clustered using the R package 'mclust' v4.0 (Yeung, et al. 2001). K=20 was chosen as the number of clusters, as Bayesian Information Content did not improve substantially with additional clusters. Clusters were grouped together based on their overall cell-type specificity.

For Principal Component Analysis (PCA) of the cell types, the quantile-normalized average values (including those below 0.5 RPKM) were used to build three principle components (linear transformations of the gene expression profile in each sample), using the 'prcomp' function in R. The first two principle components captured 99% of the variance across the cell types, and were used to map the average and individual replicate profiles onto a 2D plot.

For the comparison between marked and unmarked genes in the same cell type, we again used the average quantile-normalized values for the representative transcript of each gene, and further transformed these values to log2 scale to visualize expression for both low and high-expression genes. For each group of genes based on histone marks in a single cell type, a one-sample Wilcoxon test (R 'wilcox.test' function) was used to assess whether the distribution of expression values was significantly shifted above or below 0 on the log2 scale (corresponding to RPKM=1 on the non-log scale). A shift above this

value indicates a preference for highly expressed genes, while a shift below this value indicates a preference for weakly or non-expressed genes.

Identification of Cell Type-Specific Novel Transcripts. We first selected candidate regions for novel transcription based on the genomic coverage data output from RUM (number of RNA-seq reads overlapping each base in the genome). Specifically, we selected all regions in each alpha or beta cell RNA-seq sample where the read coverage was >1 RPM continuously for at least 200bp. We then merged these regions together into a master set of candidates regions, and filtered out all regions that overlapped any known exon in the feature quantification table used by RUM (contains RefSeq, UCSC Known Genes, and Ensembl transcripts) to rule out all known transcribed regions. Note that we still allow for any candidate region that falls entirely in a known intronic region, as some lncRNAs have been observed in these regions. We also removed any candidate region overlapping a repeat region from the UCSC Repeat Masker track as these regions are more likely to be mapping or amplification artifacts. We then compared the remaining candidate regions to the H3K4me3 peak calls from both alpha and beta cells, and limited our subsequent analysis to only those regions within 5kb of an H3K4me3 peak, as this histone mark provides additional evidence of active transcription.

The candidate selection process above resulted in 317 candidate regions with evidence for active transcription, based on both the RNA-seq and H3K4me3 data, which do not overlap known transcripts or repeat regions. We then quantified the expression level of each of these candidate regions by counting the number of uniquely aligning reads overlapping these regions, normalized to transcript length and total number of reads

in each samples (RPKM). We computed the mean normalized expression value for each novel transcript in the three alpha cell RNA-seq samples and the three beta cell RNA-seq samples, and computed a fold-change between these two cell-types. We defined alpha-specific transcripts to be those with >2x higher mean expression in alpha cells, and beta-specific transcripts to be those with >2x higher mean expression in beta cells.

Histone Modification Analysis of Functional Gene Categories within alpha- and beta-cell signature genes. Lists of strongly cell-type specific genes (Figure 2B, indicated by darker portion of bars) were analyzed for their likely gene function using DAVID. The two functional gene categories 'ion transport' and 'regulation of transcription, DNA-dependent' comprised similar numbers of genes (ion transport: 29 alpha-specific and 34 beta-specific genes; regulation of transcription, DNA-dependent: 29 alpha-specific and 31 beta-specific genes). These genes were further analyzed for their histone modification status in alpha-, beta-, and exocrine cells and the percentage of histone modification (bivalent, monovalent H3K4me3, monovalent H3K27me3, none of the above) within each gene group was calculated.

Chromatin Preparation from Small Cell Populations

Buffers

2.22% formaldehyde/PBS

1ml: 60µl 37% formaldehyde + 940µl 1XPBS

2.5M glycine

190mg in 1ml H_2O

ChIP Whole Cell Lysis Buffer

(10mM Tris-HCl, pH 8.0, 10mM NaCl, 3mM $MgCl_2$, 1%NP-40, 1% SDS, 0.5% DOC)

1M Tris-HCl (pH 8.1)	10µl
5M NaCl	2µl
1M $MgCl_2$	3µl
10% NP-40	100µl
10% SDS	100µl
10% DOC	50µl
50X Prot.Inhib.	20µl
100x Phos'tase Inhib.	10µl
H_2O	705µl
	1ml

Cross-Linking

1. Spin down cells at maximum speed for 10 sec at Room Temperature (RT), aspirate supernatant.
2. Re-suspend cell pellet in 500µl 1XPBS.
3. Add 500µl 2.22% formaldehyde/PBS.
4. Rotate 10min at RT.
5. Stop cross-linking by adding 59µl 2.5M glycine (final: 0.14M).
6. Rotate 5min at RT.
7. Spin down cells at maximum speed for 10 sec at RT, then aspirate supernatant.
8. Add 1ml PBS to wash, spin down cell, aspirate supernatant.
9. Bring up pellet in 100-200µl cold ChIP Whole Cell Lysis Buffer.

Sonication and Input Preparation

10. Sonicate using BioRuptor (conditions may vary!):
    High Setting, 5 x 5 min, cycling 30 sec on – 30 sec off
11. Spin at maximum RPM at 4C for 15 minutes.
12. Take 10µl as Input, add 90µl PBS and 3.5µl NaCl, incubate at 65C over night. Store remainder at -80C.
13. Next day: add 4µl 1M Tris-HCl (pH 7.5), 2µl 500mM EDTA, and 1µl 10mg/ml Proteinase K to the Input sample.
14. Incubate 1 hour at 45C.
15. Column purify using the QIAquick PCR Purification Kit (QIAGEN 28104).
16. Elute in 50µl EB.
17. Run 1µl on the Bioanalyzer to check for appropriate shearing. Agilent 2100 BioAnalyzer, DNA-1000 Kit (Agilent 5067-1504)

Chromatin Immunoprecipitation Assay

Day 1: Antibody Incubation and Agarose Blocking

Incubate Chromatin with Antibody:

1. Thaw chromatin and place on ice

2. Add 2-5µg of chromatin to 1 mL of **ChIP Dilution Buffer + 20 µL 50X Protease Inhibitor**

ChIP Dilution Buffer (16.7 mM Tris-HCl pH 8.1, 167 mM NaCl, 0.01%SDS, 1.1% Triton-X 100)

1M Tris-HCl, pH 8.1	835 µL
5M NaCl	1.67 mL
10% SDS	50 µL
10% Triton-X 100	5.5 mL
H2O	41.9 mL
	50 mL

