EXPLORING THE FUNCTION AND REGULATION OF HISTONE H3 LYSINE 4 TRIMETHYLATION IN MATURE BETA CELLS

by

Benjamin Geoffrey Keith Vanderkruk

BScH, The University of Victoria, 2015

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Cell and Developmental Biology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

July 2022

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the dissertation entitled:

Exploring the function and regulation of Histone H3 lysine 4 trimethylation in mature beta cells

submitted by Benjamin Geoffrey Keith Vanderkruk in partial fulfillment of the requirements for

the degree of Doctor of Philosophy

in Cell and Developmental Biology

Examining Committee:

Bradford G. Hoffman, Associate Professor, Cell and Developmental Biology, UBC
Supervisor

Francis C. Lynn, Associate Professor, Cell and Developmental Biology, UBC
Co-supervisor

James D. Johnson, Professor, Cell and Developmental Biology, UBC
Supervisory Committee Member

Matthew Lorincz, Professor, Medical Genetics, UBC
University Examiner

Michael Underhill, Professor, Cell and Developmental Biology, UBC
University Examiner

Additional Supervisory Committee Members:

C. Bruce Verchere, Professor, Pathology and Laboratory Medicine, UBC
Supervisory Committee Member

Abstract

Aims

Insufficient insulin release by β-cells is the primary etiology in type 2 diabetes and coincides with impaired expression of genes essential for β-cell function, but drivers of gene expression dysregulation are not well resolved. Alterations to the genome-wide enrichment and organization of chromatin post-translational modifications may promote gene expression dysregulation. Here, I investigate the role of H3K4me3 in mature β-cells and how its organization in chromatin is linked to the unique β-cell gene transcriptome in health and diabetes. I further test how its enrichment is altered by external challenges in the form of type 2 diabetes-like stresses or perturbation of one carbon metabolism.

Methods

To study the functional importance of H3K4me3 in mature β-cells, we depleted H3K4me3 in β-cells of mature mice using an inducible *Dpy30* deletion model under control of the *Pdx1* or *Ins1* promoter and performed a panel of metabolic, transcriptomic, and epigenetic tests. We compared H3K4me3 enrichment patterns with gene expression changes that occur in islets in a mouse model of type 2 diabetes and in human type 2 diabetes. We then examined the metabolic and transcriptomic consequences of folic acid restriction in mouse islets.

Results

H3K4me3 contributes to gene expression in mature β-cells. H3K4me3 contributes to H3K27ac levels and, in the absence of H3K4me3, promoter-associated H3K4me1 is partially sufficient to maintain expression. H3K4me3 peak breadth is correlated with gene expression dysregulation in

type 2 diabetes in mice and humans. Using a genetic mouse model to impair the methyltransferase activity of trithorax group complexes, we find that reduction of H3K4me3 reduces insulin production and glucose-responsiveness and increases transcriptional entropy. H3K4me3 in mouse β-cells is particularly required for the expression of genes that are dysregulated in a mouse model of type 2 diabetes. While locus-specific alterations are observed, global enrichment of H3K4me3 in islets is robust against external disruption of glucose homeostasis and one-carbon metabolism.

Conclusions/interpretation

Overall, this thesis shows that H3K4me3 contributes to expression of genes essential for β-cell identity and function in mature β-cells and implicates dysregulation of H3K4me3 as a factor contributing to β-cell dysfunction in type 2 diabetes by altering gene expression patterns.

Lay Summary

Type 2 diabetes (T2D) is a common disease in which blood sugar levels are consistently high because the body does not make enough insulin. Insulin tells the body to reduce blood sugar because high blood sugar can cause organ damage. Insulin is made in specialized cells in the pancreas called β-cells. To effectively make and release insulin, β-cells must use a very specific set of genes. However, the set of genes used by β-cells changes in T2D, making them less effective at releasing insulin. A protein called H3K4me3 may help cells choose which genes to use. In this thesis I studied whether H3K4me3 affects β-cell genes and insulin production in mice. I then test if H3K4me3 is altered by poor diet. I find that alteration in H3K4me3 can drive pathological changes in mouse β-cells that are similar to T2D but H3K4me3 is surprisingly resilient against changing due to poor diet.

Preface

Studies were designed by BGK Vanderkruk and BG Hoffman. All experiments and analyses were conducted by BGK Vanderkruk in the laboratory of BG Hoffman at the BC Children's Hospital Research Institute (BCCHRI) except as stated below. Animal studies were approved by the University of British Columbia's Animal Care Committee under certificates A17-0045, A18-0111, and A20-0120 and were conducted in the Animal Care Facility of BCCHRI. Technical assistance or materials were provided by the following people:

Chapter 3: BG Hoffman aided with experimental design and data analysis. FC Lynn provided *Pdx1*-CreER and *mTmG* mice. L Xu helped with fluorescence-activated cell sorting.

Chapter 4: BG Hoffman aided with experimental design. WT Gibson provided *Ins1*-Cre mice. N Maeshima assisted with *Ins1-Cre*; $Dpy30^{fl/fl}$ mouse profiling. Islet processing and imaging for TEM was performed by M Reid at the Facility for Electron Microscopy Research of McMaster University. CL McDonald quantified insulin granules in TEM images. Calcium imaging was performed by D Pasula.

Chapter 5: AM Devlin assisted with design of experiments and diets. Erythrocyte and islet one carbon-metabolites were quantified in the laboratory of J Miller in the Department of Nutritional Sciences at Rutgers University, New Brunswick, NJ, USA. Plasma one-carbon metabolites were quantified in the Analytics Core for Metabolomics and Nutrition at BCCHRI by R Dyer and J King.

A version of Chapter 1 has been published with peer review as: B Vanderkruk and BG Hoffman. Metabolism as a central regulator of β-cell chromatin state. The FEBS Journal **288**(12):2683-3693. (2021).

A version of Chapter 3-4 has been published without peer review as: B Vanderkruk, N Maeshima, DJ Pasula, M An, A Daniel, CL McDonald, DS Luciani, and BG Hoffman H3K4 Methylation in β-cells Prevents Transcriptional Downregulation and Variance Associated with Type 2 Diabetes. (2021). bioRxiv doi: https://doi.org/10.1101/2021.01.28.428651.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

10-f-THF	10-formyltetrahydrofolate
5,10-MTHF	5,10-methylenetetrahydrofolate
5-MTHF	5-methyltetrahydrofolate
Ac	Acetylation
ATP	Adenosine Triphosphate
AOC	Area over the curve
AUC	Area under the curve
BCCHRI	British Columbia Children's Hospital Research Institute
BHMT	Betaine-homocysteine methyltransferase
cAMP	Cyclic adenosine monophosphate
CBS	Cystathionine-β-synthase
ChIP	Chromatin immunoprecipitation
CoIP	Co-immunoprecipitation
COMPASS	Complex of proteins associated with SET1
CSE	Cystathionine γ-lyase
DHF	Dihydrofolic acid
DHFR	Dihydrofolate reductase
DM	Demethylase
DNA	Deoxyribonucleic acid
EdU	5-ethynyl-2'-deoxyuridine
FBS	Fetal bovine serum
FR	Folic acid-restricted diet
GSIS	Glucose-stimulated insulin secretion
Het	*Dpy30* heterozygous
KO	*Dpy30* knock-out
mESC	Mouse embryonic stem cell
me1, me2, me3	Mono-, di-, tri-methylation
MT	Methyltransferase
MAT	Methionine adenosyltransferase

MTHFD	Methylenetetrahydrofolate dehydrogenase
MTHFR	Methylenetetrahydrofolate reductase
MTR	5-Methyltetrahydrofolate-homocysteine methyltransferase
MTRR	5-Methyltetrahydrofolate-homocysteine methyltransferase reductase
MNase	Micrococcal nuclease
MODY	Maturity onset diabetes of the young
mRNA	Messenger ribonucleic acid
NAD(H)	(Reduced) nicotinamide adenine dinucleotide
NADP(H)	(Reduced) nicotinamide adenine dinucleotide phosphate
NF	Control diet with normal folic acid levels
OGTT	Oral glucose tolerance test
PcG	Polycomb group
PFA	Paraformaldehyde
RNAP2	RNA polymerase II
RPMI	Roswell Park Memorial Institute medium
SET	Su(var)3-9, Enhancer-of-zeste and Trithorax
SAH	*S*-adenosylhomocysteine
SAHH	*S*-adenosylhomocysteine hydrolase
SAM	*S*-adenosylmethionine
SD	Standard deviation
SHMT	Serine hydroxymethyltransferase
T1D	Type 1 Diabetes
T2D	Type 2 Diabetes
THF	Tetrahydrofolic acid
Trx	Trithorax
TrxG	Trithorax group
TSS	Transcription start site
TYMS	Thymidylate Synthetase
UBC	University of British Columbia
WT	*Dpy30* wild type

Acknowledgements

Many people helped me during my PhD journey. First, I thank the mentorship and support I've received from my supervisor Dr. Brad Hoffman. I am grateful for the opportunities and space I was given to explore interesting and challenging questions. I had a lot to learn but believe my experience has shaped me into a critical and independent researcher. I also thank Dr. Francis Lynn for critical insight and supports first as a lab neighbour and later co-supervisor. Thank you also to the other members of my advisory committee Bruce Verchere and Jim Johnson for providing excellent guidance throughout my degree.

I benefited from help from two very impressive trainees – Meilin An and Priya Suresh, who are on their own way to great things. Thank you also to other lab mates: Stephanie Campbell, Yuka Obayashi, Aline Daniel, Nina Maeshima, Cassie McDonald, and Jocelyn Bégin, who have all helped make the challenging moments bearable. Daniel Pasula, Lisa Xu, Cuilan Nian, Derek Dai, and others on the fourth floor of BCCHRI who participated in discussion of this work, were also great sources of technical assistance or support. I wish you all continued success.

Finally, I thank the funding agencies who supported me during this work: Canucks for Kids Foundation, BCCHRI, Cell & Developmental Biology, and the Natural Sciences and Engineering Research Council of Canada.

Dedicated to laboratory mice

Chapter 1: Introduction

1.1 Diabetes

1.1.1 Prevalence and impact of diabetes

The prevalence of diabetes is rising. The percent of the global population living with diabetes was 2.8% in the year 2000 [1], 8.3% in 2013 [2], 9.3% in 2019 [3], and is projected to reach 10.2% in 2030 [3]. Prevalence in Canada exceeds average global trends, with 10% of Canadians diagnosed with diabetes in 2020, and projected to reach 12% in 2030 [4]. A further 6.1% of Canadians currently live with prediabetes, which, if left untreated, promotes development of diabetes and its complications [4].

Along with incidence, the cost burden of diabetes is increasing year-over-year in developed and developing countries. World-wide health expenditure related to diabetes was estimated at $760 billion USD ($965 billion CAD) in 2019 and is projected to reach $825 billion ($1047 billion CAD) in 2030 [5]. Estimates which consider both health expenditure and the value of forgone production due to diabetes forecasts a total cost of $2.2 trillion ($2.8 trillion CAD) in 2030, or 1.8% of global gross domestic product [6,7]. In Canada, medical costs for the population with diabetes is three to four times greater than for the non-diabetic population [8], which leads to a significant national expense: the estimated cost in Canada was just under $30 billion in 2019 – a 2-fold increase from $14 billion in 2008 [9]. The economic burden is also borne by individuals with diabetes. Average out-of-pocket costs for diabetes management amounts to 3% of the average income [4]. The disease is four times more common among low income than high income Canadians and over half of those affected report they were not always able to meet the financial cost of their treatments [4,10]. In this way, income status is a determinant of disease severity.

Individuals with diabetes are at elevated risk for other chronic conditions including heart, kidney, and other vascular diseases, neuropathy, depression, and Alzheimer's disease [8,11,12]. Diabetes is associated with a three-fold higher rate of hospitalization for heart disease [8,13]. Hospitalization for renal disease is 12-fold higher in the diabetic population and, among patients with kidney disease, outcomes are worse for those with comorbid diabetes [8]. As many as 40% of patients with diabetes will eventually experience kidney failure [14] requiring dialysis or kidney transplant. Peripheral microvascular and nerve damage are also common, being observable in 18.8% and 60-70% of patients, respectively [15,16] and resulting in tissue damage in one third to one half of persons with diabetes [17]. Due to this, diabetes is the leading cause of both non-traumatic leg amputations [18] and blindness [19]. Another long-term complication of diabetes is central nervous system impairment, including decrements in cognitive function [20], two-fold increased risk for depression, and 1.5-fold increased risk of Alzheimer's disease [21]. Risk of stroke is 1.5- to 2-fold higher in the diabetic population [22]. Overall, diabetes is associated with a 5 to 6 year reduction in life expectancy [23] and was estimated to contribute to 11.3% of global deaths in 2020 [24], where mortality is thought to occur ultimately from comorbid conditions, chiefly cardiovascular disease [25].

1.1.2 Diagnosis of diabetes

Diabetes is a group of metabolic disorders characterized by chronic hyperglycemia, or high blood glucose [26]. Diagnosis is based on measures of fasting blood glucose concentration, short-term glucose clearance, and long-term blood glucose control which confer high risk of developing retinopathy. Diabetes Canada defines a healthy 8-hour fasted blood glucose

concentration as falling within 4.0 mM to 6.0 mM, whereas values of 7.0 mM or greater are diagnostic of diabetes [26]. Patients can alternatively be diagnosed if their blood glucose is at least 11.1 mM two hours after consuming 75 g of glucose, or if their blood glucose is at least 11.1 mM at any time [26]. Longer-term, two to three month, trends of blood glucose control are inferred by quantifying percent glycation of circulating hemoglobin (HbA1c). Non-diabetic individuals have HbA1c values up to 6.0%, and patients with diabetes have values of at least 6.5% [26]. Patients are diagnosed with prediabetes with fasting blood glucose of 6.1 to 6.9 mM, or 7.8 to 11 mM two hours after consuming 75 g of glucose, or HbA1c between 6.0 to 6.4% [26].

Most cases of diabetes can be classified as either type 1 or type 2 (T1D or T2D). In T1D, chronic hyperglycemia results from autoimmune-mediated destruction of the insulin-producing β-cells. In T2D, chronic hyperglycemia is due to a combination of β-cell dysfunction and insulin resistance [26]. In insulin resistance, insulin target tissues such as liver, adipose, and muscle become less sensitive to insulin and do not efficiently clear glucose from circulation [27]. Typical distinguishing characteristics of T1D are the presence of autoimmune markers, early age of onset (peak age of diagnosis ~14 years of age), and insulin-dependence, whereas characteristics of T2D include obesity, physical signs of insulin resistance, and later age of onset [26]. However, these characteristics are not universal and in fact diagnoses of T2D in children is on the rise [28,29]. Shared symptoms of diabetes may include polydipsia, polyurea, weight change, fatigue, reduced feeling and healing in extremities, and ketoacidosis in some circumstances. Worldwide, T1D and T2D account for about 95% of cases [30,31]. One to five percent of cases are defined as monogenic diabetes, which is caused by a mutation affecting a single gene. Defects in over 40 different genes have been identified that cause monogenic diabetes [32]. Maturity onset diabetes of the

young (MODY) is a subtype of monogenic diabetes, along with neonatal diabetes and syndromic diabetes [32]. MODY genes, such as *GCK*, *HNF1A*, and *PDX1*, tend to be involved in development or function of terminally differentiated β-cells [33]. A transient form of diabetes called gestational diabetes affects up to 20% of people during and after pregnancy [34]. Gestational diabetes shares common symptoms with T2D, including insulin resistance, but the development of insulin resistance may be related to changes in hormone levels including cortisol and progesterone [34,35], as well as the greater insulin demand caused by pregnancy [36]. Nevertheless, gestational diabetes is associated with a higher risk of later development of T2D in both mother and child [37].

1.1.3 Type 2 diabetes

T2D accounts for about 90% of diabetes cases worldwide [31]. T2D occurs when β-cells fail to secrete appropriate quantities of insulin to regulate glycemia due to both insulin resistance and progressive loss of β-cell function. Insulin resistance may be a consequence of metabolic stress due to exposure to excessive dietary nutrients and their toxic metabolites [38]. Metabolic stress also influences hormonal and inflammatory signaling and intracellular stress pathways [38,39]. For example, high fat diet suppresses adiponectin signaling from adipose cells, leading to reduction of glucose uptake and fat breakdown in muscle cells and increased glucose production in liver cells [40]. Accumulation of triacylglycerides in muscle and liver, and increased circulating free fatty acids, reduce insulin sensitivity during obesity [41,42]. At the same time, obesity is associated with reduced glucose excretion by kidneys [43], greater appetite encoded in brain activity [44], and inflammation in the intestines that promotes hyperphagia and contributes to insulin resistance [45]. Obesity is therefore a major determinant of insulin resistance and T2D [46] but overlap is imperfect. Sedentary lifestyle [47], high sugar consumption [48], and aging [49] are other environmental

risk factors for insulin resistance and T2D. Heritable genetic variations at over 150 so-far-identified genomic loci are also predicted to increase susceptibility to T2D [50].

A substantial fraction of the genetic component of T2D is thought to impact gene expression or function in β-cells rather than insulin target tissues [50,51], highlighting the necessary involvement of β-cell dysfunction in order for T2D to develop. The β-cell population has a limited capacity to expand and increase insulin output to meet metabolic challenges. In T2D these capacities are overwhelmed, leading to β-cell dysfunction and contraction of β-cell mass. By the time T2D is diagnosed, β-cell function has already halved, and continues to decline thereafter even with medical intervention [52]. Impairments in insulin production are caused by a combination of decreased β-cell mass and of reduced insulin secretion by remaining β-cells [53]. β-cell dysfunction can be caused by excess metabolic stress and high insulin demand which can lead to oxidative, hypoxic, and endoplasmic reticulum stresses [53,54]. Furthermore, β-cell function is impaired by inflammatory signaling caused by obesity or nutrient stress [55]. These stresses can drive β-cells to dedifferentiate, in which β-cells adopt a functionally immature phenotype which is ineffective at glucose-stimulated insulin secretion, and to cell death [56]. Overall, the processes leading to β-cell failure and T2D are heterogeneous, constituting many environmental and genetic variables. Research into general molecular features of diabetic β-cells and factors that cause their dysfunction is ongoing.

Lifestyle interventions that improve dietary choices, increase physical exercise, and promote weight loss are the most effective treatments to improve glucose regulation in patients with T2D [57]. In many patients, additional pharmacological treatment is necessary to reduce glycemia.

These treatments largely aim to increase either insulin sensitivity or insulin secretion. The first line medication in Canada is metformin, whose main effect is to reduce hepatic glucose production and increased glucose utilization [58]. Biguanides and thiazolidinediones are other compounds that improve insulin sensitivity, while dipeptidyl peptidase-4 inhibitors, glucagon-like peptide 1 receptor agonists, and sulfonylureas act on β-cells to increase stimulated insulin secretion [57,59,60]. Eventually, an exogenous source of insulin is necessary for many patients due to progressive decline in β-cell function despite medical interventions. This can take the form of insulin injections, a mechanical insulin pump, or transplantation of β-cells from cadaveric donors. More recently, the promise of stem-cell derived β-cells as a source of replacement β-cells has been demonstrated [61–63]. Intensive research aims to improve production and efficacy of these cells, or to identify novel approaches to forestall or reverse the underlying metabolic defects that drive β-cell dysfunction in T2D. At present, treatments remain suboptimal or unavailable for many patients and no cure is available.

1.2 β-cells

β-cells are highly specialized endocrine cells found in islets of Langerhans – endocrine micro-organs scattered within the pancreas and constituting 1-3% of its volume [64]. Each islet includes a few hundred to a few thousand cells; β-cells comprise about 60% of islet endocrine cells in humans, or 70% in mice [65]. Other islet endocrine cell types include glucagon-secreting α-cells, somatostatin-secreting δ-cells, ghrelin-secreting ε-cells and pancreatic polypeptide secreting γ- or PP-cells [65]. β-cells are notable as the source of circulating insulin in humans.

1.2.1 Insulin biosynthesis

Production of insulin by β-cells begins with transcription of the preproinsulin gene. Humans encode one gene, *INS*, while mice maintain two nonallelic copies, *Ins1* and *Ins2*. A stretch of several hundred base pairs upstream of the transcription start sites (TSS)—called the promoter—encodes binding motifs for β-cell-enriched transcription factors MAFA, PAX6, PDX1, HNF1α, and ISL1 and is sufficient to drive β-cell-specific expression [66]. Binding of some of these factors is stimulated by glucose metabolism to increase transcription [66–69]. Transcription of preproinsulin is further coupled to energy status via a cyclic adenosine monophosphate (cAMP) response element, which is activated in response to rapid metabolism of glucose and some other fuels [70], and a RIPE3b1 region, which is regulated in a redox-dependent manner [71]. In addition to increasing its transcription, glucose increases the stability of preproinsulin messenger ribonucleic acid (mRNA) [72,73], which contributes to a relatively long half-life of 77 hours [74] (versus 29 hours in low glucose conditions [74], and an average of 10 hours for all mRNA species [75]). These factors contribute to exceptionally high expression: preproinsulin mRNA constitutes nearly one third of the total mRNA pool in β-cells [76]. Ribosomes translate preproinsulin directly into the endoplasmic reticulum (ER) lumen where the signal peptide is cleaved by ER membrane-associated signal peptidase complexes to yield proinsulin [77]. Translation, too, is elevated by glucose stimulation – almost 6-fold, after accounting for the boosted rate of transcription [78]. In the ER, proinsulin folds with the help of chaperones and forms intramolecular disulfide bonds [79,80]. Proinsulin proceeds to the Golgi apparatus, where it is packaged into granules that bud off from the Golgi along with the peptides chromogranin A and B, islet amyloid polypeptide, and VGF; ions of zinc and calcium; and processing enzymes prohormone convertase 1/3, 2, and carboxypeptidase E. These enzymes further process proinsulin into insulin and C-peptide. Insulin

crystalizes with zinc and calcium ions into mature dense core granules which are transported to the cell membrane in preparation for exocytosis [81,82]. Thus, β-cells possess unique biosynthetic machinery allowing production and storage of large quantities of insulin in proportion to metabolic demand.

1.2.2 Glucose-stimulated insulin secretion (GSIS)

Insulin is a hormone: a messenger molecule secreted into circulation that influences the activity of distal target tissues. In adult humans and mice, the primary effect of insulin is to reduce circulating glucose levels. It increases glucose uptake in muscle and adipose cells, reduces glucose production in liver cells, and stimulates intracellular storage of glucose as glycogen [83]. β-cells secrete insulin in proportion to circulating glucose levels in order to maintain euglycemia (3.9 mM – 7.1 mM) via low basal secretion between meals followed by huge bursts of insulin release during meals [84]. To achieve this, mature β-cells have a unique metabolism which couples the rate of insulin secretion to the concentration of extracellular glucose. A rise in extracellular glucose triggers rapid and proportional glucose influx into β-cells by facilitated diffusion through high-capacity glucose transporters at the cell membrane. The primary glucose transporter protein in mouse β-cells is GLUT2, encoded by the *Slc2a2* gene [85,86], while human β-cells additionally express GLUT1 (*SLC2A1*) and GLUT3 (*SLC2A3*) [87]. Once inside the cell, phosphorylation of glucose to glucose-6-phosphate by glucokinase (*GCK*) prevents its export. Glucokinase may be considered the principal glucose sensor in β-cells because it catalyzes the rate-limiting step of glycolysis (and insulin secretion), is highly specific for glucose, and is not subject to feedback inhibition by its product; although it is extensively regulated by other mechanisms [88]. Glucose-6-phosphate is then processed via glycolysis and the Kreb's cycle to produce reducing equivalents

(NADH and $FADH_2$), which are used by the electron transport chain to generate ATP by oxidative phosphorylation of ADP [89,90]. A rise in the ATP:ADP ratio stimulates closure of ATP-sensitive potassium (K_{ATP}) channels at the plasma membrane, which had been maintaining a hyperpolarized membrane potential by exporting K^+ ions [91]. Closure of K_{ATP} channels therefore leads to membrane depolarization. Membrane depolarization triggers opening of voltage-gated calcium channels and influx of calcium into the cytosol, which is sensed by calcium-binding proteins to trigger insulin exocytosis. Glucose-stimulated insulin secretion (GSIS) can also be potentiated by catabolism of glucose and some other fuels through amplifying pathways which work independently of K_{ATP} channels. Known amplifying molecules include NADH, cAMP, Kreb's cycle intermediates citrate and glutamate, as well as some external metabolic and neurohormonal cues [92].

1.2.3 β-cell maturity

Robust positive glucose concentration-dependent insulin secretion is the hallmark of mature β-cells. Maturation of β-cells involves inhibition of insulin secretion at low glucose, an elevated magnitude of insulin secretion at high glucose, and an elevated threshold of stimulatory glucose concentration [93–95]. As outlined in the previous two paragraphs, insulin production, storage, and regulated secretion require coordination of complex metabolic, (post-)translational, and transcriptional processes; mature β-cells can therefore be contrasted with immature β-cells by the changes in these processes.

Mature and immature β-cells have differences in cellular metabolism which reflect differences in nutrient availability. In the fetal period, insulin secretion by immature β-cells is insensitive to

glucose but is stimulated by amino acids [96]. Due to a stable stimulus of maternally supplied amino acids, fetal β-cells secrete a steady supply of insulin which functions largely as a growth factor at this stage [96,97]. In the early postnatal period, nutrient availability becomes less stable due to a switch to intermittent feeding. This triggers a change in β-cell function from amino acid- to glucose-stimulated insulin secretion which is post-transcriptionally driven by dynamic mTORC1 signaling [96]. Later, the dietary transition at weaning from fat-rich milk to carbohydrate-rich food triggers β-cells to switch from mTORC1, a nutrient sensor, to AMPK, an energy sensor [98,99]. This is associated with the establishment of mature GSIS activity and reduced proliferative capacity [98], which appear to be mutually exclusive [100]. Establishment of maturity coincides with increased capacity of the NADH shuttle system, which transports cytosolic substrates of oxidative metabolism into the mitochondria [101], and enhanced mitochondrial oxidative ATP production [92]. Coupling of glucose metabolism to insulin secretion requires selective expression or repression of particular metabolic genes. A large number of enzymes and mitochondrial proteins are specifically "disallowed" in β-cells—that is, repressed in β-cells but otherwise widely expressed [102]. For example, mature β-cells express glucokinase while immature β-cells and other cells express additional other hexokinases which catalyze the same reaction. Glucokinase has a low affinity for glucose which prevents inappropriate triggering of insulin secretion at basal blood glucose concentration [103,104]. β-cells additionally repress genes that would cause inappropriate fuel sources to stimulate insulin secretion. For example, β-cells cannot convert circulating lactate to ATP because *Slc16a1* and *Ldha* are repressed [102,105]. Re-expression of these genes evokes lactate-stimulated insulin secretion which is undesirable since circulating lactate is increased during catabolic states like exercise and hypoglycemia [106]. Notably, the unique metabolic phenotype of mature β-cells is encoded at the level of RNA expression [107].

Control of gene transcription in mature cells tends to favour high expression of genes critical for cell function, along with low or no expression of unneeded genes [108,109]. At the level of transcription, cell maturation therefore constitutes progressive activation and upregulation of the critical gene set, along with general suppression of developmental and unneeded genes [108]. In line with this, β-cell maturation and identity are tightly regulated by gene regulatory networks. Transcription factors—DNA-binding proteins that influence gene transcription—are central to establishment and maintenance of these networks [110]. While no truly β-cell-specific transcription factor has been found, lists of essential β-cell-enriched transcription factors have been penned and continue to expand. Well-accepted members include FOXA2, ISL1, MAFA, NEUROD1, NKX2-2, NKX6-1, PAX6, and PDX1; experimental deletion of any one of these results in dysregulation of the β-cell transcriptional program and to impairment of β-cell function [111–121]. Each of these factors regulates expression of unique but overlapping gene sets in mature β-cells. Upregulation of these transcription factors occurs during development and maturation of β-cells, while reduced expression of some has been reported in islets from individuals with T2D and mouse models of T2D [114,122–124], or are mutated in some cases of monogenic diabetes [125–130]. β-cell-specific deletion of the transcription factor *Foxo1* solidified the idea of β-cell dedifferentiation, or reversion to an immature state – that is, downregulation of mature β-cell-enriched genes and upregulation of lowly expressed, progenitor cell genes and disallowed genes which compromise GSIS [124]. This, and other, work showed that direct impairment of transcription regulation using genetic mouse models, as well as T2D-like metabolic stresses, cause dedifferentiation [56,124,131]. Dysregulation of gene expression is a primary feature of dedifferentiated β-cells [132,133]. Therefore, regulation of gene expression is fundamental to β-cell

maturity in health and disease. Along with transcription factors, there is growing appreciation for the role of histones in the maintenance of β-cell function [134]; this is discussed in section 1.3.

Several genes, such as *Ucn3*, *Mafa*, and *Fltp*, are useful in the systematic identification of mature β-cells because their expression increases during development and they are reliably expressed in metabolically mature β-cells [93,135–137]. Characterization of islet cells using such markers shows that up to 30% of the β-cell population in healthy adult islets are immature [138,139]. This represents one facet of the growing appreciation for β-cell heterogeneity, in which subpopulations of β-cells exhibit distinct metabolic and transcriptional characteristics. Immature *Fltp*-negative cells have greater replicative potential and may be poised to quickly expand β-cell mass to meet metabolic demands [136]. Immature β-cells displaying lower *Ins2* gene activity are more resilient against glucose and ER stresses [140]. "Extreme" β-cells express large amount of insulin mRNAs but are metabolically immature, with a high proinsulin:insulin ratio and high secretion in basal glucose concentration [141]. "Hub" β-cells—metabolically and transcriptionally immature but highly glucose-reactive—are pacemakers for GSIS in the islet at large [142]. A recent publication identified immature β-cells using low expression of *Mafa* and *Pdx1*; forcing high expression of *Mafa* and *Pdx1* by genetic manipulation destabilized ionic fluxes and metabolism in adult mouse islets, suggesting that mature islet function depends on coexistence of mature and immature β-cells [143]. Together, various β-cell subtypes in mature islets contribute to the functionality and resilience of the β-cell population as a whole. On the other hand, gene expression variability, which is a feature of highly heterogeneous cell populations, decreases during islet functional development and is increased in T2D [108,133]. Subtypes of β-cells are dynamically interchangeable

in response to external stimuli such as metabolic stresses; disturbances in the ability of β-cells to transition between states may contribute to β-cell failure [144].

1.3 Epigenetic regulation of transcription

1.3.1 *Cis*-regulatory elements

Genomic DNA encodes genes which can be transcribed into RNA. The nucleotide sequence of genes determines their function. But what genes are transcribed, at what level, and in what conditions, is determined by other DNA-encoded features called *cis*-regulatory elements. They contain specific sequence motifs, often several, that can be bound by transcription factors or other regulatory molecules; the combination and quality of motifs determines the strength of regulation [145,146]. Promoters, enhancers, silencers, and insulators are *cis*-regulatory elements (Fig. 1.1). Promoters are an absolute requirement for transcription of a gene. They define where and in what direction RNA polymerase II (RNAP2) initiates transcription. They are located at or near the gene's transcription start site (TSS) and frequently include motifs recognized by general transcription factors which recruit RNAP2. Enhancers and silencers dynamically interact with promoters to control expression of target genes transcribed by RNAP2. They may act independently of orientation and at great distance (up to a million base pairs away [147], or on a different chromosome [148]) because three-dimensional (3D) DNA looping brings active enhancers and target promoters together [149]. Insulators are boundary elements which limit the propagation of signals on one side to promoters on the other. They are involved in the segmentation of topologically-associated domains and related structures, which restrict 3D colocalization of the DNA strand from between different domains [150].

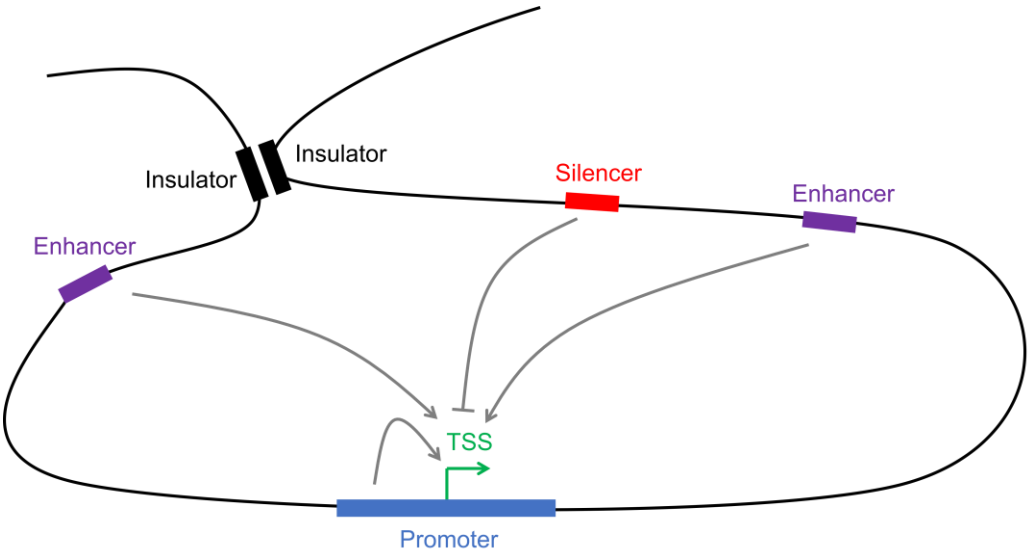

Figure 1.1 | *Cis*-regulatory elements.

1.3.2 Epigenetics

With few exceptions, all cells in a multicellular organism share the same DNA sequence, including genes and *cis*-regulatory elements. Cell-type-specific gene expression programs must therefore be specified and maintained using a regulatory system on top of the genetic sequence. This system is called epigenetics. The term was coined by geneticist Conrad H. Waddington in 1939 to describe how embryogenesis requires both "preformed qualities of the fertilized egg" and "interaction of these constituents (that) gives rise to new types of tissue and organ which were not present originally" [151]. This first definition concerns the process of one genome giving rise to multiple distinct phenotypes, such as differentiated cell states. 'Epigenetics' has been redefined over eight decades of research. In 2019 Giacomo Cavalli defined it "the study of molecules and mechanisms that can perpetuate alternative gene activity states in the context of the same DNA sequence" [152]. This definition encompasses factors governing the propagation of gene activity following mitotic cell division, inter- and transgenerational inheritance, as well as

over extended time, even without cell division, such as in post-mitotic neurons [152,153]. Notably, carriers of epigenetic memory—so-called epigenetic 'marks' or 'maintainers'—also play roles in acutely reversible biological regulation. Examples include chromatin structure and transcription factor activity that are responsive to environmental stresses and nutrient availability but may revert quickly when the environment is restored [154,155]. Nonetheless, even transient activities of epigenetic molecules and mechanisms can have meaningful consequences to a cell's identity, function, and viability [154]. Carriers of epigenetic information include transcription factors, prions, chromatin, small RNAs, DNA methylation, and 3D chromosome organization [152,154].

1.3.3 Histones, nucleosomes, and chromatin

Chromatin is a complex macromolecular assemblage of genomic DNA, proteins, and RNA. The fundamental unit of chromatin is the nucleosome, which consists of 146 ± 1 bp of DNA wound 1.65 turns around an octamer of histone proteins. The histone octamer is assembled from pairs of the proteins H2A, H2B, H3, and H4 [156] (Fig. 1.2). Arrays of nucleosomes, repeated every 200 ± 40 bp, assemble into higher-order structures which are stabilized by a fifth "linker" histone protein, H1. Histones help package the genome within the confines of the nucleus, but tight electrostatic binding between histones and DNA also imposes obstacles to DNA-binding molecules [157,158]. Nucleosome remodeling complexes enzymatically remove, place, or slide nucleosome along DNA to adjust nucleosome positioning and density [159]. Maintenance of nucleosome density within gene bodies prevents cryptic transcription initiated from non-promoter regions [160]. On the other hand, active *cis*-regulatory loci frequently contain nucleosome-depleted regions at sites where transcription factors and RNAP2 may bind DNA [161].

