HHS Public Access

Author manuscript

Cell Metab. Author manuscript; available in PMC 2024 May 02.

Published in final edited form as:

Cell Metab. 2023 May 02; 35(5): 821–836.e7. doi:10.1016/j.cmet.2023.03.008.

Epigenetic dosage identifies two major and functionally distinct β-cell subtypes

Erez Dror[1,*], **Luca Fagnocchi**[2], **Vanessa Wegert**[1,2], **Stefanos Apostle**[2], **Brooke Grimaldi**[2], **Tim Gruber**[2], **Ilaria Panzeri**[1,2], **Steffen Heyne**[1], **Kira Daniela Höffler**[1], **Victor Kreiner**[1], **Reagan Ching**[1], **Tess Tsai-Hsiu Lu**[1], **Ayush Semwal**[2], **Ben Johnson**[2], **Parijat Senapati**[3], **Adelheid Lempradl**[1,2], **Dustin Schones**[3], **Axel Imhof**[4], **Hui Shen**[2], **John Andrew Pospisilik**[1,2,5,*]

[1]Department of Epigenetics, Max Planck Institute of Immunobiology and Epigenetics, Freiburg 79108, Germany

[2]Van Andel Institute, Grand Rapids, MI 49503, USA

[3]Department of Diabetes Complications and Metabolism, Diabetes and Metabolism Research Institute, Beckman Research Institute of City of Hope, Duarte, CA, 91010, USA

[4]Biomedical Center Munich, Ludwig Maximilian University of Munich, 82152 Planegg-Martinsried, Germany

[5]Lead contact

Summary

The mechanisms that specify and stabilize cell subtypes remain poorly understood. Here, we identify two major subtypes of pancreatic β-cells based on histone mark heterogeneity (β_{HI} and β_{LO}). β_{HI}-cells exhibit ~4-fold higher levels of H3K27me3, distinct chromatin organization and compaction, and a specific transcriptional pattern. β_{HI} and β_{LO} cells also differ in size, morphology, cytosolic and nuclear ultrastructure, epigenomes, cell surface marker expression, and function, and can be FACS separated into $CD24^+$ and $CD24^-$ fractions. Functionally, β_{HI} cells have increased mitochondrial mass, activity, and insulin secretion *in vivo* and *ex vivo*. Partial loss-of-function indicates that H3K27me3 dosage regulates β_{HI}/β_{LO} ratio *in vivo,* suggesting that control of β-cell subtype identity and ratio are at least partially uncoupled. Both subtypes are conserved in humans, with β_{HI}-cells enriched in humans with type-2 diabetes. Thus, epigenetic

[*]Correspondence: andrew.pospisilik@vai.org (J.A.P.), dror@ie-freiburg.mpg.de (E.D.).

Author contribution

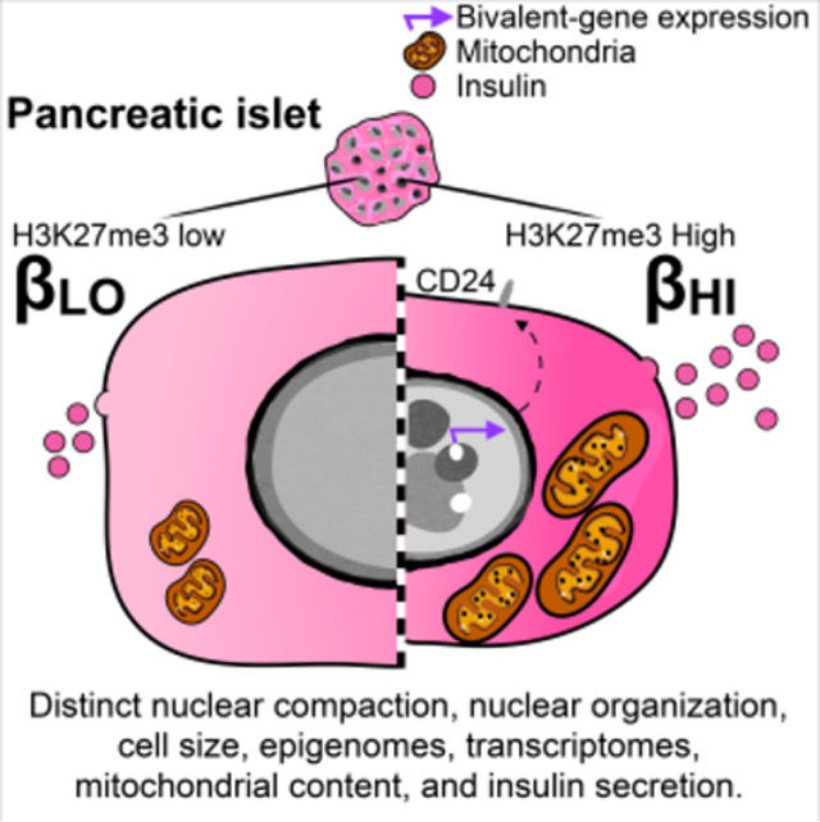
ED and JAP conceived the project. ED designed, performed, and analyzed all experiments unless stated differently. LF analyzed the ChIP-seq and RELACS datasets, and together with SA the DNA methylation array datasets. VW helped perform the *in vivo* experiments and islet isolations. SH supported the bioinformatics work and code for the *SCAN*-seq multimodal pipeline. LF, SA, BJ, AS, PS, DS provided support for the bioinformatic analysis. KDH, IP, VK, RC, TTL, AL, BG, TG, and AI helped perform experiments. JAP provided study design, analysis, interpretation and financial support. ED and JAP wrote the manuscript.

Declaration of interests

The authors declare no competing interests.

dosage is a novel regulator of cell subtype specification and identifies two functionally distinct β-cell subtypes.

Graphical Abstract



In brief

β-cells, the sole providers of insulin, are highly specialized, long lived cells that rely on specific epigenetic systems to maintain their identity. Here, by leveraging distinct differences in the histone modification H3K27me3, Dror et al. identify and characterize two subtypes of β-cells that are distinctive by their morphology, epigenomes, transcriptomes and function.

Introduction

β-cells are the sole providers of insulin in the body, acting to optimize nutrient uptake and storage, and to prevent hyperglycemia. During development, β-cells differentiate through progressive activation of transcription factor-directed gene networks and undergo functional maturation during early post-natal life[1,2]. Adult β-cells are highly specialized, quiescent and represent one of the longest-lived cell types in the body, averaging ~30–40 years in elderly humans[3,4]. β-cells therefore rely on specific epigenetic systems to stabilize and maintain cell identity over expansive time scales[5,6]. A relative loss of functional pancreatic β-cell mass results in diabetes.

Significant cell-to-cell heterogeneity has been observed within the β-cell compartment since the 1960s[7]. Early studies found heterogeneity in glucose thresholds, calcium handling, and insulin secretion[8,9], observations that were later confirmed using advanced optical and genetic tools[10,11]. Recently, specialized molecular tools identified Cfap126 as a factor that affects β-cell heterogeneity along a primary maturation gradient[12], virgin cells[13], and immune evading subsets[14]. Single-cell technologies have now re-focused attention on the origins, architecture, and potential therapeutic relevance of β-cell heterogeneity and subtypes[13,15–19], but the field has yet to assemble a universal framework for understanding β-cell subtypes and sub-states[20,21].

Challenges to establishing such a framework include a relative over-reliance on single-cell genomic techniques. These technologies overall represent shallow snapshots of the transcriptome, are almost entirely biased towards the active epigenome, involve numerous bioinformatic assumptions, and fail to distinguish ‘cell-state’ from ‘cell-type’ heterogeneity. For the purposes of this manuscript, we define cell-states as being primarily characterized by transient, periodic or progressive temporal dynamics, and for β-cells this includes circadian oscillations[22], transcriptional bursting[23], transcriptional noise[24], cell cycle[25], maturation[26], stress[17,27], and aging[24]. These properties are difficult to parse and regress out of single cell genomics data. Transgenic reporter systems also confound cross-study comparisons of β-cell heterogeneity; Cre recombinase, for example, triggers ER stress and generates artificial heterogeneity signals[28,29]. Inherently imperfect reporter expression can also generate “artificial” heterogeneity [30].

Here, we used reporter-independent approaches and found that the primary axis of β-cell heterogeneity is defined by dosage of the epigenetic silencing modification H3K27me3, and that it separates β-cells into two fundamentally distinct cell types (β_{HI} and β_{LO}) with distinct morphology, cytosolic and nuclear ultrastructure, transcriptome output, epigenome configuration, and function. In healthy adult mice, β_{HI} and β_{LO} cells comprise >90 of β-cells. They are present at an approximate 1:4 ratio (β_{HI}/β_{LO}) from pre-weaning through to old age, and can be FACS sorted live into $CD24^+$ and $CD24^-$ populations. β_{HI} and β_{LO} cells exhibit robust proliferation *in vivo* and *in vitro*. β_{HI} cells proliferate faster at baseline and their relative number is increased upon chronic high-fat diet (HFD). H3K27me3 dosage controls β_{HI}/β_{LO} ratio *in vivo*, with conditional heterozygosity of the polycomb repressive complex 2 (PRC2) core subunit Eed and the histone de-methylase Jmjd3, generating equal and opposite cell ratio skewing. Equally important, we demonstrate that β_{HI} and β_{LO} cells are conserved in humans, and that β_{HI} cells are enriched in Type 2 diabetes (T2D). These data identify two major β-cell subtypes, and identify epigenetic dosage as a novel and potentially targetable mechanism controlling β-cell compartment heterogeneity.

Results

Two common and epigenetically distinct β-cell subtypes

Historically, cell types were distinguished based on histopathological and nuclear differences[31–33]. To measure epigenetic heterogeneity in β-cells, we used FACS to quantify total H3K4me3 (active promoters), H3K27ac (active cis-regulatory elements), H3K36me3 (transcribed gene-bodies), H3K9me3 (silent constitutive heterochromatin), and

H3K27me3 (polycomb-associated heterochromatin) at single cell resolution. To avoid potential confounders of transgenic reporters, we performed antibody-based purification of freshly isolated, dissociated and fixed islet cells isolated from wildtype mice. We used insulin as a positive selection marker for all β-cells, and gated out cells that stained for CD45 (immune), CD31 (endothelial), SST (delta), GCG (alpha), and PP. Most histone marks showed robust and uniform immunoreactivity across all β-cells (Figure 1A; cell gating strategy, Figure S1A). Surprisingly, the signal for H3K27me3 appeared bimodal, suggesting two epigenetically distinct sub-populations (-LO vs -HI; Figures 1A, 1B). Averaged across independent biological replicates, -HI cells had a ~4.5-fold higher H3K27me3 mean fluorescence intensity (MFI) than -LO cells (Figure S1B). Imaging-flow-cytometry validated that the H3K27me3 signal in both -HI and -LO cells was nuclear in origin and ruled out cell-doublets, poly-nucleated cells, and cytosolic immunoreactivity as potential sources for the -HI signal (Figure 1C). H3K27me3-HI and -LO populations were consistently observed across experiments, animals, ages, and within islets of both males and females (Figure 1D; Figure S1C). In female β-cells, no difference in H3K27me3 immunoreactivity was observed on the silent X-chromosome (Barr body), highlighting the specificity of the H3K27me3 differences (Figures S1C, D). Importantly, -HI and -LO cells were found in all islets of all sizes (Figures S1C, E), arguing against inter-islet differences as the source of observed epigenetic signature. The H3K27me3 signal was validated using two independent antibodies (Figure 1A, monoclonal; and Figure S1F, monoclonal vs. polyclonal) and against β-cell-specific Eed/PRC2 knockout (KO) mouse islets that are deficient in H3K27me3 (βEedKO; Figure S1G). Parallel analyses of pancreatic islet α-, δ- and PP-cells suggested that the H3K27me3 signature was specific to β-cells (Figure S1H). Thus, β-cells exist in two common populations distinguished by their H3K27me3 levels.

Next, we used super-resolution confocal microscopy to test for differences in nuclear morphology (Figures 1E–I). H3K27me3-HI cells contained more H3K27me3-foci than -LO cells (Figures 1F, G), which might reflect compacted Polycomb-silenced genomic regions[34]. Whereas -LO cells showed H3K27me3 staining primarily at the transcriptionally silent nuclear periphery, -HI cell H3K27me3 was enriched in the active nuclear interior[35] (Figure 1F box 2, Figure 1H, and Figure S1I). Consistent with the role of H3K27me3 in chromatin silencing and compaction[36], nuclei were ~5 μm^3 smaller (on average) in the -HI relative to -LO cells (Figure 1I). Thus, pancreatic β-cells exist in two populations based on H3K27me3 level, chromatin organization, and nuclear compaction.

H3K27me3-HI cells are transcriptionally distinct and express cell surface CD24

To determine if the H3K27me3 difference between -HI and -LO β-cells translated into stable differences in transcriptome output, we FACS sorted -HI and -LO β-cells (INS+ but $GCG^-/SST^-/PP^-/CD31^-/CD45^-$) from eight wildtype mice across two age groups (4 or 10 weeks old; Figure 2A) and performed RNA-seq. By principal component analysis (PCA), -HI and -LO H3K27me3 status separated on PC1, indicating that stable and reproducible transcriptome differences exist between the -HI and -LO β-cells, and that these differences are maintained from weaning (4 weeks) into adulthood (Figure 2B). Differentially expressed genes were enriched for a set of poised or bivalent genes[5]. These H3K27me3-dependent genes were upregulated in -HI β-cells (Figure 2C; see full gene list

in Data S1), suggesting that the H3K27me3-mark plays a role in establishing differences between the two populations. The set of differentially regulated genes also included alternate islet endocrine lineages factors (*Ppy*, *Gcg*, and *Sst*), heterogeneity and plasticity markers (*Arx*, *Etv1*, *Gpx3* and *Rbp4*), and cell surface proteins (*Slc23a4* and *Cd24a*). Slc23a4 is a presumed human pseudogene, so we obtained antibodies to CD24 and performed FACS analysis to test if -HI and -LO β-cell populations could be distinguished by cell surface staining. Co-staining live cells with CD24 and H3K27me3 resolved -HI and -LO subsets into two distinct β-cell populations (Figure 2D, gating strategy in Figure S2A), where ~20% of all INS^+ β-cells were CD24-positive ($CD24^+$) and ~80% CD24-negative ($CD24^-$, Figure 2D). $CD24^+$ cells showed higher levels of H3K27me3 (Figures 2D, E). We also observed rare INS^+ cells (~1%) with extremely high CD24 levels ($CD24^{high}$ in Figure S2B, left panel). These $CD24^{high}$ cells were SST^+ (Figure S2B right panel), and are consistent with prior studies showing strong δ-cell expression of CD24[37]. These rare double-hormone positive (INS^+/SST^+) cells were excluded from all further analyses.

We validated the CD24 surface stain in several ways. By using an *Ins1*-YFP reporter mouse (*Ins1*-cre × LSL-YFP) and confocal imaging, we found CD24 expression was restricted to a subset of live β-cells, and determined that CD24 expression is cell membrane specific in single cells (Figure 2F upper panel) and in whole islets (Figure 2F lower panel, note the dim labeling compared to the YFP negative, $CD24^{high}$ delta cells). Live-sorted $Ins1\text{-}YFP^+$/$CD24^+$ double-positive cells also showed higher H3K27me3 (Figure S2C, left) and nuclear compaction (Figure S2C, right), indicating that at least in control contexts, $CD24^+$ and H3K27me3-HI β-cells are largely the same. Specificity of the CD24 antibody was confirmed using β-cells from CD24 knockout mice (Figure S2D). Thus, CD24 surface expression discriminates H3K27me3-HI and H3K27me3-LO β-cells.

To associate these findings with transcriptional heterogeneity, we modified CELseq2[38] to enable concomitant quantification of cell *S*urface (CD24), *C*ytoplasmic (Insulin) *a*nd *N*uclear (H3K27me3) protein epitopes at single-cell resolution, and applied the method to purified INS^+ β-cells isolated from wildtype mice. The new '*SCAN*-seq' method is outlined in detail in the methods. As reported elsewhere[17,19], UMAP projections based solely on the single-cell transcriptome fraction of the data, identified two major clusters (Figure 2G), with additional sub-clusters emerging as stringency for within-cluster heterogeneity is increased (Figure S2E). Consistent with the data above, one of the two main clusters exhibited elevated expression of H3K27me3 marked genes (Figure S2F). Projecting the quantitative protein measurements onto the transcriptome UMAP showed that $CD24^+$ and H3K27me3-HI β-cells are one of the two major clusters (Cluster 1, Figures 2G–I). Indeed, cells of the $CD24^+$ transcriptomic cluster were smaller (lower FSC). Interestingly, they also showed distinct granularity (SSC) relative to the $CD24^-$ cluster. We suspect that these observations would be overlooked with common scRNAseq techniques because *Cd24a* mRNA expression is low and physical characteristics and protein levels are not measured. The $CD24^+$ transcriptomic cluster showed lower *Ins1* and *Ins2* mRNA counts despite higher insulin protein staining, highlighting the added value of *SCAN*-seq over conventional scRNA-seq. Thus, H3K27me3-HI and H3K27me3-LO β-cells can be separated by single-cell transcriptomics, FSC, SSC, as well as H3K27me3, Insulin and CD24 protein expression.

