Title: Methylation of histone H3 lysine 4 is required for maintenance of beta cell function in adult mice

Authors: Ben Vanderkruk,[1,2] Nina Maeshima,[2] Daniel J. Pasula,[1,2] Meilin An,[2] Cassandra L. McDonald,[2] Francis C. Lynn, [1,2] Dan S. Luciani,[1,2] and Brad G. Hoffman[1,2]*

Affiliations: [1]Diabetes Research Group, BC Children's Hospital Research Institute, Vancouver, BC, Canada
[2]Department of Surgery, University of British Columbia, Vancouver, BC, Canada
*Correspondence. Email: brad.hoffman@ubc.ca

Keywords: H3K4me3; DPY30; COMPASS; Transcription; Beta cell; Type 2 diabetes

Abstract

Pancreatic β-cells control glucose homeostasis via regulated production and secretion of insulin. This function arises from a highly specialized gene expression program which is established during development and then sustained, with limited flexibility, in terminally differentiated β-cells. Dysregulation of this program is seen in type 2 diabetes (T2D) but mechanisms that preserve gene expression or underlie its dysregulation in mature β-cells are not well resolved. Here we show that trithorax group-dependent histone H3 lysine 4 trimethylation (H3K4me3) maintains expression of genes important for insulin biosynthesis and glucose-responsiveness in β-cells. Transcriptional changes in H3K4me3-deficient β-cells lead to severe hyperglycemia in adult mice. We show that H3K4me3 deficiency leads to a less active and more repressed epigenome profile, which locally correlates with gene expression deficits but does not globally reduce gene expression. Instead, developmentally regulated genes and genes in weakly active or suppressed states particularly rely on H3K4 methylation. We then show that H3K4me3 is re-organized in diabetic *Lepr*$^{db/db}$ mouse islets in favour of weakly active and disallowed genes at the expense of terminal β-cell markers with broad H3K4me3 peaks. Our results point to key roles of H3K4me3 in maintaining mature β-cell function and establishing a dysfunctional transcriptome in diabetic islets.

Introduction

Pancreatic islets are endocrine micro-organs that regulate glucose homeostasis. In adults, islet β-cells are the exclusive source of circulating insulin, the major blood glucose-lowering hormone. Insufficient insulin release by β-cells is the primary etiology of type 2 diabetes (T2D). The rising global incidence of T2D [1] has prompted intensive research into mechanisms underlying the progressive dysfunction of β-cells that leads to T2D. Transcriptome remodeling, involving activation of developmentally silenced genes and decreased expression of genes important for insulin biosynthesis, glucose-responsiveness, or those that enforce the β-cell identity, emerges in β-cells exposed to chronic metabolic stress [2]. While β-cell-enriched transcription factors have been the primary focus of studies defining the transcriptional program in mature β-cells, there is mounting evidence that chromatin modifications play a central role in the maintenance of mature β-cell identity and function, and that the chromatin landscape is reshaped in β-cells exposed to chronic metabolic stress and T2D [3–10].

Tissue-specific transcription programs rely on precise activation and maintenance of specific genes and stable repression of other genes. Expression patterns are established during cell differentiation and maturation, and then maintained during the balance of that cell's life, which may span decades, cell divisions, and environmental perturbations. Active and repressed genes have stereotypical patterns of chromatin modification which may stabilize their transcription state and are useful in their systematic classification [11]. For example, it has been shown that H3K4me3 is reliably enriched on nucleosomes at the 5' end of actively transcribed genes and is therefore the canonical marker of active promoters [12]. Despite its utility as a marker of promoter activity, global (i.e., genome-wide) decreases in H3K4 methylation are remarkably well tolerated in embryonic stem cells and during development in genetic knockout models [13–16]. Evidence in multicellular animal models points to a requirement for H3K4 methyltransferase activities during certain specialized developmental events, whereas evidence for continued requirement in mature tissues is rare [16–21].

Mono-, di-, and trimethylation of H3K4 in mammalian cells is predominantly catalyzed by COMPASS and COMPASS-related complexes. They are composed of the core subunits WDR5, RBBP5,

ASH2L, and DPY30, plus one of six catalytic subunits MLL1-4 or SET1A-B, and a variable retinue of accessory factors [22] . We previously defined roles for DPY30 during embryonic pancreas development. DPY30 fine-tunes cell fate decisions to decrease endocrine acinar cell specification in favour of exocrine cell specification [21] and helps to activate terminal marker genes during endocrine cell maturation [23]. By contrast, it is not known whether the continuous activity of H3K4 methyltransferases is necessary in terminally differentiated islet cells, where lineage-specific transcription programs have already been established. In islets, recruitment of H3K4 methyltransferase proteins by transcription factors—for example, recruitment of SETD7 by PDX1 [7], and MLL3/4 by MAFA [24]—support expression of insulin and other genes that are important for mature β-cell functions. Whether these functions are performed by H3K4 methylation or by independent coactivator activities of these complexes is not clear [14,21,25]. We therefore sought to define the role for H3K4 methylation in the regulation of gene expression in mature β-cells.

Results

Reduction to H3K4 methylation in beta cells of adult mice leads to glucose intolerance and hyperglycemia.

With the aim of reducing H3K4 methylation we targeted DPY30, a noncatalytic core subunit of all COMPASS complexes which is not required for complex assembly or coactivator activity but boosts methyltransferase activities [26,27]. To induce synchronized deletion of *Dpy30* in mature β-cells, mice harbouring a Pdx1:*CreER*^{TM} transgene [28] and *LoxP* sites flanking the fourth exon of *Dpy30* [21] were administered tamoxifen at eight weeks of age (herein called *Dpy30*-KO) (Fig 1A). As expected, knockout of *Dpy30* did not impact assembly of the other core subunits of the COMPASS complex WDR5, RBBP5, and ASH2L, or their association with chromatin (Fig 1B) but did lead to reductions of H3K4me3 and H3K4me1 in islets (Fig 1C).

We tracked the loss of H3K4 methylation in *Dpy30*-KO β-cells using immunofluorescence at 15-day intervals after tamoxifen administration. As seen in Fig 1D-E, H3K4me3 was stable for at least 15 days and then gradually lost during the subsequent 45 days. H3K4me1 showed a similar pattern, but its loss was even more gradual, whereas DPY30 was undetectable by 15 days after tamoxifen administration (Fig 1D). We took advantage of the slow turnover of H3K4me3 and H3K4me1 to parse the relative contribution of H3K4 methylation, versus DPY30 itself, in controlling gene expression in mature β-cells. We reasoned that defects arising in *Dpy30*-KO β-cells by 15-days post-tamoxifen (when DPY30 is lost but H3K4 methylation is unchanged) highlight functions of DPY30 itself. Defects arising later, by 45-days post-tamoxifen (when H3K4 methylation is reduced) highlight functional contributions of H3K4 methylation. To measure the β-cell transcriptome, GFP-positive islet cells labeled with the mTmG reporter [29] were enriched using fluorescence-activated cell sorting (FACS) from *Dpy30*-KO and -WT mice 15- and 45-days post-tamoxifen and processed for calibrated bulk RNA-seq analysis [30]. Remarkably, loss of DPY30 had little effect on gene expression in β-cells, with only three genes (*Dpy30*, *Edn3*, *C3*) identified as differentially expressed 15-days post-tamoxifen (Fig 1F). In contrast, 828 genes were dysregulated 45-days post-tamoxifen, the majority of which (634) showed lower expression in *Dpy30*-KO cells (Fig 1F). We conclude that transcriptome remodeling in *Dpy30*-KO β-cells results from loss of H3K4 methylation.

To determine whether loss of DPY30 and/or H3K4 methylation in β-cells impairs glucose homeostasis in adult mice, we measured glucose clearance and insulin levels after intraperitoneal

injection of glucose (2 g/kg body weight) in mice that had been fasted for 6 hrs. While glucose tolerance was no different between *Dpy30*-WT and -KO mice 15- and 30-days after tamoxifen administration, *Dpy30*-KO mice developed impaired blood glucose clearance and reduced serum insulin levels 45-days post-tamoxifen (Fig 1G). Glucose intolerance quickly progressed to hyperglycemia (Fig 1H) prompting euthanasia by 60-days.

Since the Pdx1:*CreER*^{TM} transgene is known to also drive recombination in pancreatic islet δ-cells [31] and hypothalamic neurons [32] we additionally used the *Ins1*^{Cre} mouse model [33] to confirm that *Dpy30* deletion specifically in β-cells is sufficient to drive the *in vivo* phenotype. Deletion of *Dpy30* from maturing β-cells using *Ins1*^{Cre} caused reduction of H3K4me3 and H3K4me1 in islets by 5-weeks of age (Fig S1A-B). Unfasted blood glucose was dramatically elevated from 4- to 5-weeks of age, whereas body mass was similar among *Dpy30*-KO, -WT, and -heterozygous *Ins1*^{Cre} mice (Fig S1C-D). *Ins1*^{Cre} *Dpy30*-KO mice showed impaired glucose clearance and low serum insulin following intraperitoneal glucose injection (Fig S1E-F), confirming that β-cell-intrinsic dysfunction is sufficient to drive metabolic defects in *Dpy30*-KO mice. In summary, reduction of H3K4 methylation in β-cells leads to impaired glucose tolerance, reduced serum insulin, and hyperglycemia.

H3K4me3 maintains expression of genes involved in insulin production and glucose-stimulated activity. The reduction of serum insulin in *Dpy30*-KO mice prompted us to examine insulin production and secretion. We first confirmed that H3K4me3 was lost from the *Ins1* and *Ins2* promoters in Pdx1:*CreER*^{TM} *Dpy30*-KO β-cells (Fig 2A) using calibrated chromatin immunoprecipitation sequencing (ChIP-seq) from FACS-enriched GFP-positive β-cells 45 days after tamoxifen administration. H3K4me1 also showed dramatic reduction (Fig 2A). Despite this, RNA-seq data showed a surprisingly modest reduction of *Ins1* and *Ins2* transcripts (Fig 2B). Therefore, robust expression of *Ins1* and *Ins2* RNA can occur in the absence of local H3K4me3. Notably, expression of several genes involved in insulin peptide maturation and packaging was downregulated (Fig S2A). For example *Nnat* and *Spcs3*, subunits of the signal peptidase complex which co-translationally converts preproinsulin to proinsulin [34], showed reduced expression in *Dpy30*-KO β-cells (Fig S2A-B). Proinsulin-to-insulin convertases *Pcsk1*, *Pcsk2*, and *Cpe* were not downregulated (Fig S2A). *Slc30a8*, encoding a zinc transporter required for maturation of insulin secretory granules [35] was highly downregulated in *Dpy30*-KO cells (Fig S2A). Insulin peptide content was reduced much more strongly than insulin RNA in *Dpy30*-KO islets (Fig 2B-C). Reduction of insulin content was confirmed in individual β-cells, where significant reduction to insulin granule size and density was seen in transmission electron micrographs of β-cells from *Dpy30*-KO mice 45 days after tamoxifen administration (Fig 2D-E). We conclude that loss of H3K4 methylation from mature β-cells impairs insulin biosynthesis.

Gene ontology analysis of genes showing reduced expression suggested the additional impairment of biological processes required for stimulated insulin secretion, namely vesicle fusion, calcium-dependent exocytosis, and glucose homeostasis (Fig 2F). Energy production processes including oxidation/reduction and the electron transport chain, meanwhile, were upregulated (Fig 2F). We therefore examined *ex vivo* islet activity during glucose stimulation. First, we measured insulin secretion as a fraction of insulin content to control for the 3.4-fold reduction on insulin in *Dpy30*-KO islets. Notably, the reduced insulin content of *Dpy30*-KO islets did not fully account for the impairment of insulin secretion. The persistent impairment was specific to high glucose stimulation, whereas secretion under nonstimulatory conditions, or bypassing glucose metabolism by depolarizing cells directly using KCl,

143 showed no difference in insulin secretion between *Dpy30*-WT and -KO islets (after adjusting to total insulin
144 content) (Fig 2G). Measurement of cytosolic calcium dynamics further supported an impairment of high
145 glucose stimulated activity:. consistent with lower expression of calcium signaling genes (Fig 2F), high
146 glucose-stimulated calcium influx showed a delayed response and lower magnitude in *Dpy30*-KO islets
147 (Fig 2H). Cytosolic calcium concentration was not different in nonstimulatory conditions, and direct
148 depolarization stimulated comparable calcium influx between islets from *Dpy30*-KO and *Dpy30*-WT mice
149 (Fig 2H) further supporting a specific impairment in glucose responsiveness. We therefore wondered if
150 *Dpy30*-KO impaired glucose catabolism. Expression of *Slc2a2*, encoding the glucose transporter GLUT2,
151 was dramatically reduced in *Dpy30*-KO cells (Fig S2A-B), but expression of the rate-limiting enzyme of
152 glycolysis, *Gck* [36], was not significantly altered (Fig S2A-B) nor was the rate of mitochondrial oxygen
153 consumption in high glucose media altered (Fig 2I-J). In fact, mitochondrial respiratory capacity and area
154 were increased in *Dpy30*-KO islets (Fig 2I-J, Fig S2C) , consistent with elevated expression of oxidative
155 phosphorylation genes in *Dpy30*-KO β-cells (Fig 2F). Therefore, glucose metabolism in β-cells was not
156 appreciably impaired by reduction of H3K4 methylation. These results indicate that H3K4 methylation is
157 essential for the expression of genes that couple glucose metabolism to insulin secretion in mature β-
158 cells.
159
160 **Loss of H3K4me3 leads to gene downregulation in mature beta cells.**
161 It is notable that, while H3K4 methylation is well correlated with [37] and predictive of [38] transcription, our
162 RNA-seq analysis indicates that globally only 4.3% of the 14,677 expressed genes were downregulated by
163 depletion of H3K4 methylation in β-cells. This mirrors previous findings in yeast [39], *Drosophila* embryos [15],
164 and mouse embryonic stem cells [13,14,40] that H3K4 methylation is not required for expression of most
165 genes. We therefore wondered whether gene promoters that are sensitive to H3K4 methylation in β-cells
166 have generalizable features. We first confirmed that H3K4me3 and H3K4me1 were depleted at active
167 gene promoters in *Dpy30*-KO chromatin (Fig 3A). We further confirmed genome-wide reduction of
168 H3K4me3 and H3K4me1 by comparing enrichment between WT and *Dpy30*-KO chromatin in 10-kb bins
169 spanning the genome (Fig 3B-C), confirming that *Dpy30* is necessary for maintenance of β-cell H3K4
170 methylation. Despite this, the transcriptomes of *Dpy30*-KO and -WT cells were well correlated, and the
171 total RNA expression of *Dpy30*-KO cells was not decreased (Fig 3D). Therefore, global depletion of H3K4
172 methylation does not cause a global decrease in gene expression in mature β-cells.
173
174 To explore further, we segregated genes into three groups based on residual enrichment for H3K4
175 methylation, determined using MACS2 [41]. Group (i) retained at least some H3K4me3 enrichment at the
176 transcription start site (TSS) (10,045 genes), group (ii) lost H3K4me3 but retained at least some H3K4me1
177 enrichment (2,302 genes), and group (iii) lost H3K4me3 and H3K4me1 (111 genes) (Fig 3E). As a group,
178 genes that retained at least some enrichment for H3K4me3 at the TSS were not generally downregulated
179 in *Dpy30*-KO cells (Fig 3F). Loss of both H3K4me3 and H3K4me1 from the TSS was linked to significant
180 impairment of gene expression, and genes that lost H3K4me3 but retained H3K4me1 showed an
181 intermediate phenotype (Fig 3F). To visualize this relationship in a threshold-free manner, we plotted the
182 change in H3K4me3, H3K4me1, and RNA expression for all active TSSs (Fig 3G). The majority of genes
183 clustered in the bottom left quadrant, indicating some reduction of both H3K4me3 and H3K4me1 from
184 the TSS. Strikingly, genes that were highly downregulated (red) clustered at the extreme edge of the
185 distribution, indicating that genes experiencing the greatest relative loss of both H3K4me3 and H3K4me1
186 were highly downregulated (Fig 3G). Together, these data support a role for H3K4me3 in contributing to

active gene expression: while the genome-wide reduction of H3K4me3 did not influence total cellular mRNA content, excess loss of H3K4me3 on a per-gene basis leads to downregulation of that gene.

H3K4me3 enforces expression of weakly active and repressed genes in mature beta cells.

We then compared the chromatin profile of downregulated and stably expressed genes by measuring H3K27me3, a marker of developmentally repressed chromatin, and H3K27ac, a marker of active chromatin, using ChIP-seq. The downregulated gene set had a weakly active profile even in WT cells: enrichment for H3K4me3 and H3K27ac were low and H3K4me1 and H3K27me3 were high in comparison to stably expressed genes (Fig 4A-B). Accordingly, the downregulated gene set showed lower absolute expression in WT cells than stably expressed genes (Fig 4C). The DNA sequence of downregulated genes also tended to be relatively GC rich (Fig 4E). Imprinted genes, comprising genomic loci for which one allele is silenced by DNA methylation, tended to be downregulated in *Dpy30*-KO cells (Fig 4F), raising the possibility that H3K4me3 impedes *de novo* DNA methylation at some genes in mature β-cells, as it does in mouse embryonic stell cells [42]. To summarize, genes that lose an especially large fraction of H3K4 methylation tend to be downregulated in *Dpy30*-KO cells; they tend to already be in a weakly active state in WT β-cells and may be more susceptible to silencing in the absence of H3K4 methylation.

H3K4me3 deficiency leads to a less active and more repressed epigenome in mature beta cells.

Our finding that H3K4 methylation is especially important for maintaining expression of weakly expressed genes is intriguing in light of a recent report by Dobrinić *et al.* (ref [43]) that demonstrated Polycomb-mediated gene repression especially constrains weakly expressed genes [43]. The trithorax group was originally defined as antagonistic to Polycomb-mediated repressive activity [44]. During development, repressive (H3K27me3) and active (H3K4me3) states are dynamically regulated and resolution of the genome into active and repressive domains is fundamental to the establishment of terminal cell-type-specific transcriptomes [45]. As such, it is notable that expression of genes that are dynamically regulated during, and important for, β-cell maturation [46] tend to be lower in *Dpy30*-KO cells (Fig 4F). This led us to ask whether H3K27me3 expands in H3K4me3-deficient β-cells. As shown in Fig 5A, the large-scale distribution of H3K27me3-positive compartments is unchanged in *Dpy30*-KO chromatin. H3K7me3 enrichment in 10-kb windows across the genome remained similar between *Dpy30*-KO and -WT (Fig 5B), together indicating that global reduction of H3K4me3 and H3K4me1 did not result in expansion of Polycomb-repressed compartments. However, H3K27me3 enrichment specifically increased at promoters of genes that were downregulated in *Dpy30*-KO cells (Fig 5C-E). These data indicate that genome-wide H3K27me3 organization is largely insensitive to reduction of H3K4 methylation in this model; however, maintenance of H3K4 methylation is necessary to prevent targeted Polycomb gene repression in mature β-cells.