3. Add antibody (~2 µg) and rotate in cold room overnight

Block Agarose:

4. Wash protein-A or -G agarose with 1 mL cold **ChIP Dilution Buffer**
   - add 1mL **ChIP Dilution Buffer** to appropriate amount of agarose slurry (40 uL per sample)
   - mix by inverting then pellet at minimum speed for 30 sec
   - repeat 2 additional times
   - resuspend agarose in appropriate amount of cold **ChIP Dilution Buffer + BSA**

For 1 sample:

protein-A/G agarose slurry	20 uL (from 40 µL of agarose slurry)
10mg/ml BSA	10 µL
50X Protease Inhibitor	2 µL
ChIP Dilution Buffer	68 µL
	100 µL

5. Rotate in cold room overnight

Day 2: Immunoprecipitation and Elution

Immunoprecipitate:

6. Add 100 µL of blocked protein-A or -G agarose to each chromatin sample
   - rotate in cold room for 1 hour

7. Spin down agarose 30s at minimum speed, aspirate supernatant.

Wash:

8. Add 1 mL of appropriate room temp. **Wash Buffer** to agarose pellet (see below)
   - rotate at room temp. for 5 min
   - spin down agarose 30s at minimum speed, aspirate supernatant.

9. Perform wash 1X:

1. TSE I: 20 mM Tris-HCl (pH 8.1), 150 mM NaCl, 2mM EDTA, 0.1% SDS, 1% Triton X-100

1M Tris-HC, pH 8.1	1 mL
5M NaCl	1.5 mL
0.5M EDTA	200 µL
10% SDS	500 µL
10% Triton X-100	5 mL
H2O	41.8 mL

2. TSE II: 20 mM Tris-HCl (pH 8.1), 500 mM NaCl, 2mM EDTA, 0.1% SDS, 1% Triton X-100

1M Tris-HCL, pH 8.1	1 mL
5M NaCl	5 mL
0.5M EDTA	200 µL
10% SDS	500 µL
10% Triton X-100	5 mL
H2O	38.3 mL

3. ChIP Buffer III: 10 mM Tris-HCl (pH 8.1), 0.25M LiCl, 1mM EDTA, 1% NP-40, 1% deoxycholate

1M Tris-HCl, pH 8.1	500 µL
5M LiCl	2.5 ml
0.5M EDTA	100 µL
10% NP-40	5 mL
10% deoxycholic acid	5 mL
H2O	36.9 mL

4. TE: 10 mM Tris-HCl (pH 8.1), 1mM EDTA

1M Tris-HCl, pH 8.1	500 µL
0.5M EDTA	100 µL
H2O	49.4 mL

Elute/Reverse Crosslinks:

10. Add 100 µL freshly made **Elution Buffer** to final pellet and rotate 15 min at room temp

    Elution Buffer: (1% SDS, 0.1 M $NaHCO_3$)

10% SDS	100 µL	
1M $NaHCO_3$	100 µL	(84 mg in 1 mL H_20 made fresh)
H_2O	800 µL	

11. Pellet agarose as before, transfer supernatant to new tube

12. Add an additional 100 µL **Elution Buffer** to pellet, repeat elution, and combine eluates

13. Spin tubes one more time and transfer to new tube

14. Add 8 µL 5M NaCl per 200 µL eluate (192 mM NaCl) and incubate at 65°C overnight

Day 3: Protein Digestion and DNA Purification

15. Add 8 µL 1 M Tris-HCl, pH 7.5, 4 µL 0.5 M EDTA, and 1 µL 10 mg/ml proteinase K to each tube - incubate 1 hour at 45°C

16. Purify un-crosslinked chromatin using Qiagen PCR purification kit. Elute in 50µL EB.

ChIP-Seq Library Preparation
Option 1:

Preparing Samples for ChIP Sequencing of DNA

Contents

Material

Reagents

- 10X T4 DNA Ligase Buffer with 10 mM ATP (NEB B0202S)
- T4 DNA Polymerase (NEB M0203S)
- Klenow DNA Polymerase (NEB M0210S)
- T4 Polynucleotide Kinase (PNK) (NEB M0201S)
- 10 mM dNTPs Mix
- 1 mM dATP
- Klenow Fragment (3' to 5' exo minus) (NEB M0212S)
- Quick DNA Ligation Kit (NEB M2200S)
- Oligo Only Kit (Illumina FC-102-1003)
    - Adapter Oligo Mix (final molarity: 2uM)
    - PCR Primer 1.1 (final molarity: 25uM)
    - PCR Primer 2.1 (final molarity: 25uM)
- Phusion HF DNA Polymerase (NEB M0530S/L)

Gel Purification Components

- 6X Loading Dye (NEB B7021S)
- Low Molecular Weight DNA Ladder (NEB N3233L)
- Certified Low Range Ultra Agarose (Bio Rad 161-3106)
- SYBR Gold Nucleic Acid Gel Stain (Invitrogen S-11494)
- Roche Restriction Enzyme Buffer H
- Gel x-tracta (USA Scientific 5454-2500)
- QIAquick Gel Extraction Kit (QIAGEN 28704)
- TaKara DNA-OFF (TaKara 9036)

DNA Purification Kits

- QIAquick PCR Purification Kit (QIAGEN 28104)
- MinElute PCR Purification Kit (QIAGEN 28004)

Equipment

- Heat Block/Dry Bath/Water Bath set at 20^{o}C and 37^{o}C
- PCR Machine
- Dark Reader transilluminator (ISC BioExpress U-2235-16)
- Agilent 2100 BioAnalyzer
    - DNA-1000 Kit (Agilent 5067-1504)

Perform End Repair

This step converts the overhangs into phosphorylated blunt ends.