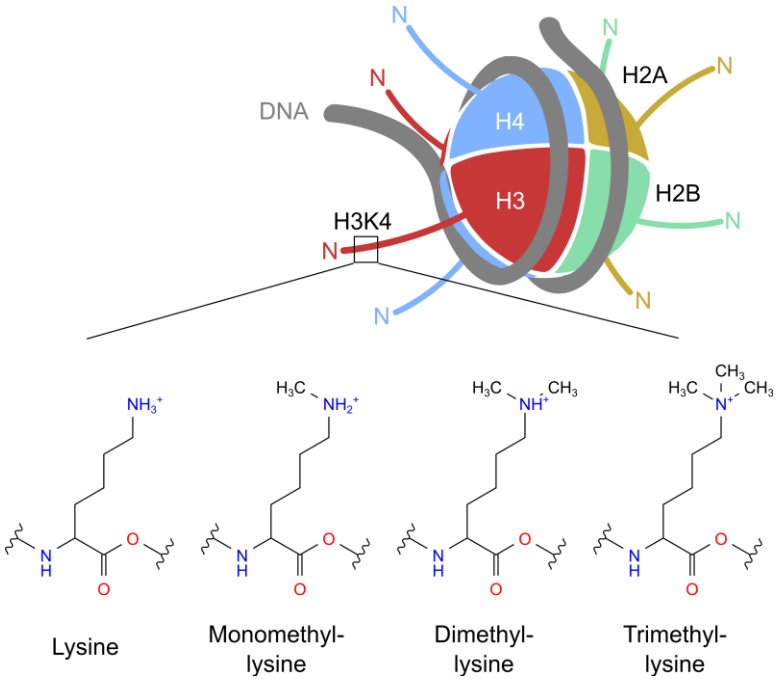

Figure 1.2 | Nucleosome structure and histone H3 methylation. Top: a mononucleosome consisting of DNA wrapped around an octamer of histone proteins comprising two H2A/H2B dimers and two H3 and two H4 monomers. Bottom: methylation products of the ϵ amine group of the fourth lysine of histone H3.

1.3.4 Histone modifications

Histones serve a much more nuanced role than simply impeding transcription, however. In 1964, Vincent Allfrey first showed that histones are post-translationally modified and that acetylation of histones alleviates their inhibition of transcription [162]. Today, over 100 chemical modifications of specific histone residues have been catalogued [163]. Many occur on the flexible amino- (N-) terminal 'tails' of histone proteins which are rich in reactive lysine and arginine residues. The best-studied modifications are lysine acetylation and methylation, and to a lesser extent lysine ubiquitylation, arginine methylation, and serine phosphorylation [164]. More exotic chemical modifications are discovered regularly: benzoylation in 2018 [165], lactylation in 2019 [166], dopaminylation in 2020 [167], isobutyrylation in 2021 [168]. While some are derived through non-

enzymatic reactions [169], cells express enzymes that catalyze addition or removal of posttranslational modifications of specific histone residues – so-called 'writers' and 'erasers', respectively. Most epigenetic writers and erasers do not bind DNA directly but are recruited to DNA through interactions with transcription factors and RNAs. In this way, histones at particular genomic loci are targeted for modification. A major advance in the ability to annotate *cis*-regulatory loci came with the discovery that histone modifications have stereotypical patterns of enrichment at functional genetic elements across the genome and across species.

1.3.5 Histone acetylation

Histone lysine acetylation is almost invariably associated with active chromatin, such as promoters and enhancers of transcribed genes, and levels correlate strongly with transcription. Promoter-associated histone lysine acetylation activates transcription by RNAP2 but is also deposited in response to RNAP2 activity in a feed-forward loop that maintains active transcription [170,171]. Acetyllysine partially neutralizes the histone's positive charge, weakening electrostatic interactions between the histone and negatively charged DNA and increasing accessibility to chromatin-binding factors [172]. Histone acetyllysines can also serve as binding sites that recruit (or repel) a variety of epigenetic 'readers', including some transcription factors, nucleosome-remodeling enzymes, and epigenetic writers and erasers [173]. Active enhancers and promoters are enriched for acetylation of histone H3 lysine 27 (H3K27ac) [174]. Acetylation of several other histone residues is enriched near active enhancers and TSS's, and others are enriched in transcribed gene bodies [174].

1.3.6 Histone methylation

Many lysines targeted by acetyltransferases are also targets of methyltransferases. Because a lysine cannot be simultaneously acetylated and methylated, histone lysine acetylation prevents histone lysine methylation, and vice versa. For example, trimethylation of H3K27 (H3K27me3) opposes H3K27ac and is broadly associated with developmentally silenced regions of chromatin [174–176]. In contrast to acetylation, histone lysine methylation does not substantially modify the electrostatic charge of histones. It can be associated with either activation or repression of transcription, depending on the residue modified and the number of methyl moieties present. Lysines can harbour zero to three methyl moieties, which impacts the binding of some epigenetic readers and their downstream effects. For example, Monomethylation of H3K27 (H3K27me1) can be found in transcribed genes where it promotes transcription, whereas H3K27me2 and H3K27me3 are widely distributed and repress enhancer activity [177]. H3K4me3 is associated with active promoters. The length of the H3K4me3-enriched region of DNA and the density of H3K4me3 in that region (i.e. H3K4me3 peak breadth and height) positively correlate with the level of transcription [178]. Broad H3K4me3 peaks are also linked to tumor suppressor and lineage-enriched genes with consistent expression between cells in a tissue [179,180]. H3K4me1 surrounds H3K4me3-positive loci and is also found at active enhancers [174]. H3K4me2 is intermediate and associated with both active promoters and enhancers [181]. Methylation of H3K9, 20, 36, 79, and others, are variously associated with gene bodies, constitutive heterochromatin, DNA replication, and other DNA features [182–190].

1.3.7 Histone code

The combinatorial potential of different histone modifications, DNA modifications, and histone variants is vast, but the set of combinations observed *in vivo* is much smaller [174]. Stereotypical combinations define chromatin "states": active promoters, containing H3K4me3, H3K27ac, and a nucleosome-depleted region; active enhancers, containing H3K4me1, H3K27ac, and a nucleosome depleted region; poised promoters and enhancers, which lack H3K27ac; bivalent promoters and enhancers, in which H3K27ac is replaced with H3K27me3, and so on [191]. Transcriptional levels are highly correlated with chromatin states [192–194]. The histone code hypothesis posits that chromatin states are not only convenient markers of regulatory loci, but are necessary for the function of those loci [195]. Experimentally parsing apart the necessary histone marks from histone marks that may be useful markers, but are functionally dispensable, is challenging. A source of uncertainty is that functions of individual histone marks are often studied by deleting, repressing, or otherwise disrupting the enzymes that add them, but this approach risks impairing essential noncatalytic functions of those enzymes that work independently of histone marks [196]. Elegant recent studies using time-resolved genomics during rapid degradation of writer enzymes (and slower loss of their products) [197], catalytically inactive but structurally stable writer enzymes [198,199], genetic perturbations of informative combinations of writer complex proteins [200,201], pulse-chase experiments by transient ectopic expression of writer enzymes [202], or ectopic expression of unmodifiable dominant-negative histone proteins [167] are defining the functional impact of histone modifications versus histone modifying enzymes on gene expression.

1.3.8 Trithorax group

The clustered homeotic (*hox*) genes direct body patterning and body part segmentation during development in metazoans. Genetic studies of *Drosophila* mutants identified a group of genes that repress Hox gene expression, referred to as the polycomb group (PcG), and a group that enhance expression or function, referred to as the trithorax group (TrxG) [203]. The eponymous founding TrxG gene *trithorax* (*trx*) was first described in 1980 by Philip Ingham [204]. A spontaneous homozygous mutation caused variegated homeotic transformations in *Drosophila* which sometimes gave rise to flies with six wings attached to three mesothorax compartments, hence the name [205]. Loss-of-function mutations in *trx* resemble loss-of-function mutations in the *hox* genes as far as their impact on body segmentation [206]. It was later discovered that *trx* is needed to maintain active *hox* expression within the correct body segments [207]. Subsequent genetic screens identified additional genes required to activate or maintain *hox* expression or that antagonized the *hox*-repressing effects of known PcG genes; these genes were referred collectively as the TrxG, without knowledge of molecular mechanism [203]. The result is that TrxG proteins are a heterogenous group that play a variety of roles in transcriptional activation and elongation. Today, TrxG can be divided into three classes according to their molecular function. The first contains Trx and other SET domain-containing factors [208]. The SET domain is a 130- to 140-amino acid motif that provides lysine methyltransferase activity. The second class includes ATP-dependent chromatin-remodeling factors such as Brm (SWI2 in mammals) [209]. The third include proteins that can directly bind DNA motifs and others that do not fit in the first two categories, such as the acetyltransferase and coactivator Fsh (BRD4 in mammals) [207,210].

1.3.9 Complex of proteins associated with Set1 (COMPASS) family of H3K4 methyltransferases

Drosophila encode three TrxG H3K4 methyltransferases, Trx, Trr, and Set1, while yeast encode just one, Set1 [208,211]. Studies in yeast first characterized the Complex of Proteins Associated with Set1 (COMPASS) [212]. In mammals there are two COMPASS complexes that associate with SETD1A or SETD1B, plus four homologous COMPASS-related complexes that associate with one of MLL1-4 [206,213] (Fig 1.3). Together, these six TrxG complexes catalyze the majority of H3K4 mono-, di-, and trimethylation in mammalian cells [214].

SETD1A complexes write most H3K4me3 in mammalian embryonic cells [215]. They are recruited via direct interaction with RNAP2 during transcriptional elongation and their methyltransferase activity is further stimulated by H2B ubiquitination, which is added to genes co-transcriptionally. This leads to broad regions of H3K4 methylation at promoters and extending into the gene body of actively transcribed genes [216,217] , and causes H3K4me3 levels to drop when transcription is inhibited [194]. This system may provide a molecular memory of gene activity [218]. Conditional deletion of SETD1B leads to downregulation of highly expressed genes with broad H3K4me3 peaks [219]. SETD1A and SETD1B show distinct enrichment across regions of active chromatin and likely regulate different genes [220]. The same is true for MLL1-4. MLL1 is enriched at the majority of active promoters with bound RNAP2, suggesting a general role in transcription, and MLL1 occupancy is strongly associated with H3K4me3 enrichment, suggesting a primary activity of H3K4 trimethylation [218]. MLL1 and MLL2 share similar domain architecture but are not interchangeable. Whereas leukemogenic chromosomal rearrangements frequently involve the *MLL1* gene, they never involve *MLL2* [221]. Knockdown of MLL2 does not appreciably reduce

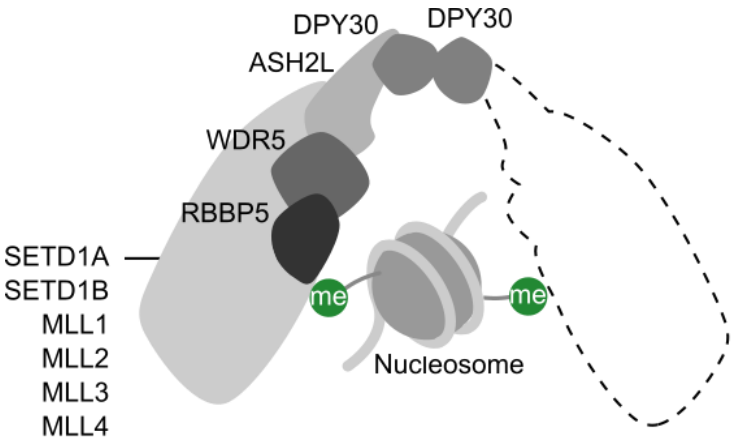

Figure 1.3 | COMPASS and COMPASS-related complexes. All H3K4 methyltransferases of the TrxG contain, at minimum, the core structural proteins WDR5, RBBP5, ASH2L, and a DPY30 homodimer, plus one of six SET domain-containing methyltransferases SETD1A-B or MLL1-4 [208]. A nucleosome symmetrically methylated at both H3K4 is shown. The dashed outline depicts a second COMPASS or COMPASS-related complex; COMPASS is dimeric in yeast [222] but this has not been demonstrated in mammals. Adapted from Choudhury *et al.* (2019) [222].

global H3K4me3 [223]. MLL2 instead establishes H3K4me3 at developmentally regulated, bivalent gene promoters [223]. MLL3 and MLL4 catalyze H3K4me1 at enhancers with partial redundancy [224]. Recruitment of MLL3 and MLL4 to enhancers is necessary for gene activation during cell differentiation [225]. Deletion of MLL3 and MLL4 decreases H3K4me1 and RNAP2 occupancy at enhancers, and reduces expression of the associated genes [226].

Nonenzymatic subunits of COMPASS and COMPASS-related complexes are variable but always include the four core structural proteins WDR5, RBBP5, ASH2L, and DPY30 (Fig. 1.3) [227]. In the intact cell, these components are involved in recruitment to *cis*-regulatory loci via interactions with transcription factors and cofactors, histone modifications, and noncoding RNAs [228–237]. Additionally, cell-free reconstitution of these proteins shows that they boost

methyltransferase activity of the MLL proteins up to several hundred fold, particularly for di- and trimethylation [227,233]. Direct binding of WDR5, RBBP5, and ASH2L to MLL1 optimizes SET domain conformation and rigidity to increase catalytic activity [228,238,239]. DPY30 homodimers bind to ASH2L at a position too distant to impact SET domain structure, yet, exclusion of DPY30 reduces *in vitro* methyltransferase activity between 1.5- and 3-fold [240–242]. Stranger still is that loss of DPY30 from intact cells causes global (i.e. genome wide) loss of H3K4 methylation, apparently exceeding what can be explained by the reduction in catalytic activity [242]. Choudhury *et al.* (2019) present a plausible explanation using yeast genetic models [222]. They show that the yeast COMPASS complex actually exists as a dimer, with Sdc1, the yeast homologue of DPY30, at the dimer interface [222]. Dimeric COMPASS introduces symmetrical H3K4 methylation on both H3 N-terminal tails of a nucleosome [222]. Yeast with mutants of Sdc1 that prevent COMPASS dimerization display dramatic loss of H3K4me3 and moderately reduced H3K4me1, not just because of reduced catalytic efficiency of Set1, but because asymmetrically methylated nucleosomes are specifically targeted for demethylation by Jhd2 [222]. While this has not been confirmed in mammals, it suggests that loss of H3K4 methylation in DPY30-KO cells is caused by both reduced methyltransferase activity and increased demethylase activity.

In the same manuscript, Choudhury *et al.* (2019) show that reduction of H3K4 methylation in Sdc1 mutant yeast imparts only a modest effect on transcription, causing dysregulation of only 4% of genes, and with no apparent relationship between the change in gene expression and H3K4 methylation at each gene [222]. Some argue that H3K4me3 is not related to transcription in yeast at all [243]. In *Drosophila* embryos, blocking H3K4 methylation by substitution of H3K4

with alanine or arginine causes variegated growth defects, but cells are still able to activate and maintain target gene expression in response to developmental signals [244]. In mouse embryonic stem cells (mESCs), deletion of a subunit of SET1A complexes causes loss of most H3K4me3 from target genes but minimal changes to transcriptional activity of those genes [245]. Therefore, in yeast, *Drosophila*, and mice, most genes do not require H3K4 methylation to be expressed. Although H3K4 methylation is well correlated with [192,193] and predictive of [246–248] transcription, it is not clear whether it is possible to predict which, if any, genes are actually regulated by H3K4 methylation.

The biological functions of TrxG H3K4 methyltransferases are not entirely related to their catalytic activities. These large multimeric complexes can also regulate gene activity by binding and bridging interactions between other chromatin factors, and this may furthermore be the more biologically impactful function. For example, MLL4 directly binds to the H3K27 acetyltransferase P300 and H3K27 demethylase UTX; the combined presence of these three factors at *cis*-regulatory loci establishes an active enhancer landscape in mESCs [249]. MLL3 and MLL4, the major writers of H3K4me1 at enhancers, are involved in RNAP2 recruitment to enhancers and downstream expression of target genes, but their methyltransferase activity, and thus H3K4me1, is not required for these functions [226,250]. Similarly, in human hematopoietic stem cells, the SET domain of MLL1 is not necessary for hematopoietic target gene expression or for the leukemogenic effect of MLL1 mutant oncoprotein [251]. Instead, the MLL1 complex appears to regulate target gene expression in these cells by recruiting MOF, a histone acetyltransferase [251]. Deletion of the SET domain of *Set1a* in mESCs reduces H3K4 methylation at bivalent genes and impairs differentiation to embryoid bodies, but has no effect on mESC self-renewal and

maintenance [252]. Deletion of the complete *Set1a* gene, meanwhile, causes mESC cell death [252]. Overall, TrxG H3K4 methyltransferases perform specialized, rather than fundamental, roles in the regulation of gene transcription that do not always rely on enzymatic activity.

1.3.10 H3K4me1 molecular functions

H3K4me1 occurs on nearly one third of all H3 histones in mESCs [226]. It is particularly enriched at enhancers, where levels are highly dynamic and well correlated with lineage-specific gene expression [253]. Conserved peptide motifs have been identified that bind preferentially to H3K4me1, such as the MBT and PHD domains [254], linking H3K4me1 to proteins that can modify gene expression. Mass spectrometry approaches have catalogued H3K4me1-associating proteins, most notably the TrxG ATP-dependent chromatin remodeling complexes [255]. Following up on one, Local *et al.* (2018) show that BAF preferentially binds to and moves nucleosomes marked with H3K4me1 [255]. H3K4me1, but not H3K27ac, is necessary to establish an open chromatin state at enhancers [226,256]. H3K4me1 is also involved in the recruitment of Cohesin, a protein complex involved in enhancer-promoter looping [199,255]. Accordingly, H3K4me1 promotes physical interactions between enhancers and promoters that regulate gene transcription levels [199]. During mESC differentiation to neural progenitor cells, MLL3 and MLL4 deposit H3K4me1 at a subset of induced enhancers; contacts between these enhancers and target promoters are severely limited in cells expressing catalytically dead point mutants of MLL3 and MLL4, which leads to impaired activation of lineage-specific genes and delays differentiation [257]. The degree to which individual genes are affected is proportional to the number of MLL3/4-regulated enhancers they contact during differentiation [257]. Furthermore, loss of H3K4me1 leads to partial depletion of H3K27ac in enhancers [226,256]. Overall, H3K4me1 fine-tunes enhancer

activity [258] by maintaining DNA accessibility and contributing to enhancer-promoter contacts. H3K4me1 is also enriched at the borders of promoters. Since H3K4 cannot be simultaneously mono- and trimethylated, H3K4me1 delimits the boundaries of H3K4me3-positive TSS-proximal regions—effectively repressing undesirable H3K4me3-dependent transcription initiation from outside of that region [259].

1.3.11 H3K4me3 molecular functions

H3K4me3 occurs on 0.35% of H3 molecules in mESCs [226] and is enriched almost exclusively near TSSs of actively transcribed or poised genes [260]. A functional role for H3K4me3 in promoters is supported by the demonstration that replacing it with H3K4me1 decreases transcription [259]. Further, ectopic addition of H3K4me3 to inactive gene promoters is sufficient to activate sustained gene expression (but is blocked by DNA methylation) [202]. Mechanistically, epigenetic readers have been catalogued that bind specifically to H3K4me3 via conserved PHD, Tudor, or chromodomain peptide motifs [261–264]. Most notable is the general transcription factor TAF3, part of the RNAP2 preinitiation complex, which is recruited to active promoter nucleosomes by selective anchoring to H3K4me3 [265]. This likely has gene-specific rather than fundamental roles in transcription since TAF3 depletion alters expression of only 5.6% of genes [178] and can be recruited in other ways, such as a TATA box DNA motif. Also notable is the SAGA complex, a multimeric transcriptional coactivator, which is recruited to H3K4me3 and acetylates nearby histones [263]. Disrupting SAGA complex activity alters expression of 8.2% of genes [266], again indicating context-specific effects. H3K4me3 may also serve anti-repressive roles. Transcriptional repression by H3K27me3 is blocked in H3K4me3-enriched loci because PcG catalytic activity is allosterically inhibited by H3K4me3 [267]. For this reason, trimethylation

of H3K4 and H3K27 almost never occur on the same histone *in vivo* [268]. Similarly, transcriptional silencing by *de novo* DNA methylation is blocked by mono, di, or trimethylation of H3K4 because DNMT3L binds specifically to the unmethylated residue [269]. In this way, H3K4me3 may maintain plasticity of reversibly repressed bivalent genes by preventing irreversible silencing [270]. Overall, H3K4me3 regulates transcription via direct and indirect recruitment of transcriptional machinery and by antagonizing deposition of repressive chromatin modifications. Due to input from other epigenetic factors, transcription can occur in the absence of H3K4me3 and it is difficult to predict what genes, under what circumstances, strictly require H3K4me3 [178]. H3K4me3 may also have a non-transcriptional role in the recruitment of DNA repair machinery to sites of DNA damage, meiotic recombination [178], or antigen receptor gene (V(D)J) recombination [264].

1.3.12 Cellular functions of H3K4 methylation in mature islets

While the genome-wide distributions of some histone marks have been characterized in adult mouse and human islets [133,271–274], it is usually in their capacity as markers to construct maps of *cis*-regulatory loci. Links to phenotypes or functions in these and other studies have been correlative and limited data is available on the function of H3K4 methylation in mature islets.

Writers and erasers of H3K4 methylation are implicated in mature β-cell metabolic functions. Islet cells express SETD7, a non-TrxG lysine methyltransferase, which is recruited to the *Ins1/2* locus by PDX1 to support *Ins1/2* gene expression, perhaps by depositing H3K4me2 [275]. Deletion of SETD7 in β-cells *in vivo* paradoxically increases *Ins1/2* expression; however, it should be noted that this function may be independent of histone methylation since SETD7 also methylates

PDX1 itself to augment its regulatory action [276]. Deletion of the H3K4 mono- and di-demethylase *Lsd1* leads to insulin hypersecretion and aberrant expression of genes involved in nutrient processing in adult β-cells; whether this effect is mediated by an increase in H3K4 methylation is not clear [277]. Recruitment of MLL3/4 COMPASS-related complexes by direct binding with MAFA and MAFB may be necessary for the transcription activation of most MAFA/B target genes [278]; again, the involvement of H3K4me1 is not clear.

H3K4me3 is implicated in maintenance of α- and β-cell identities. The α-cell lineage-determining gene *Arx* is marked by H3K4me3 in α-cells but not β-cells, in which *Arx* is instead enriched for DNA methylation [279]. Deletion of *Dnmt1*, a DNA methyltransferase, in β-cells causes loss of DNA methylation and gain of H3K4me3 at *Arx*, and β-cells to express *Arx* and transition to α-cells [279]. In α-cells, genes encoding β-cell transcriptional regulatory proteins are often in bivalent chromatin marked by both H3K4me3 and H3K27me3; unspecific chemical inhibition of histone demethylases promotes partial reprogramming of α-cells toward a β-cell fate [280]. PAX6, another islet-enriched transcription factor involved in maintaining β-cell function and identity [114], has also been shown to recruit active COMPASS and COMPASS-related complexes to target promoters and enhancers in lens epithelial cells [281] but this has not been tested in islet cells.

H3K4me3 is implicated in suppressing replication in β-cells. Increased MLL1 and H3K4me3, along with decreased H3K27me3 at the $p16^{INK4a}$ and $p19^{ARF}$ (cyclin-dependent kinase inhibitors; i.e., repressors of replication) gene loci is linked to their de-repression and reduced β-cell proliferation in aged islets [282]. Similarly, the MLL1/2 COMPASS-related subunit MEN1

maintains H3K4me3 at cyclin-dependent kinase inhibitor genes *p18*^{INK4c} and *p27*^{Kip1} [283]. Mice heterozygous for *Men1* exhibit reduction of H3K4me3 and downregulation of $p18^{INK4c}$ and $p27^{Kip1}$ in islets, β-cell hyperplasia, and hypoglycemia [283]. Promoters of genes upregulated with advanced age in islets show an increase in H3K4me3 [284]. But interestingly, genes downregulated with advanced age appear to maintain juvenile levels of H3K4me3. Downregulation does, however, coincide with increased H3K27me3 at promoters [284].

1.4 One-carbon metabolism

Methylation of H3K4 requires a supply of one-carbon units. Nutrients such as glucose, B vitamins, and certain amino acids are carriers of transferable methyl moieties that feed into a network of biochemical reactions collectively called one-carbon metabolism. A major product of one-carbon metabolism is *S*-adenosylmethionine (SAM), the methyl donor used by histone and DNA methyltransferases. The folate cycle, methionine cycle, and transsulfuration are central pathways of one-carbon metabolism (Fig. 1.4) [285].

1.4.1 Folate cycle

Folate coenzymes are a group of methyl carriers categorized based on their similar structure and chemistry to the stable synthetic species folic acid (vitamin B9) [288]. Animals cannot synthesize folates *de novo* but express transport proteins and enzymes that manipulate them in support of a variety of metabolic processes [286]. 5-methyltetrahydrofolate (5-MTHF) is the most abundant folate species in circulation and in non-hepatic tissues [289]. It is transported into cells through reduced folate carrier 1 (RCF) or through proton-coupled folate transporter (PCFT) in some tissues [290]. Once inside, 5-MTHF is retained in the cell by polyglutamation [285]. 5-MTHF serves

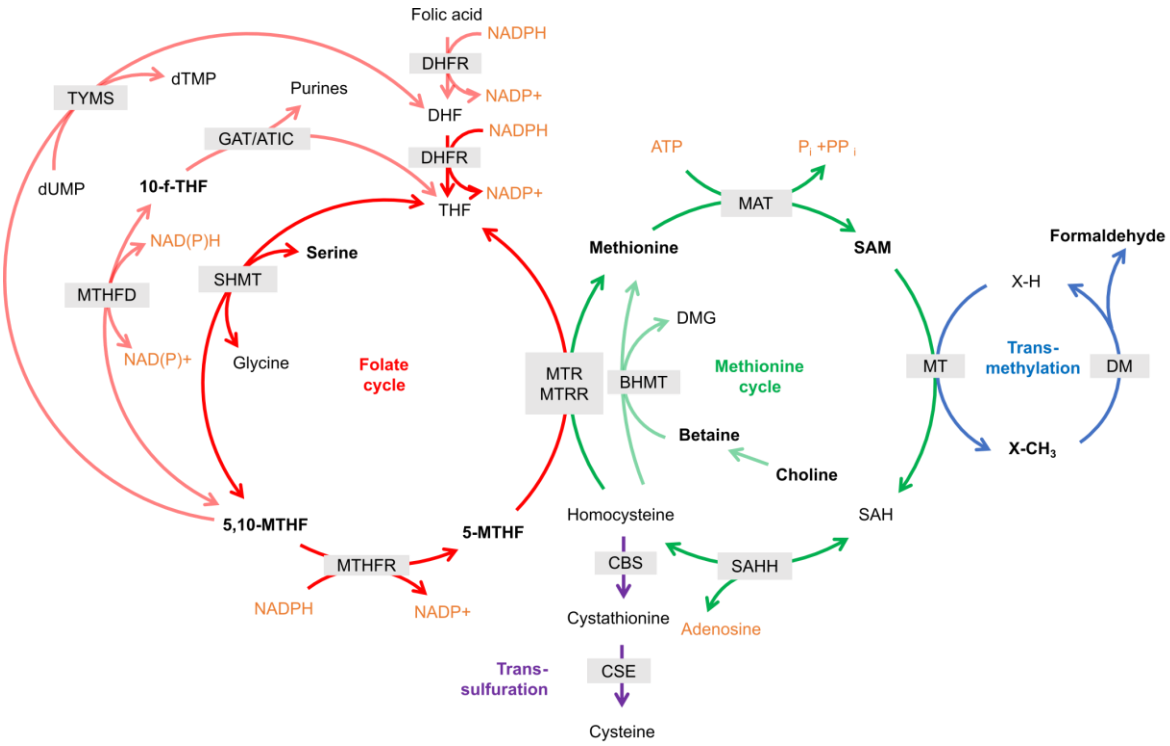

Figure 1.4 | One carbon metabolism. Black text depicts metabolites. Metabolites with transferable methyl moieties are in bold. Gray boxes depict enzymes. Some pathways are tissue-specific: DHFR activity is predominantly in the small intestine and liver, BHMT is expressed in liver and kidneys, transsulfuration occurs in liver, brain, and pancreas [286]. Adapted from Mentch *et al*. (2016) [287].

as a methyl donor for the conversion of homocysteine to methionine catalyzed by the vitamin B12-dependent enzyme methionine synthase (MTR). After catalysis, MTR reductase (MTRR) restores MTR activity. The product of 5-MTHF demethylation is tetrahydrofolate (THF), which can be reactivated by loading of a methyl unit to the N5 and/or N10 positions. The major source of these methyl groups is C3 of serine, which itself can be derived from intermediates of glycolysis and gluconeogenesis [291]. The methyl unit is transferred to THF by the B6-dependent enzyme serine hydroxymethyltransferase (SHMT), generating glycine and 5,10-methylenetetrahydrofolate (5,10-CH_2-THF), which is subsequently converted to 5-MTHF by the

B2-dependent enzyme methylenetetrahydrofolate reductase (MTHFR), completing the folate cycle, or to 10-formyl-tetrahydrofolate (10-f-THF) by the enzyme methylenetetrahydrofolate dehydrogenases (MTHFD).

In addition to their actions supporting homocysteine remethylation, folates 10-f-THF and 5,10-CH_2-THF donate one-carbon units for purine and thymidine synthesis, respectively [285]. In rapidly dividing mammalian cells in culture, the largest burden for folate one-carbon groups is in diversion to DNA and RNA synthesis [286] and essentially all *de novo* generated purines and thymidines include carbons sourced from serine via a folate intermediate [292]. Unmet demand for nucleotide synthesis may be responsible for the well-known neural tube defects caused by folate deficiency during pregnancy [293]. Furthermore, one-carbon transformations are red/ox reactions. Oxidation of 5,10-CH_2-THF to 10-f-THF is coupled to reduction of NAD^+ to NADH and represents a biologically meaningful source of reducing equivalents in mitochondria and cytosol [292]. 10-f-THF can also be fully oxidized to CO_2 to generate NADPH and ATP [291]. THF has been found tightly bound with mitochondrial demethylase enzymes; formaldehyde, the product of demethylase reactions, reacts with THF to produce 5,10-CH_2-THF [294,295]. In this way, THF protects the enzymes from oxidative cross-linking and scavenges formaldehyde for re-entry into the one-carbon pool [296].

1.4.2 Methionine cycle

Methionine is an essential amino acid, although it can be synthesized by methylation of homocysteine. Circulating levels of methionine average 30 μM in healthy adults but vary widely – more than any other coding amino acid [297]. Over half of the daily dietary intake of methionine

is converted to SAM, especially in the liver, which performs ~85% of all transmethylation reactions in the body [298]. SAM serves as the principal methyl donor for a large variety of intracellular reactions. It is between 10 μM to 90 μM intracellularly but less than 0.5 μM in serum so each cell synthesizes SAM to meet its needs [287,298]. SAM is created by the reaction catalyzed by methionine adenosyltransferase (MAT) between methionine and ATP. Mammals possess two MAT isozymes, MAT1A and MAT2A. MAT1A is expressed in hepatocytes and bile duct epithelial cells, whereas MAT2A is widely expressed in other cell types alongside a non-enzymatic accessory protein MAT2B which stabilizes MAT2A homodimers [299,300]. SAM-dependent transmethylation reactions yield the demethylated byproduct *S*-adenosylhomocysteine (SAH) [298]. SAH is a potent competitive inhibitor of SAM-dependent transmethylation reactions. For this reason, and because the physiological concentration of SAM is near the K_M of SAM-dependent methyltransferase enzymes [287], methyltransferase enzymatic activity is sensitive to both SAM and SAH concentrations. Therefore, the SAM to SAH ratio is frequently taken as an indicator of cellular "methylation potential" [301]. Differences in kinetic parameters of MAT isozymes lead cells that express MAT1A to have much higher levels of SAM, SAM to SAH ratio, and DNA methylation than cells that express MAT2A [301]. In either case, SAH is converted to homocysteine in a reversible reaction catalyzed by SAH hydrolase (SAHH). Kinetics favour the reverse reaction so rapid removal of homocysteine by export or transformation is essential [301]. Liver and kidney are major sites of homocysteine metabolism, although the gut and pancreas also express all necessary enzymes. Homocysteine can be remethylated into methionine with involvement of the folate cycle as described. In some tissues, a parallel folate-independent pathway generates methionine by transfer of a methyl group from betaine to homocysteine, catalyzed by betaine-homocysteine *S*-methyltransferase (BHMT) in complex with vitamin B6,

and generating dimethylglycine (DMG) and methionine [285]. A second potential fate for homocysteine is metabolism through the transsulfuration pathway, described very briefly below. The third potential fate is auto-oxidation with subsequent formation of reactive oxygen species like H_2O_2 [302]. A high level of circulating homocysteine is an independent risk factor of coronary, cerebral, and peripheral atherosclerosis, partly due to oxidative damage and inflammation in cells of blood vessel walls [303].

1.4.3 Transsulfuration

In cells of the liver, pancreas, and brain, homocysteine can be processed via the vitamin B6-dependent transsulfuration pathway. The first reaction, catalyzed by cystathionine β-synthase (CBS), condenses homocysteine with serine to yield cystathionine. In the second reaction, cystathionine is converted to cysteine by cystathionine γ-lyase (CSE). Cysteine can then be used in protein synthesis, production of coenzyme A, or to generate the antioxidant glutathione. Half of the glutathione in liver is derived via transsulfuration of homocysteine [304]. High levels of cellular SAM promote entry of homocysteine into the transsulfuration pathway by allosteric inhibition of MTHFR and allosteric activation of CBS [286]. In this way, the share of demethylated methionine cycle metabolites diverted away from SAM regeneration increases when SAM levels are high. The caveat is that cysteine production is limited by methylation capacity since homocysteine is an essential precursor. Thus, liver cells express glycine *N*-methyltransferase (GNMT) to decrease the SAM/SAH ratio by transferring the methyl group from SAM to glycine, generating sarcosine and SAH, which is then available for transsulfuration [305]. In yeast, products of transmethylation reactions, including histone methylations, act as “methyl sinks” to ensure

sufficient quantities of SAH are available for production of essential sulfur metabolites downstream of transsulfuration [306].

1.4.4 Regulation of chromatin methylation by one-carbon metabolism

DNA and histone modifications are metabolic products. SAM is the methyl donor for methylation; folates support SAM synthesis; α-ketoglutarate is required for demethylation; succinate and fumarate inhibit demethylases. Therefore, levels and compartmentalization of these metabolites impact levels of DNA and histone methylation. Because the majority of methyltransferase reactions occur in the nucleus, MAT2A is enriched in the nuclear compartment and interacts with several nuclear enzymes to provide SAM for local use [307–309]. Similarly, the H3K4- and H3K9-specific demethylase LSD1 is a THF-binding protein [310]. Since THF is implicated in quenching and scavenging formaldehyde evolved during demethylase reactions, folate may protect LSD1 and DNA from oxidative damage during enzymatic histone demethylation [310].

More generally, excessive intake of methionine, choline, betaine, and folate leads to DNA hypermethylation [285] and inadequate intake of methionine, choline, or folate leads to decreased levels of SAM and global DNA hypomethylation in liver cells [311]. Elevation of SAH, such as occurs in hyperhomocysteinemia, is also sufficient to induce global DNA hypomethylation [312]. Global DNA hypomethylation is a consistent feature of cancer, likely because transformed cells require large quantities of methionine to support a poorly understood methyl sink whereby most methyl groups are diverted away from other uses like DNA synthesis and methylation [313]. On the other hand, a methionine- and choline-deprived diet causes hepatic carcinoma in rats, likely due

to transcriptional destabilization downstream of DNA hypomethylation [313]. Chronic alcohol consumption (a risk factor for liver cancer) leads to hepatic DNA hypomethylation [314] caused at least in part by inhibition of MTR leading to reduction of methionine [315]. Dietary methionine restriction causes reduction of H3K4me3 in liver cells [297], and transformed (but not un-transformed) fibroblasts [316], probably related to the high methionine demand displayed by liver and transformed cells. Restriction of threonine, which provides essential one carbon units for THF remethylation in stem cells, causes dramatic and specific loss of H3K4me3 in mESCs [317]. It is not clear why methylation of H3K4 is more sensitive than other H3 lysines to methionine and threonine restriction. Also via THF remethylation, glucose-derived one carbon units fuel SAM generation and an increase in H3K36me3 that reprograms macrophages in response to inflammatory stimulation [318]. Overall, global cellular abundance of DNA and histone methylation are responsive to availability of SAM within the *in vivo* physiological range of SAM precursors.

Genetic ablation of *Mtrr* causes global DNA hypomethylation in liver and placenta but not brain or germ cells [319,320], indicating that links between DNA methylation and one-carbon metabolism show tissue-specific effects. Even if a cell does not display global reduction of DNA methylation, disruption of one-carbon metabolism can affect specific loci that are linked to gene expression changes [321]. Indeed, folate deficiency causes intergenerational epigenetic inheritance of congenital malformations that may be encoded in locus-specific changes in DNA methylation [321,322] and H3K4me3 [323].

Enzymatic activities of different histone methyltransferases are governed by unique kinetic properties. Fluctuations in substrate availability are therefore expected to impact their activities to different degrees. For example, writers of H3K27me3 have a K_M for SAM of 1.24-1.64 μM [287] while SMYD2, a writer of H3K36me3, has a K_M of 0.07 μM [324]. One would therefore predict that H3K36me3 is less altered by a decrease in SAM than H3K27me3 within the physiological range (10-90 μM).