In contrast to the locus-specific accumulation observed of H3K27me3, H3K27ac, a marker of active chromatin, was diminished throughout the genome in *Dpy30*-KO cells (Fig 5B). Like H3K4me3, H3K27ac was reduced even at promoters of genes that were stably expressed or upregulated in *Dpy30*-KO cells (Fig 5D), indicating that the generalized decrease in H3K27ac was not a consequence of gene downregulation. Despite the global reduction, however, H3K27ac was still positively associated with gene expression in *Dpy30*-KO chromatin: H3K27ac peaks gaining intensity were near upregulated genes, and H3K27ac peaks losing intensity were near downregulated genes, more frequently than expected by

230 chance (Fig S3). These observations are consistent with a model wherein acetylation of H3K27 is partially
231 a consequence of H3K4 methylation [14] and of transcription [47].
232
233 **Reduction to transcriptional consistency in *Dpy30*-KO beta cells.**
234 H3K4me3 enrichment has also been linked with transcriptional consistency—that is, low variability, or
235 "noise", in the expression of a given gene between cells [48,49]. A related metric, transcriptional entropy, is
236 a measure of the "specialization" of a cell's transcriptome, where a low entropy describes a narrow
237 distribution of highly expressed genes in each cell, versus a high entropy where gene expression becomes
238 less predictable [50]. In islets, transcriptional entropy decreases during maturation [50] and increases in T2D [4].
239 As such, we asked whether reduction of H3K4 methylation impairs transcriptional consistency in mature
240 β-cells. To measure transcriptional consistency, we processed islets from a WT and *Dpy30*-KO mouse for
241 single cell RNA-seq 45-days after tamoxifen administration, generating 2,510 WT and 2,597 KO cell
242 transcriptomes that passed quality filtering and that were identified as β-cells based on unsupervised
243 clustering and high expression of *Ins1 and Ins2* (Fig S4A). A population of β-cells expressing *tdTomato*,
244 indicating that they were not Cre-recombined, formed a distinct cluster from recombined *GFP*-positive β-
245 cells from the *Dpy30*-KO mouse (Fig 6A). To model divergence of *GFP*-positive from *tdTomato*-positive β-
246 cells (hereafter, mG and mT cells, respectively), we performed pseudotime ordering using the mT cluster
247 as the ground state (Fig 6B). As seen in Fig 6C, *Ins1* and *Ins2* expression became notably more stochastic
248 over pseudotime. To quantify variability of global transcription, we generated transcriptomic entropy
249 scores for each β-cell. In *Dpy30*-KO β-cells, there is an overall increase in transcriptional entropy along
250 pseudotime (Fig 6D) and in mG cells compared to mT cells (Fig 6E). This supports a role for H3K4
251 methylation in maintaining high transcriptional consistency in mature β-cells. Since higher entropy is a
252 feature of immature β-cells [50], we examined expression of genes associated with immaturity and
253 dedifferentiation and again did not observe any induction of these genes in Dpy30-KO cells (*c.f.* Figs 4F &
254 S4B-C). Therefore, *Dpy30*-KO in mature β-cells led to a more stochastic gene expression profile but did
255 not alter cell fate or re-activate an immature cell state.
256
257 We then asked whether the loss of transcriptional consistency was linked to the initial breadth of
258 the H3K4me3 peak, since it has been suggested that H3K4me3 peaks that span an especially large stretch
259 of DNA (i.e. broad peaks) may enforce high transcriptional consistency of associated genes [48,49]. We
260 hypothesized that genes with broader H3K4me3 peaks are more susceptible to increasingly stochastic
261 expression when H3K4me3 levels are reduced, compared to other genes with narrower peaks. To test
262 this, we ranked H3K4me3 peaks in WT cells in order of increasing breadth, grouped them into 20 quantiles,
263 each representing 5% (622) of H3K4me3 peaks with an associated gene. We then compared the variance
264 in expression, normalized to absolute expression level, of genes in each quantile between mT and mG
265 cells. Overall transcriptional variability between cells was increased in mG cells (Fig 6F). There was a
266 positive relationship between H3K4me3 peak breadth and the change in transcriptional variability, where
267 greater H3K4me3 peak breadth was linked to greater gain of variability in *Dpy30*-KO cells (Fig 6G). This
268 supports a role for H3K4me3 in constraining transcriptional variability in β-cells, which is positively linked
269 to peak breadth. Intriguingly, a similar pattern emerges in β-cells from human donors with T2D. Analysis
270 of publicly available datasets for H3K4me3 ChIP-seq from healthy human β-cells (GSE50386, ref. [51]) and
271 scRNA-seq from human donors with or without T2D (GSE124742, ref. [52]) showed that variability of gene
272 expression was higher in β-cells from donors with T2D (Fig 5H) and that the gain in variability is positively
273 correlated with initial H3K4me3 peak breadth (Fig 5I).

H3K4me3 peak breadth stratifies genes dysregulated in type 2 diabetes.

We further explored the relationship between H3K4me3 peak breadth, gene expression, and diabetes. H3K4me3 forms 18,936 distinct peaks with a median breadth of 1.2 kb in WT mouse β-cells. Ranking H3K4me3 peaks by breadth revealed a class of exceptionally broad peaks spanning up to 21 kb, including those at essential β-cell-enriched transcription factor genes (Fig 7A), consistent with the idea that genes important for maintaining cell identity are marked with broad H3K4me3 peaks [48,53]. As expected for repressed genes, the majority (29 of 39, or 74%) of islet "disallowed genes" [54], meanwhile, are H3K4me3-negative in healthy β-cells. Gene expression was moderately correlated with promoter H3K4me3 peak breadth (Spearman rank correlation R = 0.4287); however, genes with exceptionally broad peaks did not have exceptionally high RNA expression (Fig S5A). For example, Fig 7B shows the H3K4me3 profiles of a housekeeping gene with a typical H3K4me3 peak (*Rplp0*), and an expression-matched β-cell-enriched transcription factor with a broad peak (*Nkx6-1*). Therefore, β-cell lineage-enriched genes are marked by broad H3K4me3 peaks.

To explore the relationship between H3K4me3 peak breadth and transcriptome remodeling associated with diabetes, we turned to a mouse model of T2D. Leptin receptor-deficient $Lepr^{db/db}$ mice exhibit overfeeding, obesity, fasting hyperglycemia, and glucose intolerance, along with β-cell dysfunction [55]. We retrieved a published bulk RNA-seq dataset from $Lepr^{db/db}$ and $Lepr^{+/+}$ islets [56] and compared H3K4me3 peak breadth with gene expression changes. Interestingly, genes that become downregulated in $Lepr^{db/db}$ islets were explicitly enriched in the group of genes with broad H3K4me3 peaks (Fig 7C, left). Furthermore, housekeeping genes (GSEA:M11197, ref [57]), genes involved in endocrine pancreas development (GSEA:M12875), and maturity onset diabetes of the young (MODY) genes (GSEA:M18312) were enriched in the broadest H3K4me3 peaks (Fig S5B). In contrast, genes that are upregulated in $Lepr^{db/db}$ islets were uniquely enriched in the narrowest H3K4me3 peaks (Fig 7C, right). As H3K4me3 peak breadth was positively correlated with expression level (Fig S5A), we tested whether these observations could be secondary to a relationship wherein highly expressed genes become downregulated, and weakly expressed genes become upregulated, in $Lepr^{db/db}$. Ranking genes by RNA expression level instead of H3K4me3 peak breadth proved less effective at stratifying each gene set except housekeeping genes (Fig S5B), supporting a primary link with H3K4me3. Therefore, genes that are differentially expressed in $Lepr^{db/db}$ islets are diametrically stratified according to H3K4me3 breadth, with upregulated genes having narrow peaks and downregulated genes having broad peaks (in non-diabetic conditions). These observations are replicable in humans using publicly available H3K4me3 ChIP-seq data of β-cells from donors without diabetes (GSE50386, ref [51]) and RNA-seq data of islets from donors with and without T2D (GSE50398, ref [58]) (Fig S5C-E).

H3K4me3 peak breadth dynamics encode gene expression changes in a mouse model of type 2 diabetes.

The positive association between H3K4me3 peak breadth and differential gene expression in T2D prompted us to ask if H3K4me3 is altered in diabetic islets. Immunoblots showed that, overall, H3K4me3 levels were not altered in islets from $Lepr^{db/db}$ mice (Fig 7D). Calibrated ChIP-seq similarly showed that average enrichment for H3K4me3 across all active promoters is remarkably similar between $Lepr^{+/+}$ and $Lepr^{db/db}$ islets (Fig 7E). While we did not find evidence for a global shift in H3K4me3 levels, we observed contraction of H3K4me3 at downregulated genes, for example the transcription factor *Pdx1* and zinc transporter *Slc30a8*, and expansion of H3K4me3 at upregulated genes, for example the disallowed

318 *Aldh1a3* and *Ldha* (Fig 7F, Fig S5F). In general, broad and narrow H3K4me3 peaks showed significant
319 tendency to shrink or expand, respectively (Fig 7G). Similarly, genes that were transcriptionally up- or
320 downregulated in *Lepr*$^{db/db}$ islets showed corresponding expansion or contraction of HK4me3 levels,
321 respectively (Fig 7H). Therefore, a pattern emerges in *Lepr*$^{db/db}$ islets wherein weakly active genes tend to
322 gain H3K4me3 and become upregulated, while genes with broad H3K4me3 peaks tend to lose some
323 H3K4me3 and become downregulated.
324
325 Finally, we asked whether genes that are dysregulated in *Lepr*$^{db/db}$ islets are regulated by H3K4
326 methylation, i.e., are downregulated in the day-45 *Dpy30*-KO RNA-seq. To test the overlap between
327 differential gene expression in the *Dpy30*-KO and *Lepr*$^{db/db}$ models, we used a threshold-free rank-rank
328 hypergeometric overlap test (RRHO) [59]. There was significant overlap between genes downregulated in
329 *Dpy30*-KO mice with genes that were differentially expressed (either up or down) in *Lepr*$^{db/db}$ mice (Fig 7J).
330 Thus, H3K4 methylation helps maintain expression of genes that become dysregulated in the diabetic
331 *Lepr*$^{db/db}$ model. Accordingly, genes downregulated in *Dpy30*-KO mice were enriched in both the narrow
332 and broad H3K4me3 peak groups (Fig 7I-J), thus mirroring the pattern for dysregulated genes in the
333 *Lepr*$^{db/db}$ model (Fig 7C). This observation in *Dpy30*-KO cells is in agreement with observations presented
334 above: that weakly active genes are most downregulated in *Dpy30*-KO (Fig 4), and that β-cell transcription
335 factor genes have broad peaks (Fig 7A) and tend to be downregulated by *Dpy30*-KO (Fig 4F). Together,
336 this supports the idea that differential gene expression between *Lepr*$^{+/+}$ and *Lepr*$^{db/db}$ mice is partially
337 driven by changes in H3K4me3 peak breadth.
338
339 **Discussion**
340 Here we set out to address whether H3K4me3 is required for maintenance of gene expression in mature
341 β-cells. To this end we deleted a core component of the COMPASS H3K4 methyltransferase complexes,
342 *Dpy30*, leading to dramatic genome-wide reduction of H3K4me3 in mouse β-cells. Loss of H3K4
343 methylation led to a less active and more repressed epigenome profile which locally correlated with gene
344 downregulation. Gene expression became more stochastic, a feature also shared with T2D. H3K4me3-
345 dependent gene expression meaningfully contributed to maintenance of mature β-cell function and
346 maturity. H3K4me3-deficient cells showed loss of maturity including reduced glucose stimulated insulin
347 secretion, reduced expression of terminal markers, and higher transcriptome entropy. These analyses
348 support a role for H3K4me3 in maintaining β-cell gene expression and function.
349
350 In *Dpy30*-KO cells, changes to RNA expression were linked to local changes to histone methylation,
351 leading to significant dysregulation of around 5% of genes, but overall cellular RNA levels were not
352 reduced. H3K4 methylation may, therefore, be important for allocation of transcriptional machinery to
353 particular genes, rather than their absolute activity. In other words, promoter-associated H3K4me levels
354 modulate the expression of target genes but global H3K4me3 or H3K4me1 levels should not be taken as
355 a reflection of a cell's transcriptional output. We note, however, that total mature RNA content, which we
356 measured here, is not a direct measure of transcriptional activity, and measurement of nascent RNA may
357 have uncovered more widespread roles in transcription regulation.
358
359 Hinting at more general roles in transcription is the observation that *Dpy30*-KO β-cells displayed
360 a more stochastic gene expression profile. High gene expression variability between cells in a tissue is a
361 beneficial feature in some circumstances – it can increase adaptability to stress and can reflect a high

362 degree of cell subtype specialization. In other circumstances, high variability is associated with aging and
363 disease [60–63], and we find that it is increased in β-cells from donors with T2D. Very broad H3K4me3 peaks
364 are proposed to confer high transcriptional stability [48], and experimental perturbation of H3K4me3 using
365 *Dpy30*-KO supports this idea. Further, transcription of genes with H3K4me3 peaks of any breadth tended
366 to become more variable in *Dpy30*-KO, meaning that H3K4me3 also generally confers consistent mRNA
367 transcription. Defining exactly how H3K4me3 influences transcriptional activity of select genes will require
368 detailed consideration of the larger epigenetic environment and RNA polymerase activity.
369
370 Since H3K4 cannot be simultaneously mono- and trimethylated, enrichment of H3K4me3 causes
371 H3K4me1 to be generally excluded near TSSs. The degree of exclusion relates to the transcriptional output
372 of the TSS, where weakly expressed genes display closer infiltration of H3K4me1 [64]. In *Dpy30*-KO cells,
373 replacement of H3K4me3 with H3K4me1 in the TSS was associated with transcriptional downregulation,
374 consistent with a report that promoter-associated H3K4me1 is repressive [65]. Surprisingly, replacement of
375 H3K4me3 with H3K4me1 buffered gene downregulation compared to complete demethylation,
376 suggesting that promoters with H3K4me1 are not repressed but simply less active, and H3K4me1 can
377 partially compensates for loss H3K4me3 in promoters. This mirrors a finding that enhancers can function
378 with H3K4me3 instead of H3K4me1 [13]. These results argue against an immutable functional distinction
379 between H3K4me1 and H3K4me3, at least in this model.
380
381 One of the most surprising findings is that promoters in a weakly active state, with little H3K4me3,
382 show the greatest reliance on H3K4 methylation for the maintenance of gene expression. It. A possible
383 explanation is that these promoters are particularly targeted for epigenetic silencing [66] since H3K4
384 methylation sterically hinders repressive H3K27 and DNA methyltransferase activity [42,67]. Genes that were
385 downregulated in *Dpy30*-KO cells showed high initial enrichment of H3K27me3 and further accumulation
386 after loss of H3K4me3. Weakly-active genes are also the most likely to be upregulated after inactivation
387 of the Polycomb system [43] and in β-cells, inactivation of the Polycomb system leads to activation of
388 previously silenced genes [4]. Therefore, weakly active and suppressed genes may be the most responsive
389 targets in ongoing competition between Polycomb and trithorax systems in mature β-cells. Future studies
390 may unravel the degree to which H3K4me3 is an activating versus anti-repressive signal *in vivo*.
391
392 Peak breadth has emerged as an informative dimension of H3K4me3 enrichment: H3K4me3 peak
393 breadth predicts gene expression level [38]; breadth dynamics predict gene expression dynamics during
394 methionine restriction *in vivo* and *in vitro* [68]; and, exceptionally broad H3K4me3 peaks in a given cell type
395 are predictive of genes critical to the maintenance of that cell's lineage [48]. From our analyses of peak
396 breadth, we confirm that terminal markers of β-cells, including *Ins1/2* and mature β-cell enriched
397 transcription factor genes, are marked with exceptionally broad H3K4me3 peaks in mature β-cells. *Dpy30*-
398 KO led to nonspecific reduction of H3K4me3 genome-wide but transcriptional downregulation was largely
399 restricted to genes that originally had either very narrow, or very broad, H3K4me3 peaks, indicating that
400 these two classes of promoters in particular rely on H3K4me3 for the maintenance of gene expression.
401 Compared to the effect at narrow peaks described above, the transcriptional defects for genes with broad
402 H3K4me3 enrichment was more modest. For example, *Ins1* and *Ins2* were downregulated by 15 to 25%,
403 remaining the two most highly expressed genes in *Dpy30*-KO cells despite suffering complete loss of their
404 broad H3K4me3 peaks. Since broad peaks colocalize with many other activating chromatin marks [48,53], the
405 modest effect likely results from compensation from other regulatory inputs. In any case, our results show

that broad H3K4me3 increases expression of associated genes but is not strictly required for their transcription.

The *Dpy30*-KO model highlighted the central importance of relative H3K4me3 levels on a per-gene basis rather than absolute enrichment; thus, reduction of peak breadth may exert a stronger impairment on gene expression in scenarios where H3K4me3 is not simultaneously reduced at all other genes. To that point, we find that promoters with broad H3K4me3 peaks showed a unique tendency to shrink, and for the associated gene to be downregulated, in *Lepr*$^{db/db}$ islets. Meanwhile, promoters marked with narrow H3K4me3 peaks showed a significant tendency to be upregulated and gain H3K4me3 in *Lepr*$^{db/db}$ islets. It is notable that these two classes of promoters are the same that are most sensitive to H3K4me3 enrichment. More work will be required to determine what drives H3K4me3 peak dynamics in *Lepr*$^{db/db}$ islets and to what extent those mechanisms are active in human T2D. In any case, these results favour a model wherein redistribution of H3K4me3 away from promoters of genes with broad H3K4me3 peaks, that enforce β-cell identity or functions, in favour of relatively inactive, weakly-expressed genes, including developmentally repressed and disallowed genes, contributes to transcriptome remodeling in diabetes.