Cleanup Method

- QIAquick PCR Purification Kit

Sample Prep Kit Components

- T4 DNA ligase buffer with 10mM ATP
- 10mM dNTPs mix
- T4 DNA polymerase
- Klenow DNA polymerase
- T4 Polynucleotide Kinase (PNK)

Procedure

1. Prepare the following reaction mix:

	x1
Water	0.8 µl
T4 DNA ligase buffer	5 µl
10mM dNTP mix	2 µl
T4 DNA polymerase	1 µl
T4 PNK	1 µl
Klenow DNA polymerase	0.2 µl
	10 µl

2. Add 10 µl of reaction mix to 40 µl DNA
   - 40 µl of ChIP DNA
   - 10 ng of Input DNA in 40 µl of water

3. Incubate for 30 minutes at 20ºC.

4. Purify using QIAquick PCR Purification Kit:
   - Add 250 µl Buffer PB and mix by pipetting
   - Apply to column and spin at max speed 30 sec
   - Pour flowthrough back on to column and spin at max speed 30 sec
   - Discard flowthrough
   - Add 750 µl Buffer PE and spin at max speed 45 sec
   - Discard flowthrough
   - Spin empty column at max speed 1 min
   - Transfer column to new 1.5 ml tube
   - Aspirate around clear ring inside column
   - Allow to air dry 1 min
   - Apply **36 µl Buffer EB** and let stand 2 min (loss of ~2 µl volume during elution)
   - Spin at max speed 1 min and discard column

Add ‘A’ Base to the 3' End of the DNA Fragments

This protocol adds an ‘A’ base to the 3' end of the blunt phosphorylated DNA fragments and prepares the DNA fragments for ligation to the adapters.

Cleanup Method

- MinElute PCR Purification Kit

Sample Prep Kit Components

- Klenow buffer (NEB buffer #2)
- 1mM dATP
- Klenow fragment (3' to 5' exo minus)

Procedure

1. Prepare the following reaction mix:

	x1
▪ Klenow buffer (NEB buffer #2)	5 µl
▪ 1mM dATP	10 µl
▪ Klenow exo (3’ to 5’ exo minus)	1 µl
	16 µl

2. Add 16 µl of reaction mix to 34 µl DNA sample from previous step.

3. Incubate for 30 minutes at 37ºC.

4. Purify using MinElute PCR Purification Kit:
   - Add 250 µl Buffer PB and mix by pipetting
   - Apply to column and spin at max speed 30 sec
   - Pour flowthrough back on to column and spin at max speed 30 sec
   - Discard flowthrough
   - Add 750 µl Buffer PE and spin at max speed 45 sec
   - Discard flowthrough
   - Spin empty column at max speed 1 min
   - Transfer column to new 1.5 ml tube
   - Aspirate around purple ring inside column
   - Allow to air dry 1 min
   - Apply **15 µl Buffer EB** and let stand 2 min (loss of ~2 µl volume during elution)
   - Spin at max speed 1 min and discard column

Ligate Adapters to DNA Fragments

This protocol ligates adapters to the ends of the DNA fragments, allowing them to be hybridized to a flow cell surface.

Cleanup Method

- QIAquick PCR Purification Kit

Sample Prep Kit Components

- 2X Quick DNA ligase buffer
- Adapter oligo mix
- Quick T4 DNA ligase

Procedure

1. Dilute Illumina's Adapter (final Adapter mix molarity should be 2µM).

2. Prepare the following reaction mix:

	x1
2X Quick DNA ligase Buffer	15 µl
Diluted Adapter oligo mix	1 µl
Quick DNA ligase	1 µl
	17 µl

3. Add 17 µl of reaction mix to 13 µl DNA sample from previous step.

4. Incubate for 15 minutes at 20ºC.

5. Purify using QIAquick PCR Purification Kit:
   - Add 150 µl Buffer PB and mix by pipetting
   - Apply to column and spin at max speed 30 sec
   - Pour flowthrough back on to column and spin at max speed 30 sec
   - Discard flowthrough
   - Add 750 µl Buffer PE and spin at max speed 45 sec
   - Discard flowthrough
   - Spin empty column at max speed 1 min
   - Transfer column to new 1.5 ml tube
   - Aspirate around clear ring inside column
   - Allow to air dry 1 min
   - Apply **12 µl Buffer EB** and let stand 2 min (loss of ~2 µl volume during elution)
   - Spin at max speed 1 min and discard column

Size Selection: 2% Agarose Gel Size Selection

Components

- Certified Low Range Ultra Agarose
- 50X TAE buffer
- Low Molecular Weight DNA ladder
- 6X loading dye
- 10X Roche Buffer H
- SYBR Gold
- QIAquick Gel Extraction Kit
- Gel x-tracta or clean razor
- TaKara DNA-OFF

Procedure

1. Wash the gel tray, combs, cutting glass, etc. with DNA-OFF.

2. Prepare 500ml of 1X TAE Buffer.

3. Prepare 50 ml of a 2% agarose gel with 1X TAE buffer and Ultra Low Agarose. *Do not add EtBr or any DNA stain.*

4. Dilute SYBR Gold 1:1000 in water.

5. Prepare the DNA ladder mix:

	x1
Diluted SYBR Gold	10 µl
6X loading dye	3.75 µl
10X Roche Buffer H	1.25 µl
Low Mol. Weight ladder	2.5 µl
	17.5 µl

6. Prepare the loading dye mix:

	x1
6X loading dye	3 µl
10X Roche Buffer H	2 µl
	5 µl

7. Add 5 µl of loading dye mix to each DNA sample from previous step.

8. Load 17 µl of DNA ladder mix to outside well(s).

9. Load entire sample (~15 µl) on gel.

10. Run gel at 100 V for 45 minutes.

11. View the gel on a **Dark Reader transilluminator**. *Do not expose gel to UV light.*

12. Excise a region of gel just above the 200bp band in the **250 ±25** bp range using a Gel x-tracta or clean razor. *Photograph the gel after excisions for later reference.*

13. Purify DNA from gel slices using the QIAquick Gel Extraction Kit.
    - Weigh empty 1.5 ml tube to zero scale
    - Weigh each tube containing gel piece
    - To each tube add **6 volumes** of Buffer QG per volume of gel
        - ie. add 600 µl QG to a 100 mg gel piece
    - Incubate at Room Temp for 10 minutes, vortexing regularly until gel piece is dissolved.
    - To each tube add **2 volumes** of isopropanol per volume of gel and mix by pipetting.
        - ie. add 200 µl isopropanol to a 100 mg gel piece
    - Apply to column and spin at max speed 30 sec
    - Pour flowthrough back on to column and spin at max speed 30 sec
    - Discard flowthrough
    - Repeat until entire volume has been passed through column
    - Add 500 µl Buffer QG and spin at max speed 30 sec
    - Discard flowthrough
    - Add 750 µl Buffer PE and spin at max speed 45 sec
    - Discard flowthrough
    - Spin empty column at max speed 1 min
    - Transfer column to new 1.5 ml tube
    - Aspirate around purple ring inside column
    - Allow to air dry 1 min
    - Apply **38 µl Buffer EB** and let stand 1 min (loss of ~2 µl volume during elution)
    - Spin at max speed 1 min and discard column