1.4.5 One-carbon metabolism in mature islets

While there are strong links between one-carbon metabolism and glucose homeostasis in insulin target tissues [286,325–334], the field of islet biology has shown little interest in one-carbon metabolism. Mice fed a folate-restricted diet show modest impairment of glucose tolerance, insulin tolerance, as well as increased adiposity and circulating homocysteine [330]. A high-methionine/low folate diet gives a similar phenotype [329], but in both studies it is not clear how islet functions are affected. Some work using transformed β-cell lines suggest a role for one-carbon metabolites in GSIS. Homocysteine exposure reversibly impairs insulin secretion in response to a variety of insulinotropic agents, including glucose and KCl in the rat BRIN-BD11 β-cell line [335]. Addition of catalase to homocysteine-treated BRIN-BD11 cells effectively removed H_2O_2 derived from homocysteine auto-oxidation, but did not restore GSIS, suggesting H_2O_2 is not key to the detrimental actions of homocysteine [336]. Maintenance of the RINm5F rat insulinoma β-cell-line in folate-deficient media for one to two weeks causes accumulation of homocysteine and H_2O_2, impedes insulin production and secretion, and ultimately triggers apoptosis [337]. Expression of the cell surface folate transporter (FOLR1) is a marker of pancreatic endoderm cells during differentiation of human stem cell derived β-cells [338]. A recent screen for

genes that improve β-cell regeneration in zebrafish discovered that *folr1* overexpression or supplemental folate potentiates β-cell differentiation from pancreatic ductal cells [339]. Metabolomics profiling suggest folate supplementation may have decreased the burden of glycine, serine, and threonine to provide one-carbon units during β-cell regeneration [339]. Folate supplementation affected nucleotide synthesis and fatty-acid related processes, but not methionine or SAM levels [339]. Finally, human islets exposed to acute lipotoxic stress upregulate genes involved in the folate pathway, including *SHMT2* and *MTHFD2*, and exhibit slight but significant increase in global DNA methylation; whether this was due to greater methylation capacity was not determined [340].

1.5 Thesis objectives

Transcriptomic studies of tissues during health and disease uncover directional changes in gene expression that are generalizable across biological samples [341]. This is the case in islets from donors with T2D and models of T2D, which show common sets of induced and repressed genes between donors [342–347]. This suggests that the pathogenesis of metabolic diseases, including T2D, involves changes on chromatin, chromatin modifying enzymes, or metabolites used by chromatin modifying enzymes, which can contribute to persistent metabolic dysfunction via effects on gene expression [133,341,348]. While evidence is accumulating for a role of H3K4 methylation in the regulation of gene expression during specialized developmental events in mammals, few data are available for a continued role in mature tissues [215,225,349,350]. The objective of this thesis was to understand whether H3K4 methylation regulates gene expression in mature β-cells and to test how external metabolic factors—T2D or methylation potential—impact H3K4me3 in mature β-cells. In Chapter 2, I describe the experimental models and methods used in the pursuit of this

thesis. In Chapter 3, I establish that H3K4me3 is a meaningful signal for transcription in mature β-cells and present general features of gene dysregulation associated with loss of H3K4 methylation. In Chapter 4, I show that H3K4me3 peak breadth dynamics encode transcriptional information in healthy and diabetic β-cells, and link H3K4me3-dependent gene expression with the maintenance of glucose responsiveness. In Chapter 5, I test how dietary restriction of folic acid impacts systemic one-carbon metabolism and H3K4me3 in islets. In Chapter 6, I summarize the findings from Chapters 3-5 and discuss implications arising from novel findings of this research in the chromatin and islet biology fields.

Chapter 2: Materials & Methods

2.1 Animal husbandry

Mice were housed at the British Columbia Children's Hospital Research Institute (BCCHRI) Animal Care Facility and procedures were approved by the University of British Columbia (UBC) Animal Care Committee under certificates A17-0045, A18-0111, and A20-0120. Unless otherwise indicated, all *in vivo* and *ex vivo* mouse experiments and measurements were performed in adult male mice between 7- to 33-weeks of age. Up to five littermates were housed per cage with temperature and humidity control under a 12-hour light/dark cycle.

2.2 Mouse strains

Mice harboring conditional alleles of *Dpy30* (derived from EUCOMM EM:09575 [351]) and *Rosa26*^{mTmG} (Jax 007576 [352]) were crossed with *Pdx1-CreER*^{Tg} (Jax 024968 [353]) or *Ins1*^{Cre} (Jax 026801 [354]) 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*$^{flox/+}$. Genotype used for folic acid restricted diet studies was: *Ins1*$^{Cre/+}$; *Rosa26*$^{mTmG/+}$. Mice were maintained on a mixed genetic background. Genotyping primers are listed in Table 2.1. 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.

Table 2.1 Genotyping primer sequences

Genotyping PCR	Primer 1	Primer 2	Primer 3
Pdx1-CreER	AACCTGGATAGTGAAACAGGGGC	TTCCATGGAGCGAACGACGAGACC	-
Ins1-Cre	GGAAGCAGAATTCCAGATACTTG	GTCAAACAGCATCTTTGTGGTC	GCTGGAAGATGGCGATTAGC
mTmG	CTCTGCTGCCTCCTGGCTTCT	CGAGGCGGATCACAAGCAATA	TCAATGGGCGGGGGTCGTT
Floxed *Dpy30*	GTGAGTGCCAGGAACCAAAT	GTTGTGAGCTGCCATGAAGA	-

2.3 Diets

Mice were fed *ad libitum* with a standard chow diet (Teklad 2918) unless otherwise specified. Mice in folate-restriction studies were transferred to folic acid-restricted (FR) or normal-folic acid (NF) diets at 8-weeks-old. FR and NF diets are custom amino acid-defined diets formulated by Research Diets Inc (New Brunswick, NJ, USA); compositions of the diets are provided in Appendix A.

2.4 Cell culture

Primary mouse islet cells were maintained in RPMI 1640 medium containing 11 mM glucose (Gibco), 10% (v/v) FBS (Gibco), 50 U/ml penicillin, 50μg/mL streptomycin (complete RPMI) in 37°C, 5% CO_2 atmosphere. Folic acid-free RPMI (Gibco) and 10K MWCO dialyzed FBS (Gibco) were used for islets collected from NF or FR diet-fed mice. *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, with passaging once per week.

2.5 Islet and β-cell isolation

Pancreata were perfused with 3 to 5 mL 700 U/mL collagenase type XI (Sigma-Aldrich) in Hank's Buffered Saline Solution (HBSS: 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 shaking and hand-picking of islets. To isolate mouse β-cells, islets from *Pdx1-CreER*$^{Tg/0}$; *Dpy30*$^{+/+}$; *Rosa26*$^{mTmG/-}$ or *Pdx1-CreER*$^{Tg/0}$; *Dpy30*$^{flox/flox}$; *Rosa26*$^{mTmG/-}$transgenic mice were washed 3 times in 10 mL room temperature PBS + 2 mM EDTA with centrifugation at 200 × g for 1 min at room temperature, and then dispersed by constant gentle pipetting in 1 mL of 0.025% Trypsin-EDTA (Thermo Fisher Scientific) in PBS at room temperature for 8-10 minutes. Trypsin was quenched by addition of 9 mL room temperature complete RPMI, cells were centrifuged for 3-min at 600 × *g* at 4°C and resuspended in 0.3-0.5 mL ice-cold sorting buffer (PBS, 2 mM EDTA, 1% (w/v) BSA). After passing through a 70 μm cell strainer, singlet eGFP+ tdTomato- cells were enriched by fluorescence-activated cell sorting (FACS) using a FACS-Aria II with a 100 μm nozzle.

2.6 Chromatin immunoprecipitation sequencing (ChIP-seq)

A native ChIP protocol [355] was used to generate ChIP-seq libraries in biological duplicate, with modifications. For *Pdx1-CreER*$^{Tg/0}$; *Dpy30*$^{+/+}$; *Rosa26*$^{mTmG/+}$ and *Pdx1-CreER*$^{Tg/0}$; *Dpy30*$^{flox/flox}$; *Rosa26*$^{mTmG/+}$ mice, 100,000 eGFP+ tdTomato- islet cells were pooled with 50,000 *Drosophila* S2 cells by FACS. For *Lepr*$^{db/db}$ and *Lepr*$^{+/+}$ mice, 100,000 dispersed islet cells were counted by hemacytometer and pooled with 50,000 *Drosophila* S2 cells without sorting. Pooled cells were centrifuged at 600 × *g* for 5 minutes at 4°C, supernatant removed, and then cell pellets were flash-frozen and stored at -80°C for up to one month. Upon thawing, cells were permeabilized in

nuclear isolation buffer (0.1% Triton X-100, 0.1% sodium deoxycholate) containing 1 mM phenylmethylsulfonyl fluoride (PMSF) and EDTA-free protease inhibitor cocktail (Roche)) on ice. Chromatin was fragmented to mononucleosomes by micrococcal nuclease (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, then quenched by addition of 0.1 volume of a solution containing 2% Triton X-100, 2% sodium deoxycholate, and 100 mM EDTA. Soluble chromatin was pre-cleared with Protein A Dynabeads (Thermo Fisher Scientific) pre-adsorbed with normal rabbit IgG (Millipore, 12-370) for 2-hours at 4°C rotating 9 rpm, and then rotated overnight at 4°C with 1 μg of H3K4me3 (Abcam, ab1012), 2 μg H3K4me1 (Abcam, ab8895), 1 μg H3K27ac (Active Motif, 39034), or 2 μg H3K27me3 (Millipore, 07-449) antibodies pre-adsorbed to 10 μL of Protein A Dynabeads (Thermo Fisher Scientific) in 150 μL of ChIP buffer (20 mM Tris pH 8.0, 2 mM EDTA, 150 mM NaCl, 0.2% Triton X-100, 1 mM PMSF, and EDTA-free protease inhibitor cocktail (Roche)). After washing the beads twice with ice-cold wash buffer (20 mM Tris pH 8, 150 mM 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 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 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 μL 20 mg/mL proteinase K after the first 30-minutes and occasional manual inversion. DNA was purified using standard Phenol:chloroform:isoamyl alcohol extraction and ethanol precipitation and used for library preparation with the NEBNext Ultra II DNA Library Prep Kit for Illumina (NEB) with 9 (H3K4me1, H3K27me3) or 13 (H3K4me3, H3K27ac) amplification cycles. Indexed libraries were analyzed for size distribution using the Bioanalyzer High Sensitivity DNA chip (Agilent) and for concentration using the Qubit dsDNA HS Assay (Thermo Fisher Scientific), then pooled

and sequenced on the NextSeq 500 platform (Illumina; NCS v2.2.0.4) for 2 × 78 nucleotide paired end reads, targeting 40 million read pairs per sample. Sequence data is available in the Gene Expression Omnibus (GEO) under accession number GSE181951.

2.7 ChIP-seq data analysis

ChIP-seq reads were quality trimmed using Trimmomatic v0.38 [356] and aligned to concatenated mouse (GRCm38/mm10) and *Drosophila* (BDGP6.28) genome assemblies using Bowtie2 v2.3.4.1 with options '--very-sensitive --no-unal --no-discordant' [357]. Multi-mapped reads, reads with mapping quality MAPQ < 20, and suspected PCR-duplicates were discarded using Samtools [358]. To scale enrichment according to the *Drosophila* cell spike-ins, the fraction of remaining reads that mapped to the *Drosophila* genome was determined, and a scaling factor was calculated for each replicate as *(minimum* Drosophila *read fraction per histone mark)/(sample* Drosophila *read fraction)*. After determining a scaling factor, each sample was downsampled to the lowest mapped read count per each histone mark using Samtools, and then reads mapped to the *Drosophila* genome were removed. For visualization, libraries were converted to bedgraph format using the calculated scaling factor in the '-scale' argument of BEDtools v2.26.0 genomecov [359]. Genome browser views were generated for merged replicates using Spark [360] and TSS profiles using Deeptools v3.4.2 [361]. Peak calling was performed with MACS2 v2.2.6 in each biological replicate with parameters '--broad --llocal 100000' [362]. Peak regions that were detected in only one biological replicate or that overlapped with high-background blacklisted regions defined by the ENCODE project [363] were discarded. Chromosomes X, Y, and M were also excluded from the analyses. I defined TSS locations as the 5' end of the most abundant

annotated transcript of each gene detected in the RNA-seq data. H3K4me3 peaks were assigned to genes for which they overlap the region ±1 kb of the TSS.

2.8 Messenger ribonucleic acid sequencing (mRNA-seq)

eGFP+ tdTomato- β-cells (70,000-226,000 per mouse, biological triplicate) were purified and counted by FACS and supplemented with a 10% spike-in of *Drosophila* S2 cells. The cells were sorted directly into Trizol LS Reagent (Thermo Fisher Scientific) and total RNA was extracted according to the manufacturer's instructions. Residual DNA was digested by Turbo DNase (Thermo Fisher Scientific) for 30-minutes at 37°C. Purified total RNA content was measured using the Qubit RNA HS assay (Thermo Fisher Scientific). mRNA was enriched from 400 ng of total RNA per sample using the NEBNext Poly(A) mRNA Magnetic Isolation Module (NEB) and used to prepare libraries with the NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (NEB) with 11 amplification cycles with single-indexed primers. Multiplexed libraries were quantified, pooled, and sequenced for 2 × 38 nucleotide reads as described above, targeting 33 million read pairs per sample. Sequence data is available in GEO under accession number GSE181951.

2.9 mRNA-seq data analysis

Sequenced reads were aligned to the concatenated mouse (GRCm38/mm10; Gencode vM24) and *Drosophila* (BDGP6.28; Ensembl release 99) genomes and transcriptomes using STAR v2.7.3a [364] with default settings. Transcript abundance was calculated using Salmon v1.4.0 [365] in alignment mode with seqBias and gcBias options. Gene counts and differential expression were calculated using DESeq2 [366] with default parameters and implementing the apeglm method of

effect-size shrinkage [367]. Genes with fewer than two counts in any sample were filtered out and genes showing ≥ 2-fold difference in expression with $P \leq 0.01$ (Wald test with Benjamini and Hochberg adjustment for multiple comparisons) are considered differentially expressed. Genes showing <1.1-fold difference in expression are considered stable. Except in Fig. 3.6c, *Drosophila* transcripts were removed after abundance calculation; for Fig. 3.6c, *Drosophila* transcripts were used as control genes in the estimateSizeFactors function of DESeq2, and then removed.

2.10 Single-cell mRNA-seq

Islets were dissociated to single-cell suspensions as described in section 2.4 without FACS-enrichment. The cell suspensions were processed through the Chromium Single Cell 3' protocol using the Chromium Controller (firmware v4.0) with Reagent Kit v3.1 (10X Genomics) and Dual Index Kit TT Set A (10X Genomics) according to the manufacturer's instructions for a targeted 5,000-cell recovery. Total cDNA was amplified for 11-cycles, and then 13-cycles during index-ligation. cDNA concentrations were quantified by qPCR with the NEBNext Library Quant Kit for Illumina (NEB). Libraries were pooled and sequenced for 28-90 paired-end nucleotide reads to an average depth of ~70,000 mapped fragments per cell. Sequence data is available in the GEO under accession number GSE181951.

2.11 Single-cell mRNA-seq data analysis

Cellranger (v5.0.0, 10X Genomics) was used to generate FASTQ files and demultiplex reads. Cell barcode detection, read mapping, quality filtering, and transcript counting were performed using Alevin (Salmon v1.4.0) [368] against the mouse protein-coding transcriptome (Gencode

vM24), which I appended with the *Tdtomato* and *Egfp* sequences, as well as the complete mouse genome (GRCm38/mm10) as mapping decoys for selective alignment [369]. Additional cell quality filtering was performed using Seurat v4.0 [370] based on unique features (> 1,000 genes) and mitochondrial RNA content (< 25%). Putative doublets were excluded based on a high transcript count (> 90,000) or co-expression of *Tdtomato* and *Egfp* transcripts, and using DoubletFilter [371]. Genes detected in < 3 cells were excluded from analysis. After quality filtering, the dataset comprised 3,703 WT and 3,641 KO cells. Filtered data were scaled and normalized using SCTransform [372] with default parameters. PCA and UMAP dimensionality reduction using the first 60 principal components was used to cluster and visualize cell populations. Slingshot [373] was used to perform pseudotime analysis on β-cell clusters. For cell entropy analysis, I used a published R function (https://github.com/skannan4/cm-entropy-score [108]) which calculates per-cell Shannon entropy using the 1,000 most highly expressed genes.

2.12 Immunohistofluorescence

Mouse pancreata were isolated and fixed in 4% (w/v) paraformaldehyde (PFA) in PBS at 4°C overnight. Following washing with PBS and dehydration in ethanol and xylenes, tissues were embedded in paraffin and sectioned with a thickness of 5 μm. Immunofluorescent analyses of tissue sections was performed as previously described [374] in a blinded fashion from 4 to 6 sections per biological sample. Primary antibodies used are listed in Table 2.2.

2.13 Immunocytofluorescence

Islets were dissociated to single cells 3 to 4 hours after isolation and plated in technical duplicate in a 24-well cell culture dish with glass coverslips at ~50 dispersed islets/well in complete RPMI.

The next day, media was changed to 0.5 mL complete RPMI containing 10 μM EdU (Carbosynth) and either 5.5 mM or 16.7 mM D-glucose. Cumulative EdU incorporation proceeded for 48-hours in these media, with refreshment after the first 24-hours. Cells were then fixed in 2% PFA in PBS for 20-minutes, then washed once in PBS before labelling EdU for 20-minutes with 5 μM Alexa Fluor 647 azide (Thermo Fisher Scientific) in 100 mM Tris pH 8.5, 1 mM $CuSO_4$, 100 mM sodium ascorbate. Coverslips were subsequently washed 3 times with 0.5% Triton X-100 in PBS, counterstained with DAPI (Sigma Aldrich, D9542, 1:5,000) in PBS for one hour at room temperature, washed three more times in PBS, and mounted with Prolong Gold (Thermo Fisher Scientific) mounting medium onto glass slides.

2.14 Transmission electron microscopy

Islets were fixed in 2% glutaraldehyde (Sigma Aldrich) in PBS and submitted to the Facility for Electron Microscopy Research of McMaster University (Hamilton, ON, Canada). There, islets were post-fixed with 1% osmium tetroxide, dehydrated in ethanol, and embedded in Spurr's resin before sectioning with a Leica UCT ultramicrotome. Sections were stained with uranyl acetate and lead citrate and imaged with a JEOL JEM 1200 EX TEMSCAN transmission electron microscope (JEOL, Peabody, MA, USA).

2.15 Immunoblotting

Islets were lysed in Laemmli sample buffer (2% (w/v) SDS, 10% (v/v) glycerol, 60 mM Tris pH 6.8, 1mM NaF, 1 mM PMSF, protease inhibitor cocktail (Roche)) for 5-minutes at 95°C and then vigorously vortexed 10-20 sec. Protein concentration was measured by BCA assay (Thermo Fisher Scientific), then β-mercaptoethanol was added to 5% (v/v) and lysates were reboiled for

5-minutes and stored at -80°C. 10 μg lysates were resolved in polyacrylamide gels, transferred to PVDF membranes, blocked for 60-minutes in 5% (w/v) skim milk powder in TBST (20 mM Tris pH 7.6, 150 mM NaCl, 0.1% (v/v) Tween-20) at room temperature and probed with primary antibodies overnight in 5% (w/v) BSA in TBST at 4°C. The next day, after washing membranes three times with TBST, HRP-linked secondary antibodies were added for 1-hour at room temperature, membranes were washed three times with TBST, and detected using ECL system with film. When normalizing to total H3, membranes were stripped for 20-minutes in mild stripping buffer (200 mM glycine, 0.1% (w/v) SDS, 1% (v/v) Tween-20, pH 2.2) at room temperature, rinsed twice in PBS and twice in TBST for 10 minutes each, then re-blocked and probed for histone H3. For detection of (pro)insulin, islets were treated for 4-hours with 10 μM MG-132 (Sigma Aldrich) before lysing, and probed using a modified immunoblot protocol to improve peptide detection, which is described in detail by Okita, N. *et al* [375]. Primary antibodies used for immunoblots are listed in Table 2.2.

2.16 Protein co-immunoprecipitation

Immunoprecipitation of WDR5 was performed in triplicate in 750-900 dispersed islets pooled from three mice. First, nuclei were enriched from dispersed islet cells by incubation in nuclear isolation buffer for 10-minutes on ice and pelleting at 600 × *g* for 5-minutes at 4°C. Chromatin was solubilized in IP buffer (20 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 PMSF, protease inhibitor cocktail (Roche)) by MNase (NEB) digestion for 10-minutes at room temperature. EDTA was added to 5 mM and samples were centrifuged at 14,000 × *g* for 5-minutes at 4°C. The protein concentration in the supernatant was measured by BCA assay (Thermo Fisher Scientific), and an equal mass of

WT and KO nuclear lysate (~30 μg protein) was immunoprecipitated by 1 μg WDR5 (Bethyl, A302-429A) or normal rabbit IgG (Millipore, 12-370) antibodies rotating overnight at 4°C. The next day, 10 μL protein A Dynabeads (Thermo Fisher Scientific) was added to the antibody:lysate mixture and rotated for 2-hours at 4°C. Finally, bead:antibody:antigen complexes were washed six times in ice-cold IP buffer and resuspended in Laemmli sample buffer with 5% (v/v) β-mercaptoethanol for immunoblotting.

2.17 Insulin secretion assays

For intraperitoneal insulin tolerance tests, 2 g/kg glucose in a 20% (w/v) water solution was injected intraperitoneally following a 6-hour fast. Blood glucose was sampled prior to and 15, 30, 60, and 120-minutes after glucose injection from a lateral saphenous vein. Serum was collected at the 0, 15, and 60-minute time points by centrifuging blood samples at 9,000 × g for 9 minutes at 4°C and stored at -80°C. Serum insulin was measured by ELISA (Alpco) according to the manufacturer's instructions. For oral glucose tolerance tests, 2 g/kg glucose in a 20% (w/v) water solution was administered by oral gavage following a 6-hour fast. Blood glucose was sampled prior to and 5, 15, 30, and 60-minutes after glucose administration from a lateral saphenous vein. *Ex vivo* static insulin secretion assays were performed in technical duplicate in groups of 50 islets per mouse one day after isolation. Islets were preincubated in KRBH buffer (2.8 mM glucose, 20 mM HEPES, 114 mM 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 7.3) for one hour at 37°C and then sequentially for 45-minutes in KRBH, either 16.7 mM glucose or 15 mM dimethyl α-ketoglutarate (Sigma Aldrich), and finally 30 mM KCl. Conditioned medium was collected after each incubation. At the completion of the assay islets were lysed in RIPA buffer (Pierce)

containing 1 mM PMSF and EDTA-free protease inhibitor cocktail (Roche)) at 95°C for 5 minutes. Insulin concentration was measured in conditioned medium and islet lysate by ELISA (Alpco). DNA concentration of islet lysate was determined by Qubit dsDNA HS Assay (Thermo Fisher Scientific) to normalize islet insulin content.

2.18 Insulin tolerance test

Insulin (Novolin ge Toronto; Novo Nordisk Canada Inc., Mississauga, ON, CAN) was diluted to 0.1 U/mL in PBS and injected with 0.5 U/kg body mass intraperitoneally after a 2-hour fast. Blood glucose was sampled by tail prick prior to and 15, 30, 60, and 120-minutes after insulin injection.

2.19 Calcium imaging

Intact islets were cultured for two days to allow for attachment onto glass coverslips. Prior to imaging, coverslips were transferred to a 2 mL imaging chamber and stained with 5 μM Fura-2AM (Molecular Probes, F1221) for 30-minutes at 37°C in complete RPMI. The imaging chamber loaded with a coverslip was then mounted on a Leica SP8 Laser Scanning Confocal Microscope with a 10× objective lens and continuously washed with Ringer's solution (5.5 mM KCl, 2 mM $CaCl_2$, 1 mM $MgCl_2$, 20 mM HEPES, 144 mM NaCl, 2.8 mM glucose, pH 7.4) for 30-minutes. Throughout the experiment, islets were continuously perifused at a flow rate of 2.5 mL/min with Ringer's solution. Glucose and KCl concentrations were adjusted by iso-osmotic substitution of NaCl where indicated. Fura-2 was excited at 340 nm and 380 nm and emitted fluorescence was detected at approximately 510 nm. Cytosolic Ca^{2+} levels are expressed as the

mean ratio of Fura-2 fluorescence emission intensity (F_{340}/F_{380}) of 3-9 islets per biological replicate.

2.20 Respirometry

Islets were dissociated to single cells 3-4 hours after isolation and plated at 100 islets/well in a Seahorse 24-well cell culture dish in technical duplicate. Cells were cultured for 2-days in complete RPMI to allow for attachment to the culture dish and then for 1-hour in XF RPMI medium (Agilent) with 2.8 mM glucose, 1% FBS, 1 mM pyruvate, and 2 mM glutamine to reach a metabolic baseline. Oxygen consumption was then measured using a Seahorse XF24 Extracellular Flux Analyser (Agilent). The injection ports were loaded with the following compounds in XF Assay Media (Agilent): A, 167 mM glucose; B, 10 μM oligomycin; C, 15 μM FCCP; D, 10 μM antimycin A and 1 μM rotenone. Measurements were normalized to the DNA content of assayed cells, which were determined using Qubit HS DNA Assay (Thermo Fisher Scientific) after lysing in RIPA buffer upon completion of the assay.

2.21 Metabolite quantification

One hundred islets were collected from each mouse, rinsed twice in ice-cold PBS, and stored at -80°C. Blood was collected from a lateral saphenous vein after 2-hour fast into 1.7 mL Eppendorf tubes on ice that were previously rinsed with a saline/EDTA solution. Plasma and cell fractions were separated by centrifugation at 9,000×g for 9 min at 4°C. Plasma was pipetted into a clean tube and both tubes were stored at -80°C. Islets and blood cell pellets were packaged in dry ice and sent to Dr. Joshua Miller's laboratory at Rutgers University, New Brunswick, NJ, USA for quantification of SAM, SAH, methionine, cystathionine, betaine, and choline by high-performance

liquid chromatography with ultraviolet detection. Total plasma folates concentration was measured by Janette King at the Analytical Core for Metabolomics and Nutrition at BCCHRI using the ARCHITECT i1000SR immunoassay analyzer, Abbott Laboratories (Abbott Park, IL, 60004 USA). Plasma methionine, homocysteine, and cysteine concentrations were measured by Roger Dyer at the Analytical Core for Metabolomics and Nutrition at BCCHRI using ultra-high performance liquid chromatography tandem mass spectrometry.

2.22 Analysis of publicly available sequence data

Publicly available data were reprocessed for consistency. Gene expression omnibus accession numbers GSE50244, GSE50386, GSE107489, GSE124742, and GSE174194 were downloaded from the Sequence Read Archive (SRA) using the sratoolkit (https://github.com/ncbi/sra-tools) and analyzed according to the experimental application, as described above.

2.23 Statistics

Unless otherwise stated, plots show mean ± standard deviation (SD) with points representing each biological replicate. In boxplots, the central horizontal line marks median, the upper and lower limits of the box the first and third quartiles, and whiskers span 1.5× the interquartile range, in the style of Tukey. Spearman's rank coefficient was used to estimate correlation. Regression lines were calculated using a generalized additive model. *P*-values were calculated using Student's t-tests with Welch's correction, Wilcoxon rank sum test, Wald test, mixed-effect model, one- or two-way ANOVA, Fisher's exact test, or permutation test, as indicated. Correction for multiple comparison using the Benjamini-Hochberg method was applied where indicated. Statistical calculations were performed in R or Prism 8 or 9 (GraphPad Software, La

Jolla, CA, USA). $P < 0.05$ was deemed significant except in RNA-seq data, for which the cutoff $P < 0.01$ and fold-change $\geq \pm 2$ was used.

Table 2.2 Primary antibodies

Target	Host	Supplier	Cat#	Application	Dilution
ACTIN	Mouse	DSHB	JLA20-c	WB	1:10,000
ASH2L	Rabbit	Bethyl	A300-489A	WB	1:10,000
DPY30	Rabbit	Atlas	HPA043761	IHC	1:500
				WB	1:1,000
Glucagon	Mouse	Millipore	G2654	IHC	1:1,000
H3	Rabbit	Abcam	ab1791	WB	1:1,000,000
H3K4me1	Rabbit	Abcam	ab8895	ChIP	2 µg
				IHC	1:1,000
				WB	1:100,000
H3K4me3	Rabbit	CST	C42D8	IHC	1:1,000
				WB	1:1,000
	Mouse	Abcam	ab1012	ChIP	1 µg
H3K27ac	Rabbit	Active Motif	39034	ChIP	1 µg
		CST	D5E4	WB	1:1,000
H3K27me3	Rabbit	Millipore	07-449	ChIP	2 µg
				WB	1:100,000
Insulin	Guinea Pig	Agilent	IR-002	IHC	1:4
	Mouse	CST	L6B10	WB	1:10,000
Normal IgG	Rabbit	Millipore	12-370	CoIP	1 µg
RBBP5	Rabbit	Bethyl	A300-109A	WB	1:10,000
WDR5	Rabbit	Bethyl	A302-430A	WB	1:2,000
			A302-429A	CoIP	1 µg

Chapter 3: H3K4me3 regulates gene expression in β-cells

3.1 Rationale

Tissue-specific transcriptional programs rely on precise activation and maintenance of specific genes and stable repression of other genes. Active and repressed genes have stereotypical patterns of chromatin modification which are useful in the systematic classification of gene activity [376]. For example, H3K4me3 is reliably enriched on nucleosomes at the 5' end of actively transcribed genes and is therefore the canonical marker of active promoters across all eukaryotes [185,377–379]. While H3K4me3 is a well-accepted marker of active promoters, experimental evidence for the functional necessity of H3K4me3 in the regulation of transcription is rare and controversial (reviewed in ref. [380]). On one hand, H3K4 methylation may regulate gene activation via recruitment of transcription factors and cofactors [255,381–383], facilitating enhancer-promoter interactions [199], and/or by preventing repression [267] and irreversible silencing [270]. On the other hand, global (i.e., genome-wide) decreases in H3K4me1 or H3K4me3 levels are remarkably well tolerated in embryonic stem cells and during development in genetic knockout models [226,244,250], perhaps due to redundancy with other regulatory inputs and/or environmental-context-dependency [250,380]. Evidence in multicellular animal models points to a requirement for H3K4 methyltransferases during certain specialized developmental events, whereas evidence for general transcriptional functions in mature tissues is lacking [215,225,349–351]. Therefore, to what extent H3K4me3 functionally contributes to transcription is unclear, and evidence is lacking in post-development tissues where lineage-specific transcriptional programs have already been established.

H3K4 methylation is catalyzed by conserved histone lysine methyltransferases, primarily the COMPASS and COMPASS-related complexes in mammals. These consist of one MLL/SET catalytic subunit complexed with the four core structural subunits WDR5, RBBP5, ASH2L, and DPY30, as well as a retinue of auxiliary proteins that contribute to enzymatic activity and targeting [213]. In contrast to the other core subunits, DPY30 is not required for the assembly of the complex and is not required for methyltransferase activity *in vitro*, although inclusion of DPY30 boosts activity between 1.5- and 3-fold [241,242]. *In vivo* loss of DPY30 leads to reduction of H3K4me3 and H3K4me1 [222]. In this chapter, I use an inducible *Dpy30*-KO model to test whether H3K4 methylation regulates gene expression in mature mouse β-cells. I hypothesized that loss of DPY30-dependent H3K4me3 from mature β-cells would lead to dysregulation of the β-cell gene expression program by impairing expression of genes in particular epigenetic contexts and increasing transcriptional variability of genes marked by broad H3K4me3 peaks.

3.2 Conditional deletion of *Dpy30* reveals a slow turnover of H3K4me3 and H3K4me1 in mature mouse β-cells

To induce synchronized deletion of *Dpy30* in mature β-cells, mice harbouring a Pdx1-*CreER* transgene [353] and *LoxP* sites flanking the fourth exon of *Dpy30* (ref. [351], originally Knockout Mouse Programme EM:09575) were administered tamoxifen at 8-weeks of age (henceforth: "*Dpy30*-KO") (Fig 3.1a). Pdx1-$CreER^{+/0}$; $Dpy30^{wt/wt}$ and $Pdx1\text{-}CreER^{0/0}$; $Dpy30^{fl/fl}$ (*Dpy30*-WT), also administered tamoxifen, were used as controls. Removal of *Dpy30* exon 4 causes a frameshift mutation in all annotated *Dpy30* transcripts. As expected, knockout of *Dpy30* led to depletion of H3K4me3 and H3K4me1 in islets (Fig. 3.1b) but did not impact the assembly of the other core components of the COMPASS complex WDR5, RBBP5, and ASH2L, or their association with chromatin compared to WT islets (Fig. 3.1c).

The rate of loss of H3K4 methylation in KO β-cells was measured using immunofluorescence at 15-day intervals after tamoxifen administration. As seen in Fig. 3.2, H3K4me3 is stable for at least 15 days and then gradually lost during the subsequent 45 days. H3K4me1 shows a similar pattern but its loss is even more gradual. On the other hand, DPY30 is undetectable by 15 days after tamoxifen administration in *Dpy30*-KO β-cells (Fig. 3.2a). These data indicate that *Dpy30* is necessary for maintenance of H3K4me3 and H3K4me1 levels in mature mouse β-cells. The slow turnover of H3K4me3 and H3K4me1 afforded an opportunity to parse the relative contribution of H3K4 methylation, versus DPY30 itself, in controlling gene expression.

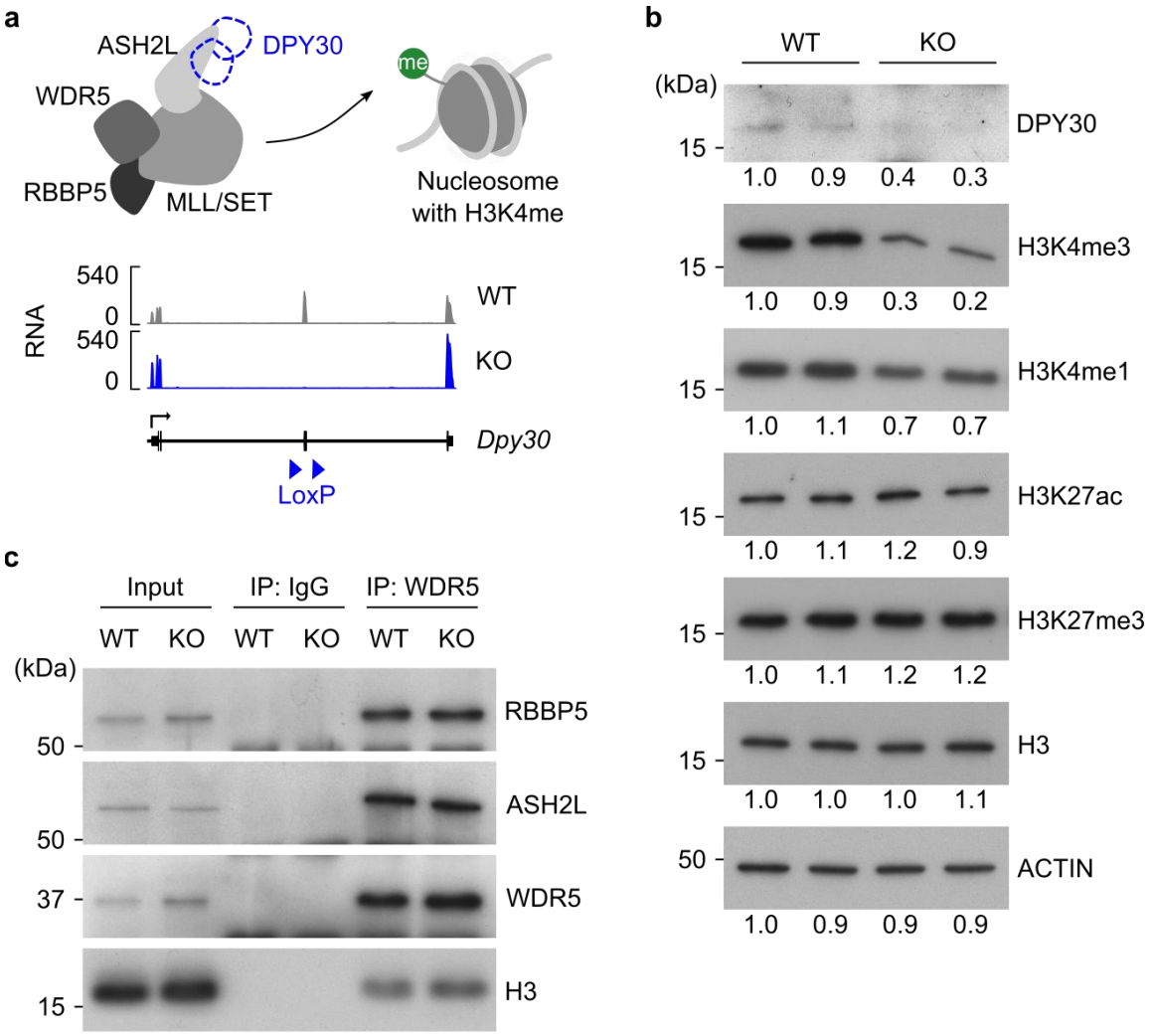

Figure 3.1 | Knockout of *Dpy30* in β-cells leads to reduction of H3K4me3 and H3K4me1. a, Top: schematic showing the core subunits of COMPASS complexes and a nucleosome methylated on H3K4. Bottom: genome-aligned RNA-seq reads at the *Dpy30* gene locus in *Dpy30*-WT and -KO cells. **b**, Immunoblots showing DPY30, H3K4me3, H3K4me1, H3K27ac, H3K27me3, total histone H3, and ACTIN 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. **c**, Immunoblots showing TrxG 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.