β_{HI} and β_{LO} cells

Reanalysis of published β-cell scRNA-seq datasets[17,39–42] validated that the CD24/H3K27me3 axis is evident a the primary axis of heterogeneity across independent studies (Figure S2F), and that expression of H3K27me3-controlled genes separates β-cells into two primary clusters in mice (Figure S2G) and humans (Figure S2H). We also validated the axis in human β-cell single nucleus (sn) ATACseq datasets[19] (which catures heterogeneity in transcriptional potential)(Figure S2I). Cell-state heterogeneity and gradients of heterogeneity markers (where detectable) appear predominantly *as gradients within* CD24/H3K27me3 discordant clusters (UMAP matrices in Figure S2K, L). The CD24/H3K27me3 axis is distinct from Beta-1/Beta-2 clusters described in human snATACseq that correlate strongly with markers of stress. Thus, the CD24/H3K27me3 axis is evident in as a primary dimension of β-cell heterogeneity across publicly available datasets.

Given the findings above and the consistency across public datasets, we named these cell types **β_{HI}** and **β_{LO}**. β_{HI} cells have HIgher nuclear compaction, H3K27me3, Insulin protein content, cell-surface CD24 expression, and as shown Figure 5 below, higher mitochondrial content and insulin secretory function. Together, they comprise ~90–95% of all insulin protein positive pancreatic β-cells in adult mice, and are detectable from before weaning to one year of age (Figure 2J). We observed a progressive increase in H3K27me3 in both cell populations with age (Figure S2J), and a slow decline in β_{HI}/β_{LO} ratio. Based on Ki67 staining of freshly isolated islets, βHI and βLO cells both proliferate with a mild but significant increase in β_{HI} cells at baseline (Figure 2K). Proliferative capacity was validated *in vivo* using Edu-incorporation and, interestingly, both β_{HI} and β_{LO} increase proliferation upon acute HFD (Figure 2L). We also find that chronic high fat diet increases β_{HI}/β_{LO} ratios (4-weeks, Figure 2M). Thus, β_{HI} and β_{LO} are proliferation competent, responsive to metabolic changes, and comprise ~90% of the β-cell compartment, across metabolic states.

β_{HI} and β_{LO} cells exhibit distinct transcriptomes

We next separated β_{HI} and β_{LO} cells by FACS (CD24 and H3K27me3), and performed bulk RNA-seq on the two populations. Leveraging two dimensions for sorting, we now identified >2500 differentially expressed genes (Figures 3A–C and S3A; see full gene list in Data S1), including pathway level enrichments for mitochondrial and amide metabolic processes, oxidative phosphorylation, nuclear RNA processing factors, and histone modification (Figure 3D). Consistent with their H3K27me3-HI phenotype, β_{HI} cells show upregulation of *Ezh2,* the main H3K27me3-depositing methyltransferase, as well as increased levels of the H3K27me3 demethylase *Kdm6b* (*Jmjd3*); active mark 'erasers' *Hdac4/5*, chromatin silencers *Cbx4*, *Suv420h2*, *Uhrf2*, and *Ehmt1/2*; and 3D looping factors *Ctcf* and *Kmt2c*/*Mll3* (Figure 3E; blue, and Figure S3B). These data suggest that β_{HI} and β_{LO} cell differences are reinforced by a persistent and complex chromatin regulatory network. We found no differences in the hallmark differentiation factors *Pdx1*, *Neurod1*, *Pax6, Nkx6.1*, *Mafa, Nkx2.2, Cfap126,* or *Cd81* (Figure 3E; beige), data that were validated at the single cell level (Figure S2K, L). We did observe modest, opposing regulation of the maturation factors *Ucn3* and *Rfx6* (Figure 3E; red). Consistent with these findings, the Rfx6 binding motif was one of several motifs enriched at promoters of β_{HI} upregulated genes (Figures 3F and S3C). In keeping with the *SCAN*-seq data (Figures 2G–I), β_{HI} cells showed lower *Ins1*

and *Ins2* transcript levels (Ins1/2; Figure 3E, black, Figure S3D) despite higher insulin levels (Figure 3G). Thus, β_{HI} and β_{LO} are both highly differentiated and exhibit distinct metabolic and chromatin regulatory gene expression.

β_{HI} and β_{LO} cells exhibit distinct epigenomes

To better understand the observed differences in nuclear ultrastructure and H3K27me3 levels, we FACS purified β_{HI} and β_{LO} cells from wildtype mice and performed H3K27me3 ChIP-seq on three paired independent biological replicates. As with their transcriptomes, β_{HI} and β_{LO} cell H3K27me3 profiles were strongly correlated (Figure 4A and Figure S4A). Consistent with literature expectations[5,43,44], H3K27me3 enriched at focal regions across the genome and as broad domains containing transcriptionally silent developmental genes such as the Hox clusters and imprinted loci (Figures 4A, B and Figures S4B, C).

PCA separated β_{HI} and β_{LO} cells on the first principal component indicating reproducible H3K27me3 differences (Figure 4C). Interestingly, H3K27me3 levels were unchanged at broad domains (Figure 4B and Figures S4B, C). Rather, H3K27me3 differences were enriched at promoters and transcriptional start sites (TSS; Figure 4D and Figure S4D), suggesting H3K27me3 might underpin differences in gene expression between the two cell types. This data highlighted a striking specificity given that H3K27me3 is primarily found *outside* genic promoters and that these represent only a fraction of the ~ 2% of the mammalian genome that is coding (Figure S4D).

We called differential H3K27me3 enrichment at TSSs (Figure 4E, Figure S4E) and explored the relationship with transcriptome output. We detected ~5200 differential peaks across ~4750 annotated TSSs at ~1550 unique genes. Approximately 80% of these genes showed relative H3K27me3 enrichment in β_{LO} cells (Figure S4E, upper portion). TSSs specifically marked in β_{LO} cells were enriched for lowly transcribed genes annotated as ‘bivalent’ in whole islet data[5], including the *Cd24a* locus (Figure 4E–G and Figures S4E–G). These data indicated that with increased cell-type resolution, β-cell bivalent domains largely resolve into cell subtype-specific active or silent states. Additional ChIP-seq confirmed this idea showing equal levels of H3K4me3 at β_{LO}-specific TSS’s in both cell types (Figure S4H, I). H3K27me3 differences correlated inversely with β_{HI} vs β_{LO} gene expression differences (Figures 4H–J) suggesting the H3K27me3 differences were causal. Consistent with higher H3K27me3 in β_{HI} cells (Figures 1E–G; -HI), β_{HI}-specific H3K27me3 deposition was much broader, extending across the entire gene-body (Figures 4K, S4E). Thus, the hallmark H3K27me3 differences between β_{HI} and β_{LO} cells appear to drive their differential transcriptional identities.

Finally, we evaluated whether silent epigenome differences were restricted to H3K27me3. We performed quantitative DNA methylation profiling on purified fresh β_{HI} and β_{LO} samples using Infinium Mouse Methylation BeadChips. Interestingly, cell subtype separated on the first PCA component indicating that β_{HI} and β_{LO} cells are also distinguished at the levels of DNA methylation (Figure 4L). Differential DNA-methylation was very specifically enriched at enhancers and H3K27me3 annotated genomic regions (Figure 4M, Figure S4J); at *Lmx1b* and three annotated pseudogenes; and at regions that bind JUNB, AEBP2, CEBPD, MAFB, ATF3, H3K3me1 and, interestingly, the developmental regulator

PROX1 (Figure 4M). The human PROX1 locus harbors a significant Type-2 diabetes variant (rs340874) suggesting the variant could exhibit cell subtype specific effects. Binding-site enrichments in regions hypermethylated in β_{LO} cells included those for NKX6.1 and NEUROD1 (Figure 4M) possibly suggesting a more complete decommissioning of NKX6.1 and NEUROD1-associated plasticity potential in β_{LO} cells. Thus, β_{HI} and β_{LO} cells exhibit distinct H3K27me3 and DNA-methylation control.

β_{HI} and β_{LO} cells are stably and functionally distinct

Interestingly, we observed a unique signature of mitochondrial genome encoded transcripts upregulated in β_{HI} cells (Figure 5A, green). A striking co-linearity and uncoupling from nuclear-encoded mitochondrial genes expression (Figure 5A orange, Figures S3D S5A–D) suggested increased mitochondrial mass in β_{HI} cells. Indeed, qPCR revealed a near 2-fold increase in β_{HI} cell mitochondrial DNA content (Figure 5B). FACS-based quantification of TOM20 validated this (Figure 5C) and showed that the difference was uniform across the entire β_{HI} and β_{LO} compartments (except a small subset of β_{HI} cells; Figure 5C). Confocal imaging revealed larger, more elongated mitochondria in β_{HI} cells (Figure 5D, E), and, increased mitochondrial activity (TMRM fluorescence[45])(Figure 5F). Thus, β_{HI} cells exhibit increased relative mitochondrial mass, transcription, elongation and TMRM-associated activity.

Mitochondria are a defining node in β-cell stimulus secretion coupling. To test for differences in mitochondrial and secretory function, we therefore FACS-purified and reaggregated β_{HI} and β_{LO} cells into uniquely β_{HI} or uniquely β_{LO}-specific spheroids (monotypic pseudo-islets; Figure S5E). No differences were observed in spheroid-forming rates (Figure S5F) or connexin 36 gap junction expression in the aggregated organoids (*Gjd2;* Figure S5G). Even after 7 days in culture, signature mRNA differences remained stable and true to the respective cell-type-of-origin. Specifically, *Ins1* and *Ucn3* were up in β_{LO} monotypic pseudo-islets, while mitochondrial, *Rfx6*, and CD24 transcripts were up in β_{HI} cells (Figure 5G, H). β_{HI} and β_{LO} cells therefore maintain their overall distinctions through dissociation, reaggregation, and culture.

We also used monotypic pseudo-islets to perform metabolic profiling under basal and glucose-stimulated conditions. β_{HI}-monotypic spheroids showed increased oxygen consumption (OCR) relative to extracellular acidification (ECAR; Figure S5H). Both spheroid types showed significant ECAR responses upon glucose stimulation. Interestingly though, β_{HI}-monotypic spheroids showed a specific and substantial OCR response (Figure 5I and S5I). β_{HI} cells are therefore more oxidatively competent in both basal and glucose-stimulated contexts. Finally, we measured GSIS in a parallel single monotypic spheroid setup (Figure S5J). β_{HI} and β_{LO} spheroids both showed robust GSIS (Figure 5K, left panel). Importantly though, β_{HI} spheroids exhibited increased GSIS with a near-doubling of insulin secretion upon high glucose challenge. Under normal culture conditions β_{LO} cells showed a modest increase in chronic insulin output (Figure 5K right panel). Thus, β_{HI} and β_{LO} cells are characterized by stable differences in mitochondrial activity, mitochondrial substrate responses, and GSIS.

H3K27me3 dosage controls overall heterogeneity and β_{HI}/β_{LO} cell ratio

Eed is a critical core subunit of the PRC2 complex that is responsible for H3K27me3 deposition[46]. We previously generated animals with a β-cell-specific loss of *Eed*, and showed that β-cells in these mice lose all detectable H3K27me3 between 2 and 8 weeks of age[5], and after several months exhibit a stark, penentrant and progressive loss of β-cell identity[5]. To test if H3K27me3 dosage itself is necessary for β-cell subtype specification and maintenance, we isolated islets from 2-month-old β-EedKO animals and performed *SCAN*-seq. As expected[5], β-EedKO β-cells were devoid of H3K27me3, exhibited normal expression of all key β-cell markers, and had normal insulin levels (Figures S6A–C). We then performed clustering analysis on cells from wild-type and β-EedKO animals, with wild-type β_{HI} and β_{LO} cells serving as the reference (Figure 6A, C and Figure S6D). In contrast to wildtype, EedKO β-cells formed a trajectory ending in a final tight cluster (Figures 6B–C). Using within cluster *sum of the squared errors* (SSE) as a readout, EedKO β-cells ultimately 'collapse' into a state of low cell-to-cell dispersion (Figures 6B, C and Figure S6B, C), lower even than the dispersion of β_{HI} or β_{LO} cells alone (Figure 6C and Figure S6E). This observation was replicated with a dedicated cluster tree analysis (Figure S6F). Thus, H3K27me3/PRC2 is necessary *in vivo* for the maintenance of β-cell transcriptional heterogeneity, including the separation of β_{HI} and β_{LO} cells.

We also wanted to understand the consequences of partial H3K27me3 dysregulation. We generated an independent cohort of mice and performed *SCAN*-seq on β-cells from heterozygous knockouts (β-Eed-Het) and their sex-matched wild-type silblings (Ins1-cre^{+}; WT), Overall the β_{HI} and β_{LO} transcriptomes for each genotype superimposed indicating that partial loss of function has no effect on identity (Figures S6G–I). Interestingly, however, β-Eed-Het animals showed an increased β_{HI}/β_{LO} ratio relative to wild-type (Figure S6J). Since single-cell transcriptomics are not designed or intended to provide accurate relative cell counts, we validated the ratio difference by FACS (n=18 mice each group; Figure 6D). Consistent with the heightened GSIS function of Type-1 cells *ex vivo* (Figure 5K), β-Eed-Het animals exhibited improved glucose tolerance *in vivo* (Figure S6K), and increased insulin secretory function *ex vivo* in isolated islets (Figure S6L). We also examined samples from β-cell specific *Jmjd3* heterozygotes, a model that *increases* H3K27me3 levels (Jmjd3 is an H3K27me3 demethylase). β-*Jmjd3*-heterozygotes showed an equal and opposite cell subtype distortion (n=9 mice each; Figure 6E). To the best of our knowledge, these represent the first genetic models that trigger β-cell subtype ratio distortion without impacting cell identity. These data also demonstrate that H3K27me3 is a critical determinant of β_{HI}/β_{LO} ratios *in vivo*, and (by extension) the primary axis of β-cell heterogeneity.

β_{HI} and β_{LO} cells are conserved in humans, and exhibit altered ratios in diabetes.

To test whether β_{HI} and β_{LO} β-cells are conserved in humans, we dispersed donor-derived islets provided by the Alberta Diabetes Institute IsletCore (Table S1) and separated them by CD24/H3K27me3 FACS. As in the mouse, human CD24^{+} β-cells were consistently H3K27me3-HI, and CD24^{-} cells were consistently H3K27me3-LO (Figures 7A, B). CD24 positive and negative cells were consistently observed within individual islet fragments (Figure 7C). Whereas mouse preparations reproducibly yielded ~20% β_{HI} cell content (Figure 2D; 19 ± 2 % of all β-cells in young adults), human islet donor preparations

exhibited β_{HI} cell numbers ranging from ~30% to ~90% of the INS^+ cell fraction (Figure 7A, Figure S7A). Donor-to-donor variability is well-acknowledged in human islet research[47]; these data demonstrate the importance of multiple donors when exploring cell-type heterogeneity in human islets.

These findings were again validated by analyzing public human scRNA-seq/snATAC-seq data, which showed that the major axis of heterogeneity is driven by H3K27me3 targeted genes (Figures S2H–I). This axis is distinct from previously reported stress-response associated heterogeneity[17] (compare Figure S7B to Figure S2H; UPR= unfolded protein response). To examine potential subtype specific regulation in the context of type-2 diabetes, we analyzed a large single β-cell RNA-seq data set that included cells isolated from humans with or without diabetes[48]. Interestingly, β-cells grouped into 3 major clusters. One cluster comprised high stress and/or low-quality cells based on gene expression signature and total per-cell transcript counts (Figure 7D, Figures S7C, D). The other 2 major clusters distributed along a β_{HI}/β_{LO} axis according to differentially expressed genes from the two subtypes (Figures S7E, F). Slingshot trajectory analysis suggests that both β_{LO} and β_{HI} cells succumb to stress (Figure 7D). β-cells isolated from humans with T2D were enriched in the stressed cluster, confirming previous observations[49] (Figure 7F, Figure S7E). β-cells isolated from humans with T2D were also enriched for β_{HI} relative to β_{LO} cells (Figures 7E–F, Figures S7E–F), suggesting a diabetes-specific skew in β_{HI}/β_{LO} ratio. Thus, β_{HI} and β_{LO} cells are conserved in humans, and their ratio affected in T2D.

Discussion

There is currently no accepted definition for what constitutes a *bona fide* cell type or subtype. Historically, stable differences in function, cell-surface protein expression, nuclear and cytological morphology, epigenome configuration, transcriptome and lineage tracing have all been used (independently) to define cell types or subtypes. We suggest using **'cell-*state*'** to define all forms of β cell heterogeneity that reflect temporally-rooted differences eg. differentiation, maturation, aging, circadian rhythm, fasting/feeding/diet, and transcriptional bursting. We propose that β-cell '**cell-*types*'** (or ***subtypes***) be reserved for cells whose distinguishing features **i.** arise normally over developmental time-scales; **ii.** are reproducibly and stably detected across a wide range of contexts (ages, circadian time, diets, disease); **iii.** exhibit wide-spread and stable differences in their active *and silent* epigenomes, transcriptomes, surface protein expression and function; and **iv.** that maintain these differences through long-term culture under identical conditions. We propose '*sub-population' or 'subset;*' be used where these distinctions are not known or intended. So, while single-cell methods like scRNA-seq are ideally suited for detecting heterogeneity (i.e., differences), they do not necessarily generate the most appropriate data for defining a cell type (or subtype). Here, we used many methods to identify a new epigenetic axis that defines two primary β-cell subtypes (βHI and βLO cells) that are distinct by at least 7 independent criteria (function, FACS markers, epigenome configuration, transcriptome, nuclear ultrastructure, cytosolic ultrastructure (mitochondria & insulin protein), and morphology), and constitute ~90% of the β-cell compartment. We also identified a $CD24^{high}$ SST/INS double-positive sub-population that deserves careful

exploration considering recent work highlighting trans-differentiation and islet endocrine cell plasticity[13,50–52].