Methods

Animal strains and care

Mice harboring conditional alleles of *Dpy30* (derived from EUCOMM EM:09575 [21]) and *Rosa26*^{mTmG} (Jax 007576 [29]) were crossed with *Pdx1-CreER*^{Tg} (Jax 024968 [28]) or *Ins1*^{Cre} (Jax 026801 [33]) lines. Genotypes used for *Pdx1-CreER* studies were: KO: *Pdx1-CreER*$^{Tg/0}$; *Dpy30*$^{flox/flox}$; *Rosa26*$^{mTmG/+}$; WT: *Pdx1-CreER*$^{Tg/0}$; *Dpy30*$^{+/+}$; *Rosa26*$^{mTmG/+}$ or *Pdx1-CreER*$^{0/0}$; *Dpy30*$^{flox/flox}$; *Rosa26*$^{mTmG/+}$. Genotypes used for *Ins1*^{Cre} studies were: KO: *Ins1*$^{Cre/+}$; *Dpy30*$^{flox/flox}$; WT: *Ins1*$^{+/+}$; *Dpy30*$^{flox/flox}$; HET: *Ins1*$^{Cre/+}$; *Dpy30*$^{flox/+}$. Mice were maintained on a mixed genetic background. At 8-weeks-old, *Pdx1-CreER* mice were administered 8 mg tamoxifen (Sigma) dissolved in 100 µL corn oil by oral gavage three times, with ~48-hours between administrations. The first day of injection is considered day 0. BKS *Lepr*$^{db/db}$ mice and BKS *Dock7*$^{m/m}$ controls (Jax 000642) were used at 12-weeks-old. Procedures were restricted to male mice and approved by the University of British Columbia Animal Care Committee under certificates A17-0045 and A18-0111. Up to five littermates were housed per cage with temperature and humidity control under a 12-hour light/dark cycle and ad libitum access to water and food (Teklad 2918).

Cell culture

Primary mouse islet cells were maintained in RPMI 1640 medium containing 11 mM glucose (Gibco), 10% (v/v) FBS, 50 U/ml penicillin, 50µg/mL streptomycin (complete RPMI) in 37°C, 5% CO_2 atmosphere. *Drosophila* S2 cells were maintained at ambient temperature and air in Schneider's Drosophila Medium (Gibco) containing 10% (v/v) heat-inactivated FBS, 50 U/ml penicillin, 50µg/mL streptomycin.

Islet and β-cell isolation

Unless otherwise indicated, mouse islets were isolated 45-days after tamoxifen administration (i.e. at 14-weeks old) and processed for analysis immediately after isolation. Pancreata were perfused with 700 U/mL collagenase type XI (Sigma-Aldrich) in Hank's Buffered Saline Solution (138 mM NaCl, 5.3 mM KCl, 0.44 mM KH_2PO_4, 0.34 mM Na_2HPO_4, 5.5 mM D-Glucose, pH 7.3) through the bile duct, excised, and held for 15-minutes at 40°C before manual disruption and hand-picking of islets. To isolate β-cells, islets from *Pdx1-CreER*$^{Tg/0}$ *Rosa26*$^{mTmG/-}$ transgenic mice were washed 3× in room temperature PBS + 2 mM EDTA and

450 then dispersed by constant gentle pipetting in 1 mL of 0.025% Trypsin-EDTA (Thermo Fisher Scientific) in
451 PBS at room temperature for 8-10 minutes. Trypsin was quenched by addition of 9 mL room temperature
452 complete RPMI, cells were centrifuged for 3-minutes at 600 × *g* at 4°C, and resuspended in 0.3-0.5 mL ice-
453 cold sorting buffer (PBS, 2 mM EDTA, 1% (w/v) BSA). After passing through a 70 µm cell strainer, singlet
454 eGFP+ tdTomato- cells were purified by fluorescence-activated cell sorting (FACS) using a FACS-Aria II with
455 a 100 µm nozzle.
456
457 **Chromatin immunoprecipitation sequencing (ChIP-seq)**
458 A native ChIP protocol [69] was used to generate ChIP-seq libraries in biological duplicate, with
459 modifications. For *Pdx1-CreER*$^{Tg/0}$; *Dpy30*$^{+/+}$; *Rosa26*$^{mTmG/+}$ and *Pdx1-CreER*$^{Tg/0}$; *Dpy30*$^{flox/flox}$; *Rosa26*$^{mTmG/+}$
460 mice, 100,000 eGFP+ tdTomato- islet cells were pooled with 50,000 *Drosophila* S2 cells by FACS. For
461 *Lepr*$^{db/db}$ and *Lepr*$^{+/+}$ mice, 100,000 dispersed islet cells were counted by hemacytometer and pooled with
462 50,000 *Drosophila* S2 cells without sorting. Pooled cells were centrifuged at 600 × *g* for 5 minutes at 4°C,
463 supernatant removed, and then cell pellets were flash-frozen and stored at -80°C for up to one month.
464 Upon thawing, cells were permeabilized in nuclear isolation buffer (0.1% Triton X-100, 0.1% sodium
465 deoxycholate) containing 1 mM phenylmethylsulfonyl fluoride (PMSF) and EDTA-free protease inhibitor
466 cocktail (Roche)) on ice. Chromatin was fragmented to mononucleosomes by micrococcal nuclease
467 (MNase) (NEB) in digestion buffer (50 mM Tris pH 7.9, 5 mM $CaCl_2$, 1.5 mM DTT) for 7.5-minutes at 37°C,
468 then quenched by addition of 0.1 volume of a solution containing 2% Triton X-100, 2% sodium
469 deoxycholate, and 100 mM EDTA. Soluble chromatin was pre-cleared with Protein A Dynabeads (Thermo
470 Fisher Scientific) pre-adsorbed with normal rabbit IgG (Millipore, 12-370) for 2-hours at 4°C rotating 9
471 rpm, and then rotated overnight at 4°C with 1 µg of H3K4me3 (Abcam, ab1012), 2 µg H3K4me1 (Abcam,
472 ab8895), 1 µg H3K27ac (Active Motif, 39034), or 2 µg H3K27me3 (Millipore, 07-449) antibodies pre-
473 adsorbed to 10 µL of Protein A Dynabeads (Thermo Fisher Scientific) in 150 µL of ChIP buffer (20 mM Tris
474 pH 8.0, 2 mM EDTA, 150 mM NaCl, 0.2% Triton X-100, 1 mM PMSF, and EDTA-free protease inhibitor
475 cocktail (Roche)). After washing the beads twice with ice-cold wash buffer (20 mM Tris pH 8, 150 mM
476 NaCl, 2 mM EDTA, 1% (v/v) Triton X-100, 0.1% (w/v) SDS) and twice with ice-cold high-salt wash buffer (20
477 mM Tris pH 8, 500 mM NaCl, 2 mM EDTA, 1% (v/v) Triton X-100, 0.1% (w/v) SDS), DNA was eluted in
478 elution buffer (100 mM $NaHCO_3$, 1% (w/v) SDS, 50 µg/mL RNase A) for 1-hour at 55°C, with addition of 1
479 µL 20 mg/mL proteinase K after the first 30-minutes and occasional manual inversion. DNA was purified
480 using standard Phenol:chloroform:isoamyl alcohol extraction and ethanol precipitation and used for
481 library preparation with the NEBNext Ultra II DNA Library Prep Kit for Illumina (NEB) with 9 (H3K4me1,
482 H3K27me3) or 13 (H3K4me3, H3K27ac) amplification cycles. Indexed libraries were analyzed for size
483 distribution using the Bioanalyzer High Sensitivity DNA chip (Agilent) and for concentration using the Qubit
484 dsDNA HS Assay (Thermo Fisher Scientific). *Dpy30*-KO samples were sequenced on the NextSeq 500
485 platform (Illumina) for 2 × 78 nucleotide paired end reads and *Lepr*$^{db/db}$ samples on the NextSeq 2000
486 platform (illumina) for 2 × 61 nucleotide paired end reads. Sequence data is available in the Gene
487 Expression Omnibus (GEO) under accession number GSE181951.
488
489 **ChIP-seq data analysis**
490 ChIP-seq reads were aligned to concatenated mouse (GRCm38/mm10) and *Drosophila* (BDGP6.28)
491 genome assemblies using Bowtie2 v2.3.4.1 with options '--very-sensitive --no-unal --no-discordant' [70].
492 Multi-mapped reads, reads with mapping quality MAPQ < 20, and suspected PCR-duplicates were
493 discarded using Samtools [71]. To scale enrichment according to the *Drosophila* cell spike-ins, the fraction

494 of remaining reads that mapped to the *Drosophila* genome was determined, and a scaling factor was
495 calculated for each sample as *(minimum* Drosophila *read fraction per histone mark)/(sample* Drosophila
496 *read fraction)*. After determining a scaling factor, each sample was downsampled to the lowest mapped
497 read count per each histone mark using Samtools, and then reads mapped to the *Drosophila* genome were
498 removed. For visualization, libraries were converted to bedgraph format using the calculated scaling
499 factor in the '-scale' argument of BEDtools v2.26.0 genomecov [72]. Genome browser views were generated
500 for merged replicates using Spark [73] and TSS profiles using Deeptools v3.4.2 [74]. Peak locations and breadth
501 were determined using MACS2 v2.2.6 with parameters '--broad --nolambda' [41]. Peaks detected in only
502 one biological replicate or that overlapped with high-background blacklisted regions defined by the
503 ENCODE project [75] were excluded. Chromosomes X, Y, and M were also excluded from the analyses.
504 DESeq2 [76] was used for differential peak width analysis using a significance cutoff of $P \leq 0.05$ (Wald test
505 with Benjamini and Hochberg adjustment), and enrichment for differential breadth in peak quantiles was
506 determined by one-sided Fisher's exact test. We defined TSS locations as the 5' end of the most abundant
507 annotated transcript of each gene detected in our RNA-seq data. H3K4me3 peaks were assigned to genes
508 for which they overlap the region ±1 kb of the TSS.
509
510 **RNA-seq**
511 eGFP+ tdTomato- β-cells (70,000-226,000 per mouse, biological triplicate) were purified and counted by
512 FACS and supplemented with a 10% spike-in of *Drosophila* S2 cells. *Drosophila* cells were included to
513 detect global changes in cell RNA content [77]. The cells were sorted directly into Trizol LS Reagent (Thermo
514 Fisher Scientific) and total RNA was extracted according to manufacturer's instructions. Residual DNA was
515 digested by Turbo DNase (Thermo Fisher Scientific) for 30-minutes at 37°C. Purified total RNA content was
516 measured using the Qubit RNA HS assay (Thermo Fisher Scientific). Mature mRNA was enriched from 400
517 ng of total RNA per sample using the NEBNext Poly(A) mRNA Magnetic Isolation Module (NEB) and used
518 to prepare libraries with the NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (NEB) with 11
519 amplification cycles. Uniquely indexed libraries were quantified, pooled, and sequenced for 2 × 38
520 nucleotide reads as described above.
521
522 **RNA-seq data analysis**
523 Sequenced reads were aligned to the concatenated mouse (GRCm38/mm10; Gencode vM24) and
524 *Drosophila* (BDGP6.28; Ensembl release 99) genomes and transcriptomes using STAR v2.7.3a [78] with
525 default settings. Transcript abundance was calculated using Salmon v1.4.0 [79] in alignment mode with
526 seqBias and gcBias options. Gene counts and differential expression were calculated using DESeq2 [76].
527 Genes with fewer than two counts in any sample were filtered out and genes showing ≥ 2-fold difference
528 in expression with $P \leq 0.01$ (Wald test with Benjamini and Hochberg adjustment for multiple comparisons)
529 are considered differentially expressed. Genes showing <1.1 fold difference in expression are considered
530 stable. Except in Fig. 6F, Drosophila transcripts were removed after abundance calculation; for Fig. 6F,
531 *Drosophila* transcripts were used as control genes in the estimateSizeFactors function of DESeq2, and then
532 removed.
533
534 **Single-cell RNA-seq**
535 Islets from one WT and one KO mouse were dissociated to single-cell suspensions as described above
536 without FACS-enrichment. The cell suspensions were processed through the Chromium Single Cell 3'
537 protocol using the Chromium Controller (firmware v4.0) with Reagent Kit v3.1 (10X Genomics) and Dual

538 Index Kit TT Set A (10X Genomics) according to the manufacturer's instructions for a targeted 5,000-cell
539 recovery. Total cDNA was amplified for 11-cycles, and then 13-cycles during index-ligation. cDNA
540 concentrations were quantified by qPCR with the NEBNext Library Quant Kit for Illumina (NEB). Libraries
541 were pooled and sequenced for 28-90 paired-end nucleotide reads to an average depth of ~70,000
542 mapped fragments per cell.
543
544 **Single cell RNA-seq data analysis**
545 Cellranger (v5.0.0, 10X Genomics) was used to generate FASTQ files and demultiplex reads. Cell barcode
546 detection, read mapping, quality filtering, and transcript counting were performed using Alevin (Salmon
547 v1.4.0) [80] against the mouse protein-coding transcriptome (Gencode vM24), which we appended with the
548 *Tdtomato* and *Egfp* sequences, as well as the mouse genome (GRCm38/mm10) as mapping decoys for
549 selective alignment [81]. Additional cell quality filtering was performed using Seurat v4.0 [82] based on unique
550 features (> 1,000 genes) and mitochondrial RNA content (< 25%). Putative doublets were excluded using
551 DoubletFilter [83]. Genes detected in < 3 cells were excluded from analysis. After quality filtering, the dataset
552 comprised 3,703 WT and 3,641 KO cells. Filtered data were scaled and normalized using SCTransform [84]
553 with default parameters. PCA and UMAP dimensionality reduction using the first 60 principal components
554 was used to cluster and visualize cell populations. We then used Slingshot [85] to perform pseudotime
555 analysis on β-cell clusters. Entropy scores are a modified Shannon entropy calculated using a published R
556 function (ref [86]).
557
558 **Immunohistochemistry**
559 Mouse pancreata were isolated and fixed in 4% (w/v) paraformaldehyde (PFA) in PBS at 4°C overnight.
560 Following washing with PBS and dehydration in ethanol and xylenes, mouse tissues were embedded in
561 paraffin and sectioned with a thickness of 5 µm. Immunofluorescent analyses of tissue sections was
562 performed as previously described[23] in a blinded fashion from 4-6 sections per biological sample. Primary
563 antibodies used were: H3K4me3 (Cell Signaling Technologies, C42D8, 1:500), H3K4me1 (Abcam, ab8895,
564 1:1,000), H3 (Abcam, ab1791, 1:1,000), DPY30 (Atlas Antibodies, HPA043761, 1:500), Insulin (Agilent, IR-
565 002, 1:4), and Glucagon (Sigma Aldrich, G2654, 1:1,000).
566
567 **Transmission electron microscopy**
568 Islets were fixed in 2% glutaraldehyde (Sigma Aldrich) in PBS and submitted to the Facility for Electron
569 Microscopy Research of McGill University (Hamilton, ON, Canada). Islets were post-fixed with 1% osmium
570 tetroxide, dehydrated in ethanol, and embedded in Spurr's resin before sectioning with a Leica UCT
571 ultramicrotome. Sections were stained with uranyl acetate and lead citrate and imaged with a JEOL JEM
572 1200 EX TEMSCAN transmission electron microscope (JEOL, Peabody, MA, USA).
573
574 **Immunoblotting**
575 Islets were lysed in Laemmli sample buffer (2% (w/v) SDS, 10% (v/v) glycerol, 60 mM Tris pH 6.8, 1mM
576 NaF, 1 mM PMSF, protease inhibitor cocktail (Roche)) for 5-minutes at 95°C. Protein concentration was
577 measured by BCA assay (Thermo Fisher Scientific), then β-mercaptoethanol was added to 5% (v/v) and
578 lysates were reboiled for 5-minutes and stored at -80°C. 10 µg lysates were resolved in polyacrylamide
579 gels, transferred to PVDF membranes, blocked for 60-minutes in 5% (w/v) skim milk powder in TBST (20
580 mM Tris pH 7.6, 150 mM NaCl, 0.1% (v/v) Tween-20) at room temperature and probed with primary
581 antibodies overnight in 5% (w/v) BSA in TBST at 4°C. The next day, after washing membranes three times

582 with TBST, HRP-linked secondary antibodies were added for 1-hour at room temperature, membranes
583 were washed three times with TBST, and detected using ECL system with film. When normalizing to total
584 H3, membranes were stripped for 20-minutes in mild stripping buffer (200 mM glycine, 0.1% (w/v) SDS,
585 1% (v/v) Tween-20, pH 2.2) at room temperature, rinsed twice in PBS and twice in TBST for 10 minutes
586 each at room temperature, then re-blocked and probed for histone H3. Primary antibodies used for
587 immunoblots were: H3K4me3 (Cell Signaling Technologies, C42D8, 1:1,000), H3K4me1 (Abcam, ab8895,
588 1:100,000), H3K27ac (Cell Signaling Technologies, D5E4, 1:1,000), H3K27me3 (Millipore, 07-449,
589 1:100,000), H3 (Abcam, ab1791, 1:1,000,000), WDR5 (Bethyl, A302-430A, 1:2,000), RBBP5 (Bethyl, A300-
590 109A, 1:10,000), and ASH2L (Bethyl, A300-489A, 1:10,000).
591
592 **Protein co-immunoprecipitation**
593 Immunoprecipitation of WDR5 was performed in triplicate in 750-900 dispersed islets pooled from 3 mice.
594 First, nuclei were enriched from dispersed islet cells by incubation in nuclear isolation buffer for 10-
595 minutes on ice and pelleting at 600 × *g* for 5-minutes at 4°C. Chromatin was solubilized in IP buffer (20
596 mM Tris pH 7.4, 150 mM NaCl, 1 mM EDTA, 5% (v/v) glycerol, 0.5% (v/v) Igepal ca-630, 5 mM $CaCl_2$, 1 mM
597 PMSF, protease inhibitor cocktail (Roche)) with MNase (NEB) for 10-minutes at room temperature. EDTA
598 was added to 5 mM to stop MNase and samples were centrifuged at 14,000 × *g* for 5-minutes at 4°C. The
599 protein concentration in the supernatant was measured by BCA assay (Thermo Fisher Scientific), and an
600 equal mass of WT and KO nuclear lysate (~30 µg protein) was immunoprecipitated by 1 µg WDR5 (Bethyl,
601 A302-429A) or normal rabbit IgG (Millipore, 12-370) antibodies rotating overnight at 4°C. The next day,
602 10 µL protein A Dynabeads (Thermo Fisher Scientific) was added to the antibody:lysate mixture and
603 rotated for 2-hours at 4°C. Finally, bead:antibody:antigen complexes were washed six times in ice-cold IP
604 buffer and resuspended in Laemmli sample buffer with 5% (v/v) β-mercaptoethanol for immunoblotting.
605
606 **Insulin secretion assays**
607 For in vivo glucose tolerance tests, 2 g/kg glucose in a 20% (w/v) water solution was injected
608 intraperitoneally following a 6-hour fast. Blood glucose was sampled immediately prior to and 15, 30, 60,
609 and 120-minutes after glucose injection from a lateral saphenous vein. Serum was collected at the 0, 15,
610 and 60-minute time points by centrifuging blood samples at 9,000 × g for 9 minutes at 4°C and stored at -
611 80°C. Serum insulin was measured by ELISA (Alpco) according to the manufacturer's instructions. Ex vivo
612 static insulin secretion assays were performed in technical duplicate in groups of 50 islets per mouse one
613 day after isolation. Islets were preincubated in KRBH buffer (2.8 mM glucose, 20 mM HEPES, 114 mM
614 NaCl, 4.7 mM KCl, 1.2 mM $MgSO_4$, 1.2 mM KH_2PO_4, 2.5 mM $CaCl_2$, 24 mM $NaHCO_3$, 0.1% (w/v) BSA, pH
615 7.3) for one hour at 37°C and then sequentially for 45-minutes in KRBH, either 16.7 mM glucose or 15 mM
616 dimethyl α-ketoglutarate (Sigma Aldrich), and finally 30 mM KCl. Conditioned medium was collected after
617 each incubation. At the completion of the assay islets were lysed in RIPA buffer (Pierce) containing 1 mM
618 PMSF and EDTA-free protease inhibitor cocktail (Roche)) at 95°C for 5 minutes. Insulin concentration was
619 measured in conditioned medium and islet lysate by ELISA (Alpco). DNA concentration of islet lysate was
620 determined by Qubit dsDNA HS Assay (Thermo Fisher Scientific) to normalize islet insulin content.
621
622 **Calcium imaging**
623 Intact islets were cultured for two days to allow for attachment onto glass coverslips. Prior to imaging,
624 coverslips were transferred to a 2 mL imaging chamber and stained with 5 µM Fura-2AM (Molecular
625 Probes, F1221) for 30-minutes at 37°C in complete RPMI. The imaging chamber loaded with a coverslip