PCR-Amplification of the Adapter-Modified DNA Fragments

Cleanup Method

- Qiagen MinElute PCR Purification Kit

Sample Prep Kit Components

- Phusion polymerase
- 5x Phusion HF buffer
- 10 mM dNTP mix
- PCR primer 1.1
- PCR primer 2.1

Procedure

1. Prepare the following PCR mix:

	x1
5x Phusion HF buffer	10 µl
dNTP mix	1.5 µl
PCR primer 1.1	1 µl
PCR primer 2.1	1 µl
Phusion polymerase	0.5 µl
	14 µl

2. Add 14 µl of PCR mix to 36 µl of DNA sample.

3. Amplify using the following PCR protocol:
   a. 30 seconds at 98°C
   b. 18 cycles of:
      - 10 seconds at 98°C
      - 30 seconds at 65°C
      - 30 seconds at 72°C
   c. 5 minutes at 72°C
   d. Hold at 4°C

4. Purify using MinElute PCR Purification Kit:
   - Add 250 µl Buffer PB and mix by pipetting
   - Apply to column and spin at max speed 30 sec
   - Pour flowthrough back on to column and spin at max speed 30 sec
   - Discard flowthrough
   - Add 750 µl Buffer PE and allow to stand 1 min
   - Spin at max speed 45 sec and discard flowthrough
   - Spin empty column at max speed 1 min
   - Transfer column to new 1.5 ml tube
   - Aspirate around purple ring inside column
   - Allow to air dry 1 min
   - Apply **17 µl Buffer EB** and let stand 2 min
   - Spin at max speed 1 min and discard column

Validate the Library

Illumina recommends using Agilent Technologies' 2100 BioAnalyzer to check the size, purity, and concentration of the library (Agilent's DNA-1000 kit).

ChIP-Seq Library Preparation
Option 2:

Preparing Samples for Multiplexed ChIP-Sequencing of DNA.

Contents

Summary of the changes/improvements:

1. Individual reagents are now available as an NEB kit.
2. Multiplexing allows for sequencing of multiple samples on one sequencing lane. Each sample on one lane needs its’ individual Index PCR primer.
3. Multiplexing adapters and primers can now be ordered and prepared separately.
4. Size Selection can be performed using a 2% Agarose gel (see protocol 1), or by using a new device called Pippin Prep.
5. A new cleanup method (Agencourt Ampure Beads) after the adapter ligation and after the PCR reaction leads to ‘cleaner’ libraries. Alternatively, the Qiagen PCR cleanup kits can still be used.

Material

Kits

ChIP Seq DNA Sample Prep Kit (Illumina IP-102-1001)

OR

NEBNext ChIP-Seq DNA Sample Prep Reagent Set 1 (NEB E6200S/L)

Adapter and Primers have to be made separately (see below)

DNA Purification Kits

- QIAquick PCR Purification Kit (QIAGEN 28104)
- MinElute PCR Purification Kit (QIAGEN 28004)
- Agencourt AMPure XP Beads (Beckman Coulter A63880)

Equipment

- Heat Block/Dry Bath/Water Bath set at 20^{o}C and 37^{o}C
- PCR Machine
- Magnetic Stand (Invitrogen's DynaMag2, 123.21D)
- Pippin Prep (Sage Science or Life Technologies)
    - With Pippin Prep 2% EF Gels (Life Technologies 4472171)
- Agilent 2100 BioAnalyzer
    - DNA-1000 Kit (Agilent 5067-1504)

Preparation of Multiplex Adapters and PCR Primers (not included in NEB kit)

For detailed information about sequences and preparation of Multiplex Adapters and PCR Primers, please contact the authors.

Perform End Repair

This step converts the overhangs into phosphorylated blunt ends.

Cleanup Method

- QIAquick PCR Purification Kit

Sample Prep Kit Components

- T4 DNA ligase buffer with 10mM ATP
- 10mM dNTPs mix
- T4 DNA polymerase
- Klenow DNA polymerase
- T4 Polynucleotide Kinase (PNK)

Procedure

1. Prepare the following reaction mix:

	x1
Water	0.8 µl
T4 DNA ligase buffer	5 µl
10mM dNTP mix	2 µl
T4 DNA polymerase	1 µl
T4 PNK	1 µl
Klenow DNA polymerase	0.2 µl
	10 µl

2. Add 10 µl of reaction mix to 40 µl DNA
   - 40 µl of ChIP DNA
   - 10 ng of Input DNA in 40 µl of water

3. Incubate for 30 minutes at 20ºC.

4. Purify using QIAquick PCR Purification Kit:
   - Add 250 µl Buffer PB and mix by pipetting
   - Apply to column and spin at max speed 30 sec
   - Pour flowthrough back on to column and spin at max speed 30 sec
   - Discard flowthrough
   - Add 750 µl Buffer PE and spin at max speed 45 sec
   - Discard flowthrough
   - Spin empty column at max speed 1 min
   - Transfer column to new 1.5 ml tube
   - Aspirate around clear ring inside column
   - Allow to air dry 1 min
   - Apply **36 µl Buffer EB** and let stand 2 min (loss of ~2 µl volume during elution)
   - Spin at max speed 1 min and discard column

Add ‘A’ Base to the 3' End of the DNA Fragments

This protocol adds an ‘A’ base to the 3' end of the blunt phosphorylated DNA fragments and prepares the DNA fragments for ligation to the adapters.

Cleanup Method

- MinElute PCR Purification Kit

Sample Prep Kit Components

- Klenow buffer (NEB buffer #2)
- 1mM dATP
- Klenow fragment (3' to 5' exo minus)

Procedure

1. Prepare the following reaction mix:

	x1
▪ Klenow buffer (NEB buffer #2)	5 µl
▪ 1mM dATP	10 µl
▪ Klenow exo (3’ to 5’ exo minus)	1 µl
	16 µl

2. Add 16 µl of reaction mix to 34 µl DNA sample from previous step.

3. Incubate for 30 minutes at 37ºC.

4. Purify using MinElute PCR Purification Kit:
   - Add 250 µl Buffer PB and mix by pipetting
   - Apply to column and spin at max speed 30 sec
   - Pour flowthrough back on to column and spin at max speed 30 sec
   - Discard flowthrough
   - Add 750 µl Buffer PE and spin at max speed 45 sec
   - Discard flowthrough
   - Spin empty column at max speed 1 min
   - Transfer column to new 1.5 ml tube
   - Aspirate around purple ring inside column
   - Allow to air dry 1 min
   - Apply **15 µl Buffer EB** and let stand 2 min (loss of ~2 µl volume during elution)
   - Spin at max speed 1 min and discard column

Ligate Adapters to DNA Fragments

This protocol ligates adapters to the ends of the DNA fragments, allowing them to be hybridized to a flow cell surface.