Prior research indicates the existence of other potential β-cell subtypes that are most certainly distinct from those described here (e.g. rare virgin cells and rare polyhormonal cells). Notably absent from our single-cell data are high-UPR clusters. In our hands, the immediate transcriptome 'freeze' that occurs in *SCAN*-seq eliminates high-UPR cells that we find in parallel CELseq2 tests from the very same islet isolation (not shown). These data suggest that a substantial fraction of the stress-response signature observed in single-cell genomic data likely results from isolation, fluidics, and sorting steps[53,54] as it can be eliminated by rapid fixation in the first step. Given the importance of UPR in β-cell biology[17,49,55,56], the experimental design for examining this process appears to be especially critical.

Populations of early responding and highly interconnected 'hub' β-cells (1–10% of total) classified by in situ Ca^{+2}-responsiveness have been reported with suggested mitochondrial functional heterogeneity. Those cells exhibited low insulin, signatures of immaturity, and increased metabolic function without a difference in TOM20[10]. A related set of 'leader' cells exhibit transcriptional enrichment of chromatin regulators[57], and derive from the smaller of two major scRNAseq clusters of β-cells[11]. PATCH-seq also identified gradients of electrophysiological responsiveness associated with physically smaller, *Rbp4*-enriched, β-cells that showed lower exocytosis under chronic conditions[48]. Our functional analyses and examination of transcriptional patterns of these factors (Figure S2K, L) suggest that the PATCH-seq *Rbp4*-enriched cells, and 'leader cells', comprise one of two subsets of β_{HI} cells.

Our findings demonstrate that the major axis of β-cell heterogeneity is formed at least in part by epigenetic silencing machinery. Interestingly, genes coding for reported markers of β-cell heterogeneity (including *Cfap126*, *Rbp4*, and *Ffar4*) and a large proportion of β-cell *disallowed* genes[58] map to H3K27me3-marked regions. Our examinations indicate that the maturation axes marked by Cfap126, Mafa, and CD81[12,18,59] exist *within both* β_{HI} and β_{LO} cell compartments, suggesting that maturation gradients exist *in vivo* for both cell-types. This raises the question of how β_{HI} and β_{LO} cells are specified, and how their maturation and maintenance couples to metabolic demands. We find both subtypes are present from early to late life in mice, they both proliferate, they are stable for at least 7 days under identical culture conditions, and they are metabolically distinct. mTOR and AMPK signaling play crucial roles in β-cell maturation process[60,61], and β-cell specific mTOR deficient mice have lower levels H3K27me3 and upregulation of a group of 'immature' genes[62]. We observed reciprocal mTOR regulation in β_{HI} and β_{LO} cells (increased expression of negative regulators of mTOR Tsc1, Tsc2, Ubr1, and Rictor in β_{HI}; upregulation of the positive regulators Lamtor2, Golph3, Rheb, Deptor, Lamtor1 in $\beta_{LO's;}$ not shown). This observation suggests that fidelity of cell subtype identity may be continuously reinforced by TOR signaling.

One of the seminal studies of β cell heterogeneity, Dorrell et al. [16], previously used surface antigens to identify four clusters of β-cells. In our *SCAN*-seq data, we found

differential expression of genes that corresponded with Dorrell's β1/β2 clusters (Figure S2K, M UMAPs panel), and identified Rfx6 and Mafb as a key transcription factors (Figure 3F) that differentiated their ST8SIA1^{-} β1/β2 cells. RFX6 also had promoter accessibility in INS^{high} state cells, as observed elsewhere[19]. These data therefore suggest that Dorrell's β1/β2 clusters may be β_{HI} cells, although our scRNAseq data show that Rfx6 also aligns with maturation gradients within βHI and βLO cells.

Next important steps for the field will be to test whether β_{HI} and β_{LO} cell differences can be harnessed for stem-cell based islet replacement strategies, and to test whether β_{HI} or β_{LO} cells are preferentially dysregulated in classically-defined and newly emerging diabetes subtypes[63]. Key steps will be to identify additional surface antigens that are robust against experimentally induced and donor-to-donor variation. An additional priority will be to understand the up- and down-stream factors that drive subtype specification, and the mechanisms that link Eed and Jmjd3 dosage to β_{HI}/β_{LO} ratio control.

To the best of our knowledge, the β-Eed-Het and β-Jmjd3-Het animals represent the first examples of reciprocal genetic models that specifically skew β-cell subtype ratios (heterogeneity) *in vivo.* In β-Eed-Het animals, scRNAseq data indicate that the skewed ratio occurs without any change in β-cell subtype transcriptomes. These data indicate that ratio control is independently regulated from identity, and that even subtle changes in the H3K27me3 levels[5] may modulate β_{HI}/β_{LO} cell ratios over the timeframes associated with human disease. Notably, the increase in β_{HI}/β_{LO} cell ratio in T2D (Figure 7F) is consistent with a model where H3K27me3 dysregulation causes skewing of cell subtype ratio. The heightened gluco-regulatory phenotype observed in β-Eed-Het animals suggests that the skewing observed in T2D constitutes a form of compensation, also that manipulating β-cell ratios could constitute a desirable therapeutic goal for metabolic disorders. The data suggest that *low-dose* or *intermittent* Ezh1/2 or Eed inhibition could serve a role in improving those methods aimed at generating β-cell replacements from stem or iPS cells. Substantial activity in the epigenetic inhibitor space has already identified a range of *in vivo* tolerated compounds with specificities for PRC2 catalytic and other subunits.

Finally, our multi-dimensional analysis does not support the prevailing view that there exists a common (>10% of β-cell compartment) immature β-cell subtype. Rather, out results support the notion that cell-*state* gradients exist across two highly specialized β-cell subtypes. β_{HI} and β_{LO} cells both exhibit robust and equal expression of essentially all known terminal differentiation markers, despite clear maturation gradients readily detectable within each. These nuances are important as the community works towards a common framework for β-cell heterogeneity.

Limitation of Study

One of the study's limitations is that our characterization of H3K27me3-axis is primarily antibody-based, and while validated across antibodies, the relative levels quantified may be affected by characteristics of the techniques such as affinity and antigen accessibility. Therefore, it will be important to characterize these differences using a more quantitative, orthogonal approach such as mass spectrometry, and extend this to all combinatorial histone modifications. High-throughput cytometry (such as CyTOF) could also be used to quantify a

larger number of chromatin marks, including histone and DNA modifications. The analysis in this study was primarily restricted to the differentiated β-cells in developed islets, developmental cues and cellular interactions that may be important for the emergence and maintenance of the HI and LO subtypes were not considered.

STAR Methods

RESOURCE AVAILABILITY

Lead contact—Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Andrew J Pospisilik.

Data and code availability

- Bulk RNA-seq, scRNA-seq, ChIP-seq/RELACS and DNA methylation array data generated in this study were deposited to Gene Expression Omnibus (GEO) repository with the GEO accession number GSE224061.
- The SeSAMe wrapper pipeline SeSAMeStr was published online in zenodo under DOI 10.5281/zenodo.7510575. No other custom code or mathematical algorithms were generated in this study. All publicly available codes and tools used to analyze the data are reported and referenced in the Methods section.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

Materials availability—This study did not generate new unique reagents.

EXPERIMENTAL MODEL DETAILS

Animal Husbandry—All animals were maintained on a normal chow diet with 15% fat (Ssniff GmbH), fed *ad libitum* with free access to water (HCl acidified, pH 2.5–3) under controlled humidity and temperature with a 12-hour light and 12-hour dark cycle. High fat diet fed mice were fed with 60% kcal% fat diet (Research Diet) for 3 days or 4 weeks. All animal studies were performed with the approval of the local authorities in Germany (Regierungspräsidium Freiburg, Germany) under license number 35-9185.81/G-16/120, or approved by the Institutional Animal Care and Use Committee at the Van Andel Research Institute, Grand Rapids, MI, USA under the animal use protocol number 21-08-023.

Genetically modified Mice—The CD24 knockout[64] mice were kindly provided by Sherri L. Christian. Breeding pairs of Ins1-cre[65] (B6(Cg)-Ins1tm1.1(cre)$^{Thor/J}$ were purchased from Jackson laboratories. Eed$^{fl/fl}$, Kdm6b$^{fl/fl}$, and YFP-reporter (B6.129X1-Gt(ROSA) 26Sortm1(EYFP)Cos/J) transgenic mouse lines (C57BL6/J) were kindly provided by Stuart Orkin and Thomas Boehm, respectively. To generate β-cell reporter mice with Eed deficiency, Eed-floxed animals were crossed with YFP harboring-Ins1-cre positive animals. All mice were backcrossed for >10 generations before any phenotyping was initiated. Experimental mice were all males unless otherwise stated. Age of the mice used for individual experiments are specified accordingly.

METHODS DETAILS

Islet Isolation—Adult pancreata were perfused through the common bile duct using a 30-gauge needle with Collagenase 4 solution (dissolved in 1x HBSS GIBCO, 10 mM HEPES at a concentration of 1 mg per mL); this step was excluded for neonatal islet isolation. The pancreata were then dissected and transferred into a 50 mL falcon tube containing and incubated 30 min in 2mL of the same collagenase solution. Islets were purified as previously described[66]. Breifly, collagenase-digested pancreata were dissociated by shaking. 40 mL of quenching buffer (HBSS, 0.5% BSA, 25mM HEPES) was added and the tubes were centrifuged at 190g, for 2 min at 4°C twice. Non-digested tissue was filtered by 500 uM mesh. The wash through was then filtered again by a 70 uM, keeping the islets on top of the mesh. Islets were then transferred to a petri dish containing pre-warmed complete media (RMPI-1640 containing 11 mM glucose, 10% FBS, 0.1% [penicillin, streptomycin, gentamicin, amphotericin B]; Thermo Fisher). The isolated islets were hand-picked and cultured in complete media, and maintained at 37°C in a 5% CO_2 environment.

Human pancreatic islets.—Human islets were obtained from the Alberta Diabetes Institute IsletCore[67] and the Clinical Islet Laboratory at the University of Alberta, respectively. They were isolated from pancreata of cadaveric organ donors in accordance with the local Institutional Ethical Approvals (Pro00013094). Islets were cultured in CMRL-1066 medium containing 5 mmol/L glucose, 100 units/mL penicillin, 100 μg/mL streptomycin, 2 mM Glutamax, and 10% FCS (Invitrogen) in a humid environment containing 5% CO_2.

Islet dispersion and single cell labeling for FACS—For islet dispersion, islets were incubated in accutase for 4 min at 37°C, then gently pipetted 20 times using a 1 mL pipette. Immediately after, single cell suspensions were examined under the microscope, and validated single cell suspensions were washed with 2 mL of ice cold FACS buffer (PBS, 0.5% BSA, 5 mM EDTA) or ice-cold PBS in case of subsequent fixable viability labeling (5 min zombie dye on ice). Cell surface CD24 (mouse: Thermo Fisher 48-0242-82, human: Biologend, 311122, 1:250), CD45 (Thermo Fisher, 11-0451-82, 1:200), and CD31 (BD pharmingen, 558738 1:200) labeling was done for 30 min on ice (diluted in FACS buffer). After washing, cells were fixed in 1% methanol free formaldehyde (Thermo fisher, 28906, 1 mL; diluted in RPMI; freshly made) for 15 min, the reaction was quenched with glycine (final concentration of 125 mM), and cells were washed with 1 mL of FACS buffer. Intracellular labeling of insulin (sc-8033, 1:100), glucagon (sc-51459, 1:100), somatostatin (sc-55565, 1:100), pancreatic polypeptide (sc-514155, 1:100), MKI67 (47-5698-82, 1:200), TOM20 (Abcam, 1:500), and chromatin labeling using the sigma ‘Mix-and-Stain’-conjugated H3K27me3 (Origene, TA347154 CF555), H3K4me3 (C15410003, CF405), H3K36me3 (C15410192, CF405), H3K9me3 (C15410193, CF488), H3K27ac (C15410196, CF488) at final concentrations of 5 μg/mL, was done in permeabilization buffer (eBioscience, 00-8333-56). For data in Figure 1A, H3K27me3 was labeled together with either H3K9me3 and H3K36me3, or H3K27ac and H3K4me3. The presented data in Figure S1F show additional labeling with H3K27me3 (C15410195, Diagenode). Unless stated differently, insulin positive β-cells were analyzed and sorted while excluding glucagon, somatostatin, pancreatic polypeptide, CD31, and CD45 positive

cells that were included in a ‘dump channel’; 488, which contained all of the antibodies that are specific for the unwanted cells. Washing between steps was determined differently; live cells were centrifuged at 190g, fixed cells at 350g, and fixed-permeabilized cells were centrifuged at 500g, all for 4 min at 4°C. For experiments with subsequent extraction of RNA, all the steps were done in the presence of RNase inhibitor (recombinant RNasein, Promega, N2511) diluted 1:4,000 for the washing steps or 1:400 for the incubation while labeling. FACS sorted cells were also kept in tubes containing 1:400 RNase inhibitor and were snap frozen and stored at −80°C.

***SCAN*-seq—**The new multi-modal *SCAN*-seq method was derived from the CEL-Seq2 method[5,68]. Insulin positive cells were index-sorted into 384 well plates containing 384 unique barcodes (Table S2). Single cells were sorted in 384-well plates (Bio-Rad Laboratories, HSP3801) containing lysis buffer and mineral oil (Sigma, M8410) using BD FACS Aria FUSION. The sorter was run on single-cell sort mode with index sorting. Doublets were excluded using pulse geometry gates (FSC-W × FSC-H and SSC-W × SSC-H). Importantly, cells from all conditions/biological replicates were equally distributed into wells of all sorted plates from the same experiment to enable optimal batch correction in case of evident plate bias in the transcriptional data. After sorting, the plates were centrifuged for 2 min at 2,200g at 4°C, snap-frozen in liquid nitrogen, and stored at −80°C for up to two weeks. 160 nL of reverse transcription reaction mix and 2.2 mL of second strand reaction mix was used to convert RNA into cDNA. cDNA from 384-cells was pooled together before the clean-up and *in vitro* transcription, generating one library from one 384-well plate. 0.8 mL of AMPure/RNAClean XP beads (Beckman Coulter GmbH, Germany) per 1 mL of sample were used during all the purification steps, including library cleanup. Libraries were sequenced on a single lane (pair-end multiplexing run, 100 bp read length) of an Illumina HiSeq system targeting 200,000 reads per cell.

Bulk-cell RNA-seq—Total RNA from 1,000 H3K27me3 HI/LO β-cells or from 50,000 sorted βHI and βLO cells was extracted using the miRNeasy FFPE Kit (QIAGEN, 217504), followed by the NEBNext® Single Cell/Low Input RNA Library Prep Kit for Illumina® (E6420L). Library fragments of 350 ± 20 bp were obtained, and the quality was assessed using a Fragment Analyzer (Advanced Analytical). Barcoded libraries were subjected to 70 bp pair-end sequencing on the Illumina HiSeq 2000.

ChIP-seq—Chromatin from snap-frozen pellets of the sorted β_{HI} and β_{LO} cells (3 biological replicates each) were prepared using the NEXSON procedure[69] to a 100–800 bp fragment size distribution. We controlled sheared chromatin size distribution and cell number by adding 2 mL of 20 mg/mL Proteinase K to a small chromatin aliquot (5 mL) of each sample. Volumes were adjusted to 20 mL using buffer EB (Qiagen). Samples were then reverse-crosslinked by incubating at 50°C for 30 min, followed by incubation at 65°C for 30 min. DNA concentration was measured using a Qubit dsDNA HS assay, and we estimated cell numbers by assuming one mouse diploid cell contains ~ 6.6 pg of DNA. Samples were then purified using Qiagen MinElute columns, and DNA fragment size distribution was checked by capillary electrophoresis (Agilent Fragment Analyzer).

Before ChIP, all chromatins were normalized to the same cell number in shearing buffer. Normalized chromatins were diluted 1:2 in 1X Buffer iC1 (supplemented with protease inhibitor cocktail) from the Diagenode iDeal ChIP-seq kit for histones (C01010173). Each chromatin sample of ~ 6,000 cells was incubated with 1 μg of anti-H3K27me3 antibody (Diagenode, C15410195, lot. A1811-001P). ChIP was performed using the automated platform SX-8G IP-Star platform (Diagenode) under the program "ChIP indirect method". The antibody-chromatin incubation lasted 10 hours, followed by 3 hours of bead incubation (protein-A conjugated), and 5 min beads washes. Ten percent of the original chromatin was used as input. After DNA elution, ChIP and input samples were de-crosslinked and purified using Qiagen MinElute columns.

Libraries were prepared on an automated liquid handler (Biomek i7) using the NEBNext Ultra II DNA library preparation kit (NEB, E7645) according to the manufacturer's instructions, and without size selection. Libraries were sequenced paired-end on the Illumina NovaSeq platform.