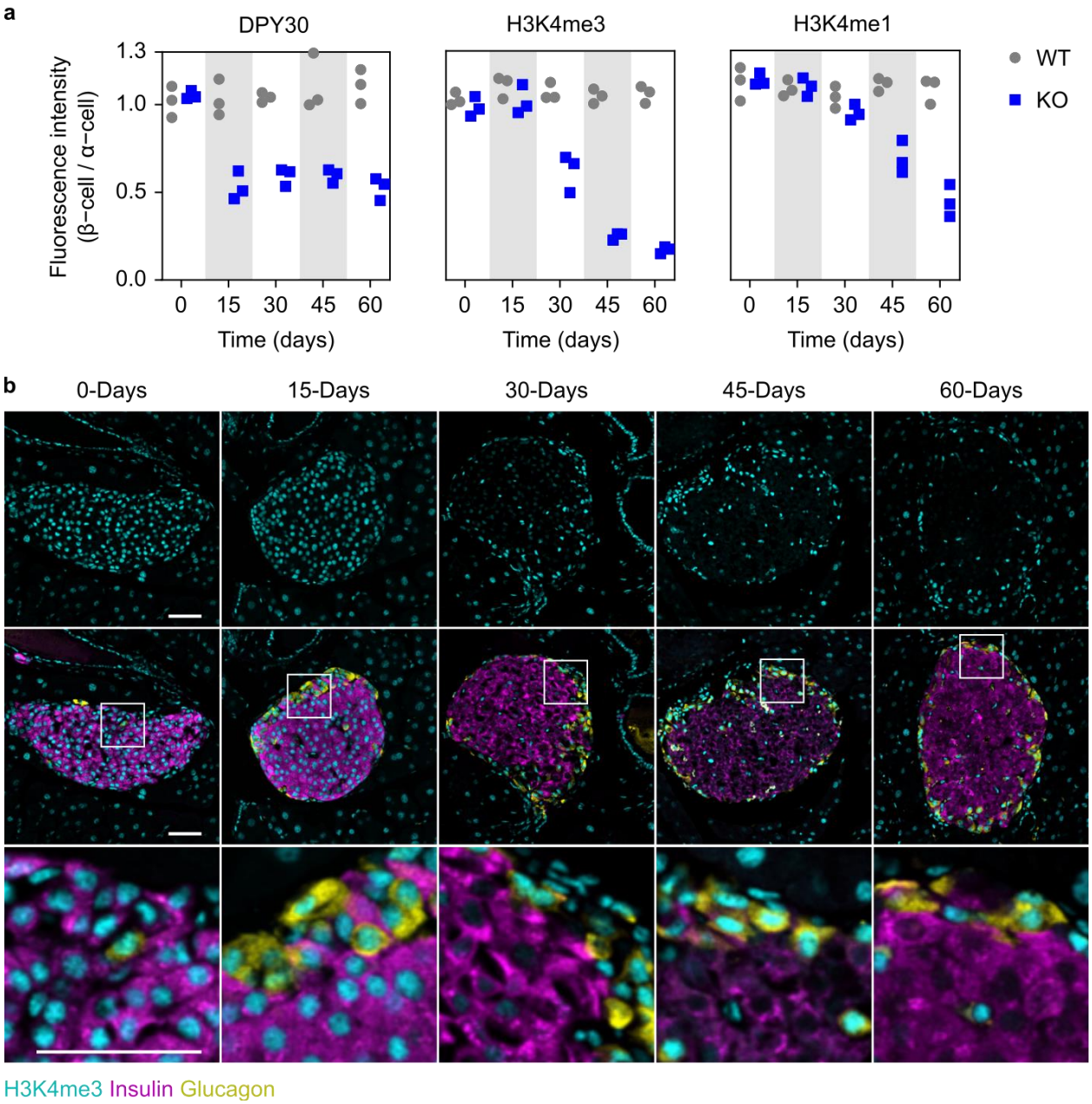

Figure 3.2 | Knockout of *Dpy30* reveals a slow turnover of H3K4me3 and H3K4me1 in mature β-cells. a, Mean immunofluorescent intensity of DPY30, H3K4me3, and H3K4me1 in *Dpy30*-WT and -KO β-cell nuclei at the indicated days after tamoxifen administration. Data are normalized to intensity of α-cell nuclei; n = 3 mice of each genotype and time-point. **b**, Example images of *Dpy30*-KO islets used for measurements in panel (a) showing H3K4me3 (cyan), insulin (magenta), and glucagon (yellow). The boxed areas are shown zoomed-in below. Scale bars: 50 μm.

3.3 Transcriptome remodeling in *Dpy30*-KO cells is secondary to reduction of H3K4 methylation

I reasoned that defects arising in *Dpy30*-KO β-cells by 15-days post-tamoxifen (when DPY30 is lost but H3K4me3 and H3K4me1 levels are unchanged) highlight functional contributions of DPY30 itself. Defects arising later, by 45-days post-tamoxifen (when H3K4 methylation is reduced) highlight functional contributions of H3K4 methylation. As H3K4 methylation and TrxG complexes are each positively associated with transcriptional activity [208,384], I assessed transcriptomes from *Dpy30*-WT and -KO β-cells mice in order to determine to what extent H3K4 methylation and DPY30 contribute to gene expression regulation. To this end, the *Rosa26*^{mTmG} reporter transgene [352] was crossed onto *Pdx1-CreER*$^{+/0}$ *Dpy30*$^{+/+}$ (WT) and *Dpy30*-KO mice to identify Cre-recombined cells. GFP+ β-cells were enriched using FACS from *Dpy30*-KO and -WT islets 15- and 45-days post-tamoxifen then processed for bulk mRNA-seq. Loss of DPY30 had little effect on gene expression in β-cells, with only three differentially expressed genes (*Dpy30*, *Edn3*, *C3*) (≥2-fold change, $P \leq 0.01$ Wald test with Benjamini-Hochberg correction) identified at 15-days post-tamoxifen (Fig. 3.3a and Appendix B). At 45-days post-tamoxifen, 828 genes were differentially expressed, with the majority (634) showing lower expression in *Dpy30*-KO cells (Fig. 3.3b and Appendix B). Together, these data support a functional role for H3K4 methylation in the regulation of gene expression.

I checked COMPASS and non-COMPASS H3K4 methyltransferase gene expression in the day-45 mRNA-seq data for evidence of compensatory upregulation. None is evident (Fig. 3.3c). In fact, expression of methyltransferases tended to be lower in *Dpy30*-KO cells; however, none reached statistical significance.

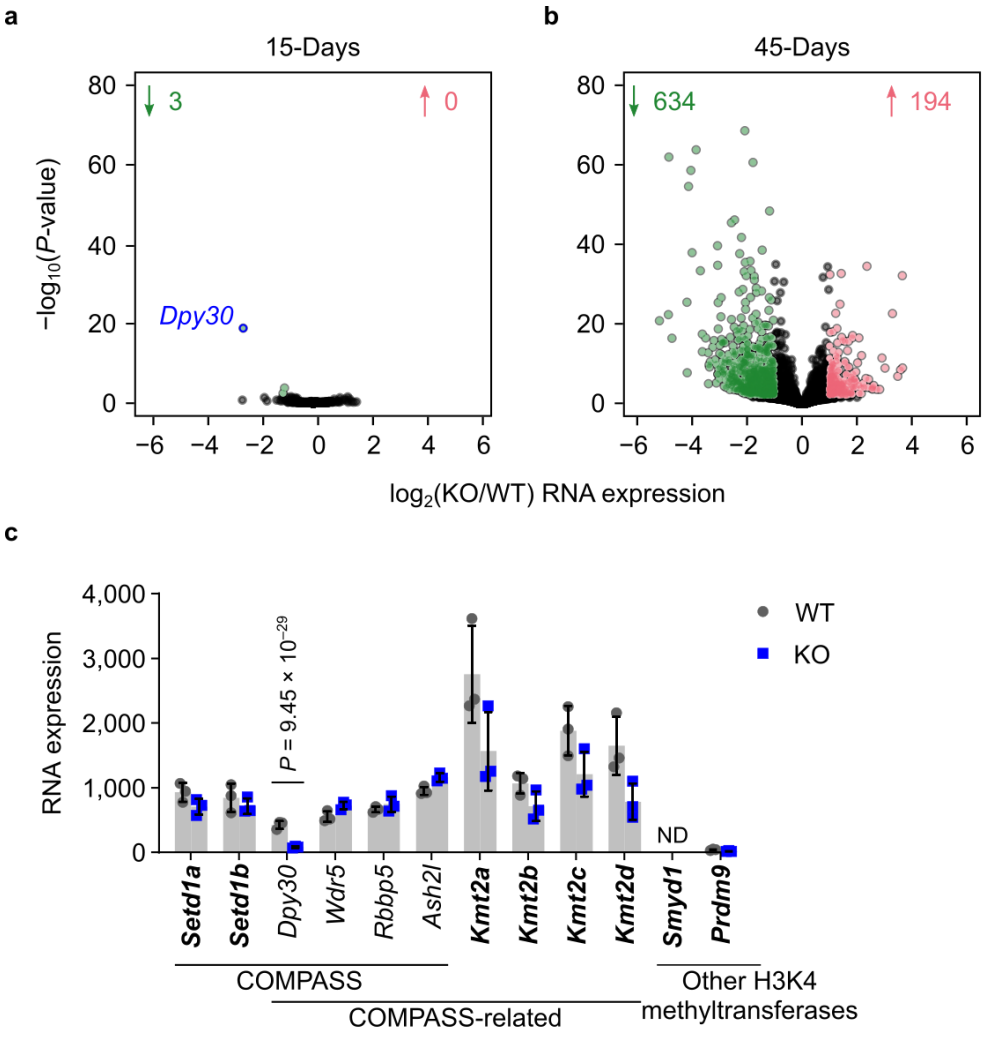

Figure 3.3 | Transcriptome remodeling downstream of loss of H3K4 methylation in *Dpy30*-KO β-cells. a-b, Scatterplots showing $\log_2$(fold-change) and $-\log_{10}$(*P*-value) of gene expression in *Dpy30*-KO versus -WT cells 15-days (a) and 45-days (b) after tamoxifen administration. Genes showing a ≥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. **c**, Expression of core COMPASS, COMPASS-related, and non-COMPASS genes in *Dpy30*-WT and -KO cells 45-days after tamoxifen administration. H3K4 methyltransferase enzymes are in bold.

3.4 Depletion of promoter-associated H3K4me3 in *Dpy30*-KO chromatin revealed using ChIP-seq with spike-in calibration

To more directly link gene expression regulation to H3K4me3 in mature β-cells, I assessed the distribution of H3K4me3 in *Dpy30*-WT and -KO chromatin using ultra-low-input native chromatin immunoprecipitation sequencing (ChIP-seq) [355] 45-days post-tamoxifen. Initially, I processed whole islets for ChIP-seq. This analysis suggested a slight, general reduction in H3K4me3 enrichment at TSSs (Fig. 3.4a, left). Examination of some notable genes revealed, for example, loss of H3K4me3 at the β-cell-enriched *Ins1* gene, but elevated H3K4me3 at the α-cell-enriched *Gcg* gene and minimal change at *Polr2d*, a housekeeping gene (Fig. 3.4b, left).

I reasoned that H3K4me3 from non-Cre-recombined cells was being preferentially immunoprecipitated in *Dpy30*-KO animal islets since there was less H3K4me3 in KO cells. To mitigate this, I FACS-purified recombined GFP+ islet cells using mTmG. It was also necessary to account for the global decrease in H3K4me3 signal in *Dpy30*-KO cells since sequencing techniques including ChIP-seq make an implicit assumption that the total signal per cell is identical between samples [385]. I therefore spiked-in *Drosophila* S2 cells during FACS isolation as an internal reference between samples [386]. ChIP-seq with these additional strategies revealed dramatic depletion of H3K4me3 in *Dpy30*-KO chromatin (Fig. 3.4a, right) and effectively removed signal from non-recombined cells (Fig. 3.4b, right). I therefore performed ChIP-seq under the same conditions for H3K4me1, H3K27ac and H3K27me3. Together, these four histone modifications distinguish active promoters, enhancers, and developmentally repressed chromatin [387].

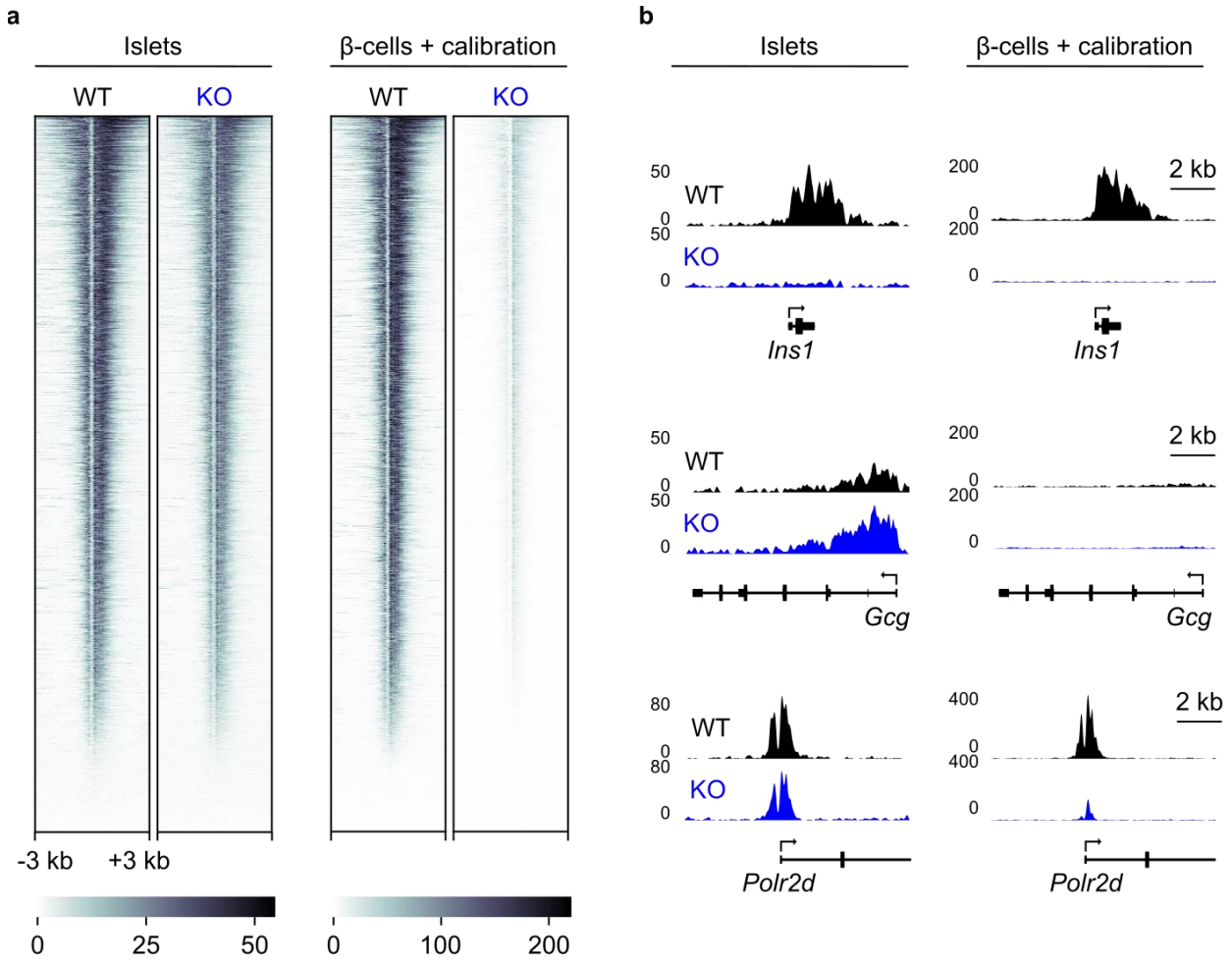

Figure 3.4 | FACS-enrichment and spike-in calibration of H3K4me3 ChIP-seq reveals differences between *Dpy30*-WT and -KO chromatin. a, Heatmaps of H3K4me3 enrichment at the TSS ±3 kb of all expressed genes in whole islets (left) or FACS-enriched β-cells with spike-in calibration (right) from *Dpy30*-WT and -KO mice 45-days after tamoxifen administration. **b**, Genome browser view of H3K4me3 enrichment of a β-cell-enriched-gene *Ins1*, α-cell-enriched gene *Gcg*, and housekeeping gene *Polr2d* for the samples described in panel (a).

3.5 The change in TSS-proximal H3K4me1 enrichment in *Dpy30*-KO chromatin is proportional to initial H3K4me3 enrichment

Like H3K4me3, H3K4me1 is strongly depleted in *Dpy30*-KO cells (Fig. 3.5a-b). Interestingly, residual H3K4me1 at promoters is enriched more closely to the TSS in *Dpy30*-KO chromatin (Fig. 3.5a-b). This led to some genomic loci actually gaining H3K4me1 in *Dpy30*-KO cells; however, this appeared to be restricted to sites previously covered with H3K4me3 (Fig. 3.5a). To quantify this relationship, H3K4me1 and H3K4me3 enrichment in the TSS ±1 kb of all active promoters was compared. In these central promoter regions, the degree to which H3K4me1 is changed in *Dpy30*-KO cells is well correlated to the initial enrichment of H3K4me3 (Spearman's correlation R = 0.8319, Fig. 3.5c). Given that H3K4me1 levels were still on a downward trajectory when ChIP-seq was done (day-45, see Fig.3.2a), this suggests that H3K4me3-positive loci proceed through a monomethylated intermediate in the inducible *Dpy30*-KO model before becoming unmethylated, and raises the question of whether H3K4me1 can support transcription at the TSS.

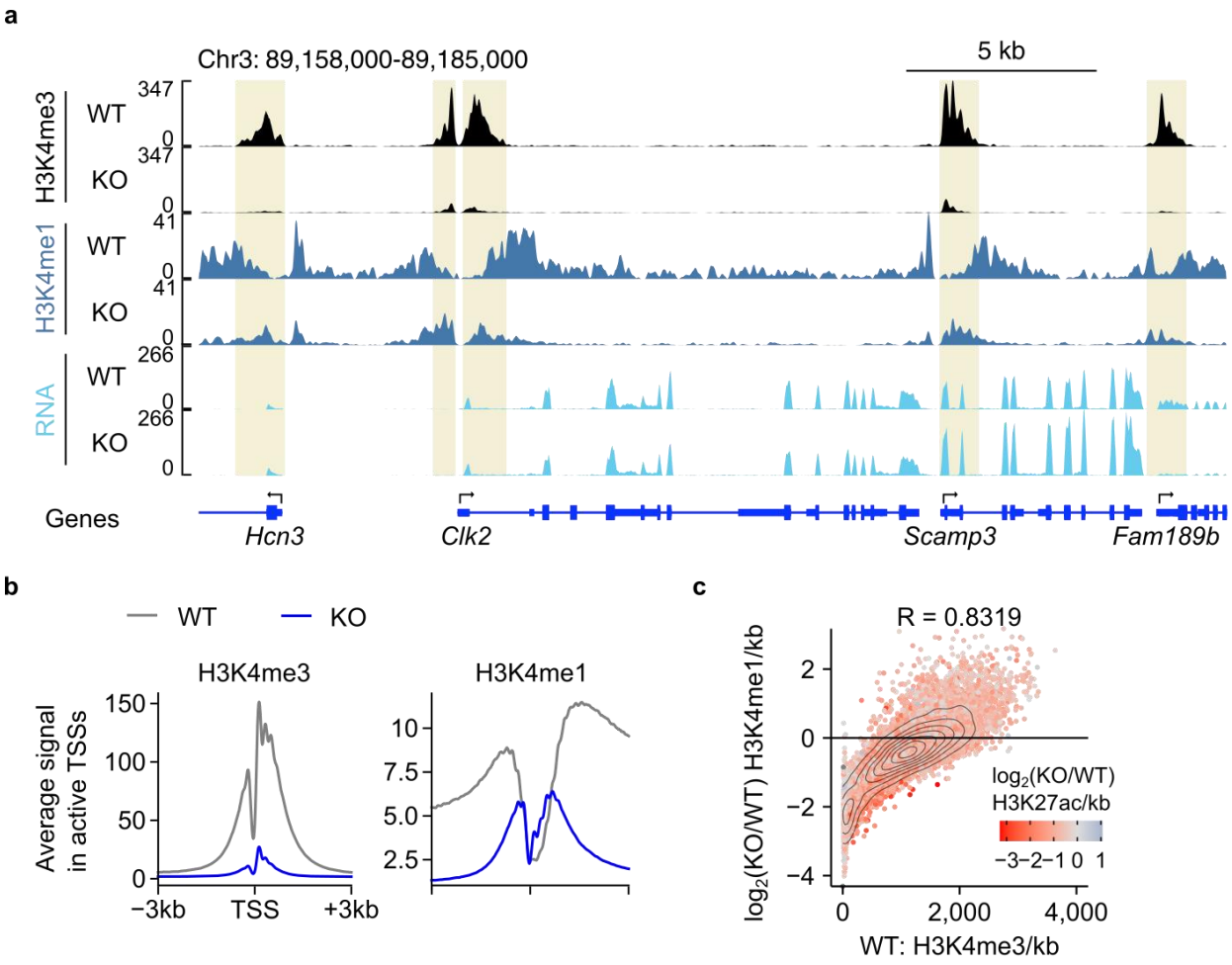

Figure 3.5 | The change in TSS-proximal H3K4me1 enrichment in *Dpy30*-KO chromatin is proportional to initial H3K4me3 enrichment. a, Genome browser representation of H3K4me3, H3K4me1, and genome-aligned mRNA in WT and KO chromatin. Regions enriched for H3K4me3 in WT chromatin are highlighted in yellow. **b**, Average enrichment profiles of H3K4me3 and H3K4me1 in the region ±3 kb with respect to the TSS of all expressed genes in WT and KO chromatin. **c**, Contour scatterplot showing H3K4me3 enrichment in WT chromatin (x-axis) plotted against the $\log_2$(fold-change) of H3K4me1 enrichment in KO versus WT (y-axis) in the regions ±1 kb with respect the TSS of all expressed genes. R indicates the Spearman's rank correlation coefficient. The $\log_2$(fold-change) of H3K27ac enrichment is shown as a colour gradient.

3.6 Genome-wide reduction of H3K4me3 and H3K4me1 does not cause a general reduction in mature mRNA expression

In the mRNA-seq dataset generated from β-cells 45 days after tamoxifen administration, the 828 significantly dysregulated genes represent only 5.6% of 14,677 total detected genes. This could indicate that DPY30 is necessary to maintain H3K4me3 at the genes that became downregulated but not at other genomic loci. To determine whether reduction of H3K4me3 in *Dpy30*-KO chromatin is global or locus-specific, I compared H3K4me3 enrichment between *Dpy30*-KO and -WT chromatin in 10-kb bins spanning the genome. H3K4me3 enrichment is reduced genome-wide in *Dpy30*-KO chromatin (Fig. 3.6a). This is also true for H3K4me1, with the exception of nine 10-kb bins (Fig. 3.6b) (examination of those bins indicate that they span large promoters that were highly H3K4me3-enriched in WT cells). Therefore, DPY30 is necessary for maintenance of H3K4 methylation throughout the genome in mouse β-cells. It follows that even TSSs associated with genes that are stably expressed (<1.1-fold change in RNA expression) or upregulated (≥2-fold, $P \leq 0.01$ Wald test with Benjamini-Hochberg correction) show significant loss of H3K4me3 and H3K4me1 in *Dpy30*-KO cells (Fig. 3.6d-e). Importantly, like the ChIP-seq data, the mRNA-seq data in *Dpy30*-KO and -WT cells were calibrated using *Drosophila* cell spike-in and so cellular mRNA content can be semi-quantitatively compared [385,388]. The transcriptomes of *Dpy30*-KO and -WT β-cells are well correlated, and the total mRNA expression of KO cells is not decreased (Fig. 3.6c, note that the regression line overlaps with the diagonal). Therefore, global depletion of H3K4 methylation does not cause a general decrease in mature mRNA expression.

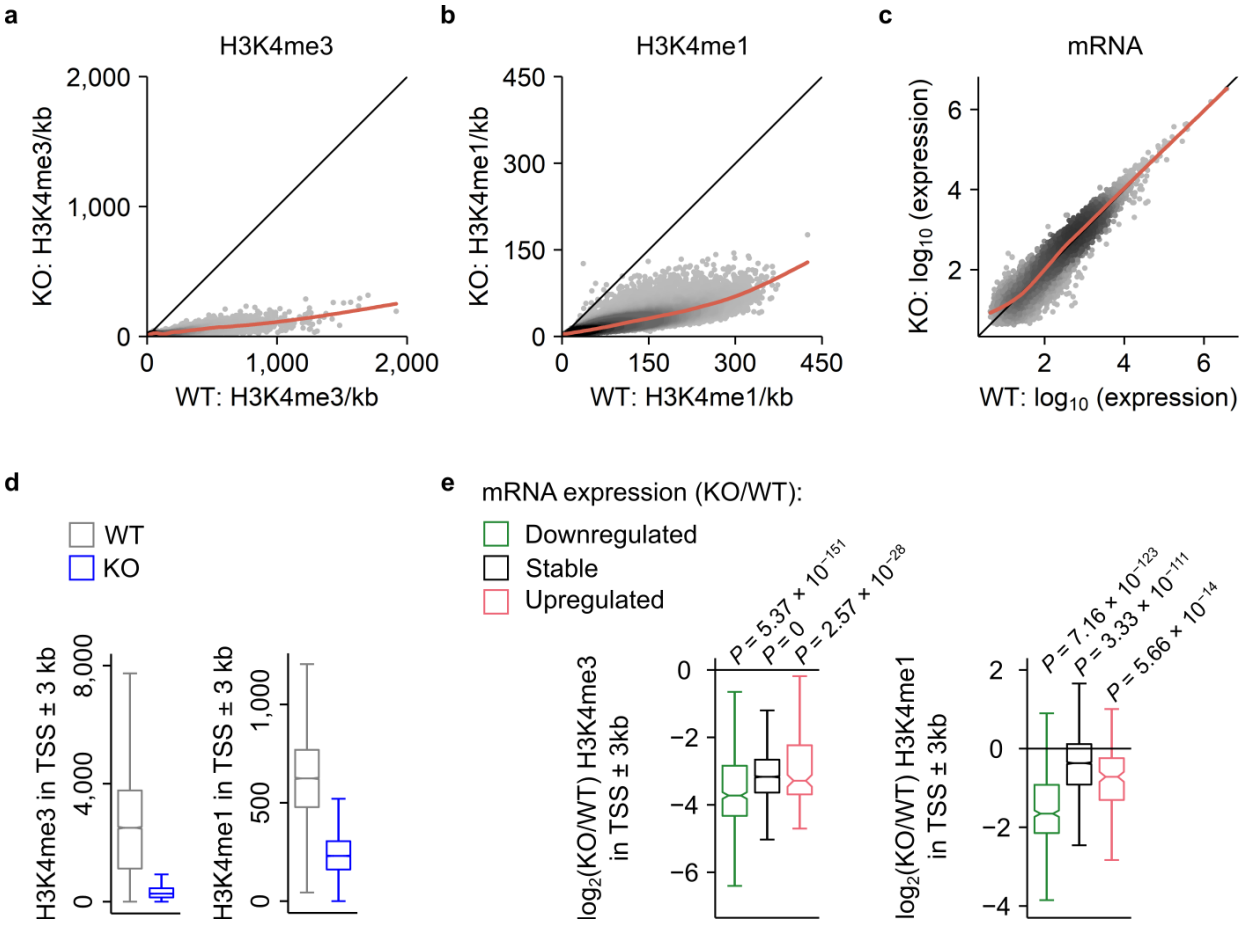

Figure 3.6 | Genome-wide reduction of H3K4 methylation does not cause a general reduction in mRNA expression. a-b, Scatterplot of H3K4me3 (a) and H3K4me1 (b) enrichment in 10-kb bins spanning the genome. A red line shows the smoothed trend generated with a generalized additive model. **c**, Scatterplot showing mean gene expression in *Dpy30*-KO versus -WT cells. A red line shows the smoothed trend generated by a generalized additive model. **d**, Box and whisker plots of H3K4me3 and H3K4me1 enrichment in the TSS ±3 kb of all expressed genes in *Dpy30*-WT and -KO cells. **e**, Box and whisker plots of the $\log_2$(fold-change) of H3K4me3 or H3K4me1 enrichment in the TSS ±3 kb of genes whose mRNA expression is downregulated, stable, or upregulated in *Dpy30*-KO compared to -WT cells. *P* indicates the Wilcoxon rank sum test with Benjamini-Hochberg correction versus no difference between WT and KO (i.e. versus y = 0).

3.7 Genes experiencing a relatively high loss of H3K4me3 and H3K4me1 in *Dpy30*-KO cells are downregulated

Next, the influence of residual H3K4me3 and H3K4me1 on expression at individual promoters was examined. Identification of areas of significant enrichment over background using MACS2 [362] indicated that H3K4 methylation is not completely lost at the majority of promoters in *Dpy30*-KO cells. I segregated genes into three groups according to residual H3K4 methylation. Group i retains at least some H3K4me3 enrichment at the TSS (10,045 genes), group ii has lost H3K4me3 but retains at least some H3K4me1 enrichment (2,302 genes), and group iii has lost H3K4me3 and H3K4me1 (111 genes) (Fig. 3.7a). As a group, genes that retain at least some enrichment for H3K4me3 at the TSS are not generally downregulated in *Dpy30*-KO cells (Figure 3.7b). Loss of both H3K4me3 and H3K4me1 from the TSS is linked to significant impairment of gene expression, and genes that lose H3K4me3 but retain H3K4me1 show an intermediate phenotype (Fig. 3.7b). To visualize this relationship in a threshold-free manner, I plotted the change in H3K4me3, H3K4me1, and mRNA expression for all active TSSs (Fig. 3.7c). As expected, most genes clustered in the bottom left quadrant, indicating reduction of both H3K4me3 and H3K4me1 from the TSS. Strikingly, genes that are highly downregulated (red) clustered at the extreme edge of this quadrant, indicating that genes experiencing the greatest relative loss of both H3K4me3 and H3K4me1 tend to be highly downregulated (Fig. 3.7c). 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. These data also suggest that H3K4me1, widely considered a mark of active enhancers but not promoters [389], can partially compensate for loss of H3K4me3 at promoters.

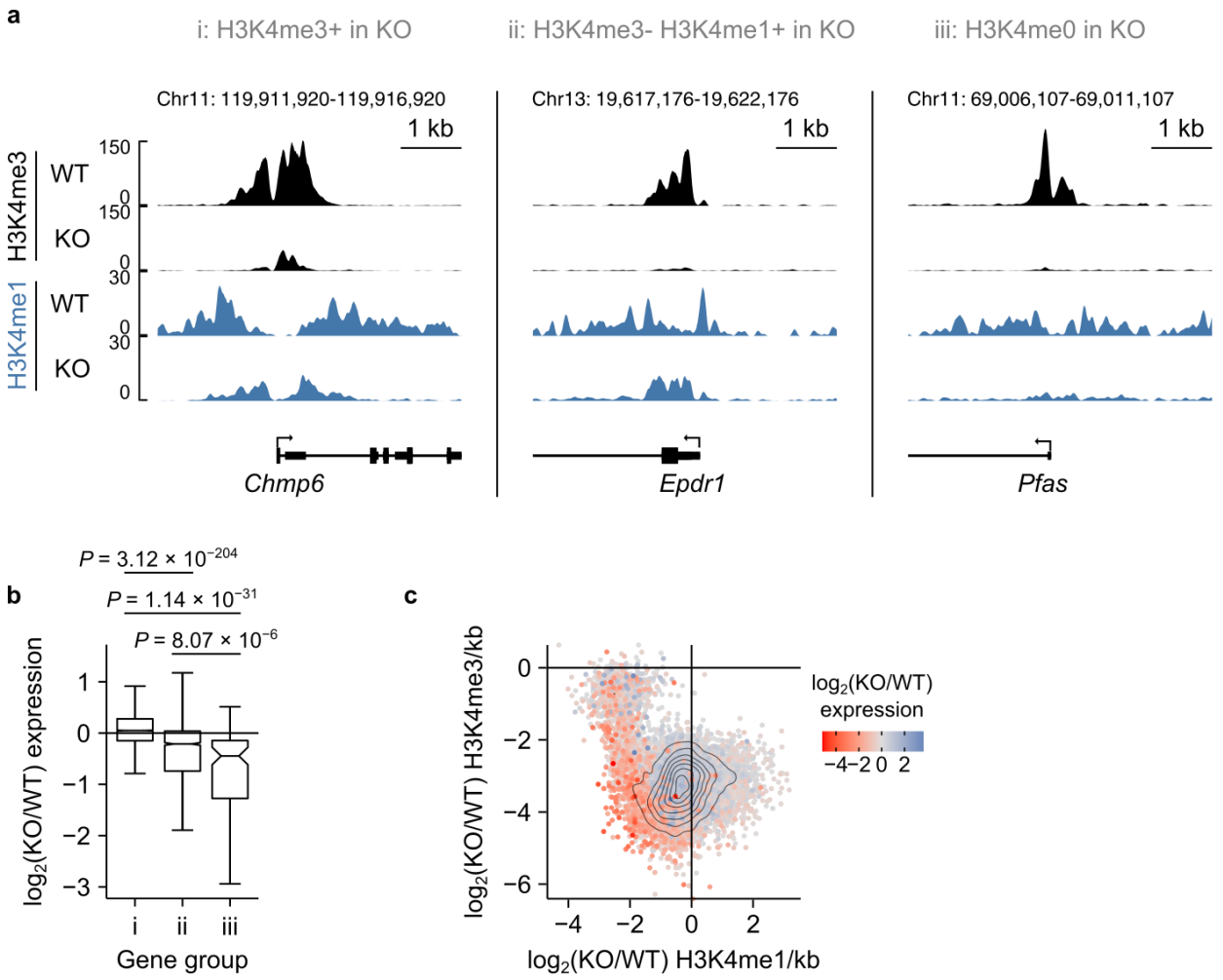

Figure 3.7 | Loss of H3K4 methylation from the TSS reduces gene expression. a, Genome browser view showing examples of genes that retain H3K4me3 (i), lose H3K4me3 but retain H3K4me1 (ii), or lose H3K4me3 and H3K4me1 (iii) from the TSS in KO chromatin. **b**, Box and whisker plot of the $\log_2$(fold-change) of RNA expression of gene groups described in panel (a). *P* indicates the Wilcoxon signed-rank test result with Benjamini-Hochberg correction. **c**, Contour scatterplot showing the $\log_2$(fold-change) of H3K4me1 enrichment in KO versus WT (x-axis) plotted against the $\log_2$(fold-change) of H3K4me3 enrichment in KO versus WT (y-axis) in the TSS ±1 kb of all expressed genes. The $\log_2$(fold-change) for RNA expression of the associated genes is shown as a colour gradient.

3.8 Genes in weakly active chromatin are susceptible to transcriptional downregulation in *Dpy30*-KO cells

To gain insight into what chromatin features cause some genes to lose H3K4 methylation and become downregulated in *Dpy30*-KO β-cells, I compared the chromatin profile of downregulated and stably expressed genes. The downregulated gene set has a relatively inactive profile even in *Dpy30*-WT cells: enrichment for H3K4me3 and H3K27ac are low and H3K4me1 and H3K27me3 are high in comparison to stably expressed genes near the TSS (Fig. 3.8a-b). Accordingly, the downregulated gene set shows lower absolute expression than stably expressed genes in *Dpy30*-WT cells (Fig. 3.8c). Furthermore, the DNA sequence of downregulated gene loci tends to be relatively G/C-rich (Fig. 3.8d-e). Imprinted genes, comprising genomic loci for which one allele is silenced by DNA methylation, tend to be downregulated in *Dpy30*-KO cells (Fig. 3.8f), raising the possibility that H3K4me3 impedes *de novo* DNA methylation at some genes in mature β-cells; however, I did not directly measure DNA methylation. To summarize, genes that are downregulated in *Dpy30*-KO cells tend to already be in a relatively inactive or repressed state in *Dpy30*-WT cells and may be susceptible to silencing in the absence of H3K4me3.