Cleanup Method (Option 2)

- Agencourt Ampure XP Beads
- Freshly prepared 70-80% Ethanol
- Alternative: QIAquick PCR Purification Kit

Sample Prep Kit Components

- 2X Quick DNA ligase buffer
- Multiplexing Adapter Mix
- Quick T4 DNA ligase

Procedure

1. Remove the Ampure XP Beads from 4^{o}C and let sit at Room Temp for ~30min.

2. Dilute Multiplexing Adapter 1:20 (final molarity: 2uM).

3. Prepare the following reaction mix:

	x1
2X Quick DNA ligase Buffer	15 µl
Multiplexing Adapter Mix	1 µl
Quick DNA ligase	1 µl
	17 µl

4. Add 17 µl of reaction mix to 13 µl DNA sample from previous step.

5. Incubate for 15 minutes at 20ºC.

6. Purify using Agencourt Ampure XP Beads: *Make sure beads came to Room Temp before proceeding*
   - Vortex/mix the beads until they are well dispersed.
   - Add 50 µl of mixed Ampure XP beads and mix by pipetting
   - Incubate the tubes at Room Temp for 10-15 minutes
   - Place tubes on magnetic stand and let sit for 2 minutes or until the liquid appears clear
   - Remove and discard 75 µl of the supernatant from each tube
   - Add 200 µl 70-80% Ethanol and let it incubate on the stand for 1 min and then discard Ethanol
   - Repeat Ethanol wash one more time
   - Let stand at Room Temp for 5-10 minutes to dry and remove all Ethanol
   - Remove from magnetic stand and resuspend the dried beads in **31 µl EB**
   - Mix by pipetting 10x and Incubate at Room Temp for 5 minutes.
   - Place tubes on magnetic stand and let sit for 2 minutes or until the liquid appears clear
   - Transfer **30 µl** of the supernatant to a new tube for **Pippin Prep** procedure

Size Selection Option 2: Pippin Prep Size Selection

Using Sage Science's Pippin Prep for size selection can help to increase the yield. For more detailed information, please check the manual.

Components

- 2% EF (Ethidium Free) Agarose Gel Cassette (Sage Science CEF-2010)
- Loading Solution (stored at 4°C)
- Marker (ladder): Marker E (stored at 4°C)
- Electrophoresis Buffer

Procedure

1. Remove marker and loading solution from refrigerator and let stand at Room Temp for 20 minutes.

2. Add 10 µl of loading solution to 30 µl DNA sample from the previous step. Mix.

3. Remove gel cassette from foil pack and **Visually Inspect the Cassette**:
   Inspect the gel columns, look for bubbles on the gel column in the optical detection region.

4. Place gel cassette into the optical nest and remove the white tabbed adhesive strips from the cassette.

5. Remove the buffer from all the elution wells and replace with 40 µl of fresh electrophoresis buffer.

6. **Perform the continuity TEST**. Check manual.

7. Replace the white tabbed adhesive strip over the elution wells.

8. Remove 40 µl of buffer from the sample wells.

9. **Set Up Size Selection Protocol:**
   - Open up the **"Protocol Editor"** tab
   - Select Cassette Type:
     - For 2% EF Agarose Gel Cassette select **2%EF Marker E V2**
   - Select Mode for each Lane:
     - **Ref** for the marker
     - **Range** for the samples
     - **Off** for the lanes not being run
   - Enter Size Range:
     - For Multiplexed Adapters; **BP Start:** 170, **BP End:** 260, **BP Pause:** 210
   - Enter Sample ID Template Names
   - Make sure the BP Range Flag is Broad.
   - Click "**Save As**" at the bottom right of the screen to save your protocol.
   - Click "**Main**" tab to get back to the main screen.

10. **Load 40 µl of sample into their respective sample well.** Press the "**Start**" button.

11. After approximately 50-55min the run should be paused and the **"RESUME"** button will be available.

12. Remove the adhesive over the elution wells and collect the entire volume (~40 µl) into new tube.

13. Add 40 µl of fresh electrophoresis buffer to the elution well and Replace Adhesive tape.

14. Click the **"RESUME"** button.

PCR-Amplification of the Adapter-Modified DNA Fragments

Cleanup Method (Option 2)

- Agencourt Ampure XP Beads
- Freshly prepared 70-80% Ethanol
- Alternative: Qiagen MinElute PCR Purification Kit

Sample Prep Kit Components

- PCR Master Mix
- Multiplex PCR Primer 1
- Multiplex PCR Primer 2 with Index

Procedure

1. Add the following to each DNA sample (40-50ul):

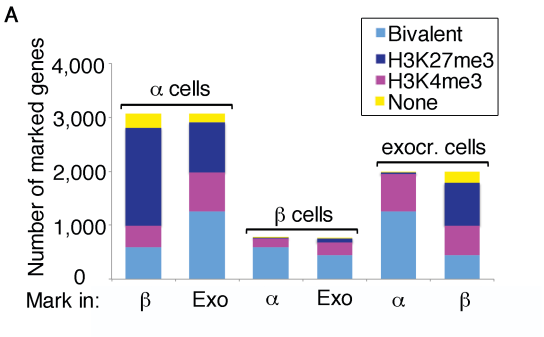
Phusion HF PCR Master Mix:	40 µl
Multiplex PCR Primer 1:	2 µl
Multiplex PCR Primer 2 (different indexes):	2 µl

3. Amplify using the following PCR protocol:
   a. 30 seconds at 98°C
   b. 18 cycles of:
      - 10 seconds at 98°C
      - 30 seconds at 65°C
      - 30 seconds at 72°C
   c. 5 minutes at 72°C
   d. Hold at 4°C