RELACS—The H3K4me3 ChIP-seq was performed using RELACS as previously described[70]. Briefly, 50,000 β_{HI} and β_{LO} cells were thawed in RELACS lysis buffer (10 mM Tris-HCl [pH 8], 10 mM NaCl, 0.2% Igepal, 1× Protease inhibitor cocktail) and the nuclei were isolated by sonication using the NEXSON procedure[69]. To digest the chromatin, 25 μL of 10× CutSmart buffer (NEB), 2.5 μL 100× Protease inhibitor cocktail and 1 μL of CviKI-1 (5 U/100,000 nuclei, NEB R0710S) were added. The digestion reaction was incubated overnight at 20°C. End repair and A-tailing was performed, and customized adapters were ligated to the fragments. Once barcoded, the samples were pooled together. Chromatin was then sheared by sonication (Covaris E220, MicroTubes, 5 min, peak power 105, duty factor 2, cycles burst 200). This chromatin was used for automated ChIP (Diagenode, C15410003) with the IP-Star Diagenode system. IPs and Inputs were de-crosslinked, DNA was purified, and libraries were prepared using the NEB Ultra II DNA Library Prep Kit for Illumina (E7645S and E6440) following the manufacturer's instructions. Integrity and size distribution of the samples was assessed before and after library preparation by running on Fragment Analyzer (Advanced Analytical).

DNA methylation array—Genomic DNA was extracted from fixed and sorted β_{HI} or β_{LO} using the Zymo Research Quick-DNA Microprep Plus Kit (Zymo Research) according to manufacturer's instructions. Purified DNA was quantified by Qubit fluorimetry (Life Technologies), and bisulfite converted using the Zymo EZ DNA Methylation Kit (Zymo Research) following the manufacturer's protocol with the specified modifications for the Illumina Infinium Methylation Assay. After conversion, the bisulfite-converted DNA was purified using the Zymo-Spin binding columns and eluted in Tris buffer. Following elution, bisulfite-converted DNA was processed through the Illumina mouse methylation array protocol. The bisulfite-converted DNA samples were first processed using the Infinium HD FFPE DNA Restore kit workflow. To perform the Infinium assay, converted DNA was denatured with NaOH, amplified, and hybridized to the Infinium bead chip. An extension reaction was performed using fluorophore-labeled nucleotides per the manufacturer's protocol. Array BeadChips were scanned on the Illumina iScan system, and signals were

assigned by using Illumina Genome Studio v2011.1 software to produce IDAT files. CpG probe selection[71] included an array of target and random controls sequences. The DNA methylation array analysis was performed using SeSAMe[72] and its wrapper pipeline SeSAMeStr (10.5281/zenodo.7510575). Nine biological replicates of βHI and βLO were compared. Data pre-processing and quality controls were performed using SeSAMe default parameters and the pre-processing code 'TQCDPB'. All samples showed a detection rate > 93% and no dye bias. PCA analysis of beta values was performed within the SeSAMeStr pipeline using the R function 'prcomp'. In all differential analysis, the effect size cutoff was set to 0.05 (i.e., 5% differential DNA methylation) and the p-value cutoff was < 0.05. Further visualization of SeSAMe/SeSAMeStr output data was perform in R using Rstudio.

Re-aggregation of islet spheroids—Islets were isolated as described above from the β-cell reporter mice that were generated by crossing Ins1-Cre mice[65] with B6.129X1-Gt(ROSA)26 Sortm1(EYFP)Cos/J floxed-stop-YFP. After overnight recovery, islets were dispersed as described above to achieve single-cell suspensions for CD24 labeling. Live, $CD24^{-}$, or $CD24^{+}$ YFP^{+} β-cells were sorted into tubes containing 1x HBSS (GIBCO) 0.5% w/v BSA (Serva) and 24 mM HEPES (Sigma). Sorted cells were centrifuged at 200g for 4 min at 4°C, and resuspended in mouse islet media (see islet isolation section above) at the concentration of 10 cells/μL. Cells were then distributed into 96 well plates (U bottom - Nunclon Sphera) to achieve 2000 cells/well. To determine spheroid formation kinetics, the plates were incubated inside a real-time quantitative cell imaging system (Incucyte®) that was set to image cells every 15 minutes for 3 days.

Microscopy and Image Quantification—Islets, whole, dispersed, or sorted β-cell cells were derived from at least three animals for each condition. Samples were stained live or fixed with 1% methanol-free formaldehyde. Cells were permeabilized with 1x permeabilization buffer (00-833-56 Invitrogen), and stained for the indicated antigens/proteins. Single cells were stained for 30 min on ice while whole islets were stained by rotating overnight at 4°C. Images were acquired on a LSM880 confocal microscope (ZEISS) using the Airyscan super-resolution (SR) mode. An identical threshold was applied to all images from the same channel to exclude background signals. Staining intensity was quantified using Imaris version 9.3.1 in a blinded manner.

Oral glucose tolerance test—For the oral glucose tolerance test (OGTT), mice were fasted for 6 hours (8:00–14:00), after which basal blood glucose was measured. Mice were given glucose (1 g/kg) by oral gavage. Blood glucose levels were measured using a OneTouch Vita blood glucose meter at 0, 15, 30, 60, and 90 min after glucose.

Measurements of Oxygen consumption rate (OCR) and extracellular acidification rate (ECAR)—An XF96e Extracellular Flux analyzer (Seahorse Biosciences) was used to determine the bioenergetic profile of single monotypic pseudo-islets. Before the assay, monotypic pseudo-islets of ~2000 cells each were incubated in unbuffered DMEM (Seahorse Biosciences). Then, single spheroids were hand-picked and added into the middle of a 96-spheroid ploy-L-lysine coated microplate (Seahorse biosciences). After two 2 min basal measurements, glucose was injected into the media

(16.7mM end concentration) and the oxygen consumption and extracellular acidification rates were measured 4 times for 2 min each time. Between every measurement, a 5 second mixing step was followed by 5 second waiting step. Wells with readouts lower than background measurements were excluded from further analysis.

Glucose stimulated insulin secretion—Single, re-aggregated, cell-type specific pseudo-islets or overnight recovered whole islets were pre-incubated for 30 min in pre-equilibrated Krebs-Ringer bicarbonate buffer (KRB; 115 mM NaCl, 4.7 mM KCl, 2.6 mM $CaCl_2$ $2H_2O$, 1.2 mM KH_2PO_4, 1.2 mM $MgSO_4$ $7H_2O$, 10 mM HEPES, 0.5% bovine serum albumin, pH 7.4) containing 2.8 mM glucose. Single β_{HI} or β_{LO} pseudo-islets were individually transferred into a V-shaped well of a 96 well plate, each well containing 50 μL of 2.8mM glucose-KRB. Pseudo-islets were incubated at 37°C for 1 hour for basal secretion. Then, individual pseudo-islets were collected, washed in PBS, and incubated with16.7 mM glucose in KRB- for 1 hour.

Five whole islets from all sizes were added to each well of a 48 well plate that contained either 2.8mM glucose (basal) or 16.7mM glucose (stimulated) KRB. Islets were incubated at 37°C for 1 hour, after which islet supernatants were collected and put on ice. Supernatants were centrifuged (2000g 5 min 4°C), transferred to new 96 well plate, and stored at −20°C for later insulin measurements using ultrasensitive insulin ELISA (Mercodia).

Mitochondrial Membrane Potential by FACS Analysis—Ins1-YFP islets were allowed to recover overnight, dissociated as described above, and washed twice with Krebs solution containing 4 mM glucose. For detection of the mitochondrial membrane potential, dissociated islet cells were incubated with 10 nM of the fluorescent probe TMRM solution (Life Technologies) containing 4 mM glucose. Cells were washed once with PBS, scored by FACS using BD Symphony, and analyzed by Flowjo.

Mitochondrial DNA quantification—Mitochondrial and genomic DNA was isolated from FACS sorted β_{HI} and β_{LO} cells according to the manufacturer's instructions. (Absolute Mouse Mitochondrial DNA Copy Number Quantification qPCR Assay Kit (AMMQ) Catalog #M8948).

QUANTIFICATION AND STATISTICAL ANALYSIS

Bulk RNA-seq analysis—RNA-seq was performed with at least three independent biological replicates. Raw sequences from the biological replicates were aligned and mapped against mouse genome version GRCm38/mm10 with the snakePipes2[73]. Differential expression analysis was performed with DEseq2[74]. Genes with counts <2 were excluded from differential analysis. Differential genes were called with an FDR threshold of 0.05 and a fold change of 1.33. After QC and exclusion of lowly expressed genes (>2 counts), differential expression of the raw counts was performed using DESeq2 v1.34.1. Samples were batch-corrected using Limma, and normalized count matrices were inspected using PCA. Gene Set Enrichment Analysis (GSEA) of DE results was performed with the fgsea R-package. Enrichment maps were generated in Cytoscape[75]. Motif enrichment analysis on β_{HI}-specific TSS was performed using HOMER v4.11[76] function 'findMotifsGenome.pl'

with '-size given-mask' options, and using the transcriptionally unchanged TSS between βHI and β_{LO} cells as background control.

ChIP-sequencing and RELACS analysis—Mouse H3K27me3 ChIP-seq (n=3 biological replicates for each β-cell type) and H3K4me3 RELACS (n=1 biological replicate for each β-cell type) data were processed and analyzed using snakePipes 2.5[73] 'DNA-mapping' and 'ChIP-seq' pipelines. Reads were trimmed and quality controlled using Cutadapt[77] and FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/), respectively. Mouse reads were mapped with Bowtie2[78] on GRCm38/mm10 genome. High quality (MAPQ>3) and properly paired mapped reads were filtered for optical/PCR duplicates using samtools view[79]. Coverage tracks for visualization in IGV or UCSC genome browsers were created with the DeepTools[80] v3.3.2 command 'bamCoverage', and normalized to sequencing depth. Spearman correlation matrices of H3K27me3 signal over the whole genome were generated with the DeepTools commands 'multiBigwigSummary' and 'plotCorrelation'. H3K27me3 peaks and broad domains were called on each single replicate using MACS2 v2.2.6[81] in 'broad' mode and epic2[82] v0.0.41 (bin size = 1000, gaps allowed = 10), respectively. PCA on counts over all identified H3K27me3 peaks among all β_{HI} and β_{LO} replicates was performed in R using the command 'prcomp'. Annotation of identified H3K27me3 peaks according to genomic regions and quantification of tag counts over specific regions were performed using the 'annotatePeaks.pl' commands under the HOMER v4.11 suite[76]. Differential H3K27me3 enrichment over annotated TSS was performed by running DESeq2[74] v1.34.0 on counts tables from biological triplicates. β_{HI} and β_{LO} cell-specific TSS were those TSS with H3K27me3 log2(fold change) > 0.2/< −0.2 and adjusted *p*-value < 0.1. Heatmap visualizations and profile plots of H3K27me3 and H3K4me3 signals over specific regions were generated using the DeepTools commands 'computeMatrix', 'plotHeatmap' and 'plotProfile'. Chromatin state annotations were based on the previously reported segmentation of the genome from whole islet's epigenetic landscapes[5]. Overlaps between annotated chromatin states and genomic regions of interest were found using the 'intersectBed' command from bedtools[83]. Further analyses (i.e., boxplots, scatterplots) were performed in a R environment using RStudio.

Quantification of transcript abundance in *SCAN*-seq—Paired end reads were processed using scRNA-seq function in snakePipes[73] (v1.3.0). Briefly, cell barcodes and UMI's from read 1 were moved into the header of read2 that was then trimmed for adaptors and polyAs using cut adapt (v2.1). The subsequent alignment to the GRCm38/mm10 reference genome was performed using STAR (v2.4.2a). Raw counts were extracted using feature counts (v1.6.4) and gene annotation version M9 of gencode, and pseudogenes were removed.

scRNA-seq and SCAN-seq analysis—These data were analyzed using the Seurat v4 algorithm[84], including standard preprocessing. In brief, the cell filtration threshold was set to unique feature counts >700 and >40% mitochondrial genes for *SCAN*-seq, or 1000 and >20% mitochondrial genes for the published droplet-based data sets. After QC filtering, we normalized the feature expression measurements for each cell to the total expression, multiplied the result by 10000, and log-transformed the product. Highly variable transcripts

were identified using the Seurat4 FindVaribleFeatures function. A batch correction was applied using the vars.to.regress option to provide an equal weight in downstream analysis, and buffer the noise of highly-expressed genes. Then, linear dimensional reduction was applied on the scaled data. To explore feature expression similarities and define cell populations, we generated the UMAP using the first 10 principal components/dimensions. Index-sorting files were used to integrate FACS parameters with the Seurat object using the CreateAssayObject function. Cluster trees were generated using the 'clustree' package. To visualize the expression of groups of genes (Figure S2), a sum of the Seurat *scaled.data* was first calculated. Senescence-associated genes (*'Cellular senescence'*) were called from the mouse Gene Ontology. Genes from Dorrell's[16] four β-cell subsets were taken from their Figure 4 ('top genes' in Figure S2K) or from their supplemental gene list (All DE genes in Figure S2M). The trajectories in Figure 6B and Figure 7D were generated using slingshot to connect the centroids of each cluster. The code to preprocess and integrate FACS data with the Seurat object is available upon request. The custom GSEA in Figure S7 was based on β_{HI}/β_{LO} signature genes (Figure 3C). The mean expression (z-score) for the two gene sets was calculated, then the magnitude and direction of differential signatures was determined by calculating the difference in expression between the two gene sets. The cells were then ranked by difference z-score. All analyses were performed in a R environment using RStudio.

Analysis of published single-cell/nucleus sequencing—Mouse and human single-cell count matrices from published islet single-cell sequencing datasets were obtained from[39–42,48]. scRNA-seq data was preprocessed as described above, except that the cell filtration threshold was set to unique feature counts >1000 and >20%. To compare β cells isolated from humans with or without T2D[48], we performed data integration[84] using sex as a covariate. 'FACS' cell data were available for 11 donors without diabetes and 7 donors with T2D. Single nucleus ATACseq data (GSE160472) was obtained from Chiou et al.[19]. Processed and de-multiplexed fastq files were reformatted for the ArchR pipeline, and reads were aligned to the reference genome h38 using Chromap with 'preset atac'. The resulting .sam files were converted to Arrow files with the createArrowFiles function, using mints = 4 and minFrags = 1000. After QC, we used replicate 2 and replicate 3 for initial clustering. We then mapped the sum of expression of H3K27me3 marked genes that were previously annotated for the mouse[5] or human[52] data sets.

Other statistical analysis—We used GraphPad prism v8 for statistical tests and comparisons, as indicated in the respective figure legends. *p*-value < 0.05 was considered statistically significant.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We would like to thank Prof. Thomas Jenuwein for stimulating interactions and enduring research support. We acknowledge excellent technical support of the MPI-IE and VAI core services (Imaging, flow cytometry, deep sequencing, and bioinformatics) and in particular, Sebastian Hobitz, Konrad Schuldes, Laura Arrigoni, Chiara

Bella, Leila Rabanni. We thank Dr. Sherri L. Christian for providing the CD24 knockout mice. Human islets for research were provided by the Alberta Diabetes Institute IsletCore (www.bcell.org/adi-isletcore) and the Clinical Islet Laboratory at the University of Alberta in Edmonton with the assistance of the Human Organ Procurement and Exchange (HOPE) program, Trillium Gift of Life Network (TGLN), and other Canadian organ procurement organizations. Islet isolation was approved by the Human Research Ethics Board at the University of Alberta (Pro00013094). All donors' families gave informed consent for the use of pancreatic tissue in research. We thank Dr. Darrell Chandler for critical evaluation of the manuscript. This work was supported by funding from the MPG, the ERC, Van Andel Institute, EFSD and NIH awards R21HG011964 and 1R01HG012444.