626 was then mounted on a Leica SP8 Laser Scanning Confocal Microscope with a 10× objective lens and
627 continuously washed with Ringer's solution (5.5 mM KCl, 2 mM $CaCl_2$, 1 mM $MgCl_2$, 20 mM HEPES, 144
628 mM NaCl, 2.8 mM glucose, pH 7.4) for 30-minutes. Throughout the experiment, islets were continuously
629 perifused at a flow rate of 2.5 mL/min with Ringer's solution. Glucose and KCl concentrations were
630 adjusted by iso-osmotic substitution of NaCl where indicated. Fura-2 was excited at 340 nm and 380 nm
631 and emitted fluorescence was detected at approximately 510 nm. Cytosolic Ca^{2+} levels are expressed as
632 the mean ratio of Fura-2 fluorescence emission intensity (F_{340}/F_{380}) of 3-9 islets per biological replicate.
633
634 **Respirometry**
635 Islets were dissociated to single cells 3-4 hours after isolation and plated at 100 islets/well in a Seahorse
636 24-well cell culture dish in technical duplicate. Cells were cultured for 2-days in complete RPMI to allow
637 for attachment to the culture dish and then for 1-hour in XF RPMI medium (Agilent) with 2.8 mM glucose,
638 1% FBS, 1 mM pyruvate, and 2 mM glutamine to reach a metabolic baseline. Oxygen consumption was
639 then measured using a Seahorse XF24 Extracellular Flux Analyser (Agilent). The injection ports were
640 loaded with the following compounds in XF Assay Media (Agilent): A, 167 mM glucose; B, 10 µM
641 oligomycin; C, 15 µM FCCP; D, 10 µM antimycin A and 1 µM rotenone. Measurements were normalized
642 to the DNA content of assayed cells, which were determined using Qubit HS DNA Assay (Thermo Fisher
643 Scientific) after lysing in RIPA buffer upon completion of the assay.
644
645 **Analysis of publicly available data**
646 Publicly available data were reprocessed for consistency. GEO: GSE50244, GSE50386, GSE107489,
647 GSE124742 were downloaded from the Sequence Read Archive (SRA) using the sratoolkit
648 (https://github.com/ncbi/sra-tools) and analyzed according to the experimental application, as described
649 above.
650
651 **Statistics**
652 Unless otherwise stated, plots show mean ± standard deviation (SD) with points representing each
653 biological replicate. In boxplots, the central horizontal line marks median, the upper and lower limits of
654 the box the first and third quartiles, and whiskers span 1.5× the interquartile range, in the style of Tukey.
655 Spearman's rank coefficient was used to estimate correlation. Regression lines were generated using a
656 generalized additive model. *P*-values were calculated using Student's t-tests with Welch's correction,
657 Wilcoxon rank sum test, Wald test, mixed-effect model, one- or two-way ANOVA, Fisher's exact test, or
658 permutation test, as indicated. Correction for multiple comparison using the Benjamini-Hochberg method
659 was applied where indicated. Statistical calculations were performed in R or Prism 8 or 9 (GraphPad
660 Software, La Jolla, CA, USA). $P < 0.05$ was deemed significant.
661
662 **Author contributions**
663 Experiments were conducted by B.V., N.M., D.J.P., M.A., and C.L.M, and conceptualized by B.V. and B.G.H.
664 D.J.P. and D.S.L. performed calcium imaging experiments and designed respirometry experiments.
665 Bioinformatics were performed by B.V. and B.G.H. Funding and supervision was provided by B.G.H. and
666 F.C.L; the manuscript was written by B.V. B.G.H. is the guarantor of this work and, as such, had full access
667 to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the
668 data analysis.
669

670 **Declaration of interests**
671 All authors declare they have no competing interests.
672
673 **Acknowledgements**
674 The authors thank the staff of the Animal Care Facility at British Columbia Children's Hospital Research
675 Institute for daily maintenance of the mouse colonies, and W. Gibson (UBC) for providing Ins1^{Cre} animals.
676 We thank M. Reid at the Facility for Electron Microscopy Research of McGill University for help with
677 transmission electron microscopy, L. Xu at the Flow Core of British Columbia Children's Hospital Research
678 Institute for help with FACS. This work was supported by operating grants from the British Columbia
679 Children's Hospital Research Institute, the Natural Sciences and Engineering Research Council of Canada
680 (RGPIN-2016-04292) and the Canadian Institute for Health Research (RN310864 – 375894). Fellowship
681 support for B.V. was provided by the British Columbia Children's Hospital Research Institute, University of
682 British Columbia, and Natural Sciences and Engineering Research Council of Canada.
683
684 **References**
685 1. Saeedi, P. *et al.* Global and regional diabetes prevalence estimates for 2019 and projections for
686 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9th edition.
687 *Diabetes Res. Clin. Pract.* **157**, 1–10 (2019).
688 2. Salinno, C. *et al.* β-Cell Maturation and Identity in Health and Disease. *Int. J. Mol. Sci.* **20**, 1–20
689 (2019).
690 3. Kim, H. *et al.* PRMT1 is required for the maintenance of mature β-cell identity. *Diabetes* **69**, 355–
691 368 (2020).
692 4. Lu, T. T. H. *et al.* The Polycomb-Dependent Epigenome Controls β Cell Dysfunction,
693 Dedifferentiation, and Diabetes. *Cell Metab.* **27**, 1294-1308.e7 (2018).
694 5. Dhawan, S., Georgia, S., Tschen, S. ing, Fan, G. & Bhushan, A. Pancreatic β Cell Identity Is
695 Maintained by DNA Methylation-Mediated Repression of Arx. *Dev. Cell* **20**, 419–429 (2011).
696 6. Dayeh, T. *et al.* Genome-Wide DNA Methylation Analysis of Human Pancreatic Islets from Type 2
697 Diabetic and Non-Diabetic Donors Identifies Candidate Genes That Influence Insulin Secretion.
698 *PLoS Genet.* **10**, e1004160 (2014).
699 7. Deering, T. G., Ogihara, T., Trace, A. P., Maier, B. & Mirmira, R. G. Methyltransferase set7/9
700 maintains transcription and euchromatin structure at islet-enriched genes. *Diabetes* **58**, 185–193
701 (2009).
702 8. Kuo, T. *et al.* Identification of C2CD4A as a human diabetes susceptibility gene with a role in β cell
703 insulin secretion. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 20033–20042 (2019).
704 9. De Jesus, D. F. *et al.* m6A mRNA methylation regulates human β-cell biology in physiological
705 states and in type 2 diabetes. *Nat. Metab.* **1**, 765–774 (2019).
706 10. Miguel-Escalada, I. *et al.* Human pancreatic islet three-dimensional chromatin architecture
707 provides insights into the genetics of type 2 diabetes. *Nat. Genet.* **51**, 1137–1148 (2019).
708 11. Jenuwein, T. & Allis, C. D. Translating the Histone Code. *Science (80-.).* **293**, 1074–1080 (2001).
709 12. Barski, A. *et al.* High-Resolution Profiling of Histone Methylations in the Human Genome. *Cell*
710 **129**, 823–837 (2007).
711 13. Rickels, R. *et al.* Histone H3K4 monomethylation catalyzed by Trr and mammalian COMPASS-like
712 proteins at enhancers is dispensable for development and viability. *Nat. Genet.* **49**, 1647–1653
713 (2017).
714 14. Dorighi, K. M. *et al.* Mll3 and Mll4 Facilitate Enhancer RNA Synthesis and Transcription from
715 Promoters Independently of H3K4 Monomethylation. *Mol. Cell* **66**, 568-576.e4 (2017).
716 15. Hödl, M. & Basler, K. Transcription in the absence of histone H3.2 and H3K4 methylation. *Curr.*

717 *Biol.* **22**, 2253–2257 (2012).
718 16. Howe, F. S., Fischl, H., Murray, S. C. & Mellor, J. Is H3K4me3 instructive for transcription
719 activation? *BioEssays* **39**, 1–12 (2017).
720 17. Ernst, P. *et al.* Definitive hematopoiesis requires the mixed-lineage leukemia gene. *Dev. Cell* **6**,
721 437–443 (2004).
722 18. Glaser, S. *et al.* The histone 3 lysine 4 methyltransferase, Mll2, is only required briefly in
723 development and spermatogenesis. *Epigenetics and Chromatin* **2**, 1–16 (2009).
724 19. Lee, J. E. *et al.* H3K4 mono- And di-methyltransferase MLL4 is required for enhancer activation
725 during cell differentiation. *Elife* **2013**, 1–25 (2013).
726 20. Bledau, A. S. *et al.* The H3K4 methyltransferase Setd1a is first required at the epiblast stage,
727 whereas Setd1b becomes essential after gastrulation. *Dev.* **141**, 1022–1035 (2014).
728 21. Campbell, S. A., McDonald, C. L., Krentz, N. A. J., Lynn, F. C. & Hoffman, B. G. TrxG Complex
729 Catalytic and Non-catalytic Activity Play Distinct Roles in Pancreas Progenitor Specification and
730 Differentiation. *Cell Rep.* **28**, 1830-1844.e6 (2019).
731 22. Piunti, A. & Shilatifard, A. Epigenetic balance of gene expression by polycomb and compass
732 families. *Science (80-.).* **352**, (2016).
733 23. Campbell, S. A. *et al.* H3K4 trimethylation is required for postnatal pancreatic endocrine cell
734 functional maturation. *bioRxiv* (2020). doi:https://doi.org/10.1101/2020.11.29.402990
735 24. Scoville, D. W. *et al.* MLL3 and MLL4 methyltransferases bind to the MafA and MAFB
736 transcription factors to regulate islet β-cell function. *Diabetes* **64**, 3772–3783 (2015).
737 25. Sze, C. C. *et al.* Histone H3K4 methylation-dependent and -independent functions of
738 set1A/COMPASS in embryonic stem cell self-renewal and differentiation. *Genes Dev.* **31**, 1732–
739 1737 (2017).
740 26. Jiang, H. *et al.* Role for Dpy-30 in ES Cell-Fate Specification by Regulation of H3K4 Methylation
741 within Bivalent Domains. *Cell* **144**, 513–525 (2011).
742 27. Haddad, J. F. *et al.* Structural Analysis of the Ash2L/Dpy-30 Complex Reveals a Heterogeneity in
743 H3K4 Methylation. *Structure* **26**, 1–10 (2018).
744 28. Gu, G., Dubauskaite, J. & Melton, D. A. Direct evidence for the pancreatic lineage: NGN3+ cells
745 are islet progenitors and are distinct from duct progenitors. *Development* **129**, 2447–2457
746 (2002).
747 29. Muzumdar, M. D., Tasic, B., Miyamichi, K., Li, L. & Luo, L. A Global Double-Fluorescent Cre
748 Reporter Mouse. *Genesis* **45**, 593–605 (2007).
749 30. Taruttis, F. *et al.* External calibration with Drosophila whole-cell spike-ins delivers absolute mRNA
750 fold changes from human RNA-Seq and qPCR data. *Biotechniques* **62**, 53–61 (2017).
751 31. Zhang, J. *et al.* The diabetes gene Hhex maintains δ -cell differentiation and islet function. *Genes*
752 *Dev.* **28**, 829–834 (2014).
753 32. Sabatini, P. V. *et al.* Neuronal PAS Domain Protein 4 Suppression of Oxygen Sensing Optimizes
754 Metabolism during Excitation of Neuroendocrine Cells. *Cell Rep.* **22**, 163–174 (2018).
755 33. Thorens, B. *et al.* Ins1 Cre knock-in mice for beta cell-specific gene recombination. *Diabetologia*
756 **58**, 558–565 (2015).
757 34. Millership, S. J. *et al.* Neuronatin regulates pancreatic β cell insulin content and secretion. *J. Clin.*
758 *Invest.* **128**, 3369–3381 (2018).
759 35. Wijesekara, N. *et al.* Beta cell-specific Znt8 deletion in mice causes marked defects in insulin
760 processing, crystallisation and secretion. *Diabetologia* **53**, 1656–1668 (2010).
761 36. Matschinsky, F. M. & Wilson, D. F. The central role of glucokinase in glucose homeostasis: A
762 perspective 50 years after demonstrating the presence of the enzyme in islets of Langerhans.
763 *Front. Physiol.* **10**, (2019).
764 37. Ernst, J. *et al.* Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature*

765 **473**, 43–49 (2011).
766 38. Karlić, R., Chung, H. R., Lasserre, J., Vlahoviček, K. & Vingron, M. Histone modification levels are
767 predictive for gene expression. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 2926–2931 (2010).
768 39. Choudhury, R., Singh, S., Arumugam, S., Roguev, A. & Stewart, A. F. The Set1 complex is dimeric
769 and acts with Jhd2 demethylation to convey symmetrical H3K4 trimethylation. *Genes Dev.* **33**, 1–
770 15 (2019).
771 40. Clouaire, T. *et al.* Cfp1 integrates both CpG content and gene activity for accurate H3K4me3
772 deposition in embryonic stem cells. *Genes Dev.* **26**, 1714–1728 (2012).
773 41. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, (2008).
774 42. Ooi, S. K. T. *et al.* DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation
775 of DNA. *Nature* **448**, 714–717 (2007).
776 43. Dobrinić, P., Szczurek, A. T. & Klose, R. J. PRC1 drives Polycomb-mediated gene repression by
777 controlling transcription initiation and burst frequency. *Nat. Struct. Mol. Biol.* **28**, 811–824
778 (2021).
779 44. Kennison, J. A. The Polycomb and trithorax group proteins of Drosophila: Trans-Regulators of
780 Homeotic Gene Function. *Annu. Rev. Genet.* **29**, 289–303 (1995).
781 45. Schuettengruber, B., Bourbon, H.-M., Di Croce, L. & Cavalli, G. Genome Regulation by Polycomb
782 and Trithorax: 70 Years and Counting. *Cell* **171**, 34–57 (2017).
783 46. Blum, B. *et al.* Functional beta-cell maturation is marked by an increased glucose threshold and
784 by expression of urocortin 3. *Nat. Biotechnol.* **30**, 261–264 (2012).
785 47. Martin, B. J. E. *et al.* Transcription shapes genome-wide histone acetylation patterns. *Nat.*
786 *Commun.* **12**, 1–9 (2021).
787 48. Benayoun, B. A. *et al.* H3K4me3 breadth is linked to cell identity and transcriptional consistency.
788 *Cell* **158**, 673–688 (2014).
789 49. Collins, B. E., Sweatt, J. D. & Greer, C. B. Broad domains of histone 3 lysine 4 trimethylation are
790 associated with transcriptional activation in CA1 neurons of the hippocampus during memory
791 formation. *Neurobiol. Learn. Mem.* **161**, 149–157 (2019).
792 50. Kannan, S., Farid, M., Lin, B. L., Miyamoto, M. & Kwon, C. Transcriptomic entropy benchmarks
793 stem cell-derived cardiomyocyte maturation against endogenous tissue at single cell level. *PLoS*
794 *Comput. Biol.* **17**, 1–21 (2021).
795 51. Bramswig, N. C. *et al.* Epigenomic plasticity enables human pancreatic a to b cell reprogramming.
796 *J. Clin. Invest.* **123**, 1275–1284 (2013).
797 52. Camunas-Soler, J. *et al.* Patch-Seq Links Single-Cell Transcriptomes to Human Islet Dysfunction in
798 Diabetes. *Cell Metab.* **31**, 1017-1031.e4 (2020).
799 53. Chen, K. *et al.* Broad H3K4me3 is associated with increased transcription elongation and
800 enhancer activity at tumor-suppressor genes. *Nat. Genet.* **47**, 1149–1157 (2015).
801 54. Pullen, T. J. *et al.* Identification of genes selectively disallowed in the pancreatic islet. *Islets* **2**, 89–
802 95 (2010).
803 55. Talchai, C., Xuan, S., Lin, H. V., Sussel, L. & Accili, D. Pancreatic β Cell Dedifferentiation as a
804 Mechanism of Diabetic β Cell Failure. *Cell* **150**, 1223–1234 (2012).
805 56. Neelankal John, A., Ram, R. & Jiang, F.-X. RNA-Seq Analysis of Islets to Characterise the
806 Dedifferentiation in Type 2 Diabetes Model Mice db/db. *Endocr. Pathol.* **29**, 207–221 (2018).
807 57. Hsiao, L. L. *et al.* A compendium of gene expression in normal human tissues. *Physiol. Genomics*
808 **2002**, 97–104 (2002).
809 58. Fadista, J. *et al.* Global genomic and transcriptomic analysis of human pancreatic islets reveals
810 novel genes influencing glucose metabolism. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 13924–13929
811 (2014).
812 59. Cahill, K. M., Huo, Z., Tseng, G. C., Logan, R. W. & Seney, M. L. Improved identification of