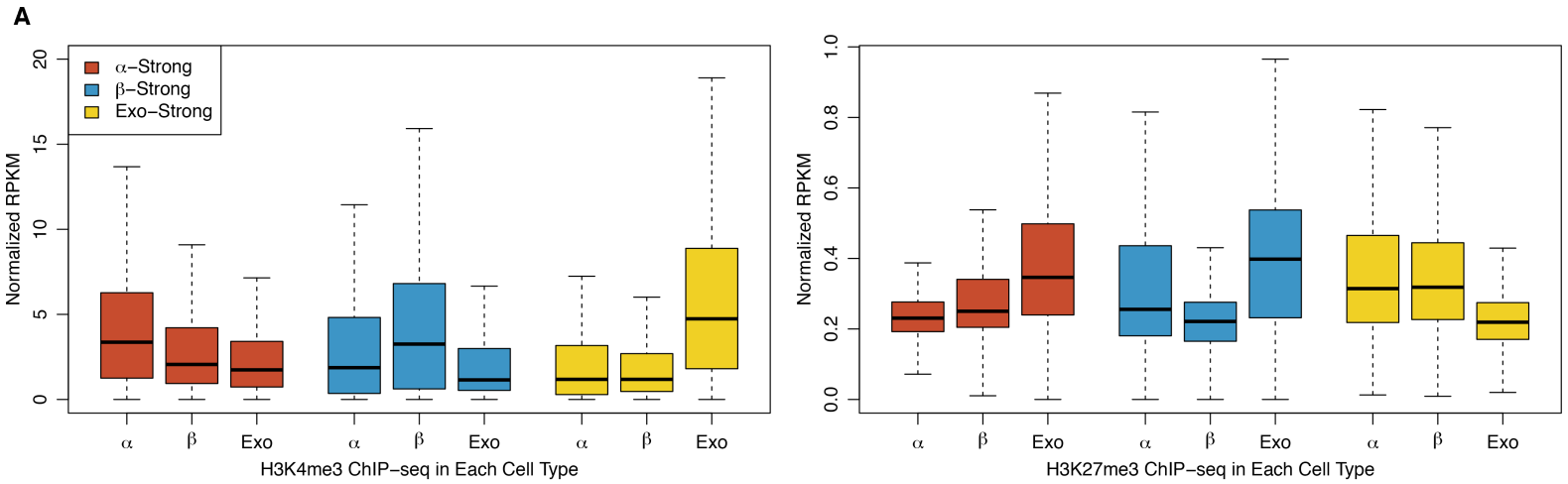
4. Purify using Agencourt Ampure XP Beads: *Make sure beads came to Room Temp before proceeding*
   - Vortex/mix the beads until they are well dispersed
   - Add 90 µl of Ampure XP beads for every 50 µl PCR volume and mix by pipetting
   - Incubate the tubes at Room Temp for 10-15 minutes
   - Place tubes on magnetic stand and let sit for 2 minutes or until the liquid appears clear
   - Remove and discard the majority of the supernatant from each tube
   - Add 200 µl 70-80% Ethanol and let it incubate on the stand for 1 min and then Discard Ethanol
   - Repeat Ethanol wash one more time
   - Let stand at Room Temp for 5-10 minutes to dry and remove all Ethanol
   - Remove from magnetic stand and resuspend the dried beads in **16 µl EB**
   - Mix by pipetting 10x and incubate at Room Temp for 5 minutes
   - Place tubes on magnetic stand and let sit for 2 minutes or until the liquid appears clear
   - Transfer **15 µl** of the supernatant to a new tube

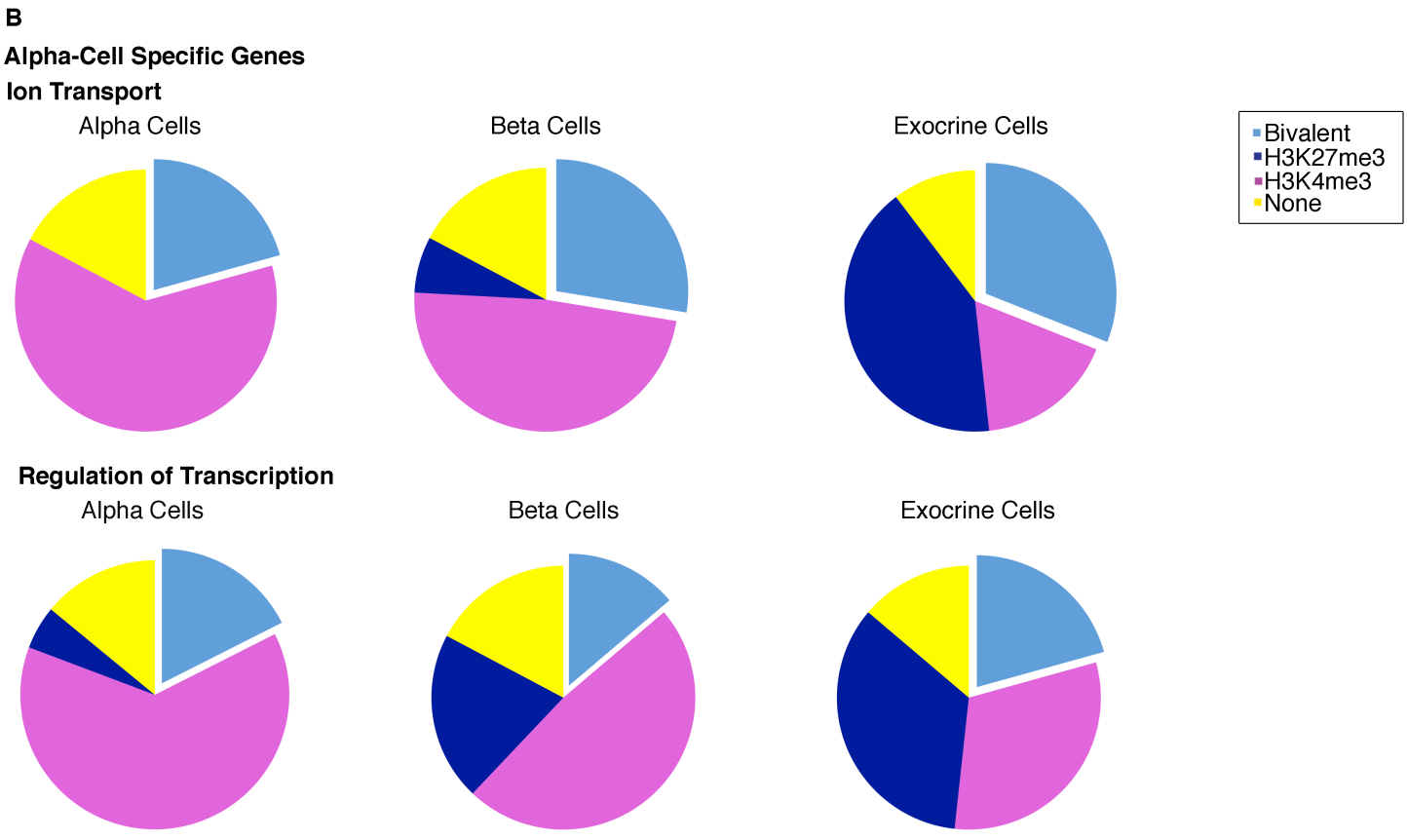
Validate the Library

Illumina recommends using Agilent Technologies' 2100 BioAnalyzer to check the size, purity, and concentration of the library (Agilent's DNA-1000 kit).