References

1. Stolovich-Rain M, Enk J, Vikesa J, Nielsen FC, Saada A, Glaser B, and Dor Y (2015). Weaning triggers a maturation step of pancreatic beta cells. Dev Cell 32, 535–545. 10.1016/j.devcel.2015.01.002. [PubMed: 25662175]
2. Salinno C, Cota P, Bastidas-Ponce A, Tarquis-Medina M, Lickert H, and Bakhti M (2019). beta-Cell Maturation and Identity in Health and Disease. Int J Mol Sci 20. 10.3390/ijms20215417.
3. Arrojo EDR, Lev-Ram V, Tyagi S, Ramachandra R, Deerinck T, Bushong E, Phan S, Orphan V, Lechene C, Ellisman MH, and Hetzer MW (2019). Age Mosaicism across Multiple Scales in Adult Tissues. Cell Metab 30, 343–351 e343. 10.1016/j.cmet.2019.05.010. [PubMed: 31178361]
4. Cnop M, Igoillo-Esteve M, Hughes SJ, Walker JN, Cnop I, and Clark A (2011). Longevity of human islet alpha- and beta-cells. Diabetes Obes Metab 13 Suppl 1, 39–46. 10.1111/j.1463-1326.2011.01443.x. [PubMed: 21824255]
5. Lu TT, Heyne S, Dror E, Casas E, Leonhardt L, Boenke T, Yang CH, Sagar, Arrigoni L, Dalgaard K, et al. (2018). The Polycomb-Dependent Epigenome Controls beta Cell Dysfunction, Dedifferentiation, and Diabetes. Cell Metab 27, 1294–1308 e1297. 10.1016/j.cmet.2018.04.013. [PubMed: 29754954]
6. Dhawan S, Georgia S, Tschen SI, Fan G, and Bhushan A (2011). Pancreatic beta cell identity is maintained by DNA methylation-mediated repression of Arx. Dev Cell 20, 419–429. 10.1016/j.devcel.2011.03.012. [PubMed: 21497756]
7. Hellerstrom C, Petersson B, and Hellman B (1960). Some properties of the B cells in the islet of Langerhans studied with regard to the position of the cells. Acta Endocrinol (Copenh) 34, 449–456. 10.1530/acta.0.xxxiv0449. [PubMed: 13852224]
8. Salomon D, and Meda P (1986). Heterogeneity and contact-dependent regulation of hormone secretion by individual B cells. Exp Cell Res 162, 507–520. 10.1016/0014-4827(86)90354-x. [PubMed: 3510882]
9. Kiekens R, In 't Veld P, Mahler T, Schuit F, Van De Winkel M, and Pipeleers D (1992). Differences in glucose recognition by individual rat pancreatic B cells are associated with intercellular differences in glucose-induced biosynthetic activity. J Clin Invest 89, 117–125. 10.1172/JCI115551. [PubMed: 1729264]
10. Johnston NR, Mitchell RK, Haythorne E, Pessoa MP, Semplici F, Ferrer J, Piemonti L, Marchetti P, Bugliani M, Bosco D, et al. (2016). Beta Cell Hubs Dictate Pancreatic Islet Responses to Glucose. Cell Metab 24, 389–401. 10.1016/j.cmet.2016.06.020. [PubMed: 27452146]
11. Salem V, Silva LD, Suba K, Georgiadou E, Neda Mousavy Gharavy S, Akhtar N, Martin-Alonso A, Gaboriau DCA, Rothery SM, Stylianides T, et al. (2019). Leader beta-cells coordinate Ca(2+) dynamics across pancreatic islets in vivo. Nat Metab 1, 615–629. 10.1038/s42255-019-0075-2. [PubMed: 32694805]
12. Bader E, Migliorini A, Gegg M, Moruzzi N, Gerdes J, Roscioni SS, Bakhti M, Brandl E, Irmler M, Beckers J, et al. (2016). Identification of proliferative and mature beta-cells in the islets of Langerhans. Nature 535, 430–434. 10.1038/nature18624. [PubMed: 27398620]
13. van der Meulen T, Mawla AM, DiGruccio MR, Adams MW, Nies V, Dolleman S, Liu S, Ackermann AM, Caceres E, Hunter AE, et al. (2017). Virgin Beta Cells Persist throughout Life at a Neogenic Niche within Pancreatic Islets. Cell Metab 25, 911–926 e916. 10.1016/j.cmet.2017.03.017. [PubMed: 28380380]
14. Rui J, Deng S, Arazi A, Perdigoto AL, Liu Z, and Herold KC (2017). beta Cells that Resist Immunological Attack Develop during Progression of Autoimmune Diabetes in NOD Mice. Cell Metab 25, 727–738. 10.1016/j.cmet.2017.01.005. [PubMed: 28190773]

15. Segerstolpe A, Palasantza A, Eliasson P, Andersson EM, Andreasson AC, Sun X, Picelli S, Sabirsh A, Clausen M, Bjursell MK, et al. (2016). Single-Cell Transcriptome Profiling of Human Pancreatic Islets in Health and Type 2 Diabetes. Cell Metab 24, 593–607. 10.1016/j.cmet.2016.08.020. [PubMed: 27667667]
16. Dorrell C, Schug J, Canaday PS, Russ HA, Tarlow BD, Grompe MT, Horton T, Hebrok M, Streeter PR, Kaestner KH, and Grompe M (2016). Human islets contain four distinct subtypes of beta cells. Nat Commun 7, 11756. 10.1038/ncomms11756. [PubMed: 27399229]
17. Xin Y, Dominguez Gutierrez G, Okamoto H, Kim J, Lee AH, Adler C, Ni M, Yancopoulos GD, Murphy AJ, and Gromada J (2018). Pseudotime Ordering of Single Human beta-Cells Reveals States of Insulin Production and Unfolded Protein Response. Diabetes 67, 1783–1794. 10.2337/db18-0365. [PubMed: 29950394]
18. Salinno C, Buttner M, Cota P, Tritschler S, Tarquis-Medina M, Bastidas-Ponce A, Scheibner K, Burtscher I, Bottcher A, Theis FJ, et al. (2021). CD81 marks immature and dedifferentiated pancreatic beta-cells. Mol Metab 49, 101188. 10.1016/j.molmet.2021.101188. [PubMed: 33582383]
19. Chiou J, Zeng C, Cheng Z, Han JY, Schlichting M, Miller M, Mendez R, Huang S, Wang J, Sui Y, et al. (2021). Single-cell chromatin accessibility identifies pancreatic islet cell type- and state-specific regulatory programs of diabetes risk. Nat Genet 53, 455–466. 10.1038/s41588-021-00823-0. [PubMed: 33795864]
20. Wang YJ, and Kaestner KH (2019). Single-Cell RNA-Seq of the Pancreatic Islets--a Promise Not yet Fulfilled? Cell Metab 29, 539–544. 10.1016/j.cmet.2018.11.016. [PubMed: 30581120]
21. Mawla AM, and Huising MO (2019). Navigating the Depths and Avoiding the Shallows of Pancreatic Islet Cell Transcriptomes. Diabetes 68, 1380–1393. 10.2337/dbi18-0019. [PubMed: 31221802]
22. Szabat M, Luciani DS, Piret JM, and Johnson JD (2009). Maturation of adult beta-cells revealed using a Pdx1/insulin dual-reporter lentivirus. Endocrinology 150, 1627–1635. 10.1210/en.2008-1224. [PubMed: 19095744]
23. Farack L, Golan M, Egozi A, Dezorella N, Bahar Halpern K, Ben-Moshe S, Garzilli I, Toth B, Roitman L, Krizhanovsky V, and Itzkovitz S (2019). Transcriptional Heterogeneity of Beta Cells in the Intact Pancreas. Dev Cell 48, 115–125 e114. 10.1016/j.devcel.2018.11.001. [PubMed: 30503750]
24. Enge M, Arda HE, Mignardi M, Beausang J, Bottino R, Kim SK, and Quake SR (2017). Single-Cell Analysis of Human Pancreas Reveals Transcriptional Signatures of Aging and Somatic Mutation Patterns. Cell 171, 321–330 e314. 10.1016/j.cell.2017.09.004. [PubMed: 28965763]
25. Dagogo-Jack I, and Shaw AT (2018). Tumour heterogeneity and resistance to cancer therapies. Nat Rev Clin Oncol 15, 81–94. 10.1038/nrclinonc.2017.166. [PubMed: 29115304]
26. Qiu WL, Zhang YW, Feng Y, Li LC, Yang L, and Xu CR (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. 10.1016/j.cmet.2017.04.003. [PubMed: 28467935]
27. Cigliola V, Thorel F, Chera S, and Herrera PL (2016). Stress-induced adaptive islet cell identity changes. Diabetes Obes Metab 18 Suppl 1, 87–96. 10.1111/dom.12726. [PubMed: 27615136]
28. Xiao Y, Karnati S, Qian G, Nenicu A, Fan W, Tchatalbachev S, Holand A, Hossain H, Guillou F, Luers GH, and Baumgart-Vogt E (2012). Cre-mediated stress affects sirtuin expression levels, peroxisome biogenesis and metabolism, antioxidant and proinflammatory signaling pathways. PLoS One 7, e41097. 10.1371/journal.pone.0041097. [PubMed: 22829911]
29. Rosenbaum M, Andreani V, Kapoor T, Herp S, Flach H, Duchniewicz M, and Grosschedl R (2014). MZB1 is a GRP94 cochaperone that enables proper immunoglobulin heavy chain biosynthesis upon ER stress. Genes Dev 28, 1165–1178. 10.1101/gad.240762.114. [PubMed: 24888588]
30. Estall JL, and Screaton RA (2020). Of Mice and Men, Redux: Modern Challenges in beta Cell Gene Targeting. Endocrinology 161. 10.1210/endocr/bqaa078.
31. Stephens AD, Banigan EJ, and Marko JF (2019). Chromatin's physical properties shape the nucleus and its functions. Curr Opin Cell Biol 58, 76–84. 10.1016/j.ceb.2019.02.006. [PubMed: 30889417]

32. Skinner BM, and Johnson EE (2017). Nuclear morphologies: their diversity and functional relevance. Chromosoma 126, 195–212. 10.1007/s00412-016-0614-5. [PubMed: 27631793]
33. Rehimi R, Nikolic M, Cruz-Molina S, Tebartz C, Frommolt P, Mahabir E, Clement-Ziza M, and Rada-Iglesias A (2016). Epigenomics-Based Identification of Major Cell Identity Regulators within Heterogeneous Cell Populations. Cell Rep 17, 3062–3076. 10.1016/j.celrep.2016.11.046. [PubMed: 27974217]
34. Boettiger AN, Bintu B, Moffitt JR, Wang S, Beliveau BJ, Fudenberg G, Imakaev M, Mirny LA, Wu CT, and Zhuang X (2016). Super-resolution imaging reveals distinct chromatin folding for different epigenetic states. Nature 529, 418–422. 10.1038/nature16496. [PubMed: 26760202]
35. Geyer PK, Vitalini MW, and Wallrath LL (2011). Nuclear organization: taking a position on gene expression. Curr Opin Cell Biol 23, 354–359. 10.1016/j.ceb.2011.03.002. [PubMed: 21450447]
36. Eskeland R, Leeb M, Grimes GR, Kress C, Boyle S, Sproul D, Gilbert N, Fan Y, Skoultchi AI, Wutz A, and Bickmore WA (2010). Ring1B compacts chromatin structure and represses gene expression independent of histone ubiquitination. Mol Cell 38, 452–464. 10.1016/j.molcel.2010.02.032. [PubMed: 20471950]
37. Berthault C, Staels W, and Scharfmann R (2020). Purification of pancreatic endocrine subsets reveals increased iron metabolism in beta cells. Mol Metab 42, 101060. 10.1016/j.molmet.2020.101060. [PubMed: 32763423]
38. Ziegenhain C, Vieth B, Parekh S, Reinius B, Guillaumet-Adkins A, Smets M, Leonhardt H, Heyn H, Hellmann I, and Enard W (2017). Comparative Analysis of Single-Cell RNA Sequencing Methods. Mol Cell 65, 631–643 e634. 10.1016/j.molcel.2017.01.023. [PubMed: 28212749]
39. Pineros AR, Gao H, Wu W, Liu Y, Tersey SA, and Mirmira RG (2020). Single-Cell Transcriptional Profiling of Mouse Islets Following Short-Term Obesogenic Dietary Intervention. Metabolites 10. 10.3390/metabo10120513.
40. Sachs S, Bastidas-Ponce A, Tritschler S, Bakhti M, Bottcher A, Sanchez-Garrido MA, Tarquis-Medina M, Kleinert M, Fischer K, Jall S, et al. (2020). Targeted pharmacological therapy restores beta-cell function for diabetes remission. Nat Metab 2, 192–209. 10.1038/s42255-020-0171-3. [PubMed: 32694693]
41. Balboa D, Barsby T, Lithovius V, Saarimaki-Vire J, Omar-Hmeadi M, Dyachok O, Montaser H, Lund PE, Yang M, Ibrahim H, et al. (2022). Functional, metabolic and transcriptional maturation of human pancreatic islets derived from stem cells. Nat Biotechnol. 10.1038/s41587-022-01219-z.
42. Avrahami D, Wang YJ, Schug J, Feleke E, Gao L, Liu C, Consortium H, Naji A, Glaser B, and Kaestner KH (2020). Single-cell transcriptomics of human islet ontogeny defines the molecular basis of beta-cell dedifferentiation in T2D. Mol Metab 42, 101057. 10.1016/j.molmet.2020.101057. [PubMed: 32739450]
43. Boyer LA, Plath K, Zeitlinger J, Brambrink T, Medeiros LA, Lee TI, Levine SS, Wernig M, Tajonar A, Ray MK, et al. (2006). Polycomb complexes repress developmental regulators in murine embryonic stem cells. Nature 441, 349–353. 10.1038/nature04733. [PubMed: 16625203]
44. Margueron R, and Reinberg D (2011). The Polycomb complex PRC2 and its mark in life. Nature 469, 343–349. 10.1038/nature09784. [PubMed: 21248841]
45. Creed S, and McKenzie M (2019). Measurement of Mitochondrial Membrane Potential with the Fluorescent Dye Tetramethylrhodamine Methyl Ester (TMRM). Methods Mol Biol 1928, 69–76. 10.1007/978-1-4939-9027-6_5. [PubMed: 30725451]
46. Xie H, Xu J, Hsu JH, Nguyen M, Fujiwara Y, Peng C, and Orkin SH (2014). Polycomb repressive complex 2 regulates normal hematopoietic stem cell function in a developmental-stage-specific manner. Cell Stem Cell 14, 68–80. 10.1016/j.stem.2013.10.001. [PubMed: 24239285]
47. Hart NJ, and Powers AC (2019). Use of human islets to understand islet biology and diabetes: progress, challenges and suggestions. Diabetologia 62, 212–222. 10.1007/s00125-018-4772-2. [PubMed: 30547228]
48. Camunas-Soler J, Dai XQ, Hang Y, Bautista A, Lyon J, Suzuki K, Kim SK, Quake SR, and MacDonald PE (2020). Patch-Seq Links Single-Cell Transcriptomes to Human Islet Dysfunction in Diabetes. Cell Metab 31, 1017–1031 e1014. 10.1016/j.cmet.2020.04.005. [PubMed: 32302527]

49. Shrestha N, De Franco E, Arvan P, and Cnop M (2021). Pathological beta-Cell Endoplasmic Reticulum Stress in Type 2 Diabetes: Current Evidence. Front Endocrinol (Lausanne) 12, 650158. 10.3389/fendo.2021.650158. [PubMed: 33967960]
50. Chera S, Baronnier D, Ghila L, Cigliola V, Jensen JN, Gu G, Furuyama K, Thorel F, Gribble FM, Reimann F, and Herrera PL (2014). Diabetes recovery by age-dependent conversion of pancreatic delta-cells into insulin producers. Nature 514, 503–507. 10.1038/nature13633. [PubMed: 25141178]
51. Thorel F, Nepote V, Avril I, Kohno K, Desgraz R, Chera S, and Herrera PL (2010). Conversion of adult pancreatic alpha-cells to beta-cells after extreme beta-cell loss. Nature 464, 1149–1154. 10.1038/nature08894. [PubMed: 20364121]
52. Bramswig NC, Everett LJ, Schug J, Dorrell C, Liu C, Luo Y, Streeter PR, Naji A, Grompe M, and Kaestner KH (2013). Epigenomic plasticity enables human pancreatic alpha to beta cell reprogramming. J Clin Invest 123, 1275–1284. 10.1172/JCI66514. [PubMed: 23434589]
53. Nguyen QH, Pervolarakis N, Nee K, and Kessenbrock K (2018). Experimental Considerations for Single-Cell RNA Sequencing Approaches. Front Cell Dev Biol 6, 108. 10.3389/fcell.2018.00108. [PubMed: 30234113]
54. Marsh SE, Walker AJ, Kamath T, Dissing-Olesen L, Hammond TR, de Soysa TY, Young AMH, Murphy S, Abdulraouf A, Nadaf N, et al. (2022). Dissection of artifactual and confounding glial signatures by single-cell sequencing of mouse and human brain. Nat Neurosci 25, 306–316. 10.1038/s41593-022-01022-8. [PubMed: 35260865]
55. Engin F, Yermalovich A, Nguyen T, Hummasti S, Fu W, Eizirik DL, Mathis D, and Hotamisligil GS (2013). Restoration of the unfolded protein response in pancreatic beta cells protects mice against type 1 diabetes. Sci Transl Med 5, 211ra156. 10.1126/scitranslmed.3006534.
56. Lee H, Lee YS, Harenda Q, Pietrzak S, Oktay HZ, Schreiber S, Liao Y, Sonthalia S, Ciecko AE, Chen YG, et al. (2020). Beta Cell Dedifferentiation Induced by IRE1alpha Deletion Prevents Type 1 Diabetes. Cell Metab 31, 822–836 e825. 10.1016/j.cmet.2020.03.002. [PubMed: 32220307]
57. Chabosseau P, Yong F, Delgadillo-Silva LF, Lee EY, Melhem R, Li S, Gandhi N, Wastin J, Noriega LL, Leclerc I, et al. (2023). Molecular phenotyping of single pancreatic islet leader beta cells by “Flash-Seq”. Life Sci 316, 121436. 10.1016/j.lfs.2023.121436. [PubMed: 36706832]
58. Pullen TJ, Khan AM, Barton G, Butcher SA, Sun G, and Rutter GA (2010). Identification of genes selectively disallowed in the pancreatic islet. Islets 2, 89–95. 10.4161/isl.2.2.11025. [PubMed: 21099300]
59. Nasteska D, Fine NHF, Ashford FB, Cuozzo F, Viloria K, Smith G, Dahir A, Dawson PWJ, Lai YC, Bastidas-Ponce A, et al. (2021). PDX1(LOW) MAFA(LOW) beta-cells contribute to islet function and insulin release. Nat Commun 12, 674. 10.1038/s41467-020-20632-z. [PubMed: 33514698]
60. Jaafar R, Tran S, Shah AN, Sun G, Valdearcos M, Marchetti P, Masini M, Swisa A, Giacometti S, Bernal-Mizrachi E, et al. (2019). mTORC1 to AMPK switching underlies beta-cell metabolic plasticity during maturation and diabetes. J Clin Invest 129, 4124–4137. 10.1172/JCI127021. [PubMed: 31265435]
61. Helman A, Cangelosi AL, Davis JC, Pham Q, Rothman A, Faust AL, Straubhaar JR, Sabatini DM, and Melton DA (2020). A Nutrient-Sensing Transition at Birth Triggers Glucose-Responsive Insulin Secretion. Cell Metab 31, 1004–1016 e1005. 10.1016/j.cmet.2020.04.004. [PubMed: 32375022]
62. Ni Q, Sun J, Wang Y, Wang Y, Liu J, Ning G, Wang W, and Wang Q (2022). mTORC1 is required for epigenetic silencing during beta-cell functional maturation. Mol Metab 64, 101559. 10.1016/j.molmet.2022.101559. [PubMed: 35940555]
63. Ahlqvist E, Prasad RB, and Groop L (2020). Subtypes of Type 2 Diabetes Determined From Clinical Parameters. Diabetes 69, 2086–2093. 10.2337/dbi20-0001. [PubMed: 32843567]
64. Nielsen PJ, Lorenz B, Muller AM, Wenger RH, Brombacher F, Simon M, von der Weid T, Langhorne WJ, Mossmann H, and Kohler G (1997). Altered erythrocytes and a leaky block in B-cell development in CD24/HSA-deficient mice. Blood 89, 1058–1067. [PubMed: 9028339]