813 concordant and discordant gene expression signatures using an updated rank-rank
814 hypergeometric overlap approach. *Sci. Rep.* **8**, 1–11 (2018).
815 60. Li, J., Liu, Y., Kim, T. H., Min, R. & Zhang, Z. Gene expression variability within and between
816 human populations and implications toward disease susceptibility. *PLoS Comput. Biol.* **6**, (2010).
817 61. Ecker, S., Pancaldi, V., Valencia, A., Beck, S. & Paul, D. S. Epigenetic and Transcriptional Variability
818 Shape Phenotypic Plasticity. *BioEssays* **40**, 1–11 (2018).
819 62. Rangaraju, S. *et al.* Suppression of transcriptional drift extends C. elegans lifespan by postponing
820 the onset of mortality. *Elife* **4**, 1–39 (2015).
821 63. Benninger, R. K. P. & Kravets, V. The physiological role of β-cell heterogeneity in pancreatic islet
822 function. *Nat. Rev. Endocrinol.* **0123456789**, (2021).
823 64. Soares, L. M. *et al.* Determinants of Histone H3K4 Methylation Patterns. *Mol. Cell* **68**, 773-785.e6
824 (2017).
825 65. Cheng, J. *et al.* A role for H3K4 monomethylation in gene repression and partitioning of
826 chromatin readers. *Mol. Cell* **53**, 979–992 (2014).
827 66. Kumar, D., Cinghu, S., Oldfield, A. J., Yang, P. & Jothi, R. Decoding the function of bivalent
828 chromatin in development and cancer. *Genome Res.* **31**, 2170–2184 (2021).
829 67. Schmitges, F. W. *et al.* Histone Methylation by PRC2 Is Inhibited by Active Chromatin Marks. *Mol.*
830 *Cell* **42**, 330–341 (2011).
831 68. Dai, Z., Mentch, S. J., Gao, X., Nichenametla, S. N. & Locasale, J. W. Methionine metabolism
832 influences genomic architecture and gene expression through H3K4me3 peak width. *Nat.*
833 *Commun.* **9**, 1–12 (2018).
834 69. Brind'Amour, J. *et al.* An ultra-low-input native ChIP-seq protocol for genome-wide profiling of
835 rare cell populations. *Nat. Commun.* **6**, 1–8 (2015).
836 70. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–
837 359 (2012).
838 71. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079
839 (2009).
840 72. Quinlan, A. R. & Hall, I. M. BEDTools: A flexible suite of utilities for comparing genomic features.
841 *Bioinformatics* **26**, 841–842 (2010).
842 73. Kurtenbach, S. & William Harbour, J. SparK: A Publication-quality NGS Visualization Tool. *bioRxiv*
843 1–4 (2019). doi:10.1101/845529
844 74. Ramírez, F. *et al.* deepTools2: a next generation web server for deep-sequencing data analysis.
845 *Nucleic Acids Res.* **44**, W160–W165 (2016).
846 75. Amemiya, H. M., Kundaje, A. & Boyle, A. P. The ENCODE Blacklist: Identification of Problematic
847 Regions of the Genome. *Sci. Rep.* **9**, 1–5 (2019).
848 76. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-
849 seq data with DESeq2. *Genome Biol.* **15**, 1–21 (2014).
850 77. Taruttis, F. *et al.* External calibration with Drosophila whole cell spike ins delivers absolute mRNA
851 fold changes from human RNA seq and qPCR data. *Biotechniques* **62**, 53–61 (2017).
852 78. Dobin, A. *et al.* STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
853 79. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-
854 aware quantification of transcript expression. *Nat. Methods* **14**, 417–419 (2017).
855 80. Srivastava, A., Malik, L., Smith, T., Sudbery, I. & Patro, R. Alevin efficiently estimates accurate
856 gene abundances from dscRNA-seq data. *Genome Biol.* **20**, (2019).
857 81. Sarkar, H., Zakeri, M., Malik, L. & Patro, R. Towards selective-alignment: Bridging the accuracy
858 gap between alignment-based and alignment-free transcript quantification. *bioRxiv* 1–7 (2017).
859 doi:10.1101/138800
860 82. Stuart, T. *et al.* Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902.e21 (2019).

83. McGinnis, C. S., Murrow, L. M. & Gartner, Z. J. DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. *Cell Syst.* **8**, 329-337.e4 (2019).
84. Hafemeister, C. & Satija, R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *bioRxiv* 1–15 (2019). doi:10.1101/576827
85. Street, K. *et al.* Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics* **19**, 1–16 (2018).
86. Kannan, S., Farid, M., Lin, B., Miyamoto, M. & Kwon, C. Transcriptomic entropy benchmarks stem cell-derived cardiomyocyte maturation against endogenous tissue at single cell level. *bioRxiv* (2020). doi:10.1101/2020.04.02.022632
87. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009).
88. Brideau, C. M., Eilertson, K. E., Hagarman, J. A., Bustamante, C. D. & Soloway, P. D. Successful Computational Prediction of Novel Imprinted Genes from Epigenomic Features. *Mol. Cell. Biol.* **30**, 3357–3370 (2010).

Figure legends

Figure 1. Reduction to H3K4 methylation in beta cells leads to glucose intolerance and hyperglycemia in adult mice.

A, Top: schematic showing the core subunits of COMPASS complexes and a nucleosome methylated on H3K4. Bottom: genome-aligned RNA-seq reads at the *Dpy30* gene locus in *Dpy30*-WT and -KO cells. **B,** Immunoblots showing COMPASS core subunits RBBP5, ASH2L, and WDR5, and nucleosome protein histone H3, coimmunoprecipitated with WDR5 or an IgG control from *Dpy30*-WT or -KO islet cell nuclei 45 days after tamoxifen administration. Representative immunoblots of three independent coIPs are shown. IP: immunoprecipitation. **C,** Immunoblots showing H3K4me3, H3K4me1, H3K27ac, H3K27me3, and total histone H3 in islets from *Dpy30*-WT and -KO mice 45 days after tamoxifen administration. Numbers beneath each band indicate the band intensity normalized to the left-most sample. **D,** Mean immunofluorescent intensity of DPY30, H3K4me3, and H3K4me1 in *Dpy30*-WT and -KO β-cell nuclei at the indicated days after tamoxifen administration. Data are normalized to fluorescence intensity of α-cell nuclei; n = 3 mice of each genotype and time-point. **E,** Example IHC images of *Dpy30*-KO islets used for measurements in panel D showing H3K4me3 (cyan), insulin (magenta), and glucagon (yellow). Scale bars: 100 µm. **F,** Scatterplots showing $\log_2$(fold-change) and -$\log_{10}$(*P*-value) of gene expression in *Dpy30*-KO versus -WT cells 15- and 45-days after tamoxifen administration. Genes showing a ≥2-fold increase or decrease in expression with $P \leq 0.01$ (Wald test with Benjamini-Hochberg correction) are coloured red or green, respectively, and enumerated above. The full-length transcript of *Dpy30* is outlined and labeled in blue. **G,** Blood glucose (top) and serum insulin (bottom) concentration during IPGTT in *Dpy30*-WT and -KO mice 15-, 30-, and 45-days after tamoxifen administration. Data presented as mean ± SD (n = 8 mice for glucose measurements, 8-15 mice for insulin measurements). *P*-values were calculated from area under the curve (AUC) using multiple two-tailed t-tests with Welch's and Benjamini-Hochberg corrections; only values ≤ 0.05 are shown. **H,** Unfasted blood glucose (top) and body mass (bottom) of *Dpy30*-WT and -KO mice during the 60-days after tamoxifen administration. Data are presented as individual measurements with mean (n = 8, however tracking was stopped after a blood glucose reading ≥ 20 mM).

Figure 2. Genes involved in insulin production and glucose-induced activity are regulated by H3K4me3.

A, Genome browser views of H3K4me3 and H3K4me1 enrichment at the *Ins1* and *Ins2* gene loci in *Dpy30*-WT and -KO β-cells 45-days after tamoxifen administration. **B,** *Ins1* and *Ins2* RNA levels in *Dpy30*-WT and

-KO cells. Data presented as mean ± SD with values from individual mice (n = 3). Expression and *P*-values were calculated from RNA-Seq data using DESeq2 (Wald test with Benjamini-Hochberg correction). **C,** Insulin content in *Dpy30*-WT and -KO islets. Data presented as mean ± SD with values from individual mice (n = 8 WT, 6 KO). *P*-value calculated by two-tailed t-test with Welch's correction. **D,** Representative transmission electron micrographs showing insulin granules in β-cells from a *Dpy30*-WT and a -KO mouse. Scale bars: 2 µm. **E,** Quantification of median insulin core granule size and density. Data presented as mean ± SD with values from individual mice (n = 3). *P*-values calculated by two-tailed t-test with Welch's correction. **F,** Overrepresentation analysis of Gene Ontology: biological process terms in differentially expressed genes 45-days after tamoxifen administration. *P*-values represent EASE scores, a modified Fisher's exact *P*-value, calculated using the Database for Annotation, Visualization, and Integrated Discovery (DAVID) v.6.8 [87]. **G,** Insulin secretion from *Dpy30*-WT and -KO islets during static *in vitro* stimulation with glucose and KCl solutions, normalized to islet insulin content. *P*-values calculated by multiple two-tailed t-tests with Welch's and Benjamini-Hochberg corrections; only values ≤ 0.05 are shown. **H,** Cytosolic Ca^{2+} concentration in islets from *Dpy30*-WT and -KO mice during *in vitro* perifusion of glucose and KCl solutions. *P*-values calculated for AUC in each time block by one-way ANOVA between genotypes (n = 4 WT, 3 KO); only values ≤ 0.05 are shown. **I,** Oxygen consumption rate in *Dpy30*-WT and -KO dispersed islet cells during treatment with the indicated compounds. n = 3 WT, 4 KO. **J,** Mitochondrial respiration in *Dpy30*-WT and -KO islet cells in 16.7 mM glucose and their maximal respiration capacity, inferred from the data shown in panel I. *P*-values calculated by two-tailed t-test with Welch's correction (n = 3 WT, 4 KO).

Figure 3. TSS-associated H3K4 methylation maintains gene expression in mature beta cells.
A, Average enrichment profiles of H3K4me3 and H3K4me1 at the TSS of all expressed genes in *Dpy30*-WT and -KO chromatin. **B-C,** Scatterplot of H3K4me3 (B) and H3K4me1 (C) enrichment in 10-kb bins spanning the genome in *Dpy30*-KO versus -WT cells. **D,** Scatterplot showing spike-in normalized gene expression in *Dpy30*-KO versus -WT cells. **E,** Genome browser views show examples of genes that retain H3K4me3 (i), lose H3K4me3 but retain H3K4me1 (ii), or lose H3K4me3 and H3K4me1 (iii) from the TSS in *Dpy30*-KO chromatin. **F,** Box and whisker plot showing $\log_2$(fold-change) of RNA expression for gene groups described in panel E. *P* indicates the Wilcoxon signed-rank test result with Benjamini-Hochberg correction. **G,** Contour scatterplot showing $\log_2$(fold-change) of H3K4me1 enrichment (x-axis) plotted against $\log_2$(fold-change) of H3K4me3 enrichment (y-axis) in the TSS ±1 kb of all expressed genes. The $\log_2$(fold-change) of RNA expression of the associated genes is shown as a colour gradient.

Figure 4. Weakly active and developmentally regulated genes are susceptible to downregulation in H3K4me3-deficient beta cells.
A, Average enrichment profiles of H3K4me3, H3K4me1, H3K27ac, and H3K27me3 in *Dpy30*-WT cells at TSSs of genes that are stably expressed or downregulated by *Dpy30*-KO. **B,** Box and whiskers plots showing quantification of data shown in panel A. **C,** Box and whiskers plot showing mRNA expression in *Dpy30*-WT cells for genes stably expressed or downregulated by *Dpy30*-KO. **D,** Average %GC content profile in the gene body ±5 kb of genes that are stably expressed or downregulated by *Dpy30*-KO. **E,** Box and whiskers plot showing quantification of data shown in panel D. **F,** Running enrichment plot for imprinted genes (ref. [88]) and genes that are up- or downregulated during mouse β-cell maturation (adults versus postnatal day 10, ref. [46]), and mature β-cell transcription factor genes. NES: normalized enrichment score. *P* values

949 calculated using Wilcoxon signed-rank test (panels B,C,E) or permutation test (F) with Benjamini-Hochberg
950 correction.
951
952 **Figure 5. H3K4me3 deficiency leads to a less active and more repressed epigenome in mature beta cells.**
953 **A,** Genome browser view of H3K4me3, H3K4me1, H3K27ac, and H3K27me3 in *Dpy30*-WT and -KO cells
954 showing preservation of mega base-scale organization of H3K27me3-positive PcG compartments. **B,**
955 Scatterplots of H3K27ac (top) and H3K27me3 (bottom) enrichment in 10-kb bins spanning the genome in
956 *Dpy30*-KO versus -WT cells. **C,** Average enrichment profiles of H3K27ac (top) and H3K27me3 (bottom) at
957 the TSS of all expressed genes in *Dpy30*-WT and -KO cells. **D,** Box and whisker plots showing the $\log_2$(fold-
958 change) of H3K27ac (top) and H3K27me3 (bottom) enrichment in the TSS of transcriptionally
959 downregulated, stable, or upregulated genes in *Dpy30*-KO compared to -WT cells. *P*-values calculated with
960 Wilcoxon rank sum test with Benjamini-Hochberg correction. **E,** Genome browser view of a
961 downregulated gene *Rab34* showing loss of H3K4me3, H3K4me1, H3K27ac, and RNA, and accumulation
962 of H3K27me3, in *Dpy30*-KO cells.
963
964 **Figure 6. Reduction of transcriptional consistency in *Dpy30*-KO and diabetic beta cells.**
965 **A,** UMAP visualization of β-cell transcriptomes from a *Dpy30*-KO mouse. Expression of *Tdtomato*, *Egfp*,
966 and their overlap are shown as a colour gradient. **B,** Pseudotime representation of A. **C,** Scatterplot
967 showing *Ins1* and *Ins2* expression in β-cells plotted against pseudotime. **D,** Scatterplot showing gene
968 expression entropy scores of β-cells plotted against pseudotime. **E,** Box and whisker plot showing the
969 transcriptional entropy of β-cells in *Tdtomato+* and *Egfp+* clusters (n = 274 and 2,323 cells, respectively).
970 **F,** Box and whisker plot showing the variance/mean ratio of gene expression across β-cells in *Tdtomato+*
971 and *Egfp+* clusters. **G,** Variance/mean ratio of β-cell gene expression in *Egfp+* versus *Tdtomato+* clusters
972 (y-axis) plotted as a function of H3K4me3 breadth in quantiles (x-axis) (622 genes/quantile). Data
973 presented as mean ± SD. **H-I,** same as F-G in β-cells from human donors with or without T2D, using
974 H3K4me3 ChIP-seq data from GSE50386 (ref. [51]) and scRNA-seq data from GSE124742 ref. [52]).
975
976 **Figure 7. H3K4me3 peak breadth encodes gene expression dynamics in a mouse model of type 2**
977 **diabetes.**
978 **A,** H3K4me3 peaks in *Dpy30*-WT β-cells ranked from narrow to broad. Peaks associated with β-cell
979 transcription factor genes are labeled. **B,** Genome browser views showing H3K4me3 enrichment in WT β-
980 cells at a housekeeping gene (*Rplp0*), and an expression-matched β-cell transcription factor gene (*Nkx6-*
981 *1*). **C,** Enrichment *P*-values (y-axis) of genes down- or upregulated in islets from $Lepr^{db/db}$ mice, where genes
982 have been ranked and grouped by H3K4me3 peak breadth from narrow to broad (x-axis). Each group
983 contains 622 genes. *P*-values were calculated using one-sided Fisher's exact test. **D,** Immunoblots of
984 H3K4me3 and total histone H3 in islet lysate from $Lepr^{+/+}$ and $Lepr^{db/db}$ mice. H3-normalized H3K4me3
985 band densities are listed. **E,** Average TSS enrichment profile of H3K4me3 in $Lepr^{+/+}$ and $Lepr^{db/db}$ islets. **F,**
986 Genome browser views of H3K4me3 in $Lepr^{+/+}$ or $Lepr^{db/db}$ islets at notable genes that are downregulated
987 (*Pdx1*, *Slc30a8*) or induced (*Aldh1a3*, *Ldha*) in $Lepr^{db/db}$ islets. **G,** Enrichment *P*-values of H3K4me3 peaks
988 showing significant reduction or expansion of peak breadth in $Lepr^{db/db}$ versus $Lepr^{+/+}$ islets (y-axis) in the
989 different gene groups ranked from narrow to broad (x-axis). Enrichment *P*-values were calculated using
990 one-sided Fisher's exact test. **H,** Box and whisker plots showing the $\log_2$(fold-change) of H3K4me3 peak
991 breadth for genes that are downregulated, stably expressed, or upregulated in $Lepr^{db/db}$ islets. *P*-values
992 calculated with Wilcoxon rank sum test with Benjamini-Hochberg correction. **I,** Stratified rank/rank

993 hypergeometric overlap plot comparing gene expression changes caused by *Dpy30*-KO (x-axis) and by
994 *Lepr*$^{db/db}$ (y-axis) mutations. Colourscale shows the enrichment *P*-value calculated using the RRHO2 R
995 package [59]**. J,** Same as panel C for genes that are differentially expressed in *Dpy30*-KO cells.
996
997 **Figure S1. Deletion of *Dpy30* in mouse β-cells using *Ins1*^{Cre} results in reduction of H3K4 methylation,**
998 **hyperglycemia, hypoinsulinemia, and impaired glucose tolerance.**
999 **A,** Immunoblots of H3K4me3, H3K4me1, and total histone H3 in islet lysate from 5-week-old *Dpy30*-WT,
1000 -KO, and -HET mice. **B,** Bar graphs showing band intensities of data in panel A. *P*-values were calculated
1001 using one-way ANOVA with Tukey's multiple comparisons correction; n = 4. **C-D,** Unfasted blood glucose
1002 concentration (C) and body mass (D) of *Dpy30*-WT, -KO, and -HET mice during 4-10 weeks of age. Values
1003 from each mouse and time point are shown (n = 10 WT, 15 KO, 7 HET, however tracking was stopped after
1004 1-2 blood glucose reading ≥ 20 mM). **E-F,** Blood glucose (E) and serum insulin (F) concentration during
1005 intraperitoneal glucose tolerance tests in 5-week-old *Dpy30*-WT, -KO, and -HET mice. *P*-values were
1006 calculated by comparison of *Dpy30*-KO versus -HET (top, teal) or -KO versus -WT (bottom, grey) AUC's
1007 using one-way ANOVA with Tukey's multiple comparisons correction; only values ≤ 0.05 are shown. WT:
1008 *Ins1*$^{+/+}$; *Dpy30*$^{flox/flox}$. KO: *Ins1*$^{Cre/+}$; *Dpy30*$^{flox/flox}$. HET: *Ins1*$^{Cre/+}$; *Dpy30*$^{flox/+}$.
1009
1010 **Figure S2. Genes involved in insulin production and glucose-induced activity are regulated by H3K4me3.**
1011 **A,** Heatmap showing expression Z-scores in RNA-seq data from *Dpy30*-WT and -KO 45-days post-
1012 tamoxifen for selected genes important for glucose metabolism, insulin production or secretion. *P*-values
1013 calculated by Wald test with Benjamini-Hochberg correction; n = 3. **B,** Genome browser views of
1014 H3K4me3, H3K4me1, and RNA in *Dpy30*-WT and -KO chromatin at selected genes. **C,** Bar graph showing
1015 the fraction of cytoplasm area occupied by mitochondria in transmission electron micrographs of *Dpy30*-
1016 WT and -KO β-cells. *P*-value calculated using two-tailed t-test with Welch's correction. **D,** Representative
1017 immunofluorescence images showing EdU incorporation in *Dpy30*-WT and -KO mouse islets that were
1018 dispersed and cultured in 5.5 mM or 16.7 mM glucose for 48-hours, as indicated, showing EdU (white)
1019 and DAPI (white) staining and eGFP (green) and tdTomato (red) endogenous fluorescence. Scale bars: 100
1020 µm. **E,** Bar graph showing quantification of EdU+ eGFP+ double positive cells as a fraction of total eGFP+
1021 cells (n = 3 per treatment and genotype). *P*-values calculated using two-way ANOVA with Tukey's
1022 correction. ns: not significant ($P > 0.05$).
1023
1024 **Figure S3. H3K27ac peak dynamics are linked to gene expression dynamics in *Dpy30*-KO cells.**
1025 **A,** Pie chart showing the number of H3K27ac peaks that are increased >1.5-fold, decreased >1.5-fold, or
1026 <1.5-fold changed for enrichment of H3K27ac in *Dpy30*-KO versus -WT chromatin. **B,** Association
1027 frequency of H3K27ac peaks classified in panel A with genes that are upregulated or downregulated in
1028 *Dpy30*-KO cells. Association was tested between H3K27ac peaks with the nearest active TSS and compared
1029 with association frequencies after randomizing peak locations 100 times. ** $P < 0.01$ by permutation test.
1030
1031 **Figure S4. *Dpy30*-KO β-cells do not activate genes associated with dedifferentiated β-cells.**
1032 **A,** UMAP visualization of islet cell transcriptomes from a *Dpy30*-WT and -KO mouse. Markers of β-cells
1033 (*Ins1*), α-cells (*Gcg*), and δ-cells (*Sst*) are shown as a colour gradient (n = 3,703 WT cells, 3,641 KO cells).
1034 **B,** Violin plots showing expression *Ins1*, *Ins2*, *Iapp*, *Glp1r*, *Slc2a2*, and *G6pc2* in *Dpy30*-WT and -KO β-cells
1035 (n = 2,510 WT, 2,597 KO β-cells). **C,** Lowess curves of gene expression of the indicated maturity and
1036 immaturity genes in *Dpy30*-KO β-cells along pseudotime defined in Fig 6B.