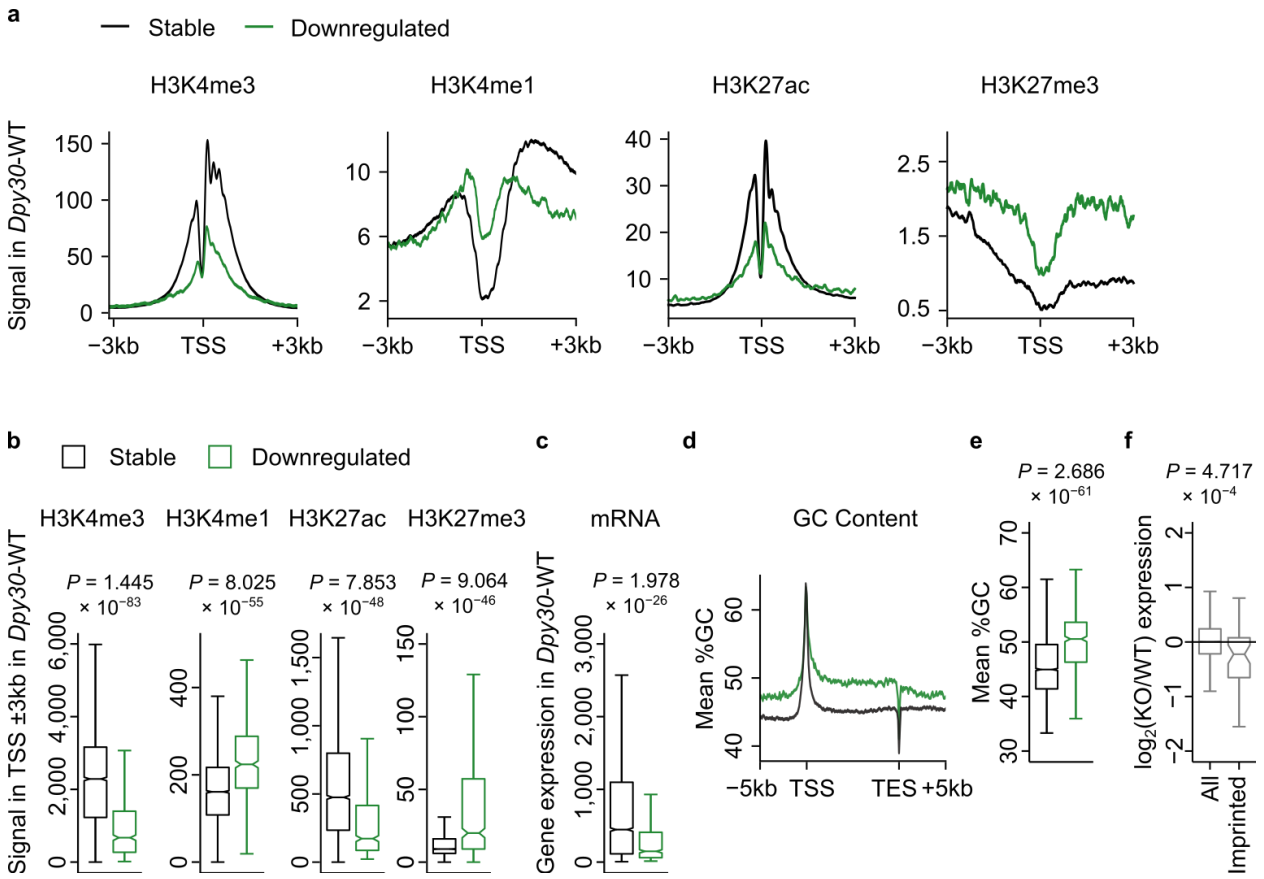

Figure 3.8 | Genes in weakly active and suppressed chromatin are susceptible to transcriptional downregulation in *Dpy30*-KO cells. a, Average enrichment profiles of H3K4me3, H3K4me1, H3K27ac, and H3K27me3 in *Dpy30*-WT chromatin for TSS ±3 kb of genes that are stably expressed (black) or downregulated (green) in *Dpy30*-KO cells. **b**, Box and whiskers plots showing quantification of data in panel (a). **c**, Box and whisker plot showing mRNA expression in *Dpy30*-WT cells for genes stably expressed or downregulated in *Dpy30*-KO cells. **d**, Average %GC content profile in the gene body ±5 kb of genes that are stably expressed (black) or downregulated (green) in *Dpy30*-KO cells. **e**, Box and whiskers plot showing quantification of data in panel (d). **f**, Box and whiskers plot showing average $\log_2$(fold change) of gene expression of all genes and of imprinted genes (GSEA:M2491; ref [390]). *P* values calculated using Wilcoxon signed-rank test with Benjamini-Hochberg correction.

3.9 H3K4 methylation prevents localized polycomb repression

The TrxG was originally defined as antagonists of PcG repressive activity [391]. I therefore examined how the genomic distribution of PcG-dependent H3K27me3 is affected by loss of TrxG-dependent methylation. As shown in Fig. 3.9, H3K27me3-positive compartments show similar distribution in *Dpy30*-WT and -KO chromatin. H3K27me3 enrichment in 10-kb windows across the genome remained similar between *Dpy30*-KO and -WT (Fig 3.10a), together indicating that global reduction of H3K4me3 and H3K4me1 did not result in expansion of polycomb-repressed compartments. Focusing in promoters, H3K27me3 enrichment remained similar overall but tended to accumulate at genes that are downregulated in *Dpy30*-KO cells (Fig. 3.10b-d). These data indicate that H3K27me3 organization is largely insensitive to reduction of H3K4 methylation in this model; however, maintenance of H3K4 methylation is necessary to prevent localized polycomb repression in mature β-cells.

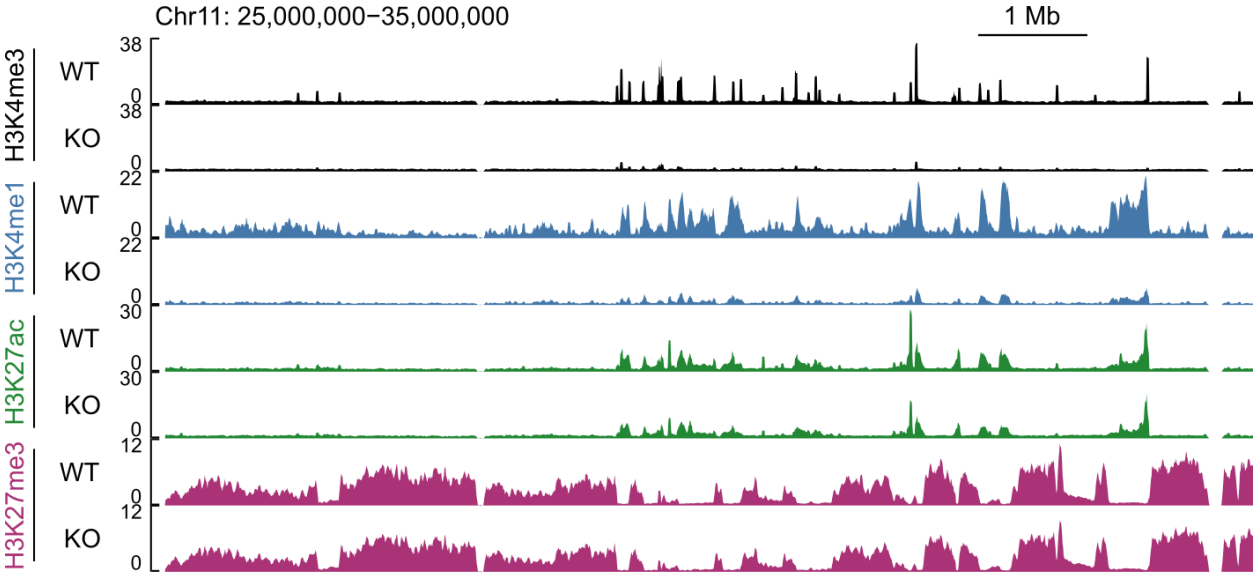

Figure 3.9 | Megabase-scale organization of H3K27me3 compartments is preserved in *Dpy30*-KO cells.

Genome browser view of H3K4me3, H3K4me1, H3K27ac, and H3K27me3 in *Dpy30*-WT and -KO chromatin.

3.10 H3K27ac is partially a consequence of H3K4me3

H3K27ac intensity is diminished from TSS's and throughout the genome in *Dpy30*-KO cells (Fig. 3.10a-c). Like H3K4me3 and H3K4me1, H3K27ac is reduced even at promoters of genes that are stably expressed or upregulated (Fig. 3.10d, right), indicating that the generalized decrease in H3K27ac in *Dpy30*-KO cells is not a consequence of gene downregulation. H3K27ac is also decreased at promoters where H3K4me1 is increased (Fig. 3.5c), supporting a role for H3K4me3 upstream of H3K27ac. Despite the global reduction, however, H3K27ac is 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 chance (Fig. 3.11). These observations are consistent with reports that H3K27ac is partially downstream of H3K4 methylation [226] and of transcription [171]. Interestingly, the correlation between gene expression level and enrichment for H3K27ac, and the anticorrelation between gene expression level and enrichment for H3K27me3 or G/C content, is strengthened in *Dpy30*-KO cells compared to -WT cells (Table 3.1). This suggests that loss of H3K4me3 allows other histone modifications and the underlying DNA sequence to have a larger influence on gene expression.

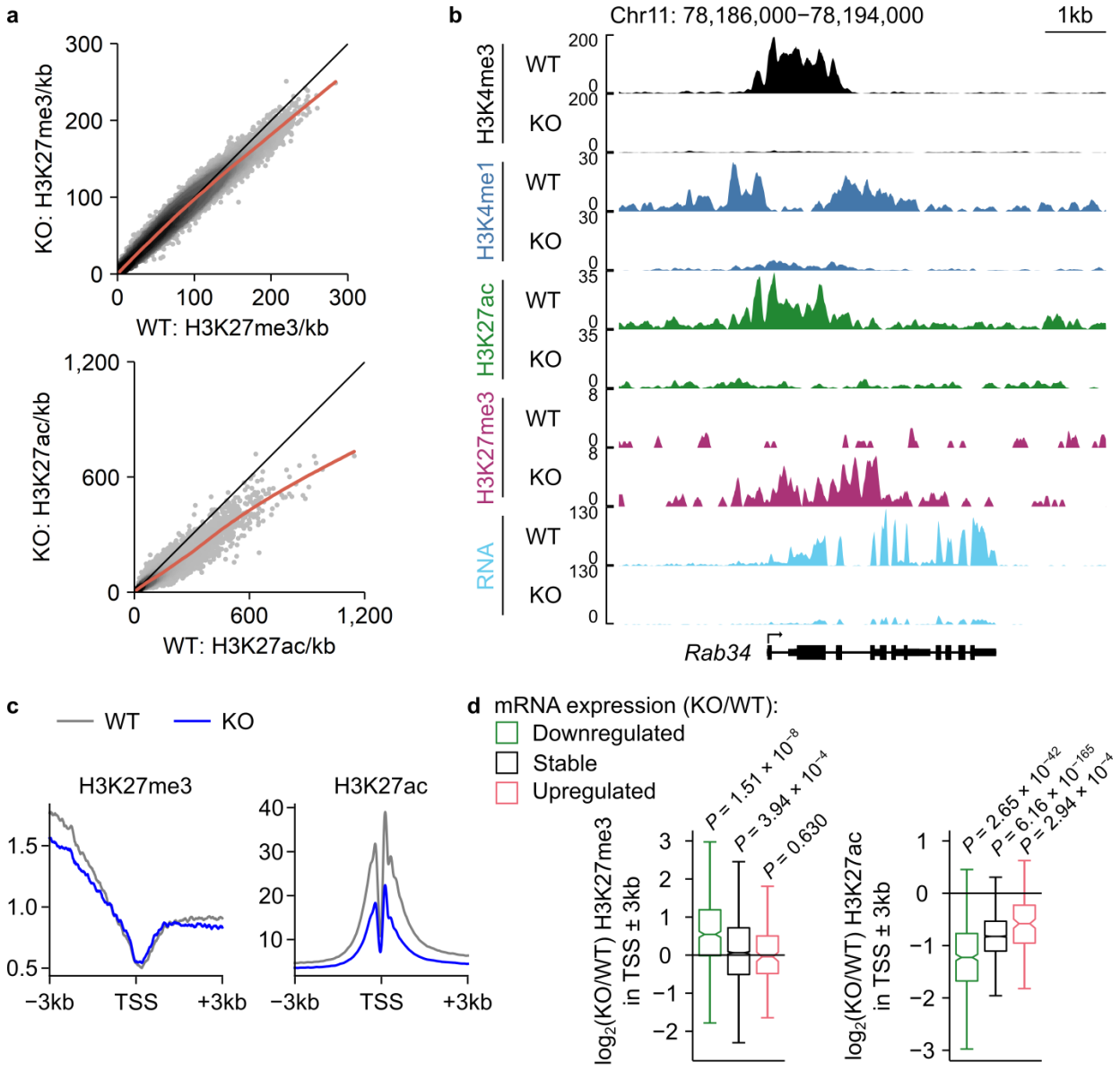

Figure 3.10 | Changes to H3K27me3 and H3K27ac across the genome and at promoters in *Dpy30*-KO cells. a, Scatterplot of H3K27me3 and H3K27ac enrichment in 10-kb bins spanning the genome. **b**, Genome browser view of a highly downregulated gene *Rab34* showing loss of H3K4me3, H3K4me1, and H3K27ac, and accumulation of H3K27me3 in *Dpy30*-KO cells. **c**, Average enrichment profiles of H3K27me3 or H3K27ac in the TSS ±3 kb of all expressed genes in *Dpy30*-WT and -KO chromatin. **d**, Box and whisker plots of the $\log_2$(fold-change) of H3K27me3 or H3K27ac enrichment in the TSS ±3 kb of transcriptionally downregulated, stable, or upregulated genes in *Dpy30*-KO compared to -WT cells. *P*-values calculated with Wilcoxon rank sum test with Benjamini-Hochberg correction versus no difference between WT and KO (i.e., versus y = 0).

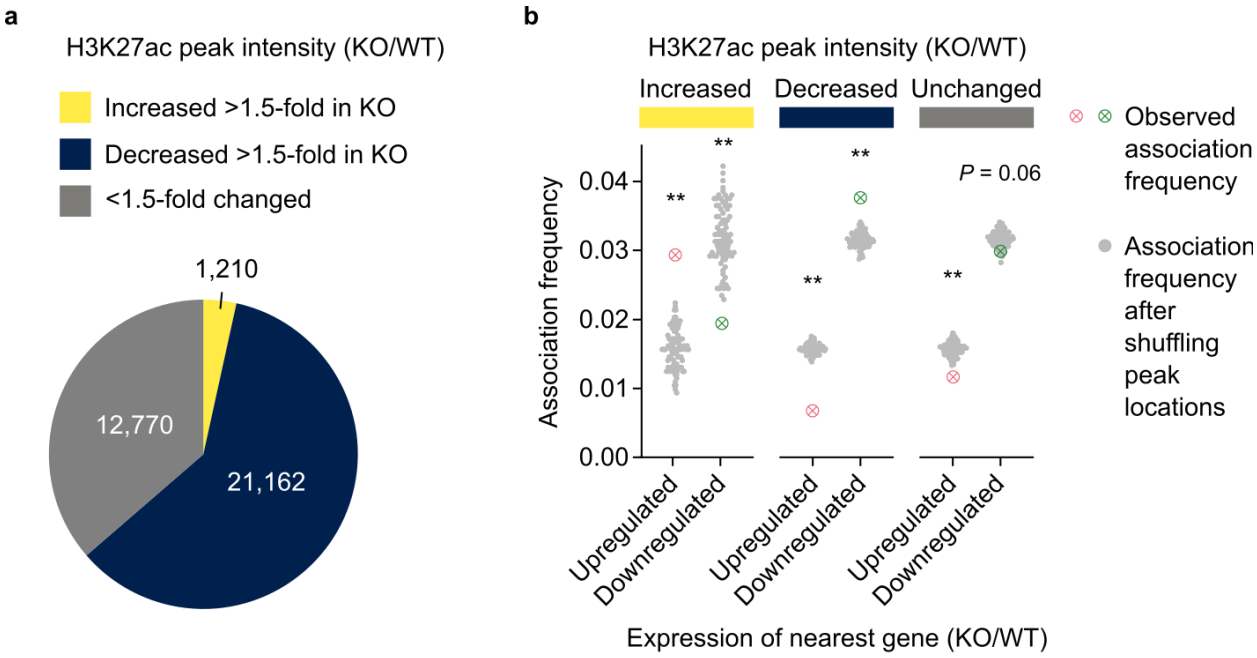

Figure 3.11 | Changes in H3K27ac enrichment in *Dpy30*-KO cells positively correlate with nearby gene expression changes. a, Pie chart showing the number of H3K27ac peaks that are increased >1.5-fold, decreased 1.5-fold, or <1.5-fold changed for enrichment of H3K27ac in *Dpy30*-KO versus -WT chromatin. **b**, Association frequency of H3K27ac peaks classified in panel (a) with genes that are upregulated or downregulated in *Dpy30*-KO cells. Association was tested between H3K27ac peaks with the nearest active TSS and compared with association frequencies after randomizing peak locations 100 times. ** $P < 0.01$ by permutation test.

Table 3.1 Relationship between gene expression level and H3K27ac, H3K27me3, or GC content is stronger after loss of H3K4 methylation. Correlations measured using Spearman's rank test.

(Epi)genetic feature	Correlation with gene expression level		
	***Dpy30*-WT**	***Dpy30*-KO**	**% Change**
H3K27ac	0.5132	0.5455	+ 6.283
H3K27me3	– 0.3819	– 0.4389	– 14.93
G/C content	– 0.0805	– 0.0957	– 18.85

3.11 H3K4 methylation maintains transcriptional consistency in β-cells

The above analyses show that H3K4 methylation contributes to maintenance of active gene expression in β-cells. Bulk downregulation of gene expression is not the only potential consequence of loss of H3K4me3, however. Transcription from genes promoters is stochastic and H3K4 methylation may impart greater consistency of gene expression between cells [180,392]. To determine if reduction to H3K4 methylation reduces transcriptional consistency, islets from a *Dpy30*-WT and a -KO mouse were isolated 45 days after tamoxifen administration and subjected to single cell RNA-seq, generating 2,510 WT and 2,597 KO cell transcriptomes that passed quality filtering and that were identified as β-cells based on unsupervised clustering and high expression of *Ins1/2*. A population of β-cells expressing *Tdtomato*, indicating that they are not Cre-recombined, formed a distinct cluster apart from the KO *Gfp*+ *β*-cells. I therefore compared *Tdtomato*+ β-cells to recombined *Egfp*+ β-cells in KO islets (henceforth mT and mG, respectively) as this is a direct comparison of *Dpy30*-expressing and -KO β-cells from the same organ (Fig. 3.12a). To model the divergence of mG from mT β-cells, I performed pseudotime ordering using the mT cluster as the ground state (Fig. 3.12b). As seen in Fig. 3.12c, *Ins1* and *Ins2* expression become notably more stochastic over pseudotime. To quantify transcriptional variability of cells, I calculated transcriptomic entropy scores for each β-cell. Transcriptomic entropy is a measure of the "specialization" of a cell's transcriptome, where a low entropy describes a narrow distribution of highly expressed genes in each cell, versus a high entropy where there is less distinction between highly- and lowly-expressed genes – a characteristic of immature cell states, including immature β-cells [108]. In *Dpy30*-KO β-cells, there is an overall increase in transcriptional entropy along pseudotime (Fig. 3.12d) and in mG cells compared to

mT cells (Fig. 3.12e). Therefore, DPY30-dependent H3K4 methylation helps to maintain low transcriptomic entropy in mature β-cells.

Finally, I asked whether reduction of gene expression consistency is linked to H3K4me3 peak breadth at the promoter, since broad H3K4me3 peaks may enforce high transcriptional consistency [180,392], that is, low cell-to-cell variability of the expression of a given gene. I hypothesized that genes with broad H3K4me3 peaks are more susceptible to increasingly stochastic expression when H3K4me3 levels are reduced, compared to other genes with narrower H3K4me3 peaks. To test this, I ranked H3K4me3 peaks in *Dpy30*-WT cells in order of increasing breadth, grouped them into 20 quantiles, each representing 622 H3K4me3 peaks with an associated gene, and compared the variance/mean ratio of gene expression in each group between mT and mG cells. Overall transcriptional variability between cells is increased in mG cells (Fig. 3.12f). As expected, there was a positive relationship between H3K4me3 peak breadth and the change in transcriptional variability, where greater H3K4me3 peak breadth was linked to greater gain of variability in *Dpy30*-KO cells (Fig. 3.12g). This supports a role for H3K4me3 in maintaining consistent gene expression.

Intriguingly, a similar pattern emerges in human β-cells from donors with T2D. Analysis of publicly available datasets for H3K4me3 ChIP-seq in healthy human β-cells (GSE50386, ref. [393]) and scRNA-seq from human donors with or without T2D (GSE124742, ref. [394]) shows that transcriptional variability is higher in donors with T2D and that the change in variability is positively related to H3K4me3 peak breadth (Fig. 3.12h-i). Transcriptional variability is also

increased in β-cells from a mouse model of T2D, *Lepr*$^{db/db}$, but a link to H3K4me3 peak breadth is not apparent in this dataset (GSE174194, ref. [395]) (Fig. 3.12j-k).

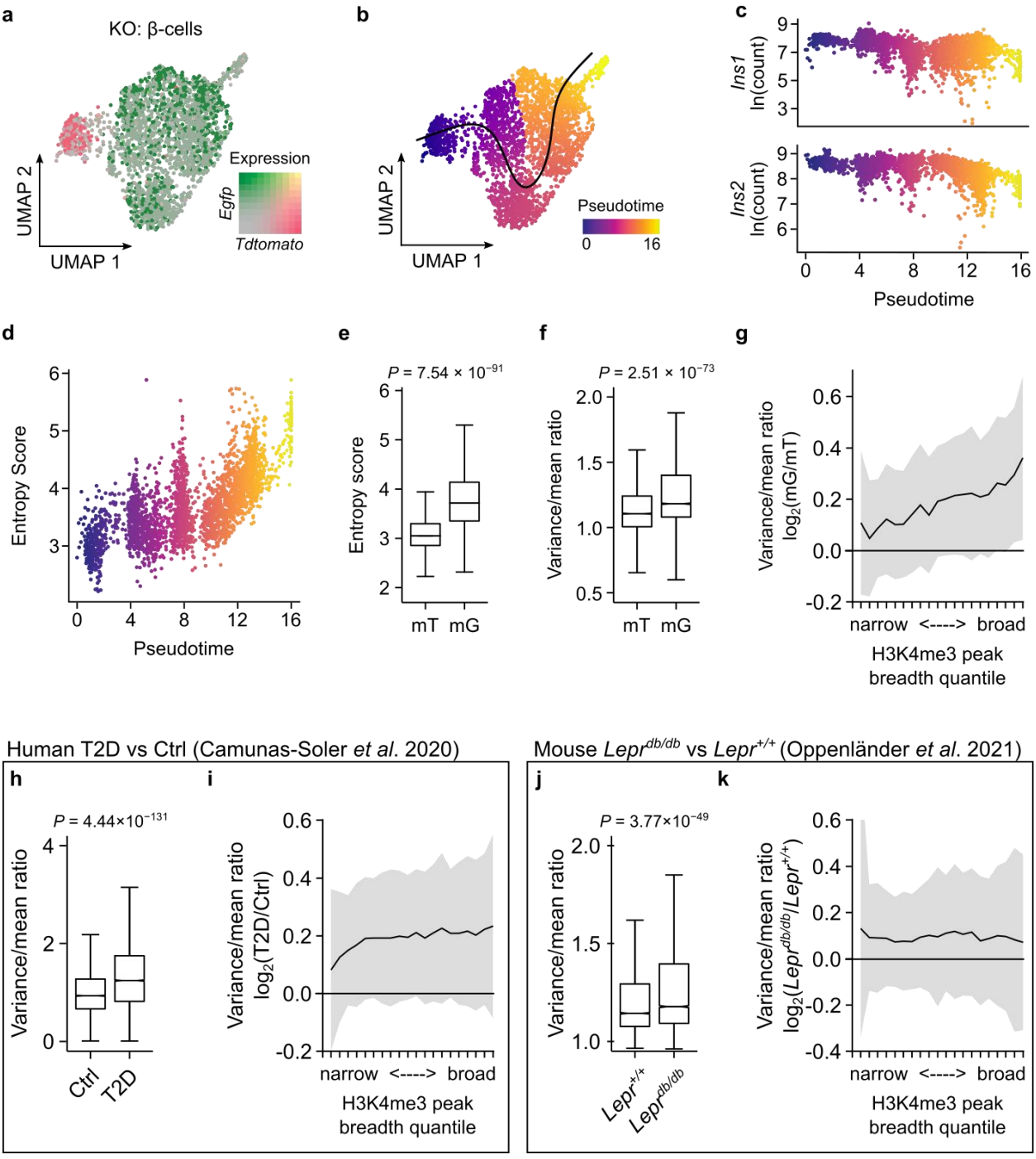

Figure 3.12 | Reduction of transcriptional consistency in *Dpy30*-KO and diabetic β-cells. a, UMAP visualization of β-cell transcriptomes from a KO mouse. Expression of *Tdtomato*, *Egfp*, and their overlap are shown as a colour gradient (n = 2,597 β-cells). **b**, Pseudotime representation of (a). **c**, Scatterplot showing *Ins1* and *Ins2* expression in β-cells plotted against pseudotime. **d**, Scatterplot showing gene expression entropy scores of β-cells plotted against pseudotime. **e**, Box and whisker plot showing the transcriptional entropy of β-cells in *Tdtomato*+ and *Egfp*+ clusters (n = 274 and 2,323 cells, respectively). **f**, Box and whisker plot showing the variance/mean ratio of gene expression across β-cells in *Tdtomato*+ and *Egfp*+ clusters (n = 274 and 2,323 cells, respectively). **g**, Variance/mean ratio of β-cell gene expression in *Egfp*+ versus *Tdtomato*+ clusters (y-axis) plotted as a function of H3K4me3 breadth in quantiles (x-axis) (622 genes/quantile). Data presented as mean ± SD. **h-i**, same as f-g in β-cells from human donors with or without T2D, using data from GSE50386 and GSE124742. **j-k**, Same as f-g in β-cells from $Lepr^{db/db}$ and $Lepr^{+/+}$ mice, using data from GSE174194. *P*-values calculated using Wilcoxon signed-rank tests.

3.12 Discussion

The complexity and inter-relatedness of TrxG activity has made confirming gene activation by H3K4me3 extremely challenging. First, there are many enzymes that write H3K4me3 which have partially overlapping genomic targets. Compensation from other enzymes therefore maintains high levels of H3K4me3 and likely obfuscate true roles of H3K4me3 if only one or a few enzymes are deleted [178]. Second, the enzymes themselves have biological functions independent of their methyltransferase activity. Transcriptional defects arising from their deletion are not necessarily downstream of methylation deficiency [226,250]. In this chapter, I deleted a core component of all six TrxG complexes, DPY30, to cause dramatic reduction of H3K4me3 genome-wide in mature murine β-cells. Using time-resolved transcriptomic analysis, transcriptome remodeling is shown to follow depletion of H3K4 methylation. Loss of H3K4 methylation led to a less active and more repressed epigenome profile which locally correlated with gene expression defects, especially at genes that may be sensitive to repression. Gene expression became more stochastic. These analyses support a role for H3K4me3 in the regulation of gene expression in mature β-cells.

Empirical normalization of cellular DNA and mRNA content by spike-in calibration [396,397] revealed a global reduction of H3K4me3 and H3K4me1 after *Dpy30* deletion. mRNA expression was linked to local changes to histone methylation but overall cellular mRNA levels were not reduced. These epigenetic modifications may, therefore, be important for allocation of transcriptional machinery to particular genes, rather than their absolute activity [398]. In other words, promoter-associated H3K4me levels modulate the expression of target genes but global H3K4me levels should not be taken as a reflection of a cell's transcriptional output (Fig. 3.13).

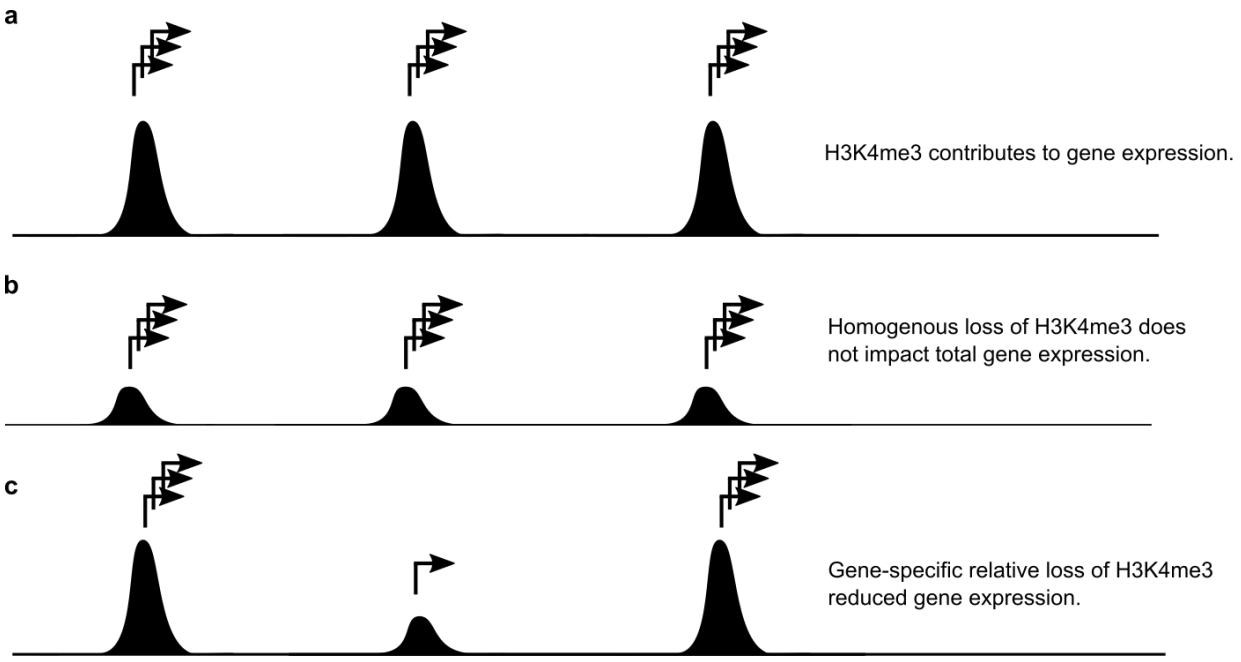

Figure 3.13 | Model for H3K4me3-dependent gene expression. a, Promoter-associated H3K4me3 is locally correlated with transcriptional output of the associated gene. **b**, Generalized reduction of H3K4me3 enrichment across all genes does not reduce total mRNA expression in a β-cell, although transcription does become more variable. **c**, Gene-specific reduction of H3K4me3, relative to other genes, leads to transcriptional downregulation.

Loss of H3K4me3 at the promoter leads to significant reduction to gene expression in this model. Targeted deposition of H3K4me3 at inactive gene promoters can activate gene expression, as shown using a dCas9-methyltransferase fusion protein [202]. Together this indicates that H3K4me3 serves as a meaningful activating signal for gene expression, even in fully differentiated cells. I noted that genes that are already in a relatively inactive or repressed chromatin state are sensitized to downregulation when H3K4 methylation is reduced. Rather than indicating a unique reliance for H3K4me3 by these genes, this may be a consequence of the model used: deletion of *Dpy30* leads to generalized slow demethylation of H3K4; genes which had little

H3K4me3 to start with are more likely to become truly demethylated within the timescale of the experiment. Therefore, although a minority of genes were downregulated in *Dpy30*-KO cells, these results do not exclude a more general role for H3K4me3 in maintenance of gene expression in β-cells. I suspect that a later time-point would allow greater depletion of H3K4 methylation and identify more H3K4me-dependent genes. Nevertheless, it is important to note that many genes are expressed in the apparent absence of H3K4me3 in *Dpy30*-KO and -WT cells, and in other models, for example ref. [244]. Among genes that completely lose H3K4me3 in *Dpy30*-KO cells, while significantly downregulated as a group, there are many genes that are not downregulated. H3K4me3 (and H3K4me1) are therefore not strictly required for all gene expression in β-cells. What causes certain genes to rely on H3K4me3, while others are unperturbed by loss of H3K4me3, is an open question, and probably related to compensation and competition from other chromatin factors.

Since H3K4 cannot be simultaneously mono- and trimethylated, enrichment of H3K4me3 causes H3K4me1 to be generally excluded near TSSs. The degree of exclusion relates to the transcriptional output of the TSS, where lowly expressed genes display closer infiltration of H3K4me1 [222,384]. In *Dpy30*-KO cells, replacement of H3K4me3 with H3K4me1 in the TSS was associated with transcriptional downregulation. Surprisingly, replacement of H3K4me3 with H3K4me1 buffered gene downregulation compared to complete demethylation, suggesting that H3K4me1 partially compensates for loss H3K4me3 in promoters. This mirrors a finding that transcriptional enhancers can function with H3K4me3 instead of H3K4me1 [250]. Together, these results argue against an immutable functional distinction between H3K4me1 and H3K4me3.

Previous studies have shown that deletion or mutation of TrxG H3K4 monomethyltransferase enzymes MLL3/4 leads to diminished H3K27ac at enhancers [399], and that the effect is partially a consequence of loss of MLL3/4, and partially of loss of H3K4me1 [226]. *Dpy30*-KO cells displayed a generalized loss of H3K27ac, including at TSSs that gained H3K4me1, consistent with the cited studies and extending them to show that H3K27ac is partially downstream of H3K4me3. Mechanistically, reduction of H3K27ac from TSSs may be caused by reduced recruitment of H3K4me3-interacting histone acetyltransferases, for example refs. [400–402]. Despite the generalized reduction of H3K27ac, its local enrichment remained linked to gene expression in *Dpy30*-KO cells, which is consistent with a report that histone acetylation is largely dependent on transcription in yeast [171]. Links between H3K27ac and H3K4 methylation and between H3K27ac and gene expression in *Dpy30*-KO cells are therefore consistent with previous reports and extends them into an *in vivo* mammalian setting.

In addition to alterations in gene expression regulation manifesting as up- or downregulation, *Dpy30*-KO β-cells display a more stochastic gene expression profile. High gene expression variability between cells in a tissue is a beneficial feature in some circumstances – it can increase adaptability to stress and can reflect a high degree of cell subtype specialization. In other circumstances, high variability is associated with aging and disease [403–406]. Very broad H3K4me3 peaks are proposed to confer high transcriptional stability [180], and experimental perturbation of H3K4me3 enrichment using *Dpy30*-KO in this chapter supports this idea. Further, transcription of genes with H3K4me3 peaks of any breadth tends to become more variable in Dpy30-KO cells, meaning that H3K4me3 generally confers consistent mRNA transcription. At the molecular level, differences in how stochastic mRNA expression is likely result from changes to

transcriptional bursting kinetics, which are uniquely shaped by promoters and enhancers by controlling aspects of transcription initiation [407,408]. This suggests that H3K4me3 partially controls bursting kinetics, as has been carefully demonstrated of H3K4me1 and H3K27me3 [197,199], and I envisage that transcription-control deficits leading to high variability in *Dpy30*-KO cells culminate into the observed changes in bulk gene expression. However, defining exactly how H3K4me3 influences transcriptional activity of select genes will require detailed consideration of the larger epigenetic environment and RNA polymerase activity. Nevertheless, the analyses presented in this chapter show that maintenance of H3K4me3 is a meaningful requirement for the regulation of gene expression in mature β-cells.

Chapter 4: H3K4 methylation influences β-cell gene expression, function, and diabetes

4.1 Rationale

Insufficient insulin release by β-cells is the primary etiology of T2D and is associated with metabolic and transcriptional remodeling in β-cells and other tissues [344,409,410]. Individual risk and prognosis of T2D is influenced by genetic and environmental factors including age, exercise, and nutrition [409]. The influence of genetic and environmental factors on the regulation of gene transcription is partially facilitated by chromatin modifications [411,412]. There is mounting experimental evidence that chromatin modifications play a central role in the maintenance of mature β-cell identity and function, and that the epigenetic landscape is reshaped in β-cells exposed to chronic metabolic stress and T2D [275,277,279,344,413–417]. However, the role of H3K4me3 in mature β-cells has not yet been explored. Studies in Chapter 3 established that H3K4 methylation contributes to the regulation of gene expression in mature β-cells but left unanswered what role H3K4me-dependent gene expression plays in the function of mature β-cells, and their dysfunction in T2D.