65. Thorens B, Tarussio D, Maestro MA, Rovira M, Heikkila E, and Ferrer J (2015). Ins1(Cre) knock-in mice for beta cell-specific gene recombination. Diabetologia 58, 558–565. 10.1007/s00125-014-3468-5. [PubMed: 25500700]
66. Dror E, Dalmas E, Meier DT, Wueest S, Thevenet J, Thienel C, Timper K, Nordmann TM, Traub S, Schulze F, et al. (2017). Postprandial macrophage-derived IL-1beta stimulates insulin, and both synergistically promote glucose disposal and inflammation. Nat Immunol 18, 283–292. 10.1038/ni.3659. [PubMed: 28092375]
67. Lyon J, Manning Fox JE, Spigelman AF, Kim R, Smith N, O'Gorman D, Kin T, Shapiro AM, Rajotte RV, and MacDonald PE (2016). Research-Focused Isolation of Human Islets From Donors With and Without Diabetes at the Alberta Diabetes Institute IsletCore. Endocrinology 157, 560–569. 10.1210/en.2015-1562. [PubMed: 26653569]
68. Hashimshony T, Senderovich N, Avital G, Klochendler A, de Leeuw Y, Anavy L, Gennert D, Li S, Livak KJ, Rozenblatt-Rosen O, et al. (2016). CEL-Seq2: sensitive highly-multiplexed single-cell RNA-Seq. Genome Biol 17, 77. 10.1186/s13059-016-0938-8. [PubMed: 27121950]
69. Arrigoni L, Richter AS, Betancourt E, Bruder K, Diehl S, Manke T, and Bonisch U (2016). Standardizing chromatin research: a simple and universal method for ChIP-seq. Nucleic Acids Res 44, e67. 10.1093/nar/gkv1495. [PubMed: 26704968]
70. Arrigoni L, Al-Hasani H, Ramirez F, Panzeri I, Ryan DP, Santacruz D, Kress N, Pospisilik JA, Bonisch U, and Manke T (2018). RELACS nuclei barcoding enables high-throughput ChIP-seq. Commun Biol 1, 214. 10.1038/s42003-018-0219-z. [PubMed: 30534606]
71. Zhou W, Hinoue T, Barnes B, Mitchell O, Iqbal W, Lee SM, Foy KK, Lee KH, Moyer EJ, VanderArk A, et al. (2022). DNA methylation dynamics and dysregulation delineated by high-throughput profiling in the mouse. Cell Genom 2. 10.1016/j.xgen.2022.100144.
72. Zhou W, Triche TJ Jr., Laird PW, and Shen H (2018). SeSAMe: reducing artifactual detection of DNA methylation by Infinium BeadChips in genomic deletions. Nucleic Acids Res 46, e123. 10.1093/nar/gky691. [PubMed: 30085201]
73. Bhardwaj V, Heyne S, Sikora K, Rabbani L, Rauer M, Kilpert F, Richter AS, Ryan DP, and Manke T (2019). snakePipes: facilitating flexible, scalable and integrative epigenomic analysis. Bioinformatics 35, 4757–4759. 10.1093/bioinformatics/btz436. [PubMed: 31134269]
74. Love MI, Huber W, and Anders S (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15, 550. 10.1186/s13059-014-0550-8. [PubMed: 25516281]
75. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, and Ideker T (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13, 2498–2504. 10.1101/gr.1239303. [PubMed: 14597658]
76. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, and Glass CK (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell 38, 576–589. 10.1016/j.molcel.2010.05.004. [PubMed: 20513432]
77. Kechin A, Boyarskikh U, Kel A, and Filipenko M (2017). cutPrimers: A New Tool for Accurate Cutting of Primers from Reads of Targeted Next Generation Sequencing. J Comput Biol 24, 1138–1143. 10.1089/cmb.2017.0096. [PubMed: 28715235]
78. Langmead B, and Salzberg SL (2012). Fast gapped-read alignment with Bowtie 2. Nat Methods 9, 357–359. 10.1038/nmeth.1923. [PubMed: 22388286]
79. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, and Li H (2021). Twelve years of SAMtools and BCFtools. Gigascience 10. 10.1093/gigascience/giab008.
80. Ramirez F, Ryan DP, Gruning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dundar F, and Manke T (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res 44, W160–165. 10.1093/nar/gkw257. [PubMed: 27079975]
81. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, and Liu XS (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol 9, R137. 10.1186/gb-2008-9-9-r137. [PubMed: 18798982]

82. Stovner EB, and Saetrom P (2019). epic2 efficiently finds diffuse domains in ChIP-seq data. Bioinformatics 35, 4392–4393. 10.1093/bioinformatics/btz232. [PubMed: 30923821]

83. Quinlan AR, and Hall IM (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841–842. 10.1093/bioinformatics/btq033. [PubMed: 20110278]

84. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM 3rd, Hao Y, Stoeckius M, Smibert P, and Satija R (2019). Comprehensive Integration of Single-Cell Data. Cell 177, 1888–1902 e1821. 10.1016/j.cell.2019.05.031. [PubMed: 31178118]

Highlights:

1. Quantitative H3K27me3 heterogeneity reveals 2 common β-cell subtypes
2. β_{HI} and β_{LO} cells are stably distinct by 7 independent sets of parameters
3. H3K27me3 dosage controls βHI / βLO ratio in vivo
4. β_{HI} and β_{LO} cells are conserved in humans and enriched in Type-2 diabetes

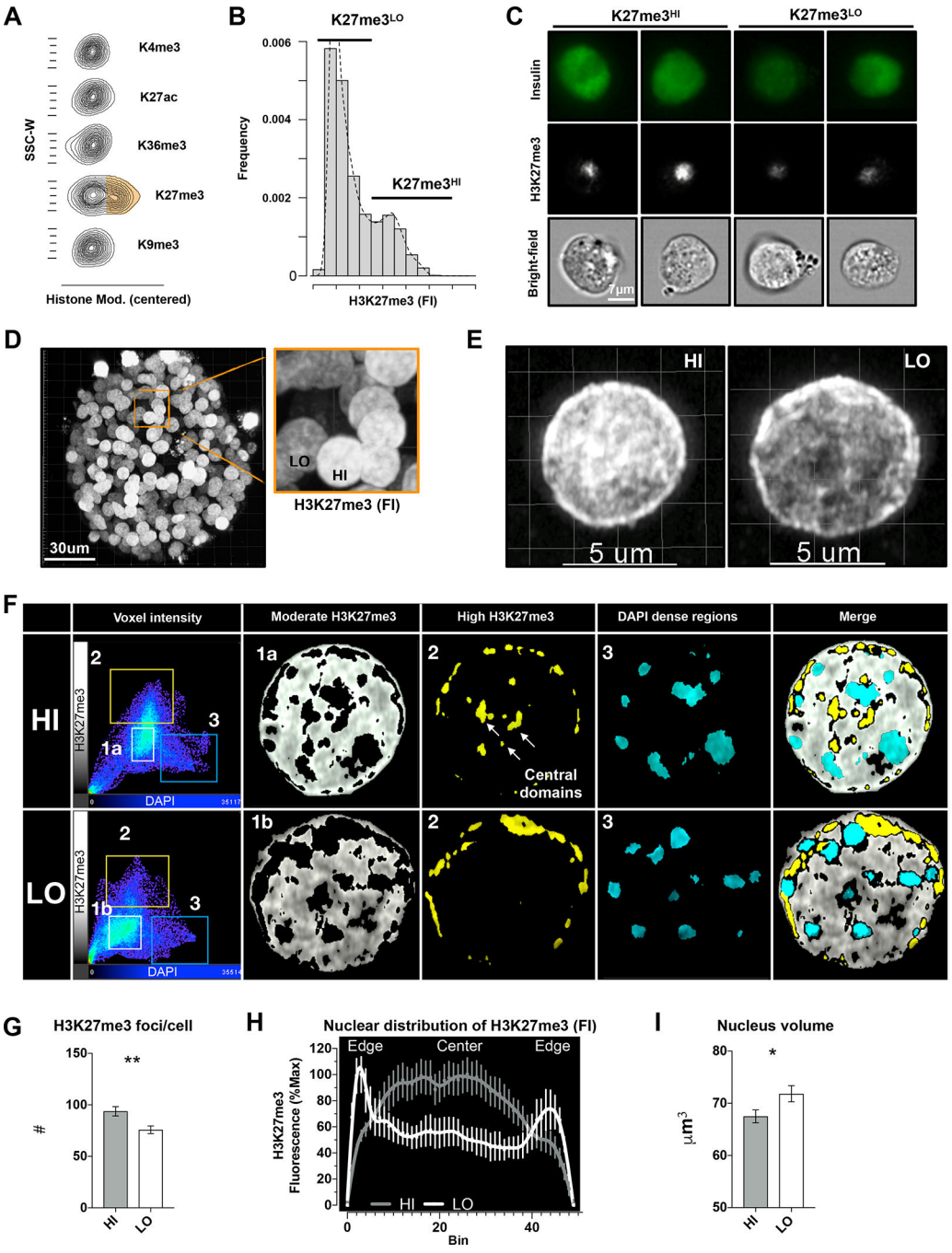


Figure 1. Two epigenetically distinct pancreatic β-cell subtypes

A. Representative contour plots of the centered intensities of the stated histone modifications in β-cells isolated from individual mice (image representative of n=6 mice from 3 experiments).

B. Representative distribution plots of H3K27me3 staining fluorescent intensities (FI) in insulin positive β-cells (representative of 5 experiments, n=4 mice each).

C. Representative ImageStream analysis of dispersed, fixed single- β-cells of islets isolated from individual mice. The different panels show the immunostaining against insulin (up)

H3K27me3 (middle), and bright field image (bottom) of the same β-cells (representative of 2 experiments).

D. Representative 3D reconstruction of one pancreatic islet isolated from male mice, immunostained against H3K27me3, and a zoomed in image of adjacent insulin positive H3K27me3 HI and LO nuclei (representative of 3 experiments).

E. 3D reconstruction of high-resolution confocal imaging from H3K27me3-HI (left) or -LO (right) sorted β-cells; one representative image out of 60 nuclei from n=4 mice.

F. Representative voxel intensities and co-localizations of H3K27me3 and DAPI in one z-plane of each of the nuclei imaged in E. Groupings of voxels was done according to their DAPI and H3K27me3 intensities (left panel). Group 1 represent low\moderate intensity voxels, are localized in the nuclear interior and are shifted when comparing -HI and -LO cells (1a and 1b). H3K27me3 high intensity voxels are in group 2 (yellow) and are localized in the nuclear periphery of both nuclei with addition of central domains in the H3K27me3-HI nucleus. DAPI high voxels are in group 3 that is unchanged.

G. Bar plot representation of the mean of numbers of H3K27me3 foci per nucleus of HI/LO β-cells isolated from 4 individual mice. Assessed by automated quantification of high-resolution images of 67 (HI) and 63 (LO) single nuclei. **= unpaired t-test, p-value<0.01. Error bars are mean ± SEM.

H. Line plot of the averaged H3K27me3 intensities across the center optical plane (binned) of HI/LO sorted β-cell nuclei. Signal is normalized per cell.

I. Bar plot representation of the Mean of the nucleus volumes of HI/LO sorted β-cells as assessed after reconstructing DAPI positive z-stacks and measuring the DAPI positive volume (analysis of high-resolution imaging of the 67 or 63 nuclei of single cells isolated from 4 individual mice). *= unpaired t-test, p-value <0.05. Error bars are mean ± SEM.

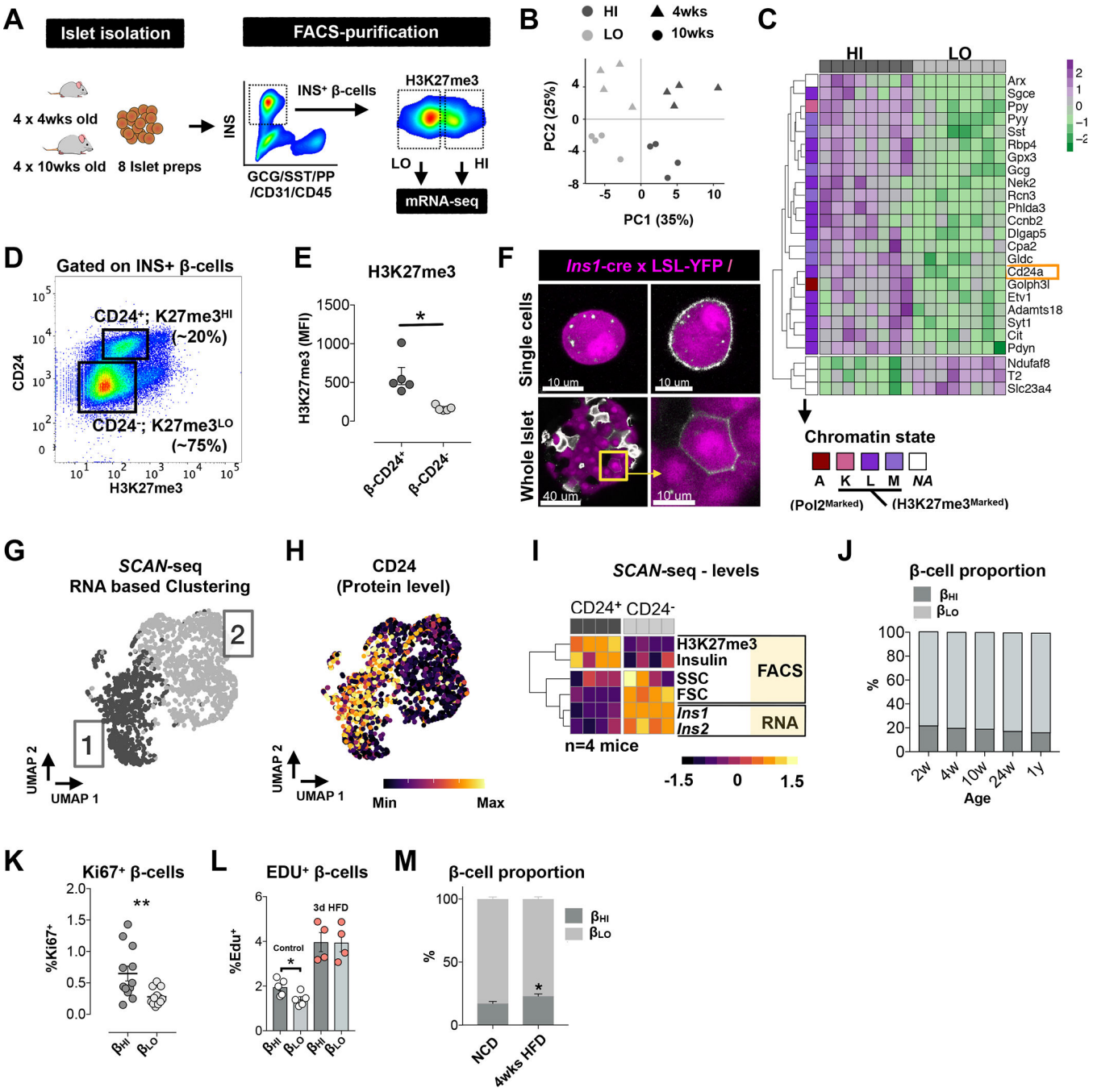


Figure 2. H3K27me3-HI cells are transcriptionally distinct and express cell surface CD24

A. Schematic of the experimental plan. Eight biological replicates of H3K27me3 HI/LO β-cells were isolated from four 4-week-old or four 10-week-old wildtype mice. One thousand H3K27me3 HI/LO cells were sorted from each mouse and low input RNA extraction and mRNA-seq was performed.

B. PCA RNA-seq signals across the β-cells from young and adult mice used in the screening study. Each data point, shown as a triangle or a circle, represents the transcriptome of HI

(dark gray) or LO (light gray) β-cells isolated from individual 4 (triangles) or 10 (circles) weeks old mice. total of n=8 mice.