1037

1038 **Figure S5. H3K4me3 peak breadth stratifies β-cell gene sets more effectively than gene expression level.**
1039 **A,** H3K4me3 peaks ranked by increasing breadth (x-axis) plotted against RNA expression of the associated
1040 gene (y-axis). R indicates Spearman's rank correlation coefficient. **B,** Overrepresentation analysis of
1041 expressed genes that are ranked and grouped by H3K4me3 peak breadth from narrow to broad (blue) or
1042 ranked and grouped by gene expression level from low to high expression (green). The number in each
1043 square indicates the number of genes of each geneset matching to each quantile. Enrichment *P*-values
1044 calculated using two-sided Fisher's exact tests. **C,** H3K4me3 peaks in human β-cells [51] ranked by peak
1045 breadth. Peaks associated with β-cell transcription factor genes are labeled. **D,** H3K4me3 peaks ranked by
1046 breadth (x-axis) plotted against RNA expression [58] of the associated gene (y-axis) in non-diabetic human
1047 islets. R indicates Spearman's rank correlation coefficient. **E,** Overrepresentation analysis of human islet
1048 genes that are ranked and grouped by H3K4me3 peak breadth from narrow to broad (blue) or ranked and
1049 grouped by gene expression level from low to high expression (green). The number in each square
1050 indicates the number of genes of each gene set matching to each quantile. Each quantile includes 783
1051 genes. Enrichment *P*-values calculated using one-sided Fisher's exact test. **F,** Running enrichment plot of
1052 genes disallowed in islets [54] and β-cell-enriched transcription factor genes in *Lepr*$^{db/db}$ RNA-seq data [56] .
1053 NES: normalized enrichment score. *P*-values calculated by permutation test.

Figure 1

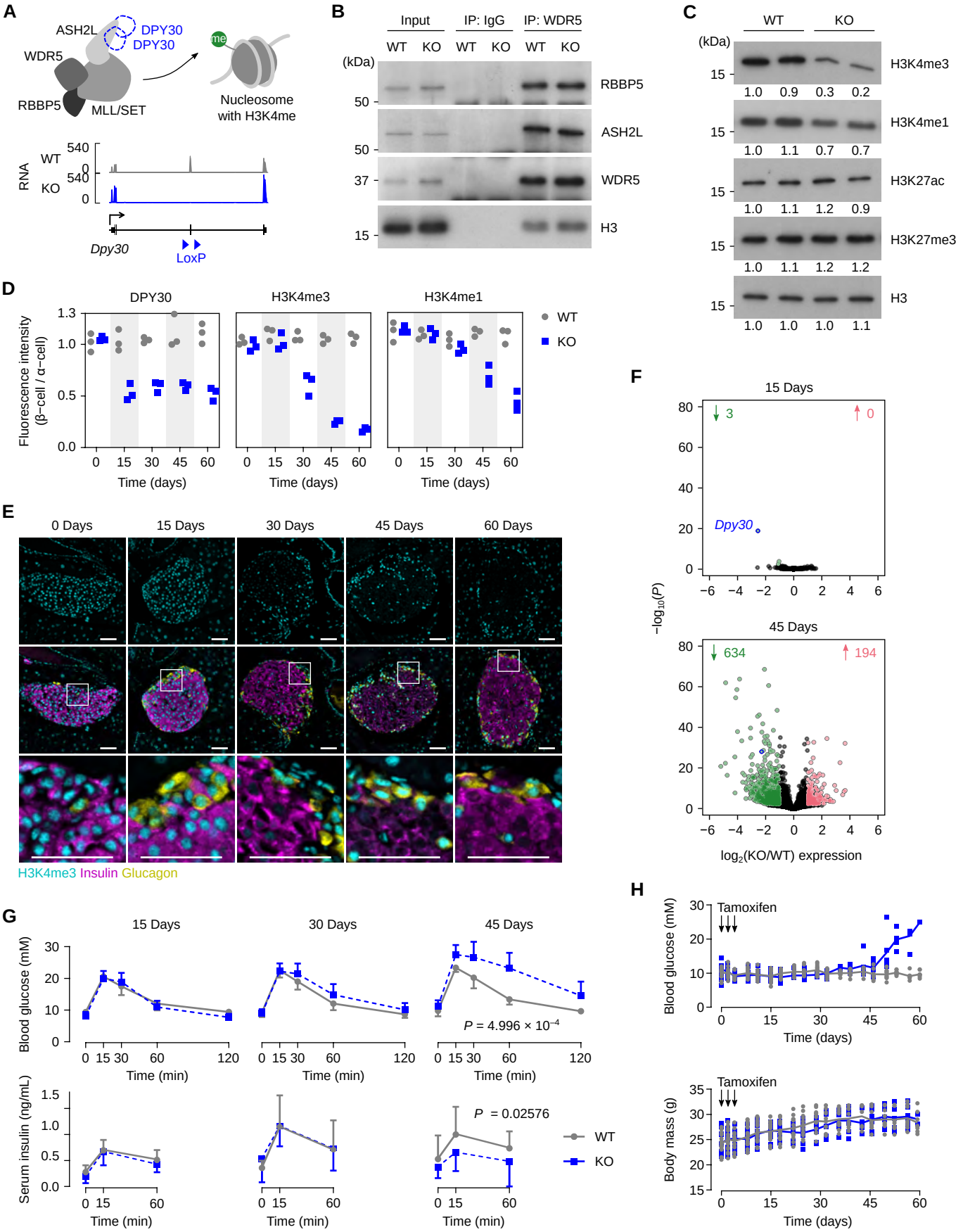

A
ASH2L
DPY30
DPY30
WDR5
RBBP5
MLL/SET
me
Nucleosome
with H3K4me
RNA
WT
KO
540
0
540
0
Dpy30
LoxP
B
Input
IP: IgG
IP: WDR5
WT KO WT KO WT KO
(kDa)
50
50
37
15
RBBP5
ASH2L
WDR5
H3
C
WT
KO
(kDa)
15
H3K4me3
1.0 0.9 0.3 0.2
H3K4me1
1.0 1.1 0.7 0.7
H3K27ac
1.0 1.1 1.2 0.9
H3K27me3
1.0 1.1 1.2 1.2
H3
1.0 1.0 1.0 1.1
D
DPY30
H3K4me3
H3K4me1
Fluorescence intensity
(β−cell / α−cell)
1.3
1.0
0.5
0.0
0 15 30 45 60
Time (days)
WT
KO
E
0 Days
15 Days
30 Days
45 Days
60 Days
H3K4me3 Insulin Glucagon
F
15 Days
3
0
Dpy30
45 Days
634
194
−log10(P)
80
60
40
20
0
−6 −4 −2 0 2 4 6
log2(KO/WT) expression
G
15 Days
30 Days
45 Days
Blood glucose (mM)
30
20
10
0
0 15 30 60 120
Time (min)
P = 4.996 × 10−4
Serum insulin (ng/mL)
1.5
1.0
0.5
0.0
0 15 60
Time (min)
P = 0.02576
WT
KO
H
Tamoxifen
Blood glucose (mM)
30
20
10
0
0 15 30 45 60
Time (days)
Tamoxifen
Body mass (g)
35
30
25
20
15
0 15 30 45 60
Time (days)

Figure 2

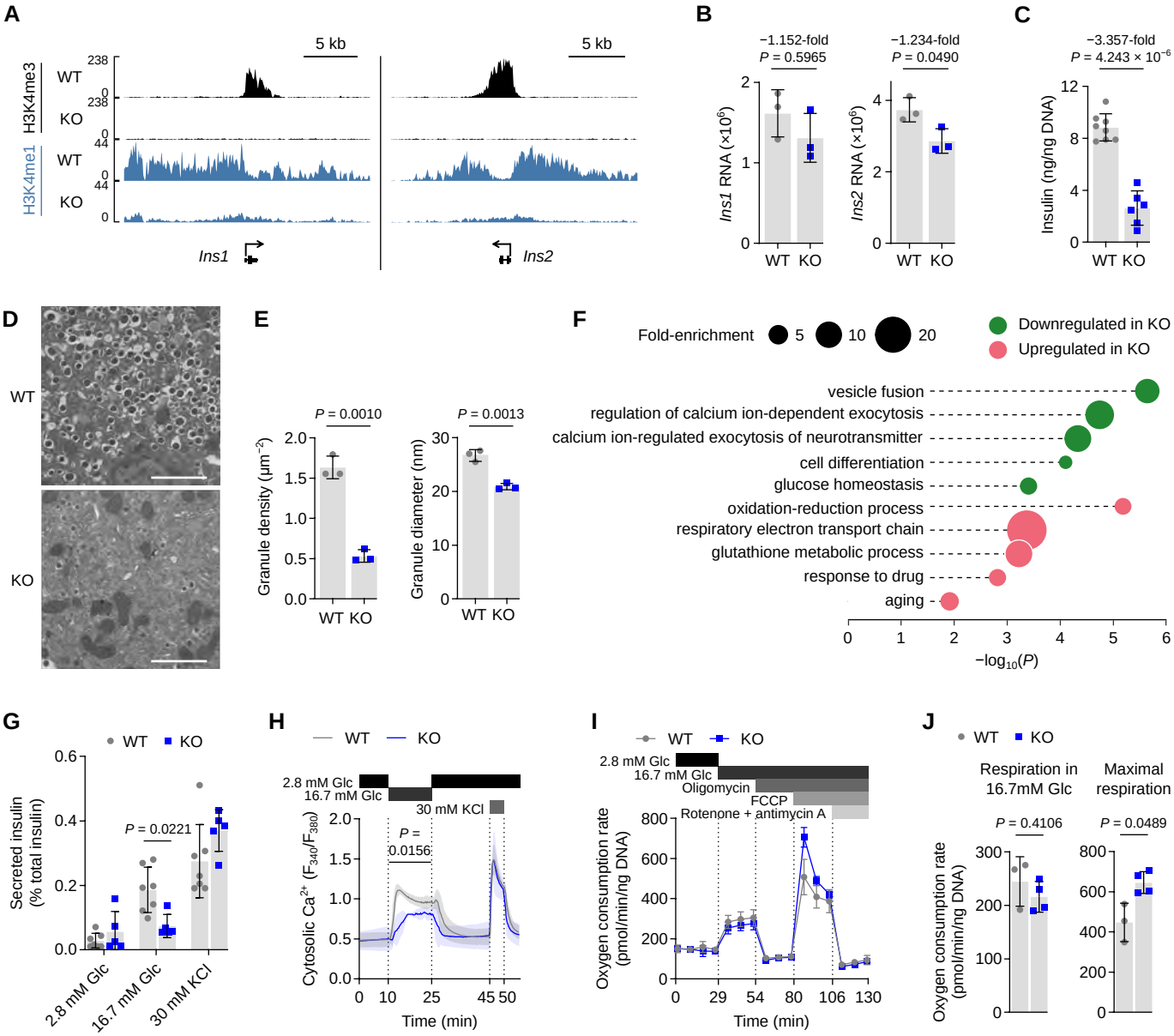

A
5 kb
5 kb
H3K4me3
H3K4me1
WT
KO
WT
KO
238
0
238
0
44
0
44
0
Ins1
Ins2
B
−1.152-fold
P = 0.5965
−1.234-fold
P = 0.0490
Ins1 RNA (×10^6)
Ins2 RNA (×10^6)
WT KO
WT KO
C
−3.357-fold
P = 4.243 × 10^−6
Insulin (ng/ng DNA)
WT KO
D
WT
KO
E
P = 0.0010
P = 0.0013
Granule density (μm^−2)
Granule diameter (nm)
WT KO
WT KO
F
Fold-enrichment 5 10 20
Downregulated in KO
Upregulated in KO
vesicle fusion
regulation of calcium ion-dependent exocytosis
calcium ion-regulated exocytosis of neurotransmitter
cell differentiation
glucose homeostasis
oxidation-reduction process
respiratory electron transport chain
glutathione metabolic process
response to drug
aging
−log10(P)
G
WT KO
Secreted insulin (% total insulin)
P = 0.0221
2.8 mM Glc
16.7 mM Glc
30 mM KCl
H
WT KO
2.8 mM Glc
16.7 mM Glc
30 mM KCl
P = 0.0156
Cytosolic Ca2+ (F340/F380)
Time (min)
I
WT KO
2.8 mM Glc
16.7 mM Glc
Oligomycin
FCCP
Rotenone + antimycin A
Oxygen consumption rate (pmol/min/ng DNA)
Time (min)
J
WT KO
Respiration in 16.7mM Glc
Maximal respiration
P = 0.4106
P = 0.0489
Oxygen consumption rate (pmol/min/ng DNA)

Figure 3

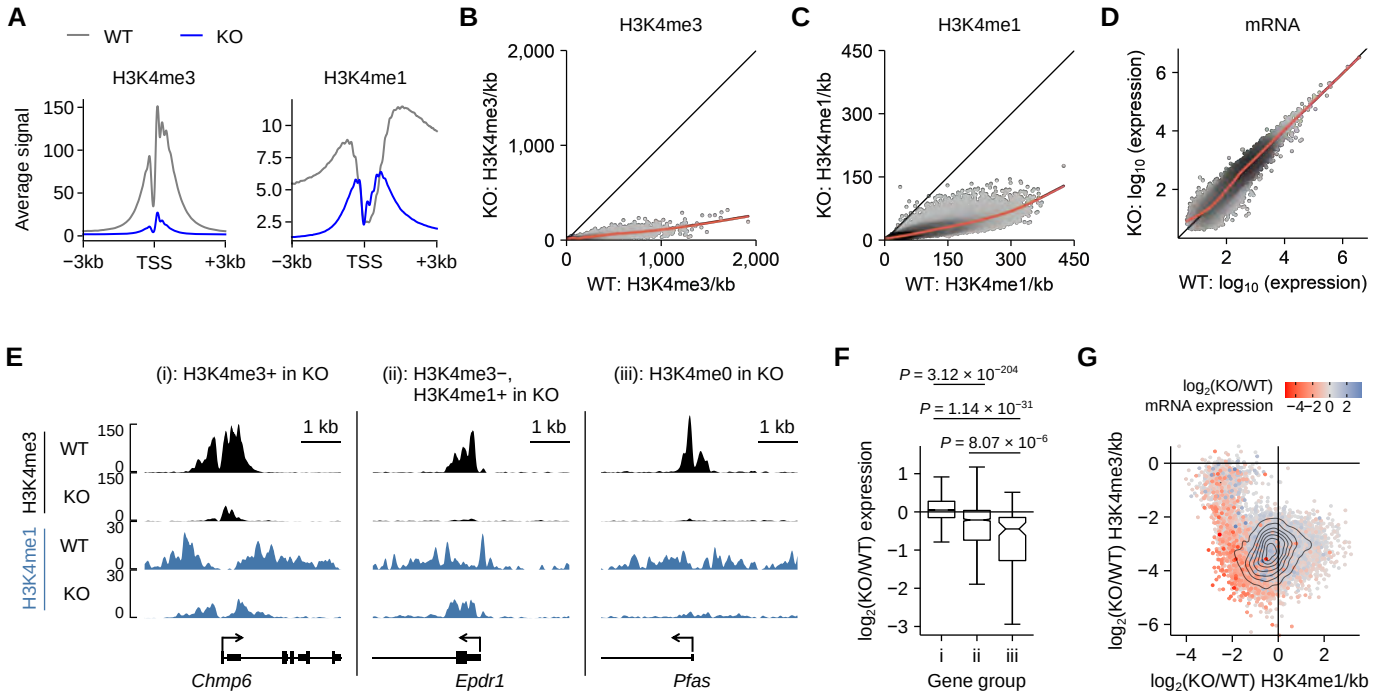

A
WT
KO
H3K4me3
H3K4me1
Average signal
150
100
50
0
10
7.5
5.0
2.5
−3kb
TSS
+3kb
B
H3K4me3
KO: H3K4me3/kb
2,000
1,000
0
WT: H3K4me3/kb
C
H3K4me1
KO: H3K4me1/kb
450
300
150
WT: H3K4me1/kb
D
mRNA
KO: log10 (expression)
6
4
2
WT: log10 (expression)
E
(i): H3K4me3+ in KO
(ii): H3K4me3−, H3K4me1+ in KO
(iii): H3K4me0 in KO
H3K4me3
H3K4me1
WT
KO
1 kb
Chmp6
Epdr1
Pfas
F
P = 3.12 × 10−204
P = 1.14 × 10−31
P = 8.07 × 10−6
log2(KO/WT) expression
i
ii
iii
Gene group
G
log2(KO/WT) mRNA expression
−4 −2 0 2
log2(KO/WT) H3K4me3/kb
log2(KO/WT) H3K4me1/kb