Most actively transcribed genes are marked by symmetrical 1-2 kb wide peaks of H3K4me3 enrichment centered at the TSS, but a small number of genes possess H3K4me3 peaks that extend up to tens of kb in length downstream into the gene body. Genes with these so-called "broad" peaks are tumor suppressors or are linked to cell identity and fulfill cell-type specific functions. Proposed functions of the broad H3K4me3 peaks themselves include enforcement of high transcription levels, high transcriptional consistency, and regulation of enhancer activity

[179,180]. The relative breadth of H3K4me3 peaks likely encodes transcriptional information for other genes as well. Expansion and contraction of H3K4me3 peak breadth, rather than peak height or area, are most predictive of differential RNA expression levels when H3K4me3 levels are perturbed using neural stimulation, conditional *Kmt2a*/*b* (MLL1/2) knockout, dietary methionine restriction in mice, or methionine-restricted culture medium in a human cell line [392,418]. In light of these reports and my finding in Chapter 3 that H3K4me3 functionally contributes to mRNA expression in healthy β-cells, I wondered if H3K4me3 peak breadth dynamics encode gene expression changes associated with T2D. In this chapter, I test the hypotheses that genes that are sensitive to H3K4me3 in β-cells contribute to β-cell function and that the genome-wide distribution of H3K4me3 is altered in a mouse model of T2D.

4.2 β-cell lineage-enriched transcription factor genes have broad H3K4me3 peaks

In healthy mouse β-cells, H3K4me3 is enriched in 18,936 distinct peak regions with a median breadth of 1.2 kb, and with 95% of peaks spanning less than 3.7 kb. Ranking H3K4me3 peaks by breadth revealed a class of exceptionally broad peaks of up to 21 kb, which included H3K4me3 peaks covering gene promoters of hallmark β-cell-enriched transcription factors (Fig. 4.1a), consistent with the idea that genes important for cell identity are marked with broad H3K4me3 peaks [179,180]. The majority (29/39, or 74%) of genes disallowed in islets [102], meanwhile, do not have any detectable H3K4me3. Gene expression is moderately correlated with promoter H3K4me3 peak breadth (Spearman rank correlation R = 0.4287); however, genes with broad peaks (henceforth, the top 5%; breadth of 3,718 – 21,304 bp) do not have exceptionally high RNA expression (Fig. 4.1b). For example, Fig. 4.1c shows the broad H3K4me3 profile of a β-cell transcription factor gene, *Nkx6-1*, contrasted with an expression-matched housekeeping gene, *Rplp0*, which has a typical H3K4me3 profile. These data confirm that genes critical to the β-cell lineage are marked by broad H3K4me3 enrichment in β-cell chromatin, but this does not simply reflect high expression of these genes.

4.3 H3K4me3 peak breadth stratifies genes dysregulated in *Lepr$^{db/db}$* and human T2D

I examined the relationship between H3K4me3 enrichment and transcriptional changes that occur in 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 and dedifferentiation [124]. I compared the H3K4me3 chromatin profiles in WT mouse β-cells for genes upregulated, downregulated (≥2-fold changed, $P \leq 0.01$, Wald test with Bejamini-Hochberg correction), or stably expressed (≤1.1-fold changed) in a publicly available *Lepr$^{db/db}$* versus

Lepr$^{+/+}$ RNA-seq dataset [419]. Genes that are downregulated in *Lepr*$^{db/db}$ islets show an exaggerated shoulder of H3K4me3 enrichment extending into the gene body in WT (non-diabetic) β-cells (Fig. 4.2a). Genes upregulated in *Lepr*$^{db/db}$ islets displayed weak enrichment for H3K4me3 (Fig. 4.2a). To conduct a quantitative analysis, I ranked all TSS-associated H3K4me3 peaks by breadth and grouped them into 20 quantiles, so that each quantile captures 5% of all genes with similar H3K4me3 breadth. Enrichment analysis shows that genes downregulated in *Lepr*$^{db/db}$ mice are strongly enriched for broad H3K4me3 peaks, while upregulated genes are strongly enriched for narrow H3K4me3 peaks (Fig. 4.2b). Furthermore, housekeeping genes (GSEA:M11197, [420]), genes involved in endocrine pancreas development (GSEA:M12875), and genes related to maturity onset diabetes of the young (GSEA:M18312) tend to have broad peaks (Fig. 4.2b). As H3K4me3 peak breadth is positively correlated with expression (Fig. 4.1b), I tested whether this observation could be secondary to a relationship wherein highly expressed genes become downregulated, and lowly expressed genes become upregulated. Ranking genes on the basis of RNA expression indicates that each gene set except housekeeping genes is more effectively stratified by H3K4me3 peak breadth than by expression (Fig. 4.3), supporting a primary link with H3K4me3. These data show that genes that are differentially expressed in *Lepr*$^{db/db}$ islets are diametrically stratified according to H3K4me3 breadth, with upregulated genes tending to have narrow peaks and downregulated genes tending to have broad peaks (in non-diabetic mice). Importantly, these observations are repeatable in humans using publicly available H3K4me3 ChIP-seq data of non-diabetic human β-cells [393] and islet RNA-seq data from non-diabetic (HbA1c < 6, n = 51) and type 2 diabetic (HbA1c > 6.5, n = 11) human donors [410] (Fig. 4.4).

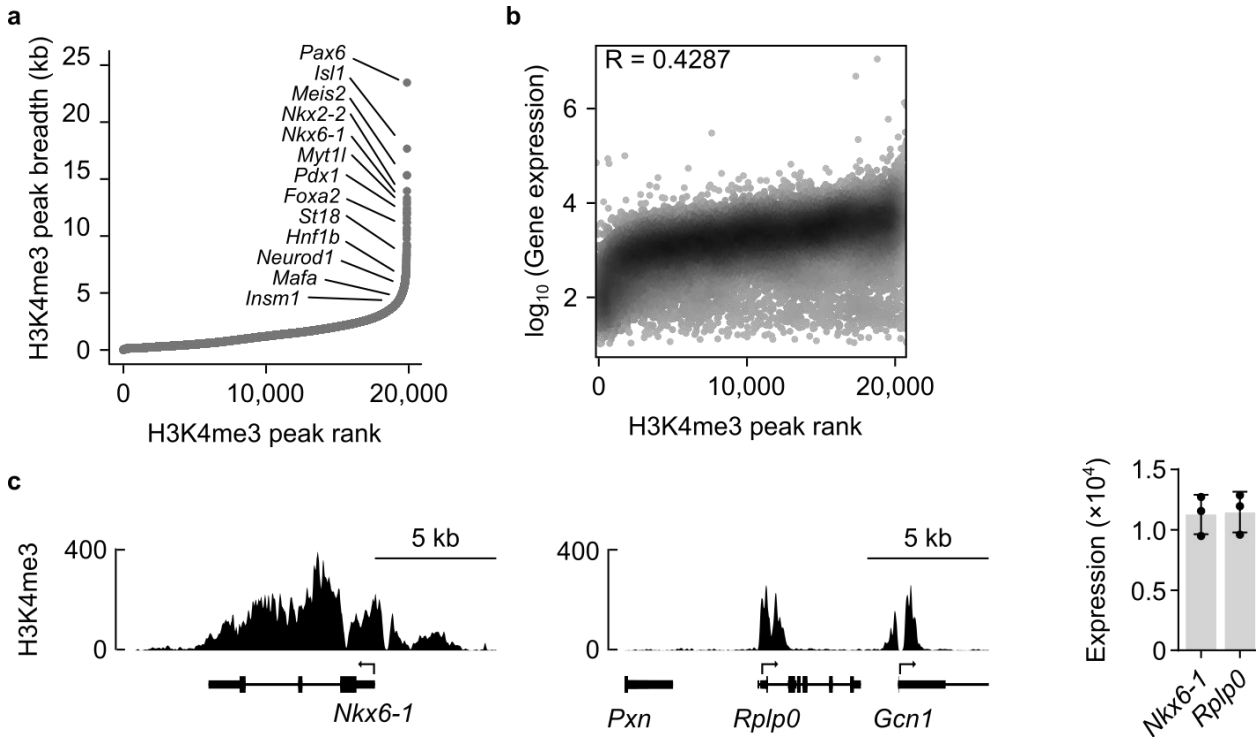

Figure 4.1 | β-cell lineage-enriched transcription factor genes are marked by broad H3K4me3 peaks. a, H3K4me3 peaks in WT mouse β-cells ranked by peak breadth. Peaks associated with hallmark β-cell transcription factor genes are labeled. **b**, H3K4me3 peaks ranked by breadth (x-axis) plotted against RNA expression of the associated gene (y-axis). R indicates Spearman's rank correlation coefficient. **c** Example of H3K4me3 density at a β-cell transcription factor gene *Nkx6-1* and an expression-matched housekeeping gene *Rplp0*.

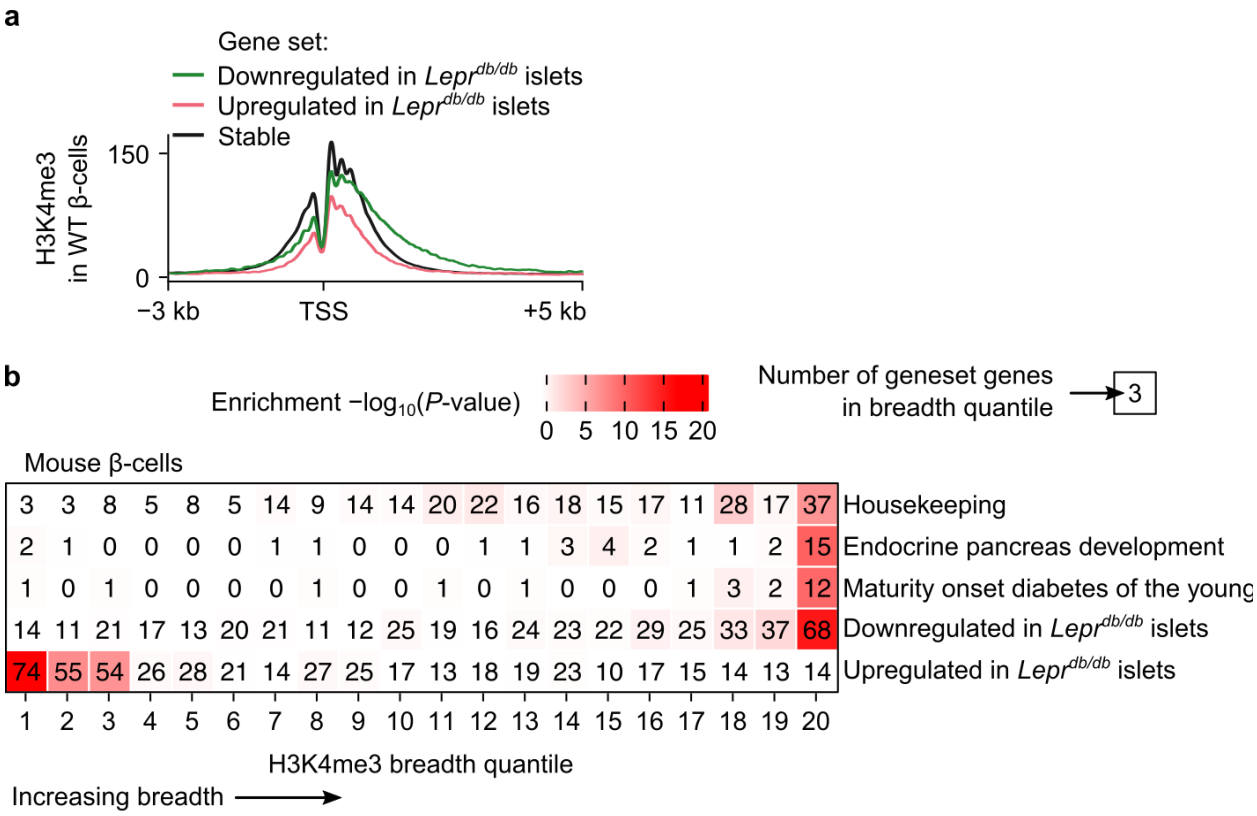

Figure 4.2 | Genes dysregulated in *Lepr*$^{db/db}$ islets are stratified by their H3K4me3 peaks. a, Average enrichment profile of H3K4me3 in *Dpy30*-WT mouse β-cells in the regions spanning –3 kb to +5 kb with respect to the TSS of genes that are downregulated (green), upregulated (red), or stably expressed (black) in *Lepr*$^{db/db}$ versus *Lepr*$^{+/+}$ islet cells. **b**, Heatmap showing the enrichment of gene sets in *Dpy30*-WT mouse β-cell transcriptome data that has been grouped into 20 quantiles based on promoter H3K4me3 peak breadth. Each quantile includes 5% (622) of genes with similar promoter H3K4me3 peak breadth, with narrowest H3K4me3 peaks at left and broadest H3K4m3 peaks at right. The number in each square indicates the number of genes of each gene set matching to each quantile. *P*-values were calculated using one-sided Fisher's exact tests.

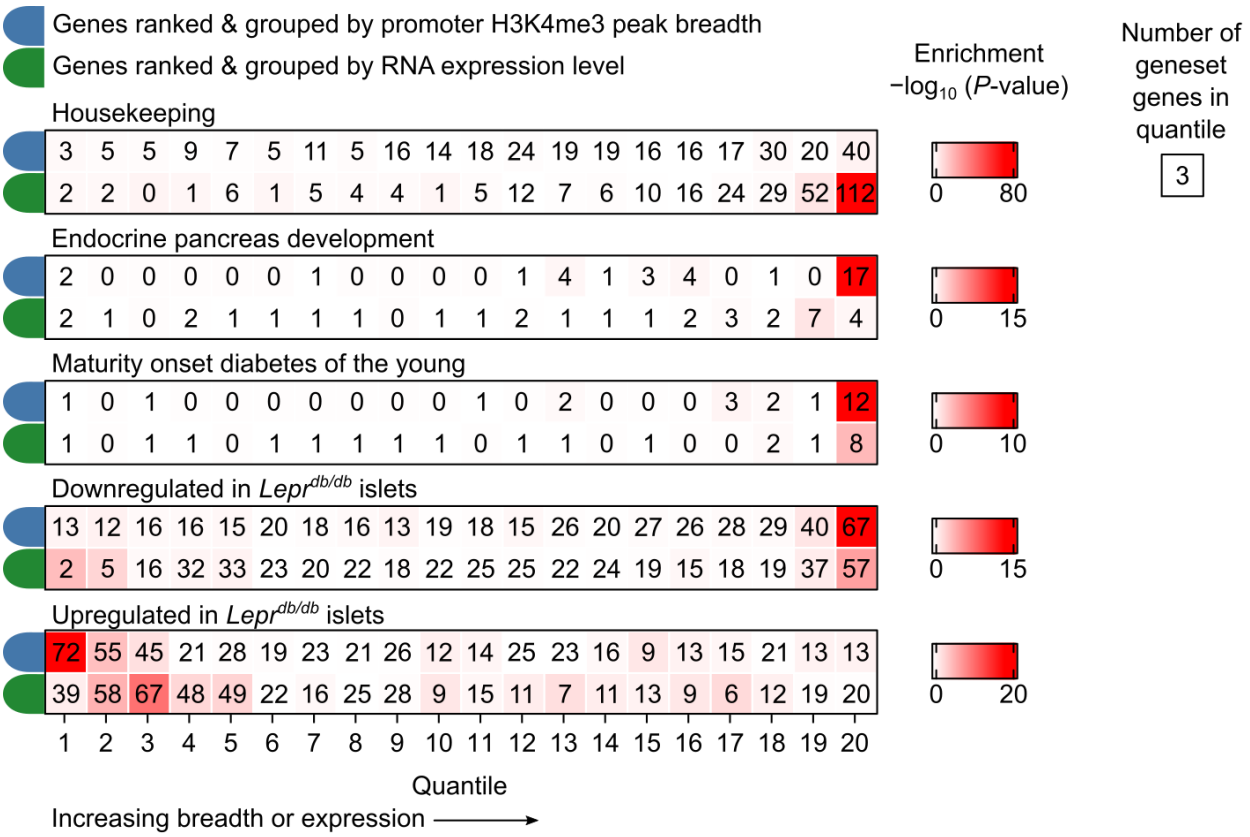

Figure 4.3 | β-cell-enriched gene sets are more strongly stratified by H3K4me3 peak breadth than by RNA expression level. Heatmap showing enrichment analysis of mouse β-cell genes that are ranked and grouped by H3K4me3 peak breadth from narrow to broad (as in Fig. 4.2b), or ranked and grouped by gene expression level from low to high expression. The number in each square indicates the number of genes of each geneset matching to each quantile. Each quantile includes 622 mouse genes. *P*-values were calculated using two-sided Fisher's exact tests.

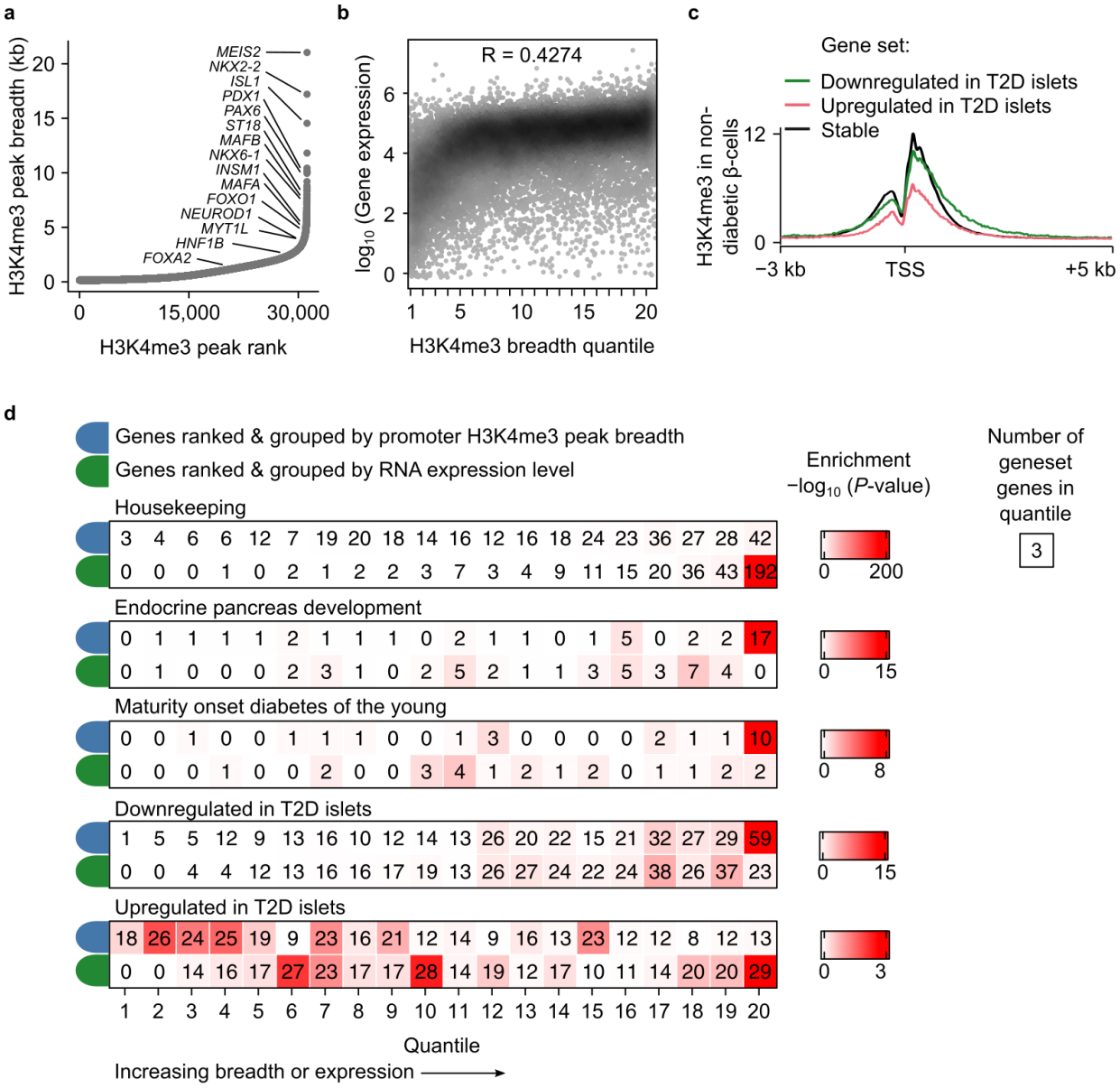

Figure 4.4 | In human β-cells, lineage-enriched genes and genes dysregulated in T2D are stratified by the breadth of promoter-associated H3K4me3 peaks. a, H3K4me3 peaks in human β-cells ranked by peak breadth. Peaks associated with hallmark β-cell transcription factor genes are labeled. **b**, H3K4me3 peaks ranked by breadth (x-axis) plotted against RNA expression of the associated gene (y-axis) in non-diabetic human islets. R indicates Spearman's rank correlation coefficient. **c**, Average enrichment profile of H3K4me3 in non-diabetic human β-cells in the regions spanning –3 kb to +5 kb with respect to the TSS of genes that are downregulated (green), upregulated (red), or stably expressed (black) in T2D. **d**, Enrichment analysis of human islet genes that are ranked and grouped

by H3K4me3 peak breadth from narrow to broad (as in Fig. 4.2b), or ranked and grouped by gene expression level from low to high expression. The number in each square indicates the number of genes of each geneset matching to each quantile. Each quantile includes 783 human genes. *P*-values calculated using one-sided Fisher's exact tests.

4.4 Expression of genes dysregulated in *Lepr*$^{db/db}$ islets is sensitive to H3K4me3 enrichment

I wondered if genes that are up- or downregulated in *Lepr*$^{db/db}$ mice require H3K4me3 for expression, i,e, are downregulated in *Dpy30*-KO mice. Gene set enrichment analysis within the H3K4me3 peak breadth quantiles shows that two populations of genes tend to be downregulated in *Dpy30*-KO cells: those at the two extremes of H3K4me3 breadth (Fig. 4.5a). In other words, genes that have very narrow or very broad H3K4me3 peaks require H3K4me3 to maintain their expression. In *Lepr*$^{db/db}$ islets, genes with very narrow peaks tend to be upregulated, and genes with very broad peaks tend to be downregulated (Fig. 4.5a). For example, hallmark β-cell transcription factor genes which have broad H3K4me3 peaks tend to be downregulated, while genes disallowed in islets [102] which have small or no H3K4me3 peaks are upregulated in *Lepr*$^{db/db}$ islets (Fig. 4.5b). Disallowed genes are not upregulated in the *Dpy30*-KO model (Fig. 4.5b), which could suggest they require H3K4me3 to be induced. Stratified rank-rank hypergeometric overlap comparison [421] reveals significant overlap between genes that are either up- or downregulated in *Lepr*$^{db/db}$ cells with genes that are downregulated in *Dpy30*-KO cells (Fig. 4.5c), further supporting that H3K4me3 is required for the expression of genes dysregulated in *Lepr*$^{db/db}$. Gene ontology analysis of overlapping downregulated genes shows a reduction to insulin secretion, response to glucose, and regulation of transcription in both models (Fig. 4.5d), suggesting that H3K4me3 is necessary to maintain these critical β-cell functions. Genes that are downregulated in *Dpy30*-KO, and upregulated in *Lepr*$^{db/db}$, models are enriched for regulation of

cell growth, transport, and unfolded protein/ER stress response (Fig. 4.5e), suggesting that H3K4me3 is necessary to activate genes involved in islet stress response and compensation in the *Lepr*$^{db/db}$ model. Therefore, H3K4me3 is necessary for the expression of genes which become up- or downregulated in *Lepr*$^{db/db}$ islets and perform contextually important biological functions.

4.5 H3K4me3 dynamics are linked to differential gene expression in *Lepr*$^{db/db}$ islets

I next asked whether H3K4me3 is altered in *Lepr*$^{db/db}$ islets. Surprisingly, the H3K4me3 landscape is relatively undisturbed. Western blot shows that H3K4me3 levels are not globally altered in islets from *Lepr*$^{db/db}$ islets (Fig. 4.6a-b). ChIP-seq similarly shows no global change in average TSS-associated H3K4me3 enrichment in *Lepr*$^{db/db}$ islets (Fig. 4.6c). However, comparison of H3K4me3 peaks at differentially expressed genes uncovered a positive association between breadth and gene expression dynamics, wherein genes that are transcriptionally up- or downregulated show corresponding expansion or contraction in HK4me3 peak breadth, respectively (Fig. 4.6d-e). This finding is in line with a previous report that changes in H3K4me3 peak breadth predict changes in gene expression [418]. In light of my finding in Chapter 3 that H3K4me3 regulates gene expression in β-cells, and in section 4.4 that H3K4me3 regulates expression of genes that are up- and downregulated in *Lepr*$^{db/db}$ mice, this supports the idea that differential gene expression between *Lepr*$^{+/+}$ and *Lepr*$^{db/db}$ mice is partially driven by changes in H3K4me3 peak breadth.

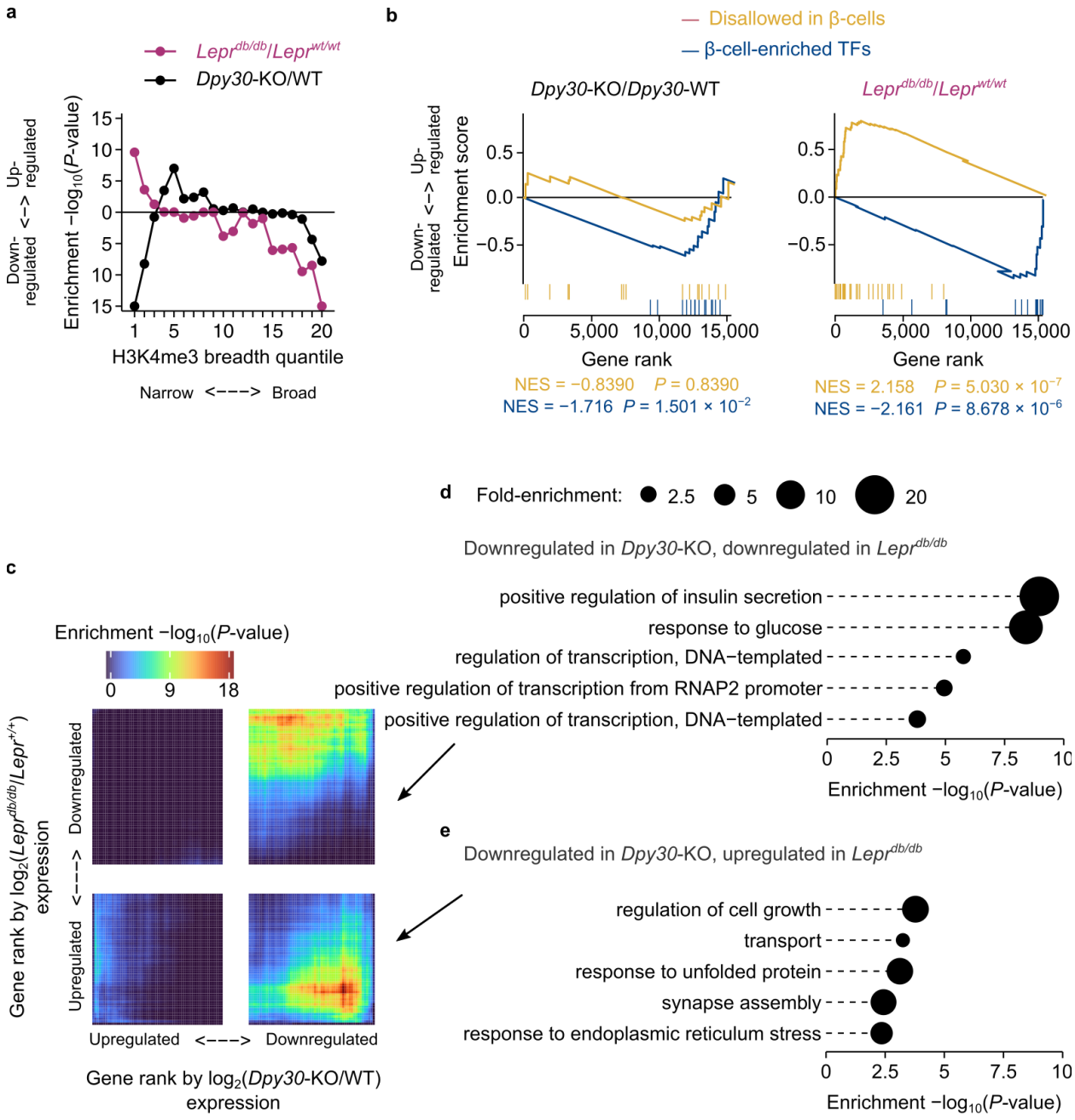

Figure 4.5 | Genes that are dysregulated in *Lepr*$^{db/db}$ islets are sensitive to loss of H3K4me3. a, Promoter H3K4me3 peak breadth quantiles ranked from narrow to broad (x-axis) plotted against the *P*-value for genes in each quantile being generally upregulated or downregulated (y-axis). *P*-values were calculated by permutation tests in gene set enrichment analysis using RNA-seq data from *Dpy30*-KO/WT (black) and *Lepr*$^{db/db}$/*Lepr*$^{+/+}$ (purple) islets. H3K4me3 peak breadth was measured in WT mouse β-cells ChIP-seq data; 622 genes per quantile. **b**, Gene set

enrichment analysis of genes disallowed in islets and β-cell-enriched transcription factor genes in RNA-seq data from *Dpy30*-KO/WT and *Lepr*$^{db/db}$/*Lepr*$^{+/+}$ islets. * $P < 0.05$, Wald test with Benjamini-Hichberg correction. **c**, Stratified rank-rank hypergeometric overlap plot comparing gene expression differences between *Dpy30*-KO versus -WT (x-axis) and *Lepr*$^{db/db}$ versus *Lepr*$^{+/+}$ [347] (y-axis). Colourscale shows the hypergeometric *P*-value calculated using the RRHO2 R package [421]. **d-e**, Gene ontology biological processes enriched among genes that are downregulated in both *Dpy30*-KO and *Lepr*$^{db/db}$ (c) or downregulated in *Dpy30*-KO and upregulated in *Lepr*$^{db/db}$ (d) islet cells. The five terms with smallest *P*-values are shown. *P*-values represent EASE scores calculated using DAVID v6.8.

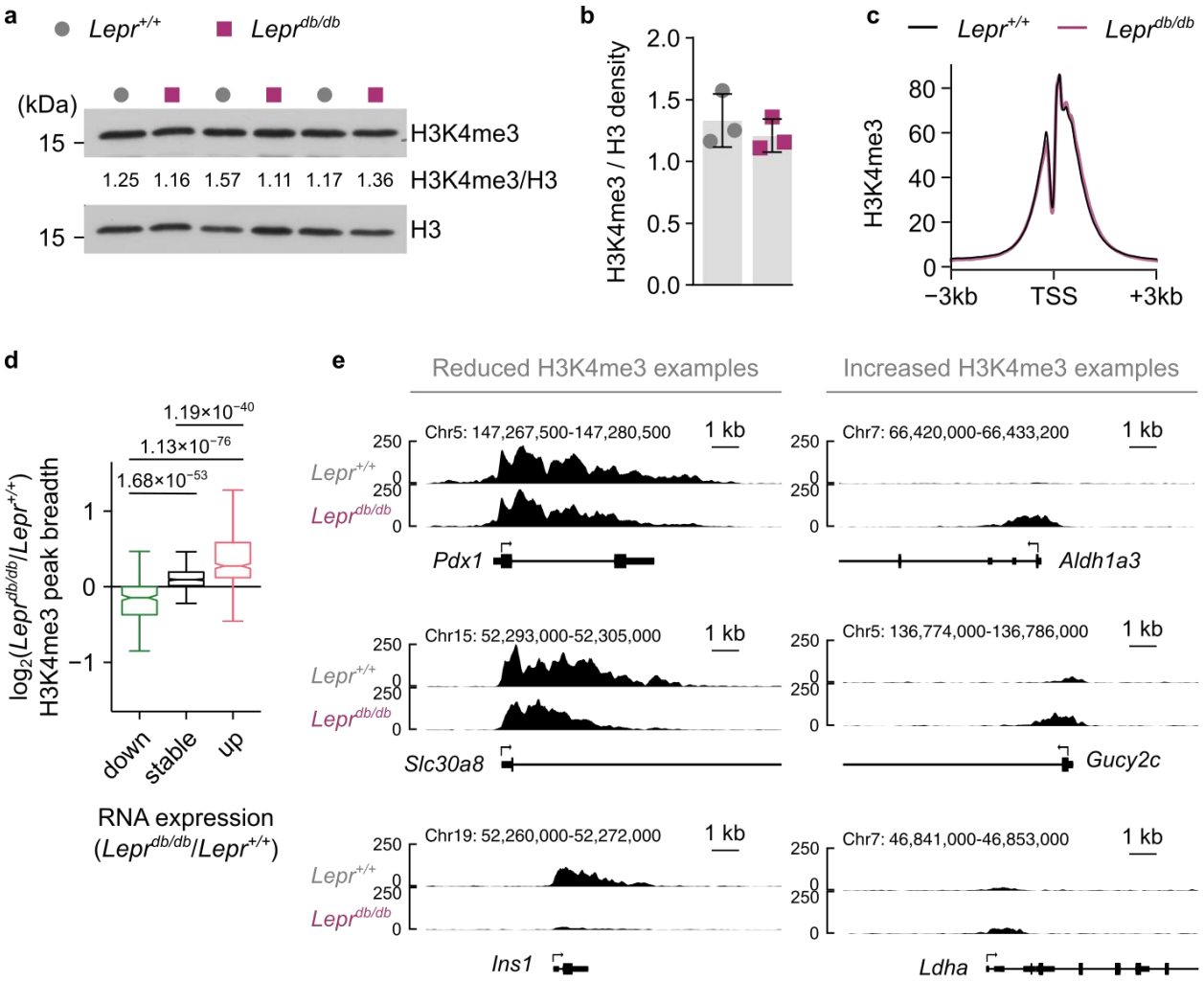

Figure 4.6 | H3K4me3 peak breadth dynamics are linked to differential gene expression in *Lepr*$^{db/db}$ islets. a, Immunoblots of H3K4me3 and total histone H3 in islet lysate from *Lepr*$^{+/+}$ and *Lepr*$^{db/db}$ mice. **b**, Bar graph showing band intensities of data in panel (a). **c**, Average enrichment profile of H3K4me3 at the TSS of all expressed genes in *Lepr*$^{+/+}$ and *Lepr*$^{db/db}$ islets. **d**, Box and whisker plots of the $\log_2$(fold-change) of H3K4me3 peak breadth for transcriptionally downregulated, stable, or upregulated genes in *Lepr*$^{db/db}$ compared to *Lepr*$^{+/+}$ islets. *P*-values calculated with Wilcoxon rank sum test with Benjamini-Hochberg correction. **e**, Examples of H3K4me3 peaks in *Lepr*$^{+/+}$ and *Lepr*$^{db/db}$ islets for genes that are downregulated (*Pdx1*, *Slc30a8*, *Ins1*) and upregulated (*Aldh1a3*, *Gucy2c*, *Ldha*).

4.6 Reduction to H3K4me3 in β-cells leads to hyperglycemia and glucose intolerance

These observations motivated me to test whether reduction of H3K4me3 from β-cells leads to a diabetes-like phenotype. I first examined *in vivo* metrics of β-cell function in the *Pdx1-CreER Dpy30*-KO mouse model. Blood glucose concentration of unfasted *Dpy30*-KO mice is dramatically elevated from ~45-days after tamoxifen administration without difference in body mass (Fig. 4.7). *Dpy30*-KO mice display glucose intolerance and impaired insulin secretion in glucose tolerance tests 45-days after *Dpy30* inactivation, whereas no significant defect is seen at 15- or 30-days (Fig. 4.8). Notably, the delay between Cre induction and appearance of gross phenotype supports a primary effect of H3K4 methylation rather than DPY30 in this model, as discussed in section 3.3.

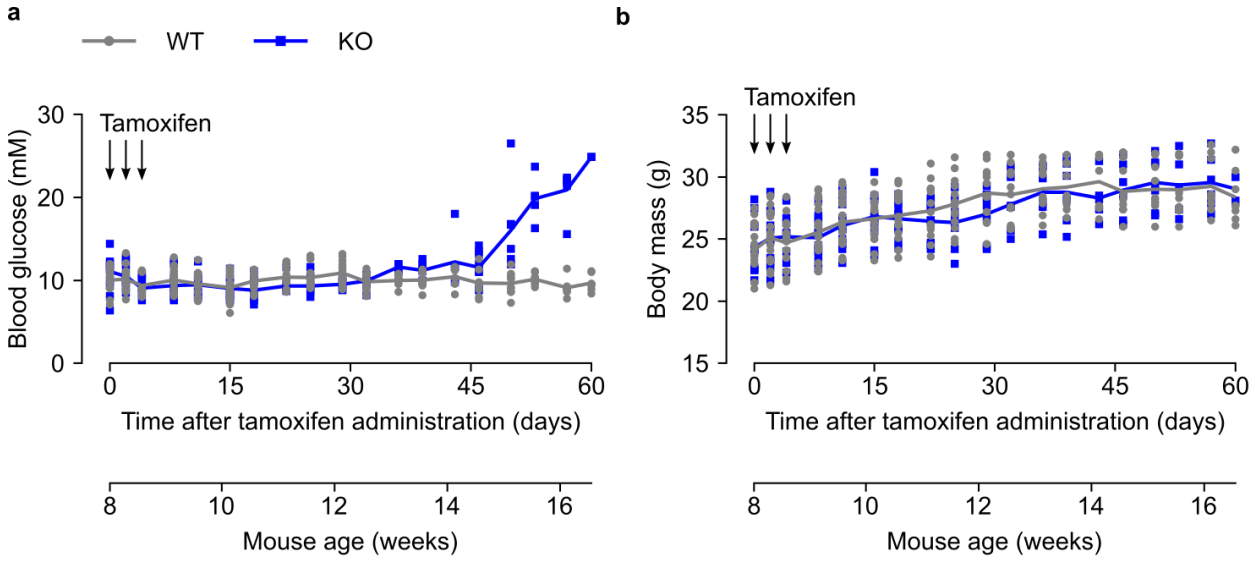

Figure 4.7 | *Dpy30*-KO mice develop hyperglycemia. a-b, Unfasted blood glucose concentration (a) and body mass (b) of *Dpy30*-WT and -KO mice during the 60-days following tamoxifen administration. Data presented as individual measurements and mean (n = 7 *Dpy30*-WT, 8 *Dpy30*-KO, however tracking was stopped after a blood glucose reading ≥ 20 mM. Days that tamoxifen was administered are indicated by arrows.