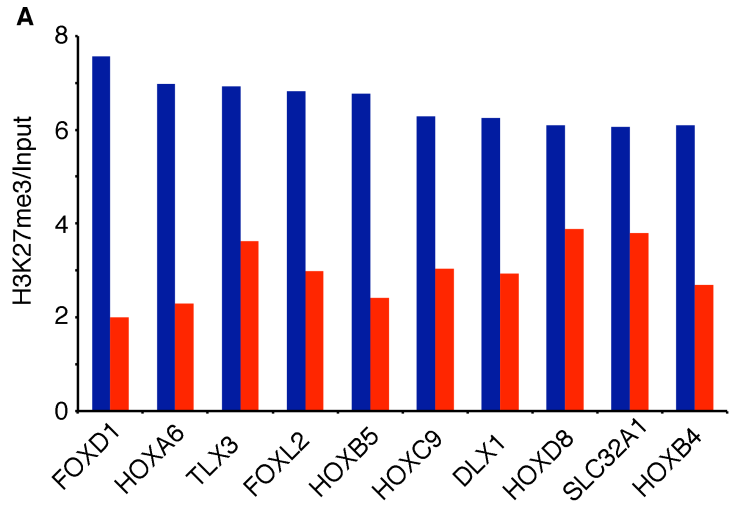


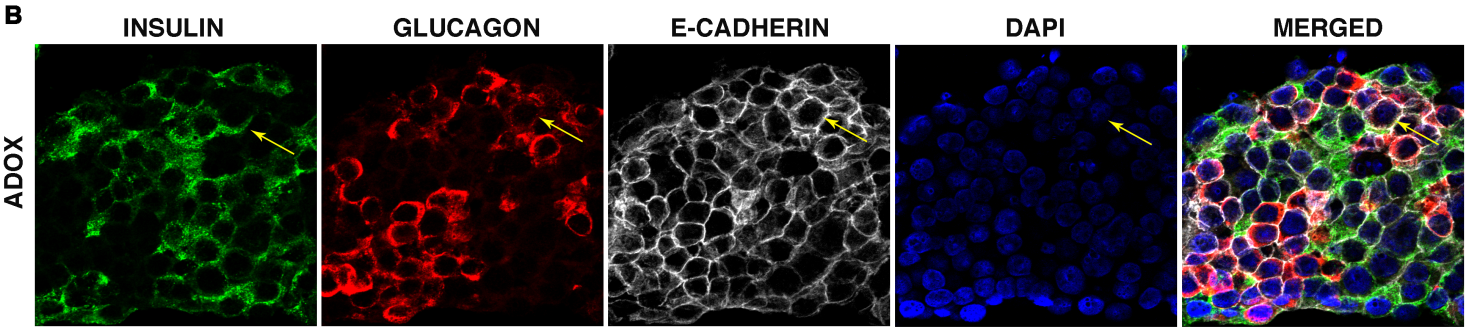
Supplemental Figure 1: (A) Separate analysis of bivalent marks in one donor (CITH068) confirms the higher number of bivalent marks in alpha cells.

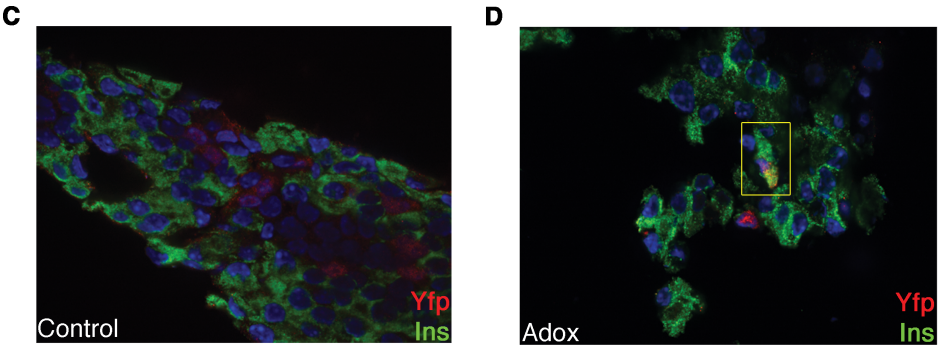




Supplemental Figure 2. Integrative analysis of α- and β-cell signature genes with our ChIP-Seq data sets. (A) Quantitative analysis of H3K4me3 and H3K27me3 levels in our strongly cell-type specific signature genes (α-strong (orange), β-strong (blue), exo-strong (yellow), as coloring of bars in Figure 2B). As expected, signature genes show increased H3K4me3 levels in their respective cell type. H3K27me3 enrichment of α-specific genes were found at comparable levels in α- and β-cells, but increased levels in α–cells and decreased levels in β-cells were found in β-cell specific genes. This indicates higher prevalence of H3K27me3 levels in α-cells, repressing β-cell signature genes, and not vice versa. (B) Analysis of the histone modification landscape in α-cell signature genes revealed comparable percentages of bivalently marked genes across α-, β- and exocrine cells. Of the alpha-cell enriched ion transport genes, 21%, 28% and 31% were marked bivalently in α-, β-, and exocrine. For the α-cell signature genes involved in transcriptional regulation, 17%, 14% and 21% were marked as bivalent in α-, β-, and exocrine cells, respectively.







Supplemental Figure 3, related to Figure 6: (A) Treatment of human islets with the histone methyltransferase inhibitor Adox leads to decrease in H3K27me3-enrichment. (B) Treatment of human islets with the histone methyltransferase inhibitor Adox leads to co-localization of glucagon (red) and insulin (green) granules in pancreatic cells (yellow arrows). Orginial magnification 63x. (C and D) Adox-treatment of islets from GlucagonCre;Rosa26EYFP mice results in the occurrence of insulin granules (green) in YFP+ cells (red) (yellow box), indicating partial α- to β-cell fate conversion, which was not observed in control islets (C). Original magnification 60x.