Figure 4

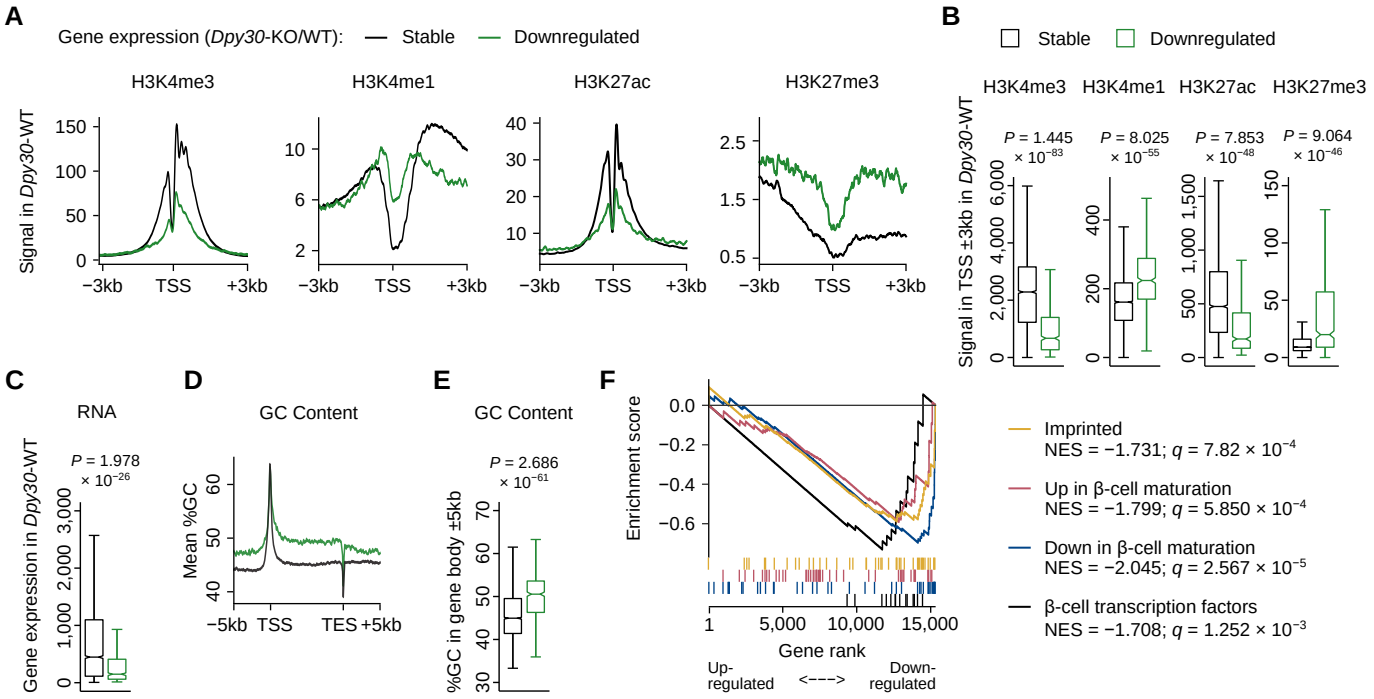

A
Gene expression (Dpy30-KO/WT): Stable Downregulated
H3K4me3
H3K4me1
H3K27ac
H3K27me3
Signal in Dpy30-WT
−3kb TSS +3kb
B
Stable Downregulated
H3K4me3 H3K4me1 H3K27ac H3K27me3
P = 1.445 × 10−83
P = 8.025 × 10−55
P = 7.853 × 10−48
P = 9.064 × 10−46
Signal in TSS ±3kb in Dpy30-WT
C
RNA
P = 1.978 × 10−26
Gene expression in Dpy30-WT
D
GC Content
Mean %GC
−5kb TSS TES +5kb
E
GC Content
P = 2.686 × 10−61
%GC in gene body ±5kb
F
Enrichment score
Gene rank
Up-regulated <---> Down-regulated
Imprinted
NES = −1.731; q = 7.82 × 10−4
Up in β-cell maturation
NES = −1.799; q = 5.850 × 10−4
Down in β-cell maturation
NES = −2.045; q = 2.567 × 10−5
β-cell transcription factors
NES = −1.708; q = 1.252 × 10−3

Figure 5

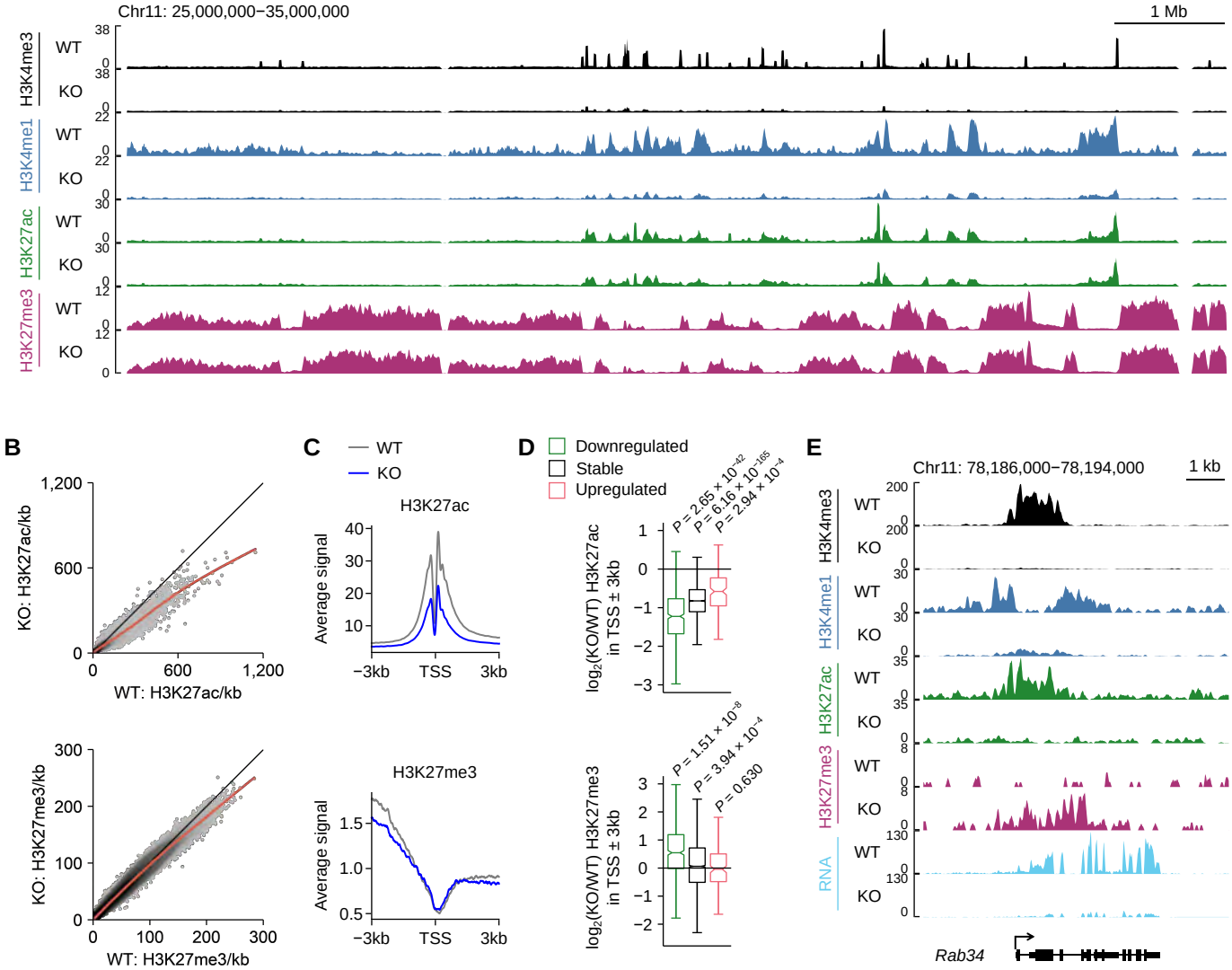

A
Chr11: 25,000,000−35,000,000
1 Mb
H3K4me3
H3K4me1
H3K27ac
H3K27me3
WT
KO
B
KO: H3K27ac/kb
WT: H3K27ac/kb
KO: H3K27me3/kb
WT: H3K27me3/kb
C
WT
KO
H3K27ac
H3K27me3
Average signal
−3kb
TSS
3kb
D
Downregulated
Stable
Upregulated
log2(KO/WT) H3K27ac in TSS ± 3kb
log2(KO/WT) H3K27me3 in TSS ± 3kb
P = 2.65 × 10−42
P = 6.16 × 10−185
P = 2.94 × 10−4
P = 1.51 × 10−8
P = 3.94 × 10−4
P = 0.630
E
Chr11: 78,186,000−78,194,000
1 kb
RNA
Rab34

Figure 6

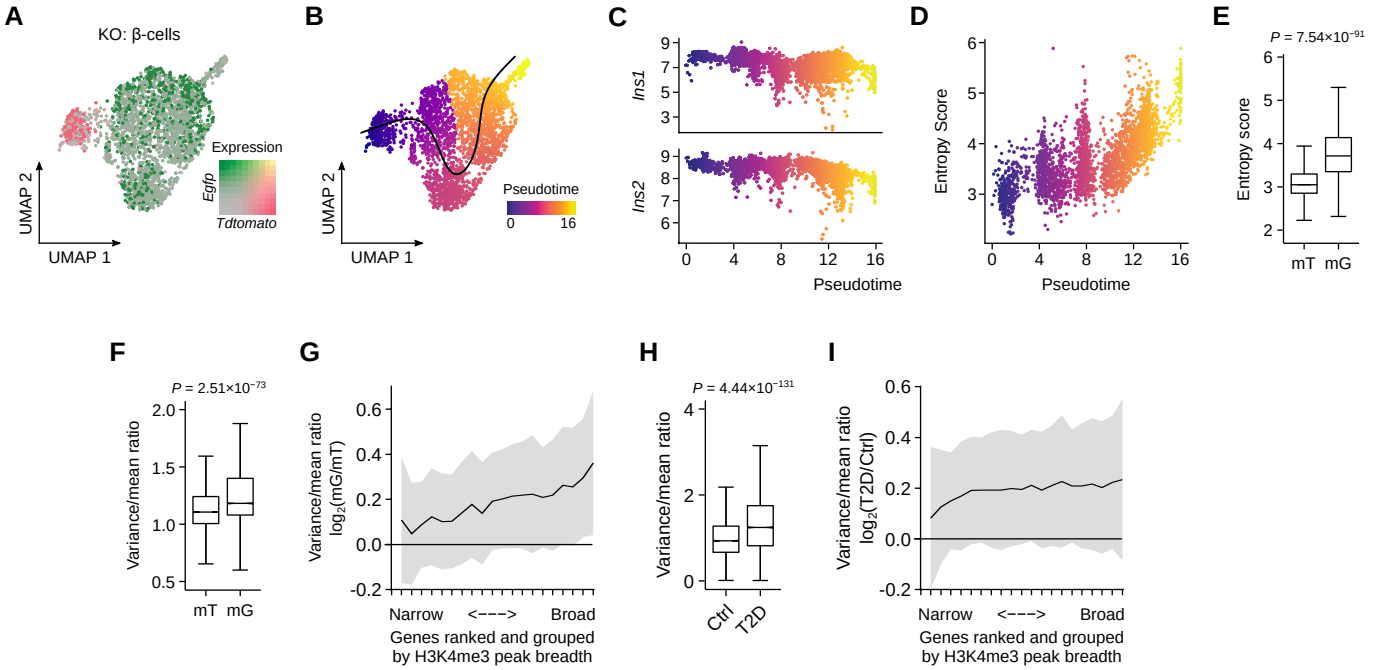

A
KO: β-cells
Expression
Egfp
Tdtomato
UMAP 2
UMAP 1
B
Pseudotime
0
16
UMAP 2
UMAP 1
C
Ins1
Ins2
Pseudotime
D
Entropy Score
Pseudotime
E
P = 7.54×10−91
Entropy score
mT
mG
F
P = 2.51×10−73
Variance/mean ratio
mT
mG
G
Variance/mean ratio
log2(mG/mT)
Narrow
<---->
Broad
Genes ranked and grouped
by H3K4me3 peak breadth
H
P = 4.44×10−131
Variance/mean ratio
Ctrl
T2D
I
Variance/mean ratio
log2(T2D/Ctrl)
Narrow
<---->
Broad
Genes ranked and grouped
by H3K4me3 peak breadth

Figure 7

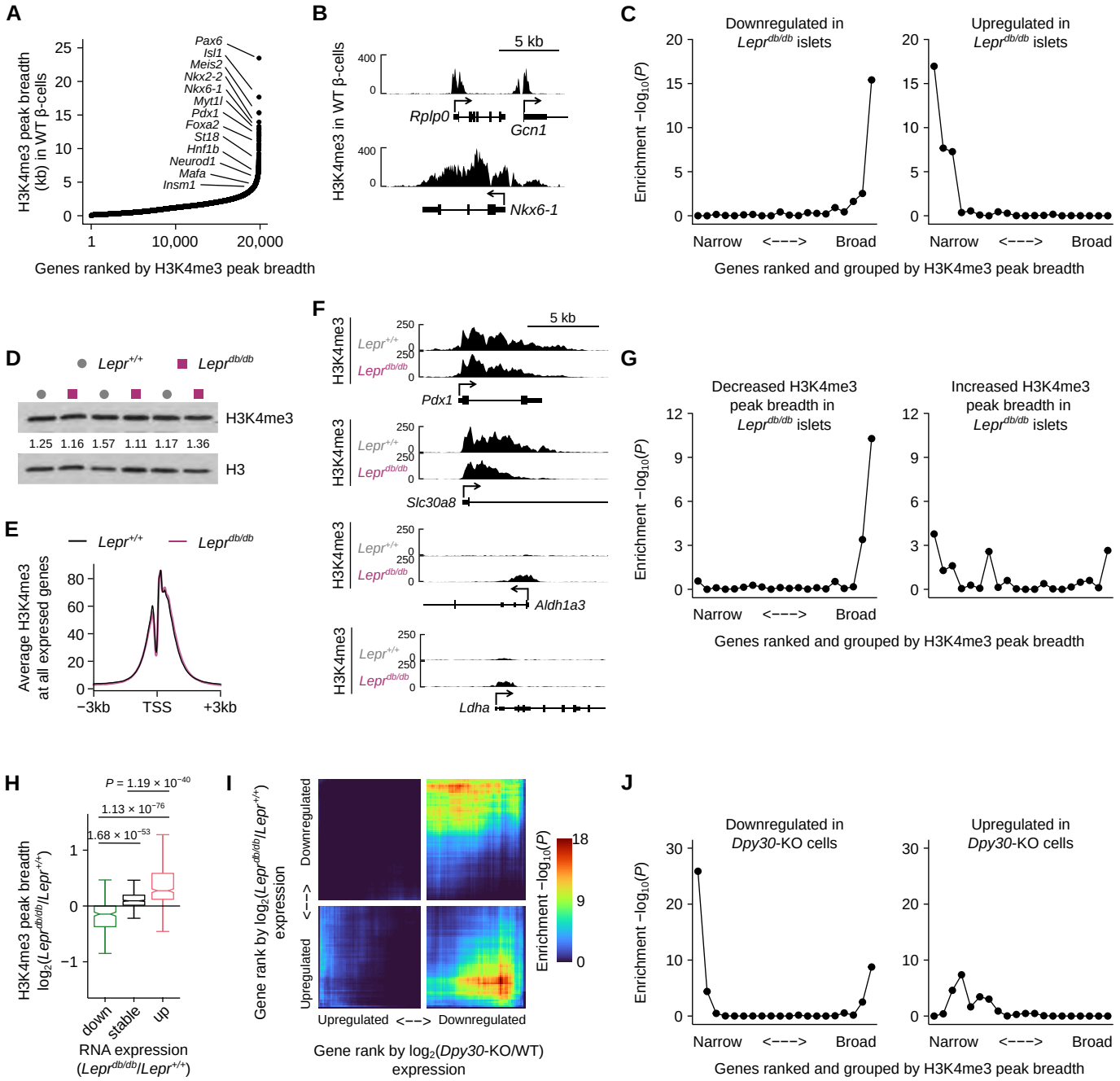

A
H3K4me3 peak breadth (kb) in WT β-cells
Pax6
Isl1
Meis2
Nkx2-2
Nkx6-1
Myt1l
Pdx1
Foxa2
St18
Hnf1b
Neurod1
Mafa
Insm1
Genes ranked by H3K4me3 peak breadth
B
H3K4me3 in WT β-cells
5 kb
Rplp0
Gcn1
Nkx6-1
C
Downregulated in Lepr^db/db islets
Upregulated in Lepr^db/db islets
Enrichment −log10(P)
Narrow <---> Broad
Genes ranked and grouped by H3K4me3 peak breadth
D
Lepr+/+
Lepr^db/db
H3K4me3
1.25 1.16 1.57 1.11 1.17 1.36
H3
E
Average H3K4me3 at all expressed genes
−3kb TSS +3kb
F
H3K4me3
5 kb
Pdx1
Slc30a8
Aldh1a3
Ldha
G
Decreased H3K4me3 peak breadth in Lepr^db/db islets
Increased H3K4me3 peak breadth in Lepr^db/db islets
Genes ranked and grouped by H3K4me3 peak breadth
H
P = 1.19 × 10^−40
1.13 × 10^−76
1.68 × 10^−53
H3K4me3 peak breadth log2(Lepr^db/db/Lepr+/+)
down stable up
RNA expression (Lepr^db/db/Lepr+/+)
I
Gene rank by log2(Lepr^db/db/Lepr+/+) expression
Downregulated <---> Upregulated
Enrichment −log10(P)
Gene rank by log2(Dpy30-KO/WT) expression
J
Downregulated in Dpy30-KO cells
Upregulated in Dpy30-KO cells
Genes ranked and grouped by H3K4me3 peak breadth