C. Heatmap of the differentially expressed genes between H3K27me3-HI/LO murine β-cells, and their chromatin-states as previously annotated[5]. Log(normalized counts), z-scored per row.

D. Representative example of CD24 expression versus H3K27me3 intensities in β-cells isolated from 10-week-old wildtype mice. representative of n=5 experiments.

E. H3K27me3 mean fluorescence intensities (MFI) in $CD24^{-/+}$ β-cells; each dot represents a population from an individual mouse. Paired t-test, * represent *p*-value<0.05. n=5 experiments. Error bars are mean± SEM.

F. Representation of the heterogeneity in CD24 expression in live single β-cells or whole islets isolated from β-cell reporter mouse line (YFP is expressed upon Ins1-promoter driven CRE expression).

G. UMAP visualization of sorted mouse β-cells that underwent *SCAN-seq* protocol. Colors represent the two major clusters of β-cells. n=2,156 cells

H. UMAP map overlaid with the FACS-recorded levels of CD24 protein of each cell.

I. Heatmap representation of SCAN-seq-scaled and averaged values (FACS-recorded intensities of the depicted parameters or RNA expression levels; Z-scored per row) from single β-cells negative or positive for CD24 from n=4 individual mice (columns).

J. Representation of the proportion of H3K27me3-HI\$CD24^{+}$ β cells through the life-span of mice. 8–12 mice per age group from n=4 experiments. Error bars are mean± SEM

K. Representation of the proliferating cell fraction of H3K27me3-HI\$CD24^{+}$ and H3K27me3-LO\$CD24^{-}$ β cells. Paired t-test, * represent *p*-value<0.05. Each dot represents one mouse, 12 mice from a total n=4 experiments. Error bars are mean± SEM

L. Representation of the proliferation in the H3K27me3-HI\$CD24^{+}$ or the H3K27me3-LO\$CD24^{-}$ β cell compartment during 3 days of normal chow diet (control) or high fat diet (HFD) feeding. Mice were injected with Edu once per day. Paired t-test, * represent *p*-value<0.05. Each dot represents one mouse, 4–5 independent mice. Error bars are mean± SEM

M. Representation of the proportion of H3K27me3-HI\$CD24^{+}$ β-cells upon 4 weeks of high fat diet feeding. Unpaired t-test, * represent *p*-value<0.05. 10–11 mice per treatment group from n=3 experiments. Error bars are mean± SEM.

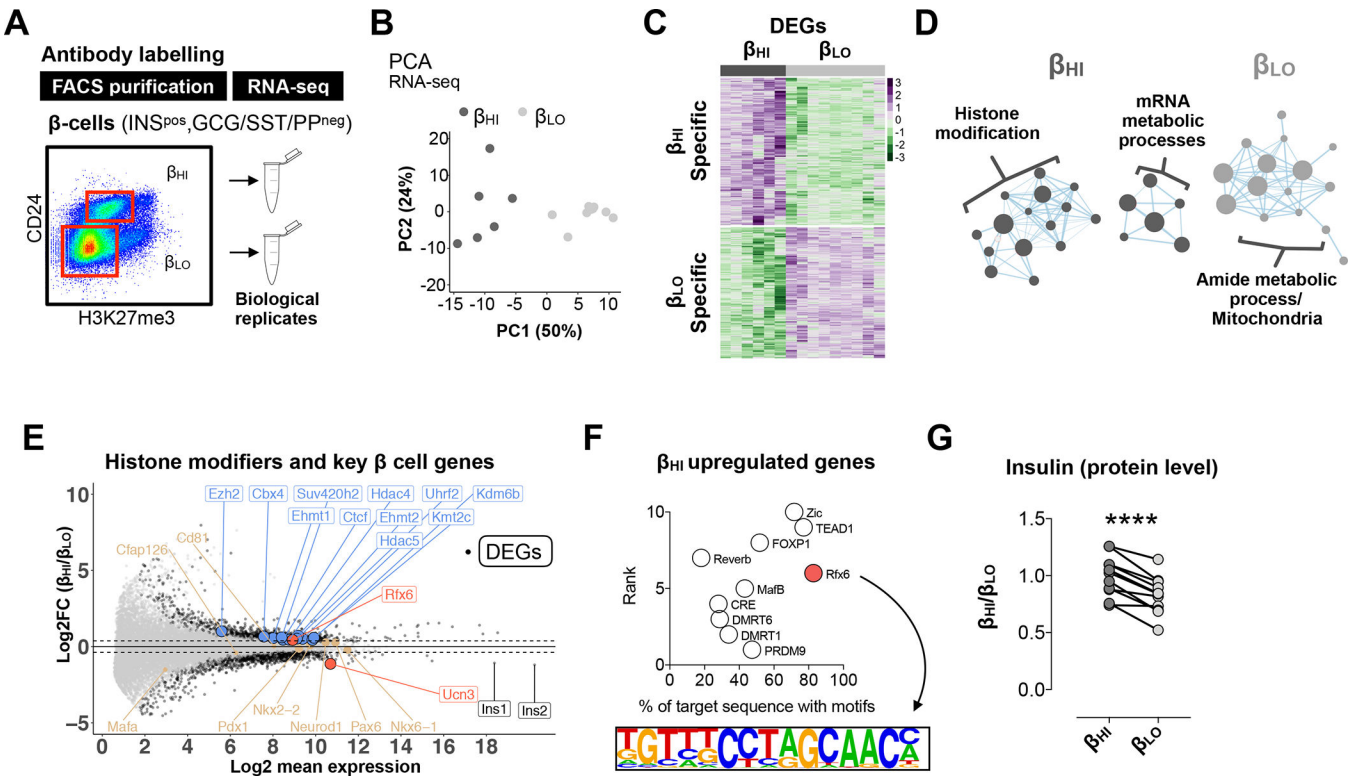


Figure 3. β_{HI} vs β_{LO} cells are functionally distinct and specialized

A. Schematic of the experimental plan. Two dimensions, CD24 and H3K27me3 allows clean separation of β_{HI}/β_{LO} for RNA sequencing analysis.

B. PCA of H3K27me3 RNA-seq signals, showing reproducible separation of β_{HI} and β_{LO} β-cells, each dot represents one biological replicate from 3 independent experiments.

C. Clustered heatmap representation of the log(normalized) expression of all differentially expressed genes (n=~2500) across all replicates, Z-score was calculated per gene (row).

D. A Cytoscape plot of GSEA pathways represents the β_{HI} (dark gray) or β_{LO} (light gary) enriched gene sets. Dot size is proportional to the false discovery rate *q*-value.

E. MA plot showing the fold change in expression generated by comparing β_{HI} over β_{LO} β-cells. Black dots represent significantly deregulated genes, that are also boxed when labeled and highlighted (histone modifiers-blue; genes associated with β-cells and their maturation-red/beige; Ins1/2 genes-black). Black or Boxed genes are statistically significant (*P*-value adjusted for multiple testing < 0.05, with fold change cutoff of 1.33).

F. Top 10 significant transcription factor motifs enriched within +/− 2kb from TSS of upregulated genes in β_{HI} cells. Rfx6 transcription factor and its binding motif are highlighted.

G. Fold increase in insulin protein levels of β_{HI} cells. Connected dots represent cells from each of the types isolated from an individual mouse. **** = paired t-test, *p*-value <0.0001.

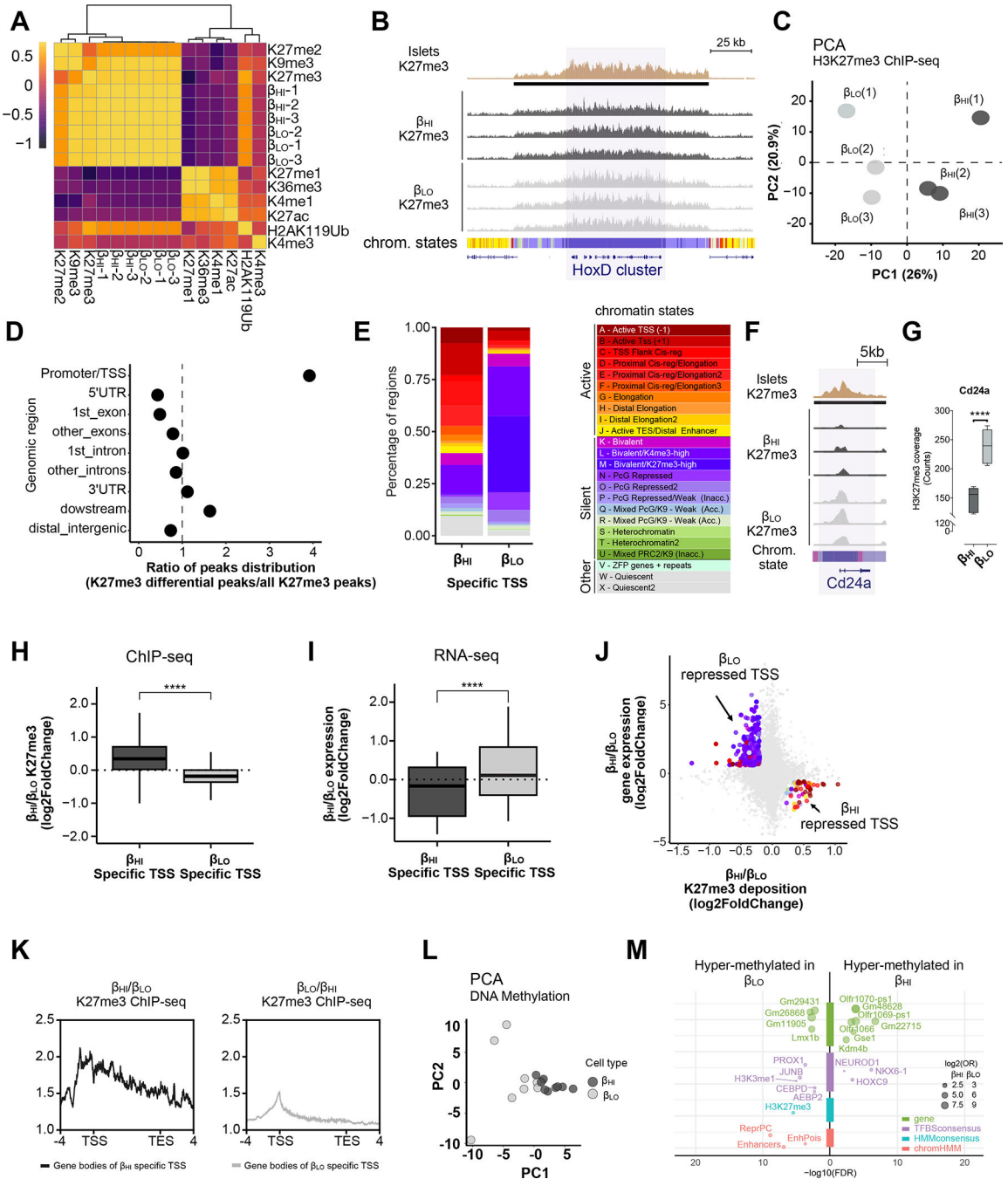


Figure 4. β_{HI} and β_{LO} cells exhibit distinct epigenomes

A. Heatmap showing Spearman correlations of ChIP-seq signals of the indicated histone marks from whole islets, compared to H3K27me3 signals from triplicate experiments of β_{HI} and β_{LO} β-cells.

B. Genomic snapshots showing H3K27me3 ChIP-seq tracks from whole islets and purified β_{HI} and β_{LO} cells, as indicated. The HoxD cluster of genes is represented. Horizontal black bars represent H3K27me3 covered broad regions. Colored horizontal bars represent chromatin states, as previously described[5] and reproduced in panel (E).

C. PCA of H3K27me3 ChIP-seq signals over all identified H3K27me3 peaks, showing reproducible separation of β_{HI} and β_{LO} β-cells.

D. Genomic regions' enrichment among H3K27me3 differential peaks between βHI and β_{LO} β-cells. The dot-plot shows a specific enrichment on transcription start sites (TSS) for H3K27me differential peaks. The distribution of annotated genomic regions over H3K27me3 differential peaks was compared to the same distribution of all identified peaks and plotted as a ratio of percentages (i.e. values >1 mean relative enrichment of H3K27me3 differential peaks over the overall peaks' distribution, while values <1 mean relative depletion).

E. Chromatin states distribution on β_{HI} (left) and β_{LO} (right) H3K27me3-enriched TSS; relative gain of H3K27me3 on active genes (red hues) and relative loss on bivalent genes (purple hues), characterize β_{LO} β-cells. Color-code for chromatin states as previously described[5] is reported here

F. Genomic snapshots showing H3K27me3 ChIP-seq tracks from whole islets and purified β_{HI} and β_{LO} cells, as indicated. The Cd24a gene is represented. Horizontal black bars represent H3K27me3 covered broad regions. Colored horizontal bars represent chromatin states, as previously described[5] and reproduced in panel (E).

G. Box plot representation of the Cd24a gene coverage in β_{HI} and β_{LO} cells

H. Boxplot showing the ratio of the normalized K27me3 ChIP-seq signal between β_{HI} and β_{LO} cells, on β_{HI} (left, dark-grey) and β_{LO} (right, light-grey) K27me3-enriched TSS. **** = *p*-value < 0.0001, as assessed by *t*-test.

I. Boxplot showing the ratio of the normalized RNA-seq signal between β_{HI} and β_{LO} β-cells, on β_{HI} (left, dark-grey) and β_{LO} (right, light-grey) K27me3-enriched TSS. The transcriptional regulation is in line with the reciprocal K27me3 enrichment in panel G. **** = *p*-value < 0.0001, as assessed by *t*-test.

J. Scatter plot showing the correlation between β_{HI}/β_{LO} gene expression and H3K27me3 ChIP-signal. Only β_{HI} vs β_{LO} -specific TSS are colored by their chromatin states.

K. β_{HI} (left) and β_{LO} (right) β-cells H3K27me3 ChIP-seq signal over the gene bodies of related βHI and βLO-specific TSS's. The signals are from merged triplicate experiments, and visualized as gene bodies +/− 4 Kb. The coverage profiles show a reciprocal enrichment/depauperation of the K27me3 signal on TSS vs the gene bodies, in β_{LO} and β_{HI} cells, respectively.

L. PCA of DNA methylation array signals, showing reproducible separation of β_{HI} and β_{LO} β-cells.

M. Enrichment analysis of Differentially Methylated Loci (DMLs) between β_{HI} and β_{LO} within the indicated dataset.

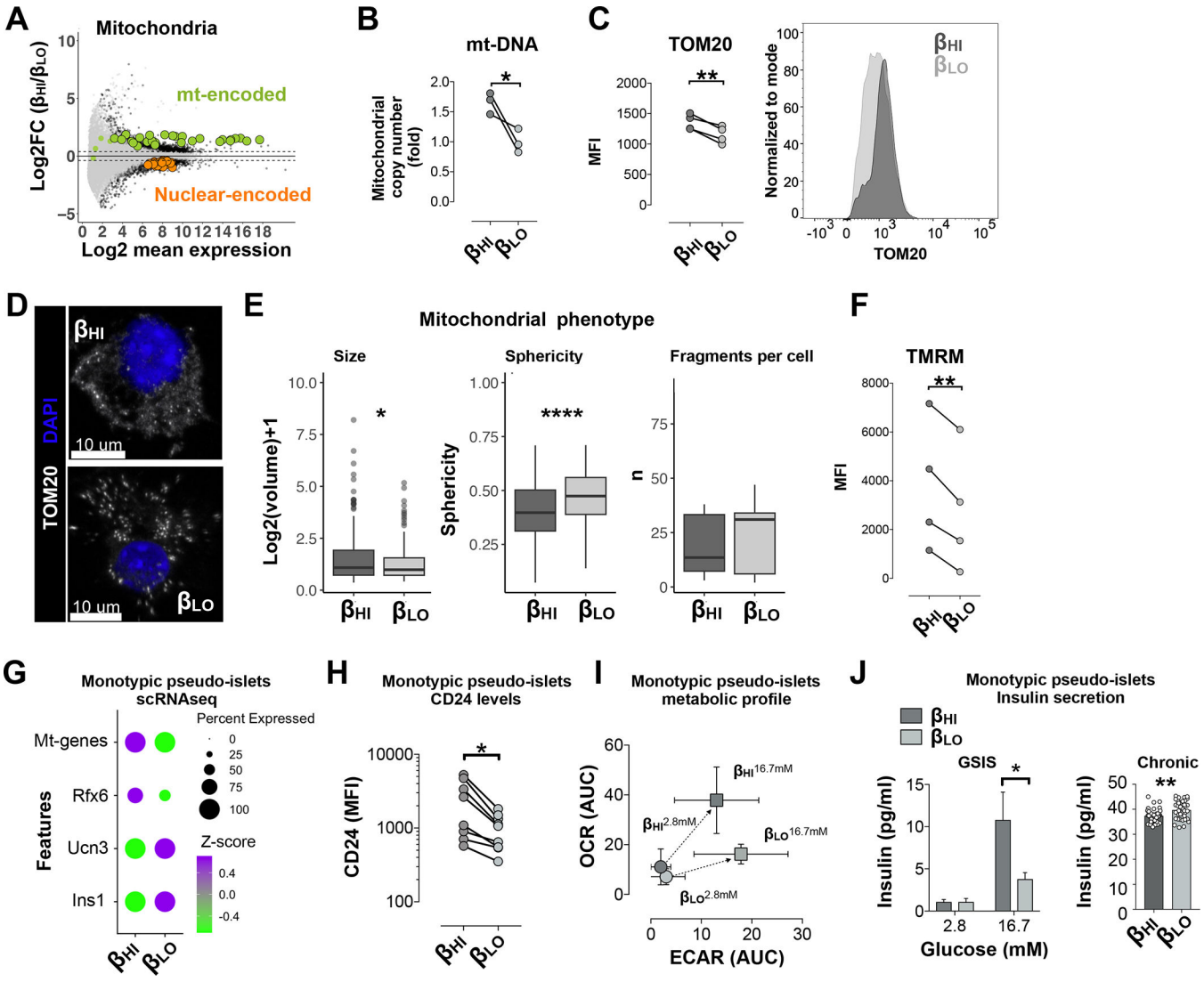


Figure 5. β_{HI} and β_{LO} cells are stably and functionally distinct

A. MA plot showing the fold change in expression generated by comparing β_{HI} over β_{LO} β-cells. Green dots represent mt-encoded mitochondrial genes, orange dots represent nuclear encoded mitochondrial genes listed in S3F. Differentially expressed genes are surrounded by black borders; (*p*-value adjusted for multiple testing < 0.05, with fold change cutoff of 1.33).