ID	Sex	BMI	Age	Cause of death	Race	Experiments
CITH053	F	24.5	49	CVA, ICH	AA	H3K4me3, RNA-Seq
CITH068	M	26.8	52	HT	C	H3K4me3, H3K27me3, RNA-Seq
XHP472	M	25.9	52	CVA	H	H3K4me3
CITH070	M	29.2	28	Anoxia	C	H3K4me3, RNA-Seq
ICRH025	M	28.3	47	CVA, ICH	AA	H3K27me3
XK2367	F	22	54	GSWH	C	H3K27me3
CITH090	M	33.8	22	CVA	C	Adox treatment of cultured islets, IF
ZAN423	M	33	27	HT	C	Adox treatment of cultured islets, sort
ZIQ120	F	25.4	57	CVA	H	Adox treatment of cultured islets, IF

Table S1, related to Figure 1: Islet Donor Information.

AA = African American
C = Caucasian
H = Hispanic

CVA = Cerebrovascular accident
GSWH = Gun shot wound to head
HT = Head trauma
ICH = Intracerebral hemorrhage

Region	Beta	Alpha	Beta/Alpha
chr22:45828637-45828942	14.54	2.46	5.91
chr15:32834228-32834476	11.05	3.13	3.53
chr7:98085143-98085366	43.10	12.86	3.35
chr7:98085583-98085801	47.10	16.87	2.79
chr16:546970-547176	7.28	2.78	2.62
chr7:97681668-97681869	14.41	5.77	2.50
chr22:17401525-17401730	12.20	4.99	2.45
chr7:97680610-97680816	11.32	4.72	2.40
chr22:28447815-28448044	11.87	5.41	2.20
chr10:71888615-71888815	14.43	7.05	2.05
chr12:107438050-107438287	19.60	9.63	2.04
chr2:91141148-91141355	17.53	8.65	2.03
chr10:100010310-100010623	4.48	39.59	-8.83
chrX:129343875-129344132	2.63	13.44	-5.11
chr1:146614868-146615074	8.57	18.98	-2.22
chr13:57102889-57103187	5.25	11.11	-2.12
chrX:110229332-110229577	15.61	31.49	-2.02

Table S4: Identification of non-coding human endocrine cell-type specific transcripts. Analysis of our RNA-seq data revealed twelve β-cell specific and five α-cell specific non-coding RNA transcripts not previously annotated in the genome (>2x higher mean expression than the opposite cell type). Stringent filtering criteria included a minimum length of 200bp with at least 1RPM (read per million) coverage and removal of all regions overlapping a repeat region from the UCSC Repeat Masker track, which includes other non-coding RNAs, such as ribosomal RNAs, and small nucleolar RNAs. In addition, regions were required to be within 5kb of a H3K4me3 peak for additional evidence for transcription.

GO category	Gene count	Benjamini p-value	FDR
Pattern specification process	57	1.3E-21	9.0E-22
Embryonic morphogenesis	59	2.5E-20	3.3E-20
Neuron differentiation	69	2.4E-19	4.7E-19
Embryonic organ development	43	7.4E-19	2.0E-18
Regionalization	42	9.7E-16	3.2E-15
Embryonic organ morphogenesis	35	8.4E-16	3.4E-15
Sensory organ development	44	5.7E-15	2.7E-14
Neuron development	52	1.1E-13	5.9E-13
Skeletal system development	48	3.0E-12	1.8E-11
Tube morphogenesis	30	4.7E-12	3.2E-11
Tube development	39	5.8E-12	4.3E-11
Cell morphogenesis involved in differentiation	41	6.6E-12	5.3E-11
Cell morphogenesis	50	8.4E-12	7.3E-11
Cell fate commitment	30	4.0E-11	3.8E-10

Table S6, related to Figure 4: Top GO categories from DAVID analysis on the genes marked bivalently by H3K4me3 and H3K27me3 in alpha cells and marked monovalently by the repressive H3K27me3 mark in beta cells show strong enrichment in developmentally relevant processes.

Gene	Alpha	Beta	Exocrine	T-Cells
HNF1A	H3K4me3	H3K4me3	H3K4me3	-
PCSK2	H3K4me3	H3K4me3	H3K27me3	H3K27me3
IRX2	H3K4me3	H3K27me3	H3K27me3	H3K27me3
GCG	H3K4me3	H3K27me3	H3K27me3	H3K27me3
DPP4	H3K4me3	Bivalent	Bivalent	H3K4me3
PTPRD	Bivalent	Bivalent	H3K27me3	H3K27me3
ARX	Bivalent	H3K27me3	H3K27me3	H3K27me3
IRX1	Bivalent	H3K27me3	H3K27me3	H3K27me3
MAFA	Bivalent	H3K4me3	H3K27me3	H3K27me3
PCSK1	Bivalent	H3K4me3	H3K27me3	H3K27me3
KCNQ2	Bivalent	H3K4me3	H3K27me3	H3K27me3
IAPP	Bivalent	H3K4me3	H3K27me3	-
GLP1R	Bivalent	H3K4me3	H3K4me3	H3K27me3
PDX1	Bivalent	H3K4me3	H3K4me3	Bivalent
HDAC9	Bivalent	Bivalent	H3K27me3	H3K4me3
INS	H3K27me3	H3K4me3	H3K27me3	-
INS-IGF2	H3K27me3	H3K4me3	H3K27me3	H3K27me3
CDKN1C	H3K4me3	H3K4me3	H3K4me3	H3K4me3
KCNJ11	H3K4me3	H3K4me3	H3K4me3	H3K4me3
SLC30A8	H3K4me3	H3K4me3	H3K27me3	H3K27me3
NKX6-1	H3K4me3	H3K4me3	Bivalent	H3K27me3
AMY1A	-	-	-	-
AMY1B	-	-	-	-
AMY1C	-	-	-	-
AMY2A	H3K27me3	H3K27me3	-	-
AMY2B	H3K27me3	H3K27me3	H3K27me3	-
PNLIP	H3K27me3	H3K27me3	-	H3K27me3
CTRB1	-	-	H3K4me3	-
CTRB2	-	-	H3K4me3	-
SPINK1	-	H3K27me3	H3K4me3	H3K27me3
PRSS3	H3K4me3	Bivalent	-	H3K27me3
JAG1	H3K4me3	Bivalent	H3K4me3	Bivalent
SOX9	Bivalent	Bivalent	H3K4me3	H3K27me3
PTF1A	Bivalent	H3K27me3	H3K4me3	H3K27me3

Table S8, related to Figure 5: Histone modification profiles of selected α-, β-, and exocrine-specific genes in α-, β-, exocrine and $CD4^+$ T-Cells.