Figure S1

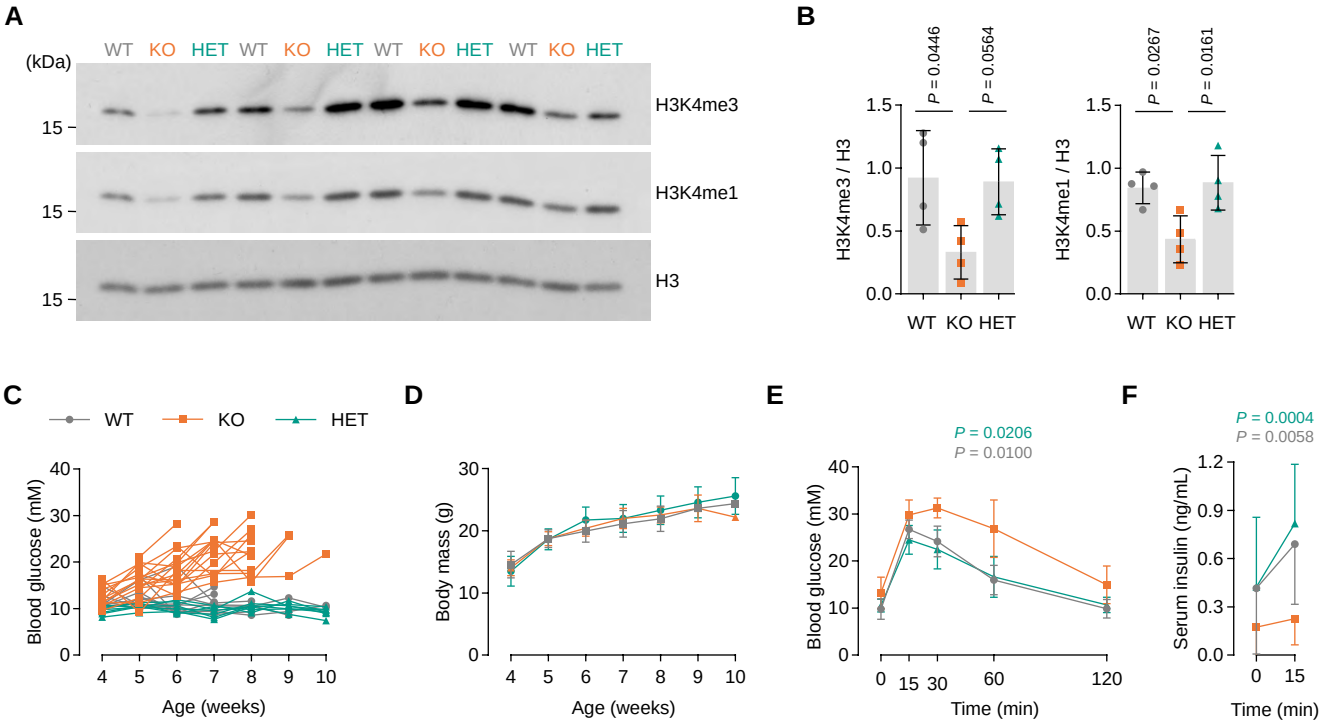

A
WT KO HET WT KO HET WT KO HET WT KO HET
(kDa)
15
H3K4me3
H3K4me1
H3
B
P = 0.0446
P = 0.0564
P = 0.0267
P = 0.0161
H3K4me3 / H3
H3K4me1 / H3
WT KO HET
C
WT KO HET
Blood glucose (mM)
Age (weeks)
D
Body mass (g)
Age (weeks)
E
P = 0.0206
P = 0.0100
Blood glucose (mM)
Time (min)
F
P = 0.0004
P = 0.0058
Serum insulin (ng/mL)
Time (min)

Figure S2

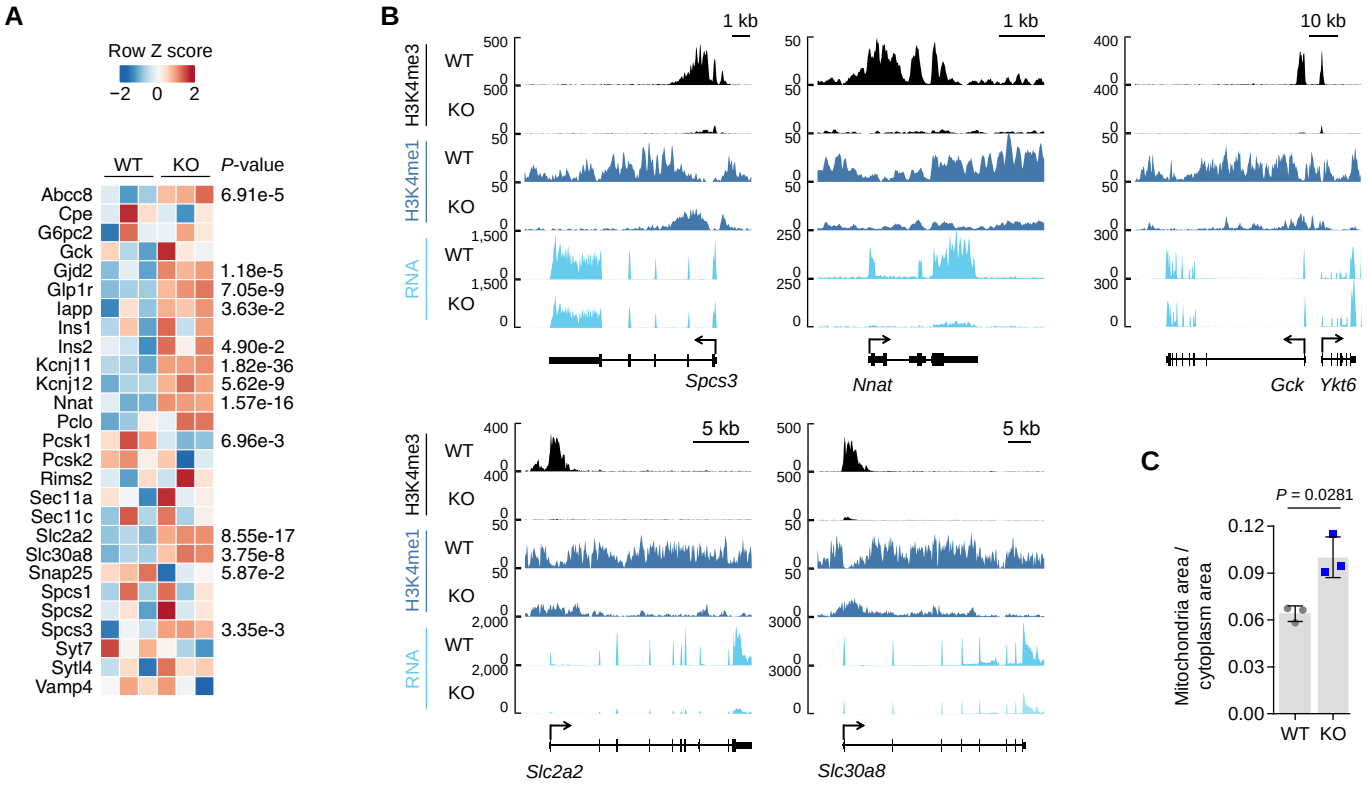

A
Row Z score
−2 0 2
WT KO P-value
Abcc8 6.91e-5
Cpe
G6pc2
Gck
Gjd2 1.18e-5
Glp1r 7.05e-9
Iapp 3.63e-2
Ins1
Ins2 4.90e-2
Kcnj11 1.82e-36
Kcnj12 5.62e-9
Nnat 1.57e-16
Pclo
Pcsk1 6.96e-3
Pcsk2
Rims2
Sec11a
Sec11c
Slc2a2 8.55e-17
Slc30a8 3.75e-8
Snap25 5.87e-2
Spcs1
Spcs2
Spcs3 3.35e-3
Syt7
Sytl4
Vamp4
B
H3K4me3
H3K4me1
RNA
WT
KO
1 kb
10 kb
5 kb
Spcs3
Nnat
Gck
Ykt6
Slc2a2
Slc30a8
C
P = 0.0281
Mitochondria area / cytoplasm area
0.12
0.09
0.06
0.03
0.00
WT KO

Figure S3

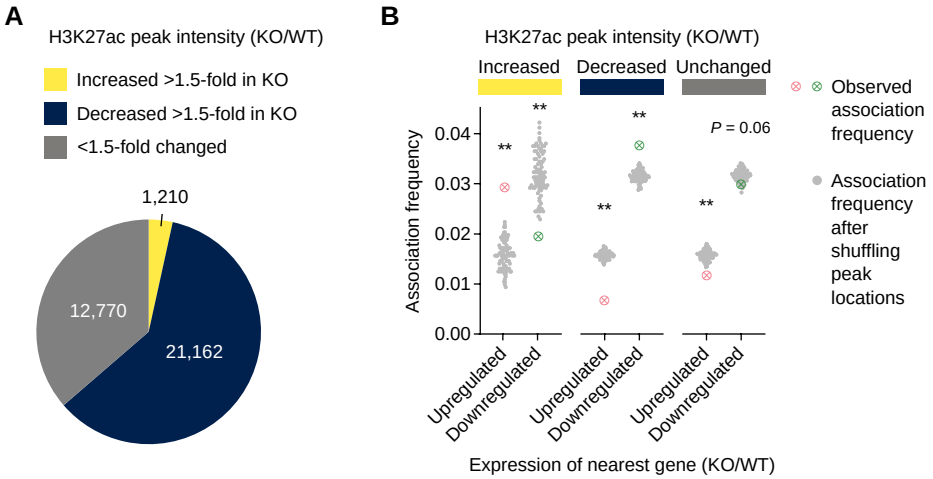

A
H3K27ac peak intensity (KO/WT)
Increased >1.5-fold in KO
Decreased >1.5-fold in KO
<1.5-fold changed
1,210
12,770
21,162
B
H3K27ac peak intensity (KO/WT)
Increased
Decreased
Unchanged
Observed association frequency
Association frequency after shuffling peak locations
Association frequency
0.04
0.03
0.02
0.01
0.00
**
**
**
**
P = 0.06
Upregulated
Downregulated
Upregulated
Downregulated
Upregulated
Downregulated
Expression of nearest gene (KO/WT)

Figure S4

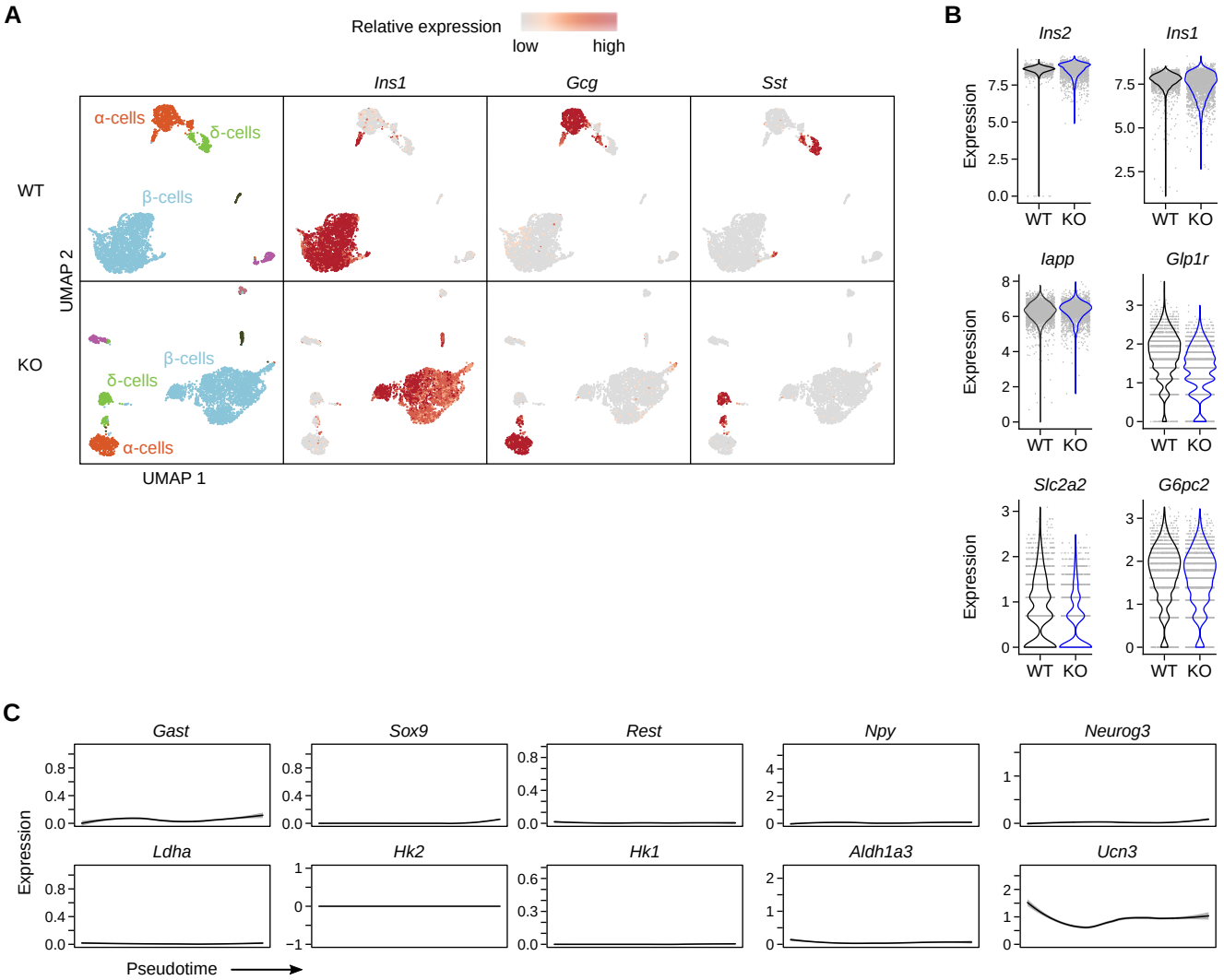
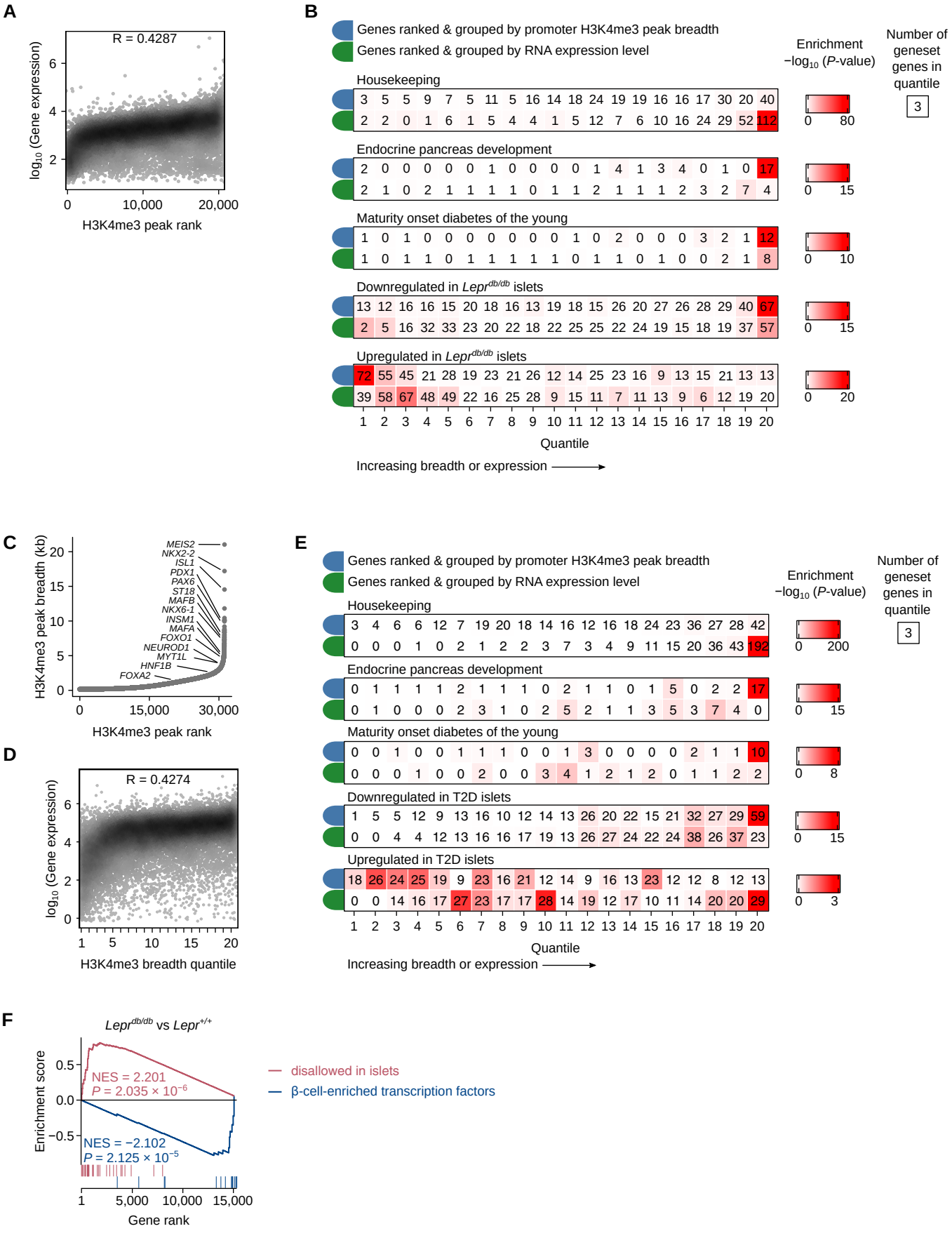

A
Relative expression
low
high
Ins1
Gcg
Sst
WT
KO
α-cells
δ-cells
β-cells
UMAP 2
UMAP 1
B
Ins2
Ins1
Iapp
Glp1r
Slc2a2
G6pc2
Expression
WT
KO
C
Gast
Sox9
Rest
Npy
Neurog3
Ldha
Hk2
Hk1
Aldh1a3
Ucn3
Expression
Pseudotime

Figure S5


A
R = 0.4287
log10 (Gene expression)
H3K4me3 peak rank
B
Genes ranked & grouped by promoter H3K4me3 peak breadth
Genes ranked & grouped by RNA expression level
Enrichment −log10 (P-value)
Number of geneset genes in quantile
Housekeeping
Endocrine pancreas development
Maturity onset diabetes of the young
Downregulated in Leprdb/db islets
Upregulated in Leprdb/db islets
Quantile
Increasing breadth or expression
C
H3K4me3 peak breadth (kb)
MEIS2
NKX2-2
ISL1
PDX1
PAX6
ST18
MAFB
NKX6-1
INSM1
MAFA
FOXO1
NEUROD1
MYT1L
HNF1B
FOXA2
H3K4me3 peak rank
D
R = 0.4274
log10 (Gene expression)
H3K4me3 breadth quantile
E
Genes ranked & grouped by promoter H3K4me3 peak breadth
Genes ranked & grouped by RNA expression level
Enrichment −log10 (P-value)
Number of geneset genes in quantile
Housekeeping
Endocrine pancreas development
Maturity onset diabetes of the young
Downregulated in T2D islets
Upregulated in T2D islets
Quantile
Increasing breadth or expression
F
Leprdb/db vs Lepr+/+
Enrichment score
NES = 2.201
P = 2.035 × 10−6
NES = −2.102
P = 2.125 × 10−5
Gene rank
disallowed in islets
β-cell-enriched transcription factors