B. Fold increase in mitochondrial DNA content (copy number normalized to genomic DNA, as measured by qPCR). Each dot represents an independent experiment, n=3. *= unpaired t-test, *p*-value<0.05

C. Dot plot representation of the MFI of TOM20 in the β-cell types, A Representative flow cytometer histogram of TOM20 labeling in the β-cell types. The connected dots represent cells from n=4 individual mice. **= paired t-test, *p*-value<0.01.

D. Representative images of TOM20 antibody labeling of one βHI and one βLO cells. fixed β-Cells were first sorted according to their insulin, H3K27me3, and CD24 levels and then labelled with antibody against TOM20 (white) and analyzed at high resolution confocal microscopy. DAPI (blue) was used as counter staining.

E. Box plot representations of mitochondrial size, sphericity, and number of fragments per cell. 16 cells were analyzed from n=3 independent mice. *= paired t-test, *p*-value<0.05, ****= paired t-test, *p*-value<0.0001.

F. Mean fluorescent intensities (MFI) of TMRM in the β-cell types, connected dots represent cells from n=4 individual mice. **= paired t-test, *p*<0.01.

G. Dot plot representation of gene expression levels (z-scored) from scRNAseq of dissociated monotypic β_{LO} or β_{HI} pseudo-islets after 7 days in culture.

H. Dot plot representation of FACS measurements of CD24 protein levels in single cells from monotypic β_{HI} or β_{LO} pseudo-islets after 7 days in culture.

I. Single spheroid metabolic profiling via Seahorse extracellular flux analysis in basal glucose (2.8mM) and glucose stimulated (16.7mM) conditions. Oxygen consumption rate (OCR) extracellular acidification rate (ECAR) Area under the curves (AUC) are shown in Figure S5I.

J. Glucose stimulated insulin secretion (GSIS) and 48 hours, chronic, insulin secretion in single pseudo-islets generated by aggregating 2000 of βHI or of βLO cells. Insulin levels were measured for one hour before stimulation (2.8mM glucose), followed by another hour after stimulation (16.7mM glucose). 25–40 single spheroids were analyzed from n=5 independent experiments. *= two-way ANOVA with multiple comparison correction, *p*-value<0.05.

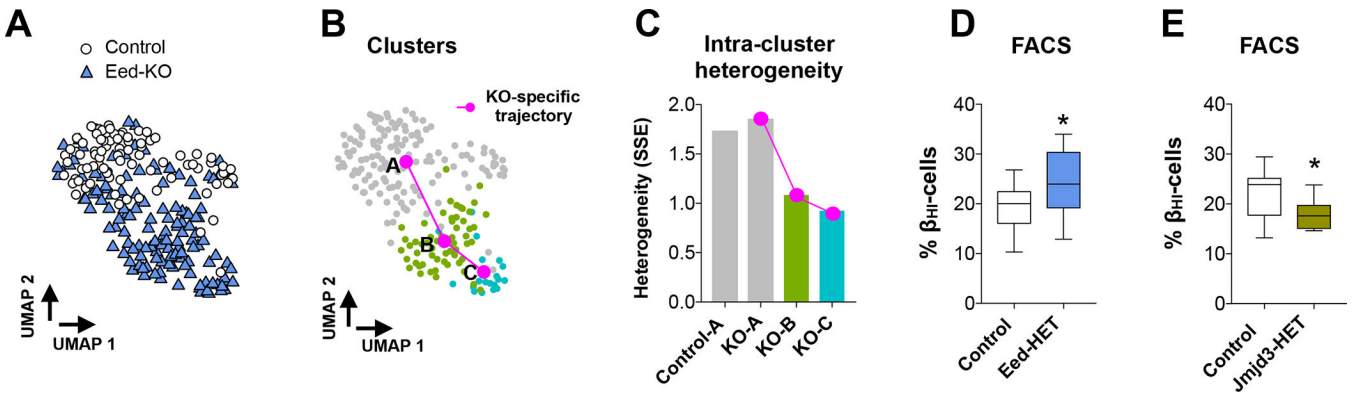


Figure 6. H3K27me3 dosage controls β_{HI}/β_{LO} β-cell ratios and overall heterogeneity

A. UMAP visualization of sorted mouse β-cells that underwent *SCAN*-seq protocol. Colors and shape represent mouse genotypes Eed KO (n= 131 cells) or wild-type (Control; n=83 cells).

B. Cluster topology for the data set in (A). Trajectory was inferred by slingshot. Initial clustering was done on all cells, splitting KOs from Controls. KO cluster was further divided into 2 clusters.

C. Bar plot showing the intra-cluster sum of squared errors (SSE) per the indicated cluster of cells. As in (B), the magenta line connects the 3 KO groups.

D. Box plot representation of the percentage of β_{HI} cells per genotype. Data are medians of Control or Eed-HET mice, n=18 mice each group from 10 or 12 experiments (correspondingly). *= unpaired t-test, *p*-value<0.05. box plots show the median and whiskers indicate min and max values.

E. Box plot representation of the percentage of β_{HI} cells per genotype. Data are medians of Control or Jmjd3-HET mice, n=9 mice each group from 6 experiments. *= unpaired t-test, *p*-value<0.05. box plots show the median and whiskers indicate min and max values.

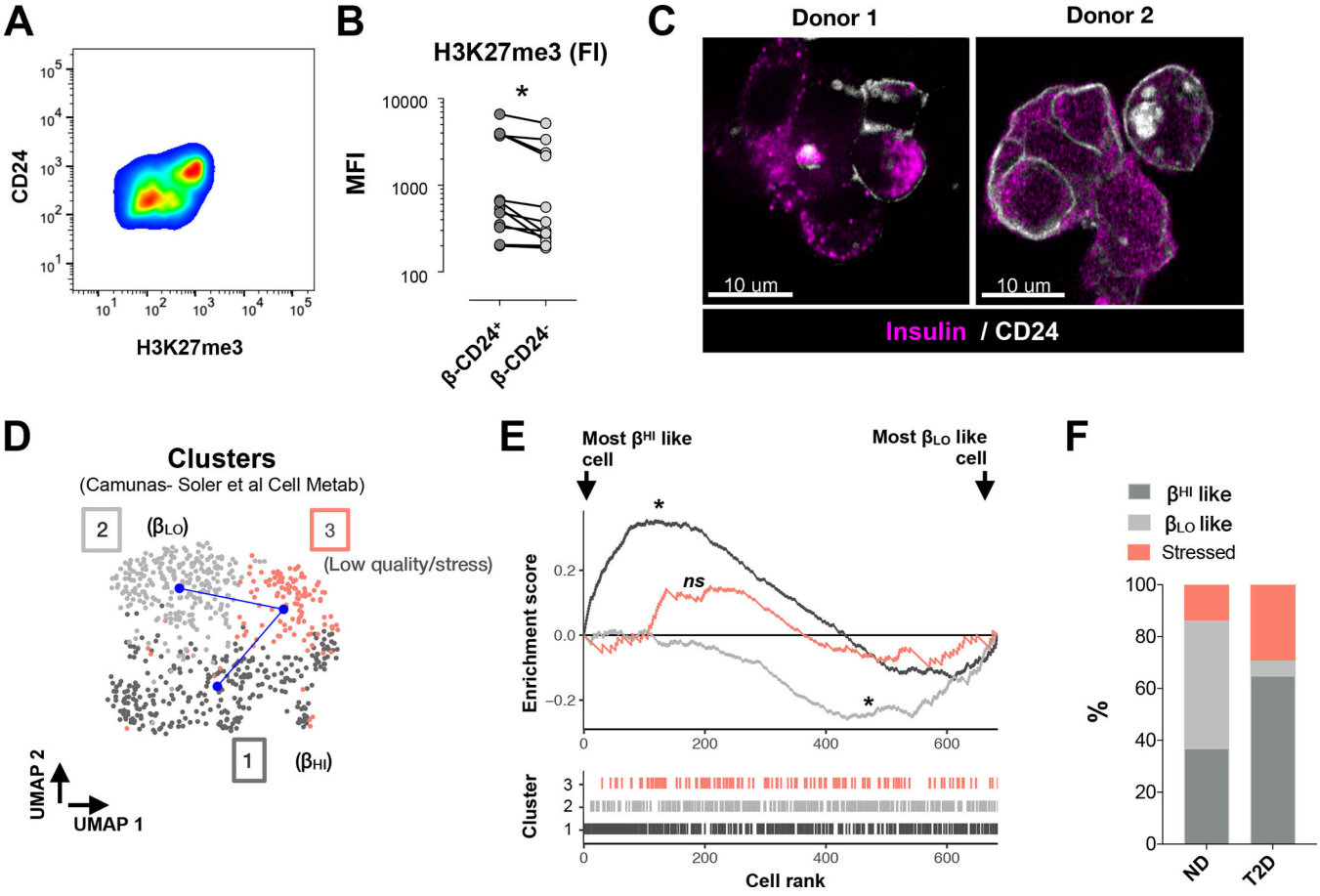


Figure 7. β_{HI} and β_{LO} cells are conserved in humans and their ratio altered in diabetes.

A. FACS plot of the fluorescence intensities of CD24 and H3K27me3 in human β-cells isolated from one donor.

B. FACS fluorescence intensities of H3K27me3 levels in $CD24^+$ compared to $CD24^-$ human β-cells, each connected pair of dots represents the mean fluorescent intensity (MFI) from one donor, n=12 donors. *= paired t-test, *p*-value<0.05.

C. Representation of the β-cell surface labeling of CD24 (white) in sub-optimally dispersed, adjacent human islet cells from 2 independent donors. Counter staining of insulin is shown in magenta.

D. UMAP representation of the cluster topology of human beta cells. β_{HI}/β_{LO} clusters were determined after assessment of expression of the signature, genes reported in Figure 3C. Trajectory was inferred by slingshot. (n=638 β-cells from 11 donors without T2D and 7 donors with T2D)

E. Custom gene set enrichment analysis (GSEA) representation of β_{HI}/β_{LO} signature genes (see Figure 3C). The mean expression (z-score) for the two gene sets was calculated, then the magnitude and direction of differential signatures was determined by calculating the difference in expression between the two gene sets. The cells were then ranked by difference z-score. Plots of cells from all clusters are shown. Cluster 3 had no enrichment. Significant enrichments had *p*-value<0.05.

F. Stacked bar plot representation of the percentage of β-cells in each of the clusters shown in (D). Bars split the cluster distributions of the cells that were isolated from humans without diabetes (ND) or with T2D.

REAGENT OR RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rat monoclonal anti CD24	Thermo Fisher	48-0242-82 RRID:AB_1311169
Mouse monoclonal anti CD24	Biologend	311122 RRID: AB_2561691
Rat monoclonal anti CD45	Thermo Fisher	11-0451-82 RRID: AB_465050
Rat monocolnal anti CD31	BD pharminfen	558738 RRID:AB_397097
Mouse monoclonal anti insulin	Santa Cruz	sc-8033 RRID:AB_627285
Mouse monocolonal anti glucagon	Santa Cruz	sc-514592 RRID:AB_2629431
Mouse monoclonal anti somatostatin	Santa Cruz	sc-55565 RRID:AB_831726
Mouse monoclonal anti pancreatic polypeptide	Santa Cruz	sc-514155 RRID: NA
Rat monoclonal anti MKI67	eBioscience	47-5698-82 RRID:AB_2688065
Mouse monoclonal anti TOMM20	Abcam	ab205487 RRID: NA
Mouse monoclonal anti H3K27me3	Origene	TA347154 RRID: NA
Rabbit polyclonal anti H3K4me3	Diagenode,	C15410003 RRID: AB_2616052
Rabbit polyclonal a nti H3K36me3	Diagenode	C15410192, RRID:AB_2744515
Rabbit polyclonal anti H3K9me3	Diagenode	C15410193 RRID: AB_2616044
Rabbit polyclonal anti H3K27ac	Diagenode	C15410196 RRID:AB_2637079
Rabbit poylclonal anti H3K27me3 (Figure S1F)	Diagenode	C15410195 RRID: AB_2753161
Chemicals		
HBSS	Gibco	14025092
RPMI 1640	Gibco	11875093
Accutase	Sigma	A6964
Collagenase 4	Worthington	LS004189
RNaseIN	Promega,	N2511
PBS	Gibco	14190
AMPure XP beads	Beckman Coulter	A63881
glutamax	Gibco	35050
HEPES	Gibco	1563049
NEBNext Second Strand Synthesis (dNTP-free) Reaction Buffer	NEB	B6117S
DNA Polymerase I	NEB	M0209L
RNase H, recombinant	NEB	M0297L
Murine RNase Inhibitor	NEB	M0314L

REAGENT OR RESOURCE	SOURCE	IDENTIFIER
ProtoScript II Reverse Transcriptase	NEB	M0368L
Deoxynucleotide Solutions, Mix	NEB	N0447L
E. coli DNA Ligase	NEB	M0205L
Critical commercial assays		
Mix-n-Stain™ CF™ 405 Antibody Labeling Kit	Sigma	MX405SS50
Mix-n-Stain™ CF™ 448 Antibody Labeling Kit	Sigma	MX488AS50
Mix-n-Stain™ CF™ 555 Antibody Labeling Kit	Sigma	MX555S20
Insulin ELISA	Mercodia	10-1247-10
In vivo EDU Click-iT AF488	Sigma	BCK488-IV-IM-M
NEBNext® Single Cell/Low Input RNA Library Prep Kit for Illumina®	NEB	E6420L
miRNeasy FFPE Kit	QIAGEN,	217504
iDeal ChIP-seq kit for histones	Diagenode	C01010173
NEBNext Ultra II DNA library preparation kit	NEB	E7645
Quick-DNA Microprep Plus Kit	Zymo Research	D4074
Zymo EZ DNA Methylation Kit	Zymo Research	D5001
Absolute Mouse Mitochondrial DNA Copy Number Quantification qPCR Assay Kit (AMMQ)	ScienCell	M8948
MEGAscript T7 Transcription Kit	Thermo Fisher Scientific	AM1334
Experimental models: Organisms/strains		
Mouse: C57BL/6J	In house breeding	
Mouse: $Cd24^{-/-}$	Gift from Dr. Sherri Christian	Nielsen P.J et al. referenced here
Mouse: $Eed^{fl/fl}$	Gift from Dr. Stuart Orkin	Jax strain #022727; RRID:IMSR_JAX:022727
Mouse: $Kdm6b^{fl/fl}$	Gift from Dr. Stuart Orkin	Jax strain#:029615; RRID:IMSR_JAX:029615
Mouse: R26R-EYFP	Gift from Dr. Thomas Boehm	Jax strain #006148; RRID:IMSR_JAX:006148
Biological samples		
Isolated human islets (Imformation about the donos is in Supplementary Table 2)	Alberta Diabetes Institute IsletCore	www.bcell.org/adi-isletcore
Oligonucleotides		
GCCGGTAATACGACTCACTATAGGGCTACACGACGCTCTTCCGATCTNNNNNNN[6 base barcode]TTTTTTTTTTTTTTTTTTTTTTTTTVN	Hashimshony et al., 2016	Full list of barcode is found in S1 table
Deposited Data		
Bulk RNA-seq, scRNA-seq, ChIP-seq/RELACS and DNA methylation array	This study	GSE224061.
Software and algorithms		
Graphpad Prism 8	Graphpad	https://www.graphpad.com
Flowjo_V10	BD Biosciences	https://www.flowjo.com
Imaris V9.3.1	Oxford instruments	RRID:SCR_007370
R (v4.0.4)		RRID:SCR_001905
Seurat (v4.0.2)	Stuart, T. et al 2019	RRID:SCR_016341

REAGENT OR RESOURCE	SOURCE	IDENTIFIER
DEseq2 (v1.34.1)	https://genomebiology.biomedcentral.com/articles/10.1186/s13059-014-0550-8	RRID:SCR_015687
DeepTools v3.3.2	Ramirez, F et al 2016	RRID:SCR_016366
HOMER (v4.11)	Heinz, S. et al 2010	SCR_010881
Other		
High Fat Diet	Reseach Diet	#D12492i