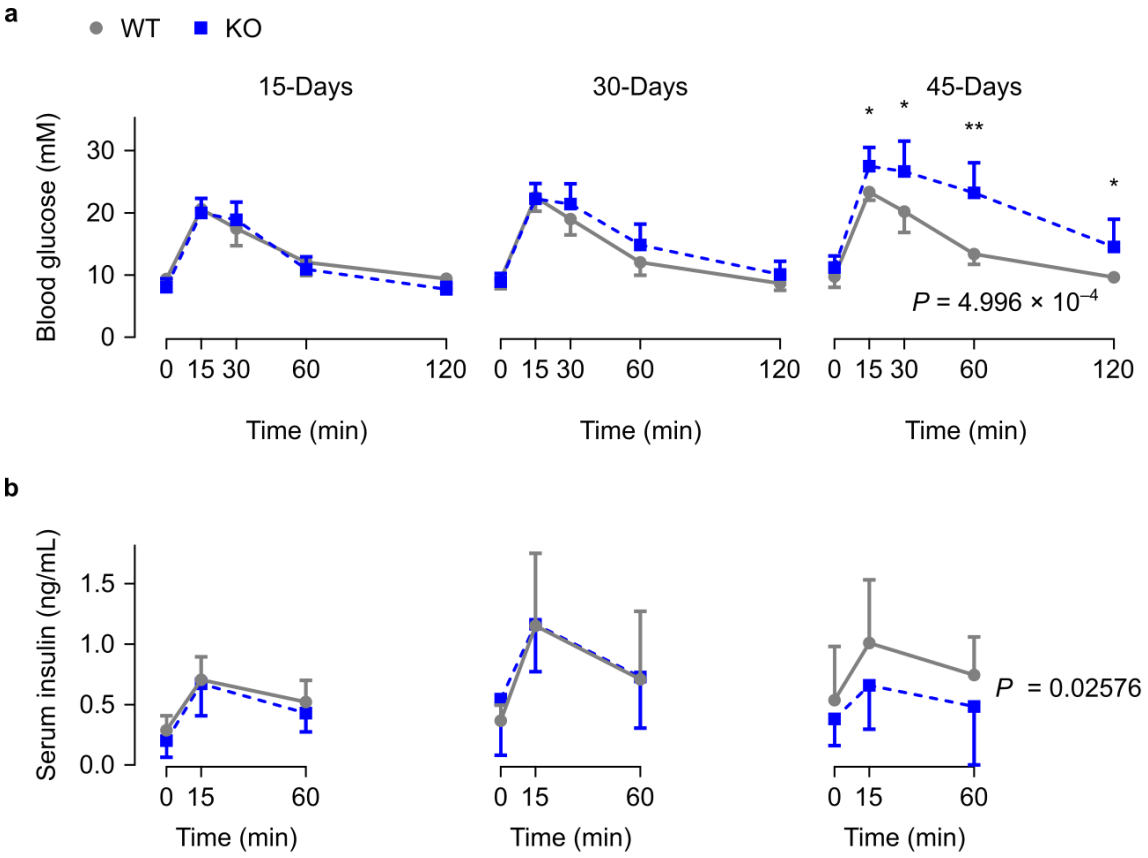

Figure 4.8 | *Dpy30*-KO mice develop glucose intolerance. a-b, Blood glucose (a) and serum insulin (b) concentration during intraperitoneal glucose tolerance tests in *Dpy30*-WT and -KO mice 15-, 30-, and 45-days after tamoxifen administration. Data presented as mean + or – SD; n = 8 mice for glucose measurements, 8-15 for insulin measurements. *P*-values were calculated using two-way ANOVA with Geisser-Greenhouse and Benjamini-Hochberg corrections. * < 0.05 ** < 0.01.

Functional annotation of differentially expressed genes in β-cells 45-days after tamoxifen administration highlights a reduction to exocytosis, calcium signaling, and glucose homeostasis processes (Fig. 4.9), supporting a defect in insulin secretion in these animals. However, the *Pdx1-CreER* transgene is known to drive recombination in pancreatic islet δ-cells [422] and hypothalamic neurons [423]; so to confirm that *Dpy30* deletion in β-cells is sufficient to drive the *in vivo* phenotype I additionally used the *Ins1*^{Cre} model [354]. Deletion of *Dpy30* from maturing β-

cells using *Ins1*^{Cre} causes reduction of H3K4me3 and H3K4me1 by 5-weeks of age (Fig. 4.10a-b). Unfasted blood glucose is dramatically elevated from 4-5 weeks of age, whereas body mass is similar among *Dpy30*-KO, WT, and heterozygous *Ins1*^{Cre} mice (Fig. 4.10c-d). *Ins1*^{Cre} *Dpy30*-KO mice show impaired glucose clearance and low serum insulin during glucose tolerance tests (Fig. 4.10e-f), confirming that β-cell-intrinsic dysfunction is sufficient to drive metabolic defects in conditional *Dpy30*-KO mice. In summary, reduction of H3K4me in β-cells leads to a diabetes-like phenotype of impaired glucose tolerance, reduced serum insulin, and hyperglycemia.

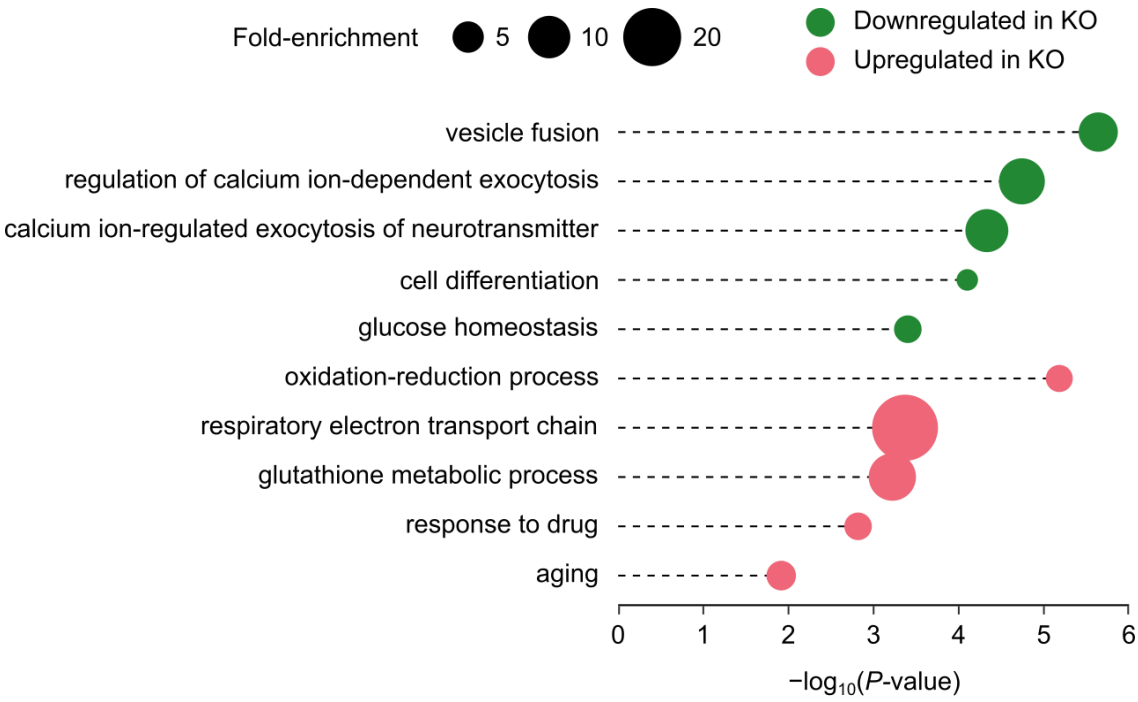

Figure 4.9 | Gene ontology biological processes enriched among genes downregulated or upregulated in *Dpy30*-KO cells. Differentially expressed genes (fold-change >2, $P < 0.01$, Wald test with Benjamini-Hochberg correction) were identified using mRNA-seq 45-days after tamoxifen administration. The five terms with lowest *P*-values are shown. *P*-values represent EASE scores calculated using the Database for Annotation, Visualization, and Integrated Discovery (DAVID) v.6.8 [424].

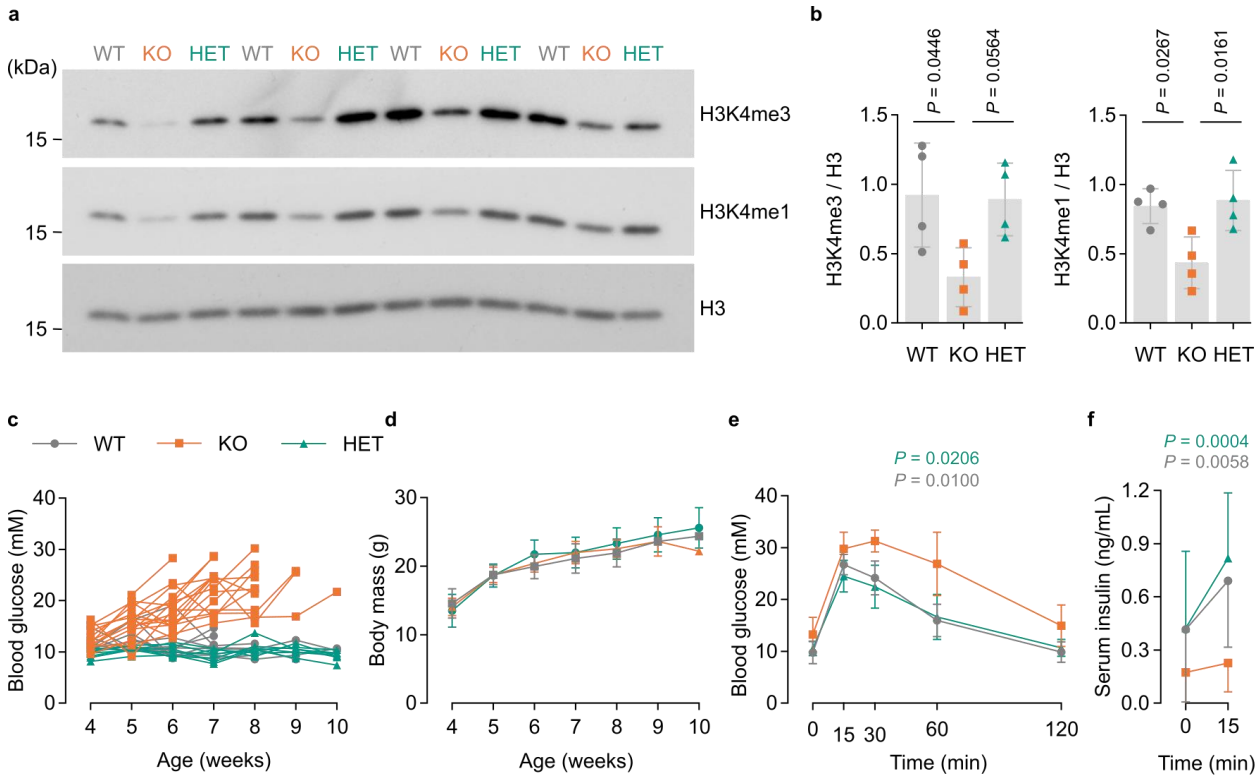

Figure 4.10 | Deletion of *Dpy30* in β-cells using *Ins1*^{Cre} results in reduction of H3K4 methylation, hyperglycemia, hypoinsulinemia, and impaired glucose tolerance. a, Immunoblots of H3K4me3, H3K4me1, and total histone H3 in islet lysate from 5-week-old *Dpy30*-WT, -KO, and -HET mice. **b**, Bar graph showing band intensities of data in panel (a). *P*-values were calculated using one-way ANOVA with Tukey's multiple comparisons correction; n = 4. **c-d**, Unfasted blood glucose concentration (c) and body mass (d) of *Dpy30*-WT, -KO, and -HET mice during 4-10 weeks of age. Values from each mouse and time point are shown (n = 10 WT, 15 KO, 7 HET, however tracking was stopped after 1-2 blood glucose reading ≥ 20 mM). **e-f**, Blood glucose (e) and serum insulin (f) concentration during intraperitoneal glucose tolerance tests in 5-week-old *Dpy30*-WT, -KO, and -HET mice. *P*-values were calculated by comparison of *Dpy30*-KO versus -HET (top, teal) or -KO versus -WT (bottom, grey) AUC's using one-way ANOVA with Tukey's multiple comparisons correction; only values ≤ 0.05 are shown. WT: *Ins1*$^{+/+}$ *Dpy30*$^{flox/flox}$. KO: *Ins1*$^{Cre/+}$ *Dpy30*$^{flox/flox}$. HET: *Ins1*$^{Cre/+}$ *Dpy30*$^{flox/+}$.

4.7 Reduction of H3K4 methylation impairs insulin gene transcription and peptide processing in β-cells

To understand why *Dpy30*-KO animals develop hypoinsulinemia, I examined insulin production and secretion. Interrogation of the *Ins1* and *Ins2* gene loci in β-cells revealed broad H3K4me3 enrichment extending several kb around each gene in *Pdx1-CreER*; *Dpy30*-WT chromatin. Strikingly, *Pdx1-CreER*; *Dpy30*-KO chromatin shows complete loss of H3K4me3 and reduction of H3K4me1 and H3K27ac at these loci (Fig. 4.11a). Despite the reduction of active promoter marks, mRNA-seq data shows only modest reduction of *Ins1* and *Ins2* transcripts (–1.152 and –1.234-fold, respectively, Fig. 4.11b), indicating that while it contributes to their full expression, robust expression of these genes can occur in the absence of local H3K4me3. Despite the modest reduction in *Ins1* and *Ins2* RNA, insulin peptide content is reduced 3.4-fold in *Dpy30*-KO islets (Fig. 4.12a). Reduction of insulin content was confirmed in individual β-cells, where reduction of insulin granule size and density was apparent in transmission electron micrographs (TEM) (Fig. 4.12b-c), suggesting a defect in insulin translation, storage, and/or maturation. Two genes downregulated in KO cells, *Nnat* and *Spcs3* (Fig. 4.12d), are components of the signal peptidase complex which converts preproinsulin into proinsulin in β-cells [77]. Both genes are also downregulated in *Lepr*$^{db/db}$ islets (data not shown), and *Nnat* is an imprinted gene repressed in human obesity [425]. Reduction of Nnat causes preproinsulin to be redirected for cytosolic degradation, which suggests that an impairment of proinsulin-to-insulin processing contributes to reduced insulin content in *Dpy30*-KO β-cells [77]. I performed immunoblots using an anti-insulin antibody that detects immature and processed insulin species. To improve detection of preproinsulin, I treated islets *ex vivo* with MG-132 to inhibit proteasomal degradation of mistargeted preproinsulin before lysis (as in ref. [77]) and used a modified immunoblot method that

improves peptide retention on the membrane (ref. [375]). While I was nevertheless unable to detect preproinsulin in islet lysate, insulin and proinsulin bands were resolved. After adjusting lysate loading volumes for differences in insulin content between *Dpy30*-WT and -KO islets, the ratio proinsulin:insulin band density is similar in *Dpy30*-WT and -KO islets, consistent with a defect in insulin production upstream of proinsulin-to-insulin processing (Fig. 4.12e-f). However, accurate quantitation from immunoblot is challenging due to the large difference in signal intensity between insulin and proinsulin.

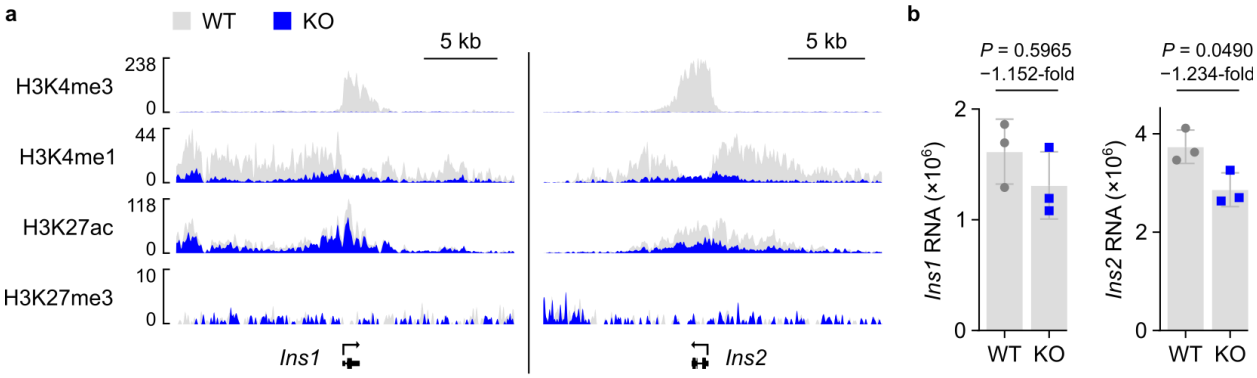

Figure 4.11 | Loss of H3K4me3 from *Ins1* and *Ins2* gene loci modestly impairs their RNA expression. a, Genome browser representation of the levels of H3K4me3, H3K4me1, H3K27ac, and H3K27me3 at the *Ins1* and *Ins2* gene loci in *Dpy30*-WT and KO chromatin. **b**, *Ins1* and *Ins2* mRNA levels in *Dpy30*-WT and -KO cells. Data presented as mean ± SD with values from individual mice (n = 3). Expression and *P*-values were calculated from mRNA-seq data using DESeq2 (Wald test with Benjamini-Hochberg correction).

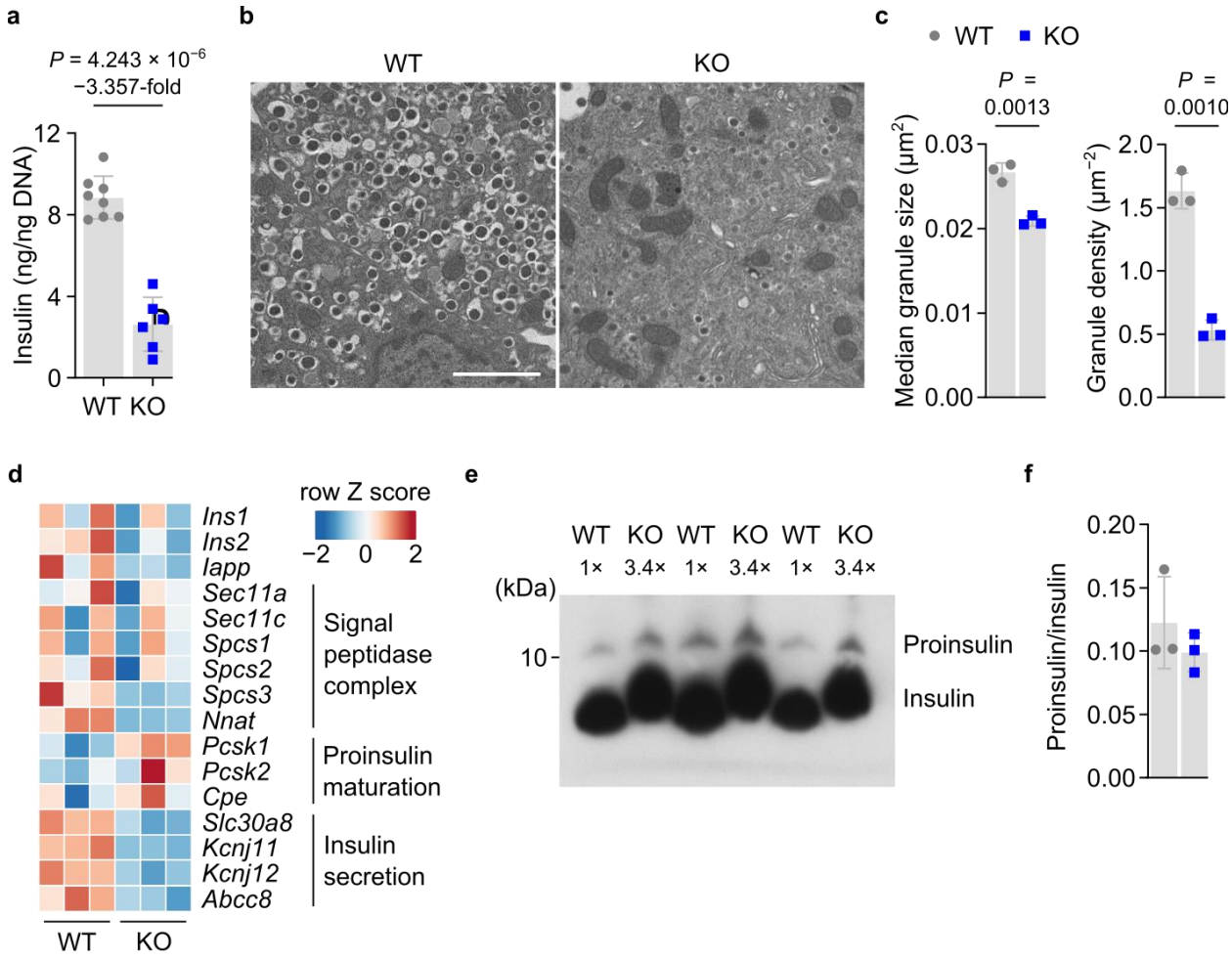

Figure 4.12 | Impairments upstream of proinsulin processing drive reduction to insulin content in *Dpy30*-KO cells. a, Insulin content in *Dpy30*-WT and -KO islets normalized to DNA content. Data presented as mean ± SD with values from individual mice (n = 8 WT, 6 KO). *P*-value calculated with two-tailed t-test with Welch's correction. **b**, Representative transmission electron micrograph showing insulin granules in β-cells from a *Dpy30*-WT and -KO mouse. Scale bar: 2 μm. **c**, Quantification of median insulin core granule size and density in TEM images exemplified in panel (b). Data presented as mean ± SD with values from individual mice (n = 3), *P*-values calculated using two-tailed t-test with Welch's correction. **d**, Heatmap showing expression Z-scores for genes involved in insulin processing. **e**, Immunoblot showing insulin and proinsulin peptide levels in islet lysate from three *Dpy30*-WT mice (10 μg/lane) and three *Dpy30*-KO mice (34 μg/lane). **f**, Quantification of the ratio proinsulin/ insulin band intensity shown in panel (e). Data presented as mean ± SD with values from individual mice (n = 3).

4.8 H3K4 methylation contributes to β-cell glucose-responsiveness via regulation of calcium signaling genes

I confirmed that glucose-stimulated insulin secretion is deficient in *ex vivo* islets (Fig. 4.13a). Notably, accounting for the reduced insulin content of *Dpy30*-KO islets did not fully explain the impairment of insulin secretion. The persistent impairment was specific to high glucose stimulation, whereas secretion under basal conditions, stimulated with the membrane-permeable Krebs cycle metabolite analogue dimethyl α-ketoglutarate, or depolarizing cells directly using KCl, showed no difference between *Dpy30*-WT and -KO insulin secretion (after controlling for insulin content) (Fig. 4.13a), suggesting impaired coupling of glucose sensing to insulin secretion. Gene set enrichment analysis suggests a defect in calcium-mediated signaling, whereas expression of genes involved in oxidative phosphorylation is elevated in *Dpy30*-KO cells (Fig. 4.13b). To confirm a functional defect in stimulus-secretion coupling I measured cytosolic Ca^{2+} during islet stimulation. High glucose-stimulated Ca^{2+} influx showed a delayed response and lower magnitude in *Dpy30*-KO islets, whereas KCl infusion stimulated normal Ca^{2+} influx (Fig. 4.13c), consistent with a defect in metabolic coupling involving glucose catabolism and/or Ca^{2+} signaling upstream of membrane depolarization. Respirometry of dispersed islet cells suggests that *Dpy30*-KO cells do not display defects in mitochondrial glucose catabolism (Fig. 4.13d-e). Consistent with the enrichment of oxidative phosphorylation genes, *Dpy30*-KO cells display elevated maximal respiration capacity (Fig. 4.13d-e), as well as expansion of mitochondria area (Fig. 4.13f), which could indicate compensation for higher metabolic demand and/or lower metabolic efficiency. Overall, reduction to H3K4me3 in β-cells impairs insulin release at multiple steps, including *Ins1* and *Ins2* transcription, preproinsulin processing, and glucose-stimulated insulin secretion upstream of mitochondrial respiration.

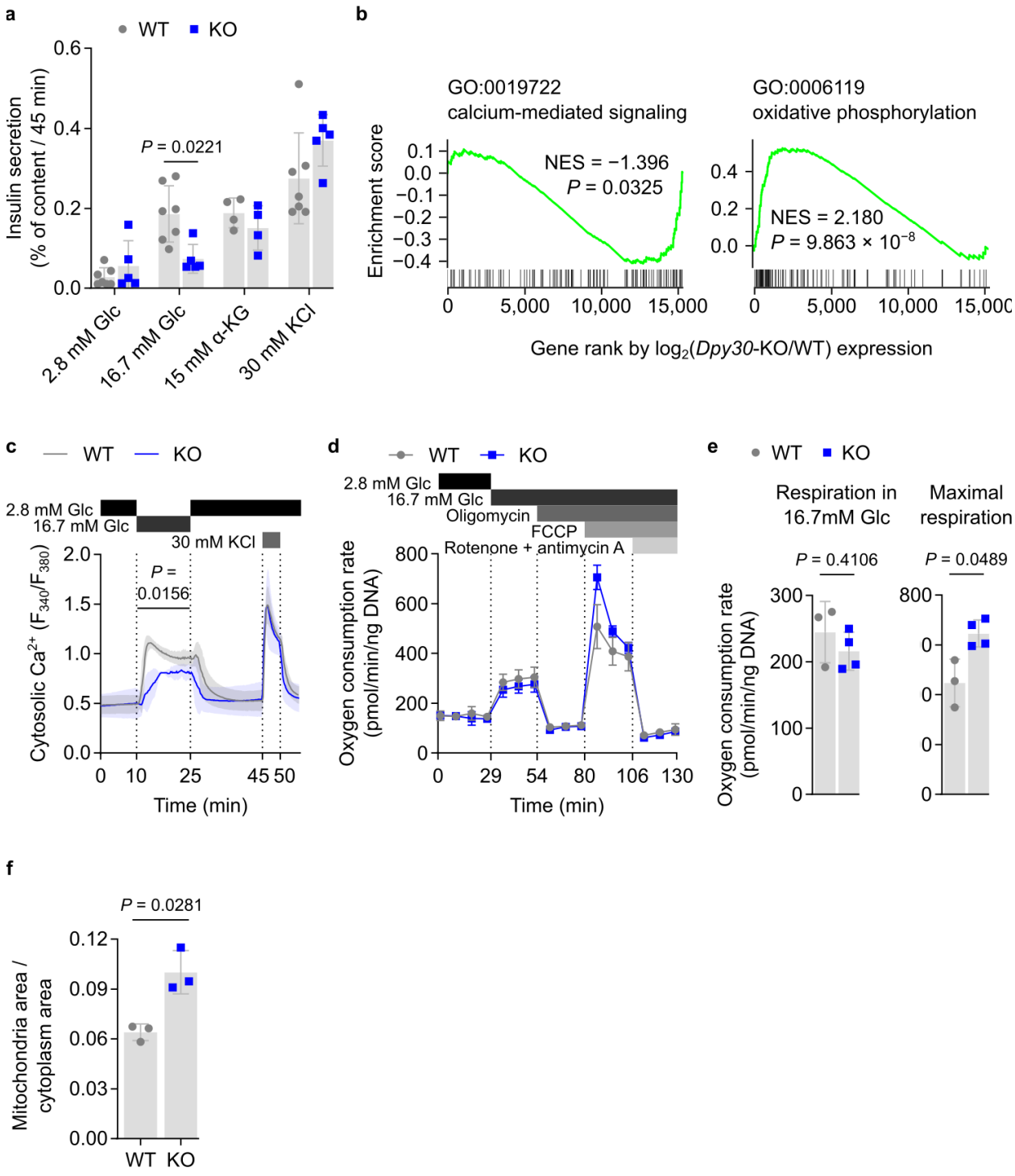

Figure 4.13 | Impaired glucose-stimulated calcium signaling, but not mitochondrial respiration, reduces insulin secretion from *Dpy30*-KO islets. a, Insulin secretion from *Dpy30*-WT and -KO islets during static stimulation, normalized to islet insulin content. *P*-values calculated by multiple two-tailed t-tests with Welch's and

Benjamini-Hochberg corrections; only values < 0.05 are shown. **b**, Gene set enrichment analysis of calcium-mediated signaling and oxidative phosphorylation genes in *Dpy30*-KO versus -WT cells. *P*-values calculated by permutation test with Benjamini-Hochberg correction. **c**, Cytosolic Ca^{2+} concentration in islets from *Dpy30*-WT and -KO mice during 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. **d**, Oxygen consumption rate of *Dpy30*-WT and -KO dispersed islets during treatment with the indicated compounds. n = 3 *Dpy30*-WT, 4 *Dpy30*-KO mice. **e**, Respiration of *Dpy30*-WT and -KO islet cells in 16.7 mM glucose and their maximal respiration capacity, inferred from the data shown in panel (d). *P*-values calculated by two-tailed t-test with Welch's correction (n = 3 WT, 4 KO). **f**, Bar graph showing the fraction of cytoplasm area occupied by mitochondria in TEM images of *Dpy30*-WT and -KO β-cells. *P*-value calculated using two-tailed t-test with Welch's correction. All panels show mean ± SD. Glc: Glucose. α-KG: dimethyl α-ketoglutarate. FCCP: Carbonyl cyanide 4-(trifluoromethoxy)phenylhydrazone. NES: Normalized enrichment score.

As a final functional test of *Dpy30*-KO cells, I measured their propensity to replicate in response to short-term high glucose stimulation because reduced H3K4me3 has been proposed to drive greater β-cell replication [282–284]. I quantified replication in *Dpy30*-WT and -KO cells by measuring EdU incorporation of islet cells cultured *ex vivo* in 5.5 mM versus 16.7 mM glucose for 48 hours. While 16.7 mM glucose media induced a ~22-fold increase of replication in *Dpy30*-WT β-cells, *Dpy30*-KO β-cells were induced less than 2-fold, and not statistically significantly differently than 5.5 mM (Fig. 4.14). The basal replication rate appeared similar to *Dpy30*-WT cells and was non-zero, indicating that KO cells are still capable of replication. Therefore, reduction to H3K4me3 leads to loss of acute glucose-stimulated insulin secretion and compensatory replication of β-cells.

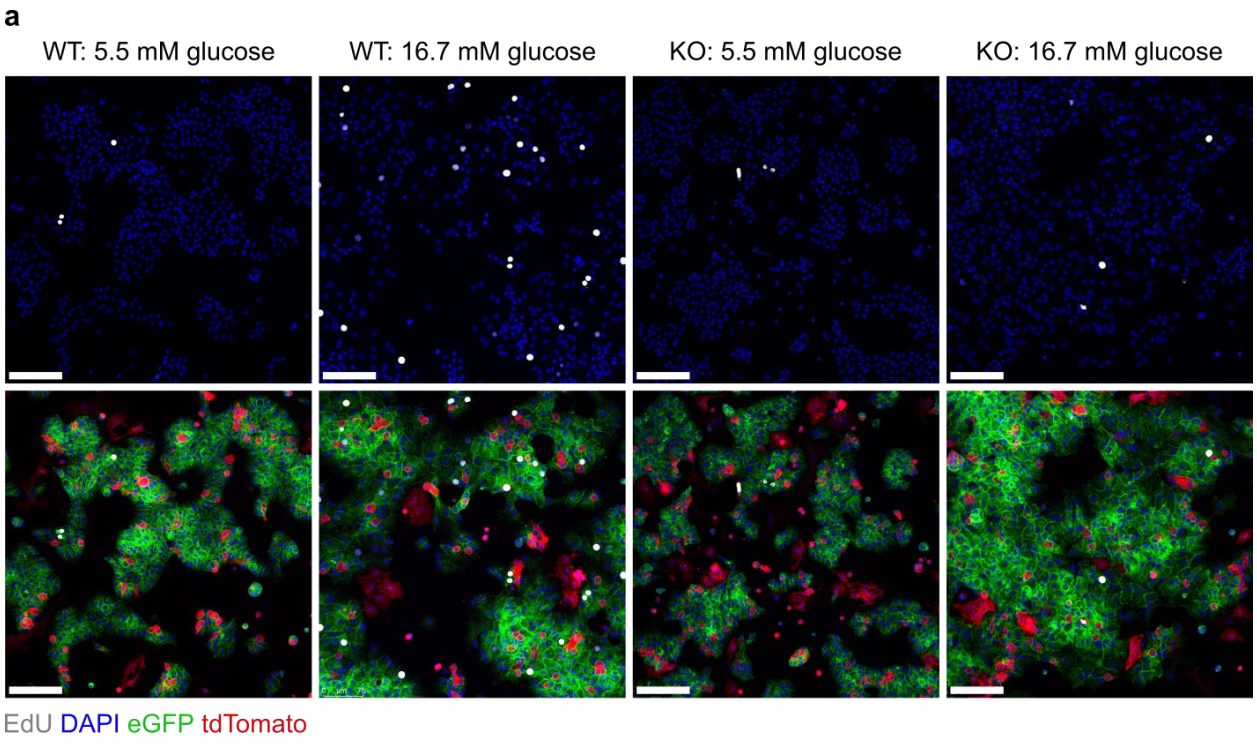

Figure 4.14 | H3K4 methylation is required for high glucose-stimulated β-cell proliferation. a, Representative immunofluorescence images of EdU incorporation in *Dpy30*-WT and -KO mouse islets that were dispersed and cultured in 5.5 mM or 16.7 mM glucose for 48-hours, as indicated, showing EdU (white) and DAPI (white) staining and eGFP (green) and tdTomato (red) endogenous fluorescence. Scale bars: 100 μm. **b**, Bar graph showing quantification of EdU+ eGFP+ double positive cells as a fraction of total eGFP+ cells (n = 3 per treatment and genotype). *P*-values calculated using two-way ANOVA with Tukey's correction. ns: not significant ($P > 0.05$).

4.9 Discussion

In this chapter, I show that H3K4me3 peak organization in mature β-cell chromatin is linked to transcription of critical β-cell genes and to their dysregulation in T2D. Transcriptome remodeling in diabetic *Lepr*$^{db/db}$ mouse islets, a model of hyperphagia-induced T2D, is encoded by changes in H3K4me3 peak breadth. Using a genetic model to deplete H3K4 methylation from mature β-cells, I show that H3K4me3 regulates the expression of genes necessary for mature β-cell function and that are dysregulated in *Lepr*$^{db/db}$ islets. Depletion of H3K4 methylation is deleterious to multiple aspects of insulin production and glucose-responsiveness in β-cells and causes *Dpy30*-KO mice to develop hyperglycemia, impaired glucose clearance, and reduced insulin secretion.

This work adds to existing evidence that maturation of β-cells and the maintenance of mature β-cell function relies on epigenetic control of gene expression [426]. While epigenetic features of fully differentiated cells have limited flexibility, their genome-wide enrichment and organization can be altered by environmental factors including nutrient stress [427]. In particular, the availability of metabolites required by histone methyltransferases (e.g., SAM) or demethylases (e.g., α-ketoglutarate) may shift, owing to metabolic rewiring or changes in nutrition [412]. In the case of *Lepr*$^{db/db}$ mice, however, my observation that global enrichment of H3K4me3 is not altered *en masse* implies that changes in H3K4me3 are not caused by a deficiency or surfeit of metabolites required for methylation or demethylation reactions. Rather, changes to H3K4me3 are locus-specific and linked to transcriptional up- or downregulation (Fig. 4.15), suggesting that targeted recruitment of histone methyltransferase or demethylase enzymes to particular genes drives changes in H3K4me3 between *Lepr*$^{+/+}$ and *Lepr*$^{db/db}$ mice. Two potential candidate

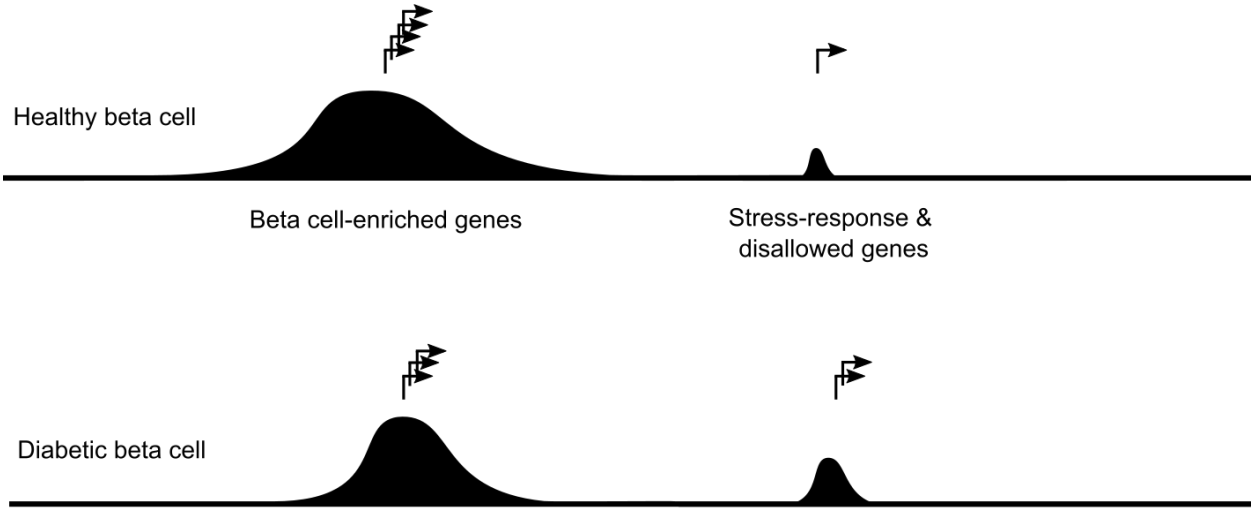

Figure 4.15 | Model for H3K4me3 in healthy and diabetic mouse β-cells. Contraction of H3K4me3 at broad-peaked, β-cell enriched gene promoters, and expansion of H3K4me3 at narrow-peaked, stress-response and disallowed gene promoters, regulates transcriptome remodeling in diabetic mouse β-cells.

transcription factors, MAFA and PAX6, are known interactors of TrxG complexes [278,281], are inactivated in *Lepr*$^{db/db}$ islets [123], and their deletion leads to a similar phenotype as *Dpy30*-KO [114,118]. Therefore, it is possible that TrxG-dependent H3K4me3 contributes to gene expression mediated by these transcription factors, and that H3K4me3 is reduced at targeted loci owing to reduced recruitment of TrxG complexes by these and other transcription factors. More work will be required to identify the mediators of H3K4me3 peak size in β-cells.

Insulin peptide content is reduced 3.357-fold in *Dpy30*-KO islets. Surprisingly, the reduction of insulin could not be fully explained by reduction to *Ins1* and *Ins2* mRNA expression, which were reduced only 1.152 and 1.234-fold, respectively. They remained the two most highly expressed genes in *Dpy30*-KO cells despite suffering complete loss of H3K4me3 from their promoters, indicating that H3K4me3 is dispensable for the expression of insulin mRNAs. Further, the

outstanding deficit in insulin content could not be attributed to impairment of proinsulin-to-insulin processing, since *Pcsk1*, *Pcsk2*, and *Cpe* are upregulated, and proinsulin did not buildup, in *Dpy30*-KO cells. Rather, H3K4 methylation is essential for the expression of signal peptidase complex subunits *Spcs3* and *Nnat*, loss of which is known to impair preproinsulin processing and redirect it for degradation [77]. H3K4me3-regulated gene expression therefore contributes to mRNA transcription and indirectly to peptide processing of preproinsulin genes.

The insulin deficit in *Dpy30*-KO mice is compounded by an impairment in high glucose-stimulated insulin secretion. Mitochondrial dysfunction does not appear to be responsible for the impairment, since high glucose-stimulated oxygen consumption is unchanged, genes involved in oxidative phosphorylation are upregulated, and mitochondrial area and oxidative capacity are increased in *Dpy30*-KO β-cells. *Dpy30*-KO islets also retained a similar capacity as WT islets to secrete insulin during KCl-mediated membrane depolarization, suggesting insulin exocytosis machinery remained functional. Instead, this impairment is at least partially caused by downregulation of calcium signaling genes leading to reduction in the rate and magnitude of calcium influx during high glucose challenge.

Short term high glucose exposure induces mouse β-cells replication [428]. I show that *Dpy30*-KO blocks high glucose-induced β-cell replication. Notably, replication in basal glucose concentration is non-zero and comparable between *Dpy30*-KO and -WT β-cells, indicating that *Dpy30*-KO cells retain the ability to replicate. It has been noted previously that knockdown or knockout of DPY30 causes a reduction in the proliferation rate in hematopoietic progenitor cells [429], fibroblasts [430], and pancreas progenitors [351], but that proliferation of mouse embryonic stem

cells [431], differentiating hematopoietic progenitor cells [429], K562 leukemia cells [429], or pancreatic acinar cells [351] is not inhibited by DPY30 knockdown or knockout. This shows that, rather than a direct involvement in cell replication, DPY30 likely influences replication through context-specific regulation of other genes involved in DNA replication and cell cycle progression [429]. This suggests that DPY30-dependent H3K4me3 in mature β-cells is necessary for the induction of replication in response to high glucose stimulus. Overall, evidence in this chapter demonstrates a requirement for H3K4me3 in the maintenance of glucose responsiveness by β-cells.

Chapter 5: Dietary folic acid restriction in mature mice does not reduce methylation potential or H3K4me3 in pancreatic islets

5.1 Rationale

Among enzymatic cofactors, the intracellular abundance of *S*-adenosylmethionine (SAM) is second only to ATP [286]. Generated by the activities of the folate and methionine cycles, SAM is an essential donor of methyl groups for myriad processes including nucleic acid and amino acid synthesis, and modifications of proteins, nucleic acids, and lipids. SAM is the source of methyl groups for histone methylation [432].

In contrast to ATP whose intracellular concentration far exceeds the K_M of most enzymes that use it, moderate fluctuations to SAM content can have a meaningful impact on enzymatic activity [287]. Blood SAM levels are variable in human populations. In those with T2D, SAM levels are inversely correlated with severity of the disease, such that the ratio of SAM versus its demethylated form, *S*-adenosylhomocysteine (SAH) (SAM/SAH, henceforth, "methylation potential"), is almost five times greater in healthy individuals versus the most brittle diabetics [433]. Further, a sulfur-poor diet, which redirects SAM toward cysteine biosynthesis, may increase susceptibility to T2D in humans and mice [434]. Reduction to methylation potential has been noted in peripheral nerves of *Lepr*$^{db/db}$ mice [435], as has a reduction in circulating methionine, the immediate precursor of SAM [436]. Notably, methionine restriction in healthy mice is sufficient to reduce methylation potential and H3K4me3 levels in liver [297,418]. When I began work summarized in this chapter, I had preliminary evidence that H3K4me3 is reduced in islets during diabetes, which I later disproved (Fig. 4.6a-c). At that time reduction of islet methylation

potential was therefore a good candidate to describe why H3K4me3 levels were reduced. I therefore sought to test whether islet methylation potential is sensitive to systemic methylation potential, and whether H3K4me3 in islets is sensitive to islet methylation potential.

To experimentally reduce systemic methylation potential, dietary restriction of essential methyl donors such as methionine or folic acid is a convenient method [297,437]. Since methionine restriction causes a potent amino acid response [438] and liver damage [439], I chose to reduce methylation potential using a folic acid-restricted diet.

Folic acid, and the folate cycle, facilitate regeneration of SAM (Fig. 1.4). A deficit of folic acid is therefore expected to reduce methionine and SAM and increase homocysteine and SAH, itself a competitive inhibitor of SAM in methyltransferase reactions [432]. Studies in rats show that dietary folic acid restriction is sufficient to reduce methylation potential in liver [437] and in whole pancreas [440,441], and folic acid restriction in mice reduces DNA methylation in brain, kidney, and liver tissues [442]. Reduced DNA methylation in liver of people with T2D may be explained by reduced folate levels [443]. Other studies have reported that changes in methylation potential induce global changes in the levels of histone modifications [306,317,418,444–448] but a connection between folate and H3K4me3 has not been reported, nor have islets been studied under conditions of low methylation potential. I hypothesized that dietary folic acid restriction in mature mice leads to reduced methylation potential and H3K4me3 in islets.

5.2 Dietary folic acid restriction causes excess weight gain and impairs glucose homeostasis in male mice

At 8-weeks-old, male and female mice began diets of either normal folic acid ("NF", 2 mg folic acid/3987 kcal, equivalent to the American Institute for Nutrition's recommended folic acid intake of 2 mg/kg diet for rodents [449]) or folic acid restricted ("FR", 0.2 mg folic acid/3987 kcal, 10% of recommended dose) diets. Diets were otherwise identical and included 100 mg/kg succinylsulfathiazole to inhibit folate synthesis by gut bacteria. Complete diet compositions are listed in Appendix A. Mouse body mass and unfasted blood glucose concentration were measured every four weeks. From week-16 on diet, male FR mice had greater body mass compared to NF male mice (Fig. 5.1a). Dietary folic acid dose had no effect on female body mass or unfasted blood glucose concentration in male or female mice (Fig. 5.1c-d).

Mice were subjected to glucose tolerance tests every four weeks. Male FR mice displayed reduced glucose tolerance from 16-weeks on diet compared to male NF mice (Fig 5.2a-b). Female FR mice displayed a trend toward reduced glucose tolerance at the end of the study period, which was not statistically significant after adjusting for multiple comparisons (P = 0.1023, two-way ANOVA with Benjamini-Hochberg correction of AUC at 24-weeks of diet) (Fig. 5.2c-d). In insulin tolerance tests at 22-weeks of diet, male FR mice displayed reduced insulin sensitivity compared to NF mice, whereas female FR mice showed no effect of folic acid (Fig. 5.2e-f). Therefore, dietary folic acid restriction starting from 8-weeks-old disrupts glucose homeostasis in male, but not female, mice.

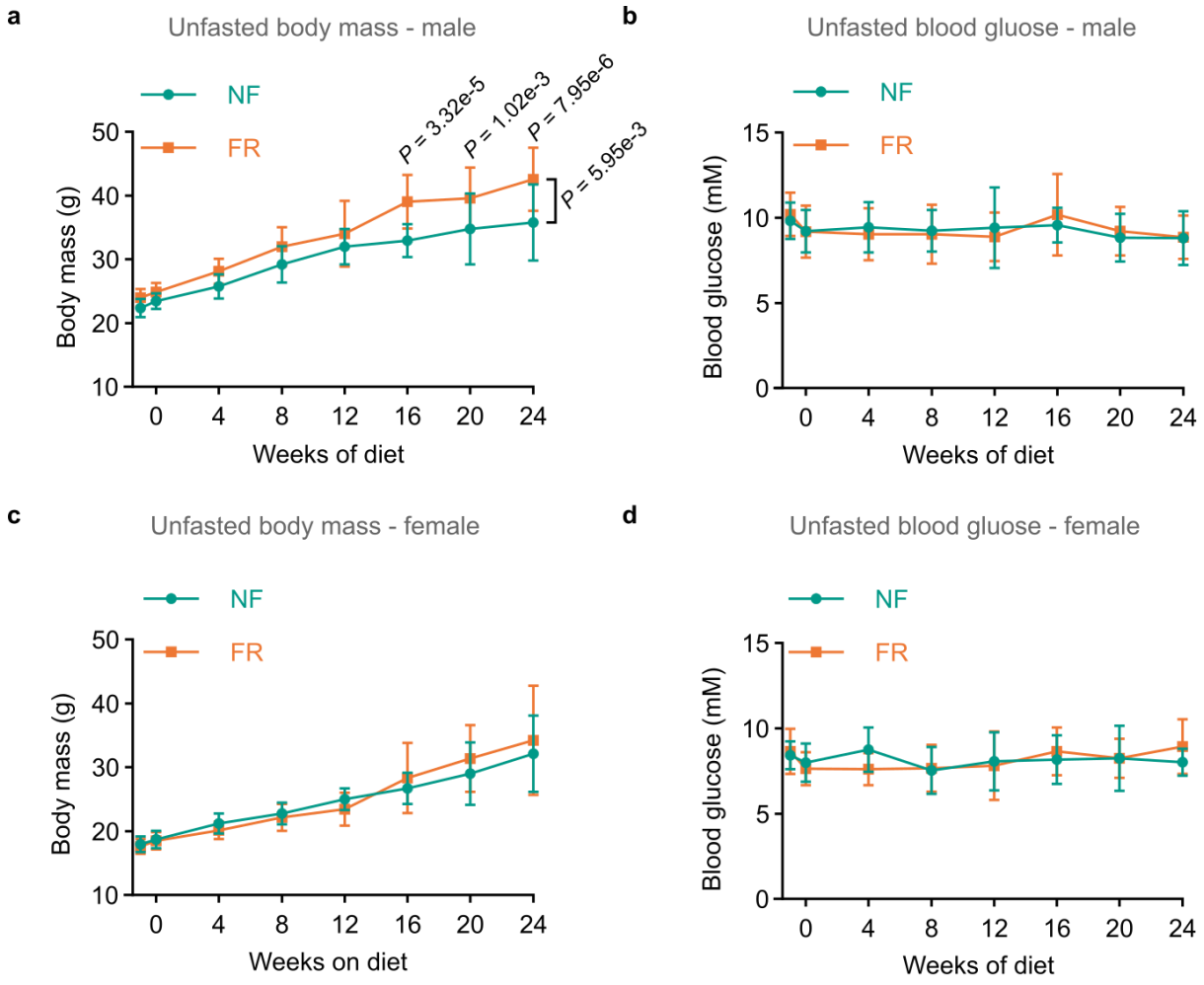

Figure 5.1 | Body mass and unfasted blood glucose concentration of NF and FR mice. a-b, Unfasted body mass (a) and blood glucose concentration (b) of male mice at four-week intervals of diet (n = 14 NF, 13 FR). **c-d**, Same as (a-b) for female mice (n = 13 NF, 12 FR). *P*-values calculated with a mixed-effect model with Benjamini-Hochberg correction; only values < 0.05 are shown.

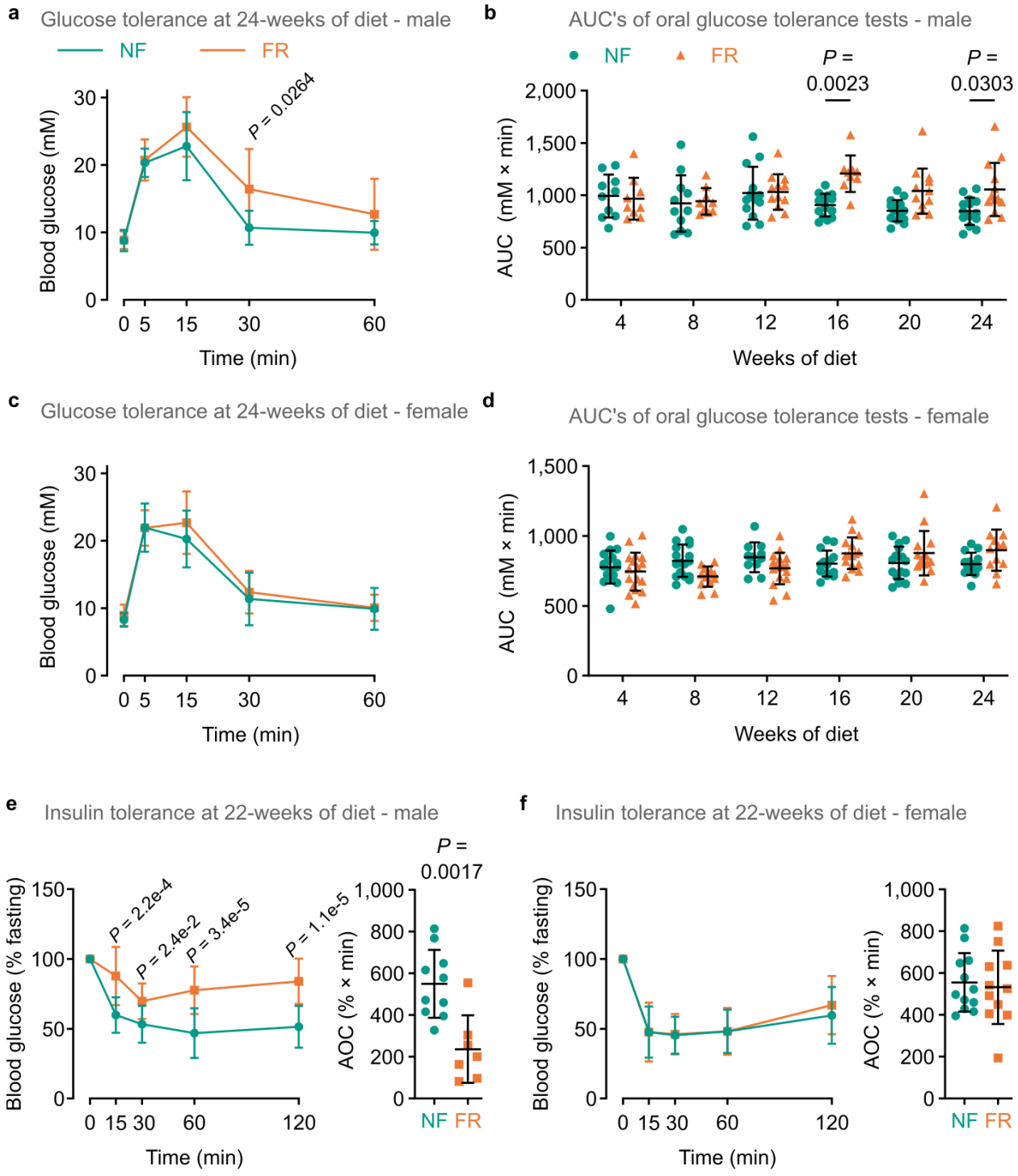

Figure 5.2 | Dietary folic acid restriction impairs glucose homeostasis in male, but not female, mice. a, Blood glucose concentration during OGTTs of male mice after 24-weeks of diet. n = 14 NF, 13 FR. *P*-values calculated using two-way ANOVA with Benjamini-Hochberg correction; only values < 0.05 are shown. **b**, AUC of OGTTs of

male mice at 4-week intervals after starting diets. n = 10-15. *P*-values calculated with a mixed-effect model with Benjamini-Hochberg correction; only values < 0.05 are shown. **c-d**, Same as panels (a-b) for female mice. n = 11-17. **e**, Blood glucose concentration expressed as a percent of fasting concentration after IP insulin injection showing plasma glucose over time (left) and area over the curve (AOC) of the ITT (right). n = 10 NF, 7 FR. *P*-values calculated using two-way ANOVA with Benjamini-Hochberg correction (left) or t-test with Welch's correction (right). **f**, Same as panel (e) for female mice. n = 12 NF, 11 FR.

5.3 Folic acid increases *in vitro* high glucose-stimulated insulin secretion

Since the impairment of insulin tolerance indicates that dietary folic acid restriction likely affects the function of insulin target tissues, I tested whether folic acid restriction can directly impact islets insulin secretion. First, examination of healthy mouse β-cell RNA-seq data introduced in Chapter 3 confirmed expression of genes in the folate and methionine cycles, at least at the RNA level (Fig. 5.3). This provides support for the ability of β-cells to process folates. I next performed *in vitro* insulin secretion assays. Islets from male chow-fed mice were cultured in RPMI 1640 medium with either 1 or 0 mg/L folic acid for 48-hrs before static insulin secretion assay in folic acid-free buffer. *In vitro* folic acid restricted culture reduced high glucose-stimulated insulin secretion (Fig. 5.4a), confirming that folic acid contributes to islet function. Similar results were obtained by recovering chow-fed mouse islets overnight in RPMI 1640 medium with 1 mg/L folic acid, and then performing the static insulin secretion assay in buffer containing either 1 or 0 mg/L folic acid (Fig. 5.4b) – suggesting that folic acid signaling is directly involved in high glucose-stimulated insulin secretion by mouse islets.

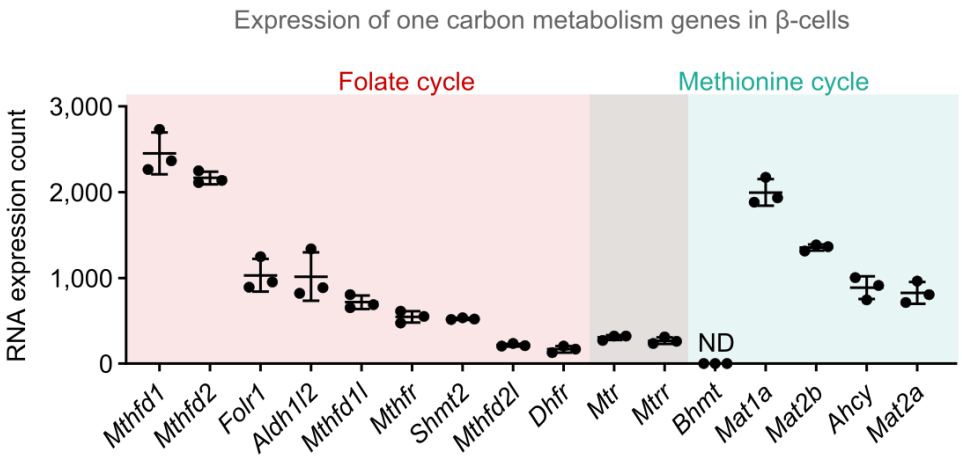

Figure 5.3 | Folate and methionine cycle genes are expressed in mature mouse β-cells. Expression of genes encoding proteins involved in the folate and methionine cycles, in bulk RNA-seq of β-cells from 10-week-old male *Pdx1-CreER*$^{Tg/0}$ *mTmG*$^{+/-}$ mice. Note that the enzyme encoded by *Bhmt* provides an alternative (folate-independent) route to regenerate methionine from homocysteine. ND: not detected.

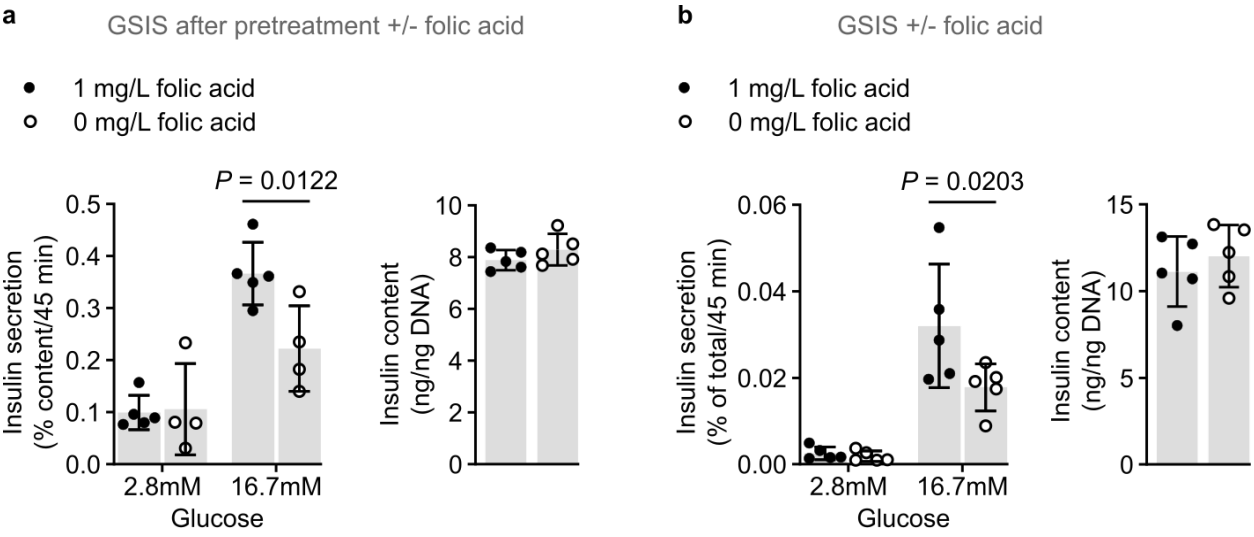

Figure 5.4 | Folic acid increases *in vitro* high glucose stimulated insulin secretion. a, Insulin secretion and content from chow-fed male mouse islets incubated *ex vivo* for 48-hrs with RPMI media containing 1 or 0 mg/L folic acid, then stimulated with glucose solutions without folic acid. **b**, as (a) but islets were recovered for 16-hrs in RPMI with 1 mg/L folic acid, then stimulated with glucose solutions containing 1 or 0 mg/mL folic acid. *P*-values calculated by two-way ANOVA with Benjamini-Hochberg correction. Displaying mean ± SD; n = 4 to 5.

5.4 Acute folate restriction does not cause transcriptome remodeling in mouse islets

After confirming that islet function is acutely sensitive to folic acid availability, I wondered if folic acid restriction elicits an acute transcriptional response in islet cells. I performed mRNA-seq in islets from chow-fed male mice that were cultured for 24-hours *ex vivo* in media containing 1 or 0 mg/L folic acid (n = 3). Only one gene—*Lrrk2*, a cytosolic kinase and GTPase [450] of unknown significance in islets—was identified as differentially expressed (Fig. 5.5). I concluded that gene expression is not acutely sensitive to folic acid availability in islets and returned to examining the *in vivo* dietary model.

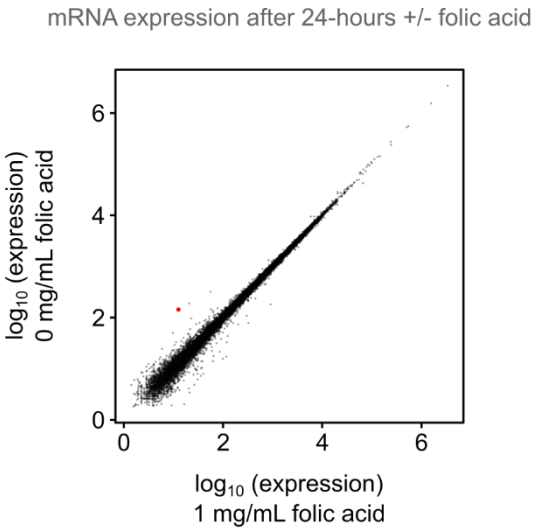

Figure 5.5 | *Ex vivo* folic acid restriction for 24-hours does not cause transcriptome remodeling in mouse islets. Differentially expressed genes ($P < 0.01$, Wald test with Benjamini Hochberg correction) are red.

5.5 Dietary folic acid restriction did not alter relative fat, liver, or pancreas mass

Organ mass data were collected after 24-weeks of diet. As shown in Fig. 5.6, gonadal and inguinal fat depots, liver, and pancreas masses were not altered by folic acid restriction in male or female mice.

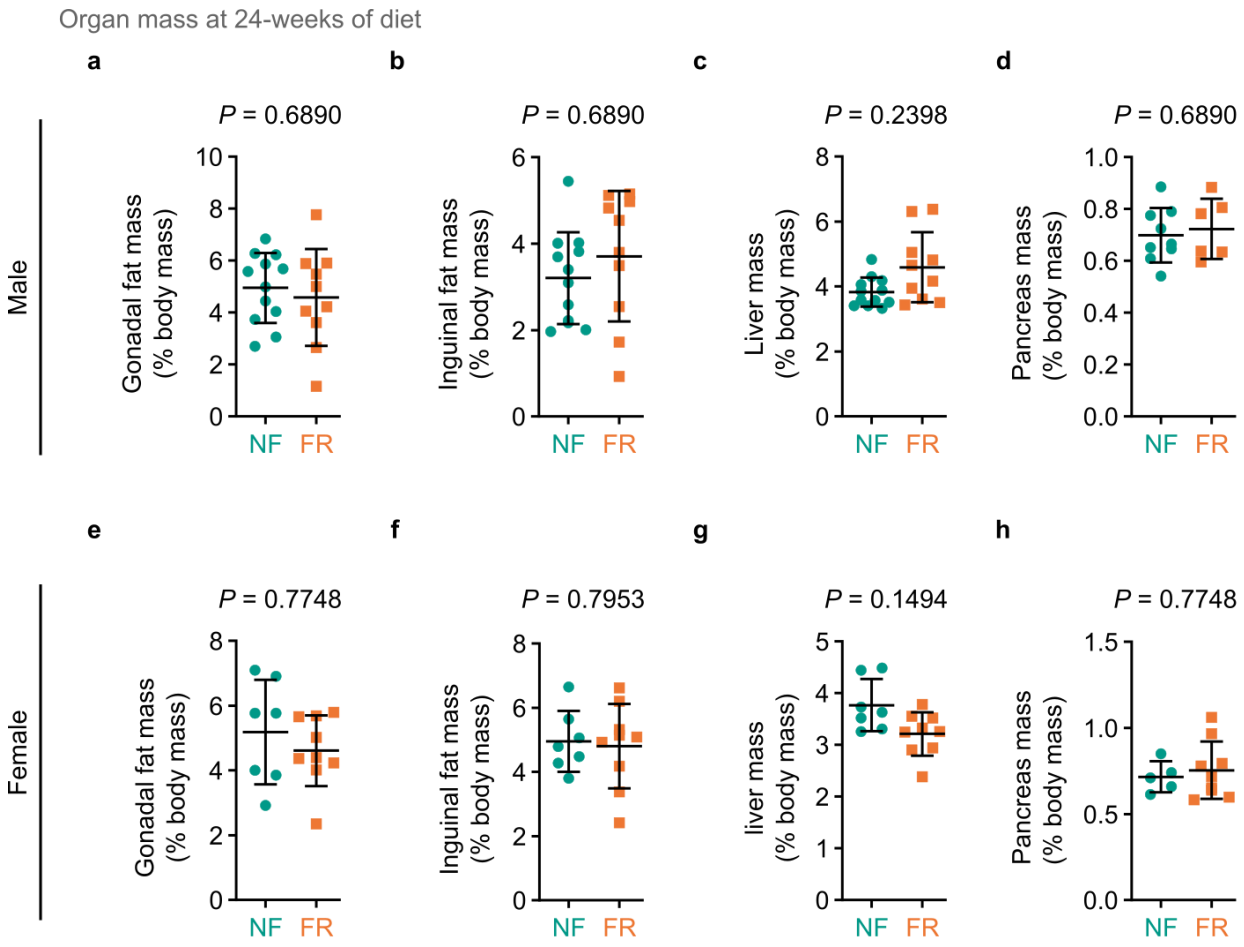

Figure 5.6 | FR diet does not alter relative fat, liver, or pancreas mass. a-d, Gonadal fat (a), inguinal fat (b), liver (c), and pancreas (d) mass relative to body mass in male mice after 24-weeks of NF or FR diets. **e-f**, Same as (a-d) for female mice. *P*-values calculated using multiple two-tailed t-tests with Welch's and Benjamini-Hochberg corrections. Showing mean ± SD with values from individual mice; n = 5-12.

5.6 Dietary folic acid restriction did not reduce methylation potential in blood or islets

As expected, the FR diet led to a lower plasma folate concentration by 24-weeks of diet (Fig. 5.7a). The folate cycle facilitates regeneration of SAM via methylation of homocysteine into methionine, the immediate precursor of SAM [432,451]. Homocysteine can alternatively be processed into the transsulfuration pathway to exit the methionine cycle. For this reason, circulating biomarkers of the methionine cycle and transsulfuration pathway were quantified. Plasma homocysteine concentration was elevated in FR mice, supporting an impairment of homocysteine remethylation (Fig. 5.7f). Despite this, no differences in the circulating concentration of the other methionine cycle metabolites methionine, SAM, or SAH were observed (Fig. 5.7b-e). Cystathionine and cysteine, which are intermediates of transsulfuration, were also not different between NF and FR mice (Fig. 5.7g, h). Two other dietary methyl donors, betaine and choline, were also not different. In islets from NF and FR mice, no difference was observed in any measured one carbon metabolite including SAM or SAH (Fig. 5.8). Altogether, plasma homocysteine is the only one-carbon metabolite that was disturbed in the FR diet. Therefore, dietary folic acid restriction to 10% of the recommended dose was not sufficient to reduce methylation potential in circulation or in islets.

5.7 Dietary folic acid restriction did not reduce H3K4me3 in islets

Histone methylation reactions source methyl groups from SAM, and SAH inhibits methyltransferase reactions [286]. Since the FR diet did not lead to a decrease in SAM or an increase in SAH in islets, it is unsurprising that islet H3K4me3 levels were also similar between NF and FR mice (Fig. 5.9).

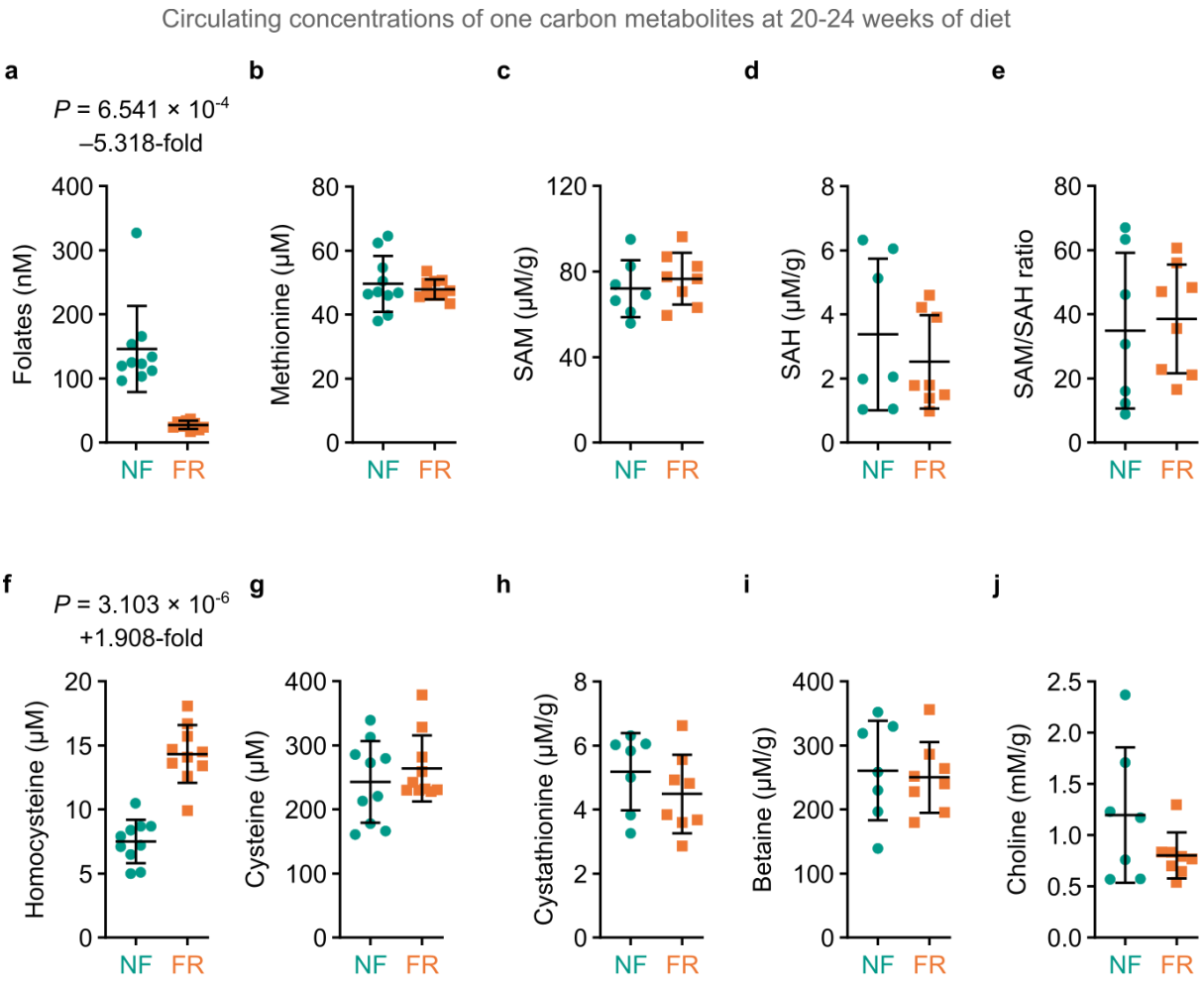

Figure 5.7 | FR diet has minimal impact on circulating one-carbon metabolite concentrations. a-j, Blood concentration of total folates (a), methionine (b), SAM (c), SAH (d), SAM/SAH ratio (e), total homocysteine (f), cysteine (g), cystathionine (h), betaine (i), and choline (j) in female mice after 20-24 weeks of diet. a, b, f, and g were measured in plasma fraction, c, d, e, h, i, and j were measured in blood cell fraction. *P*-values calculated using multiple two-tailed t-tests with Welch's and Benjamini-Hochberg correction; only values < 0.05 are displayed. Showing mean ± SD and values from individual mice; n = 7-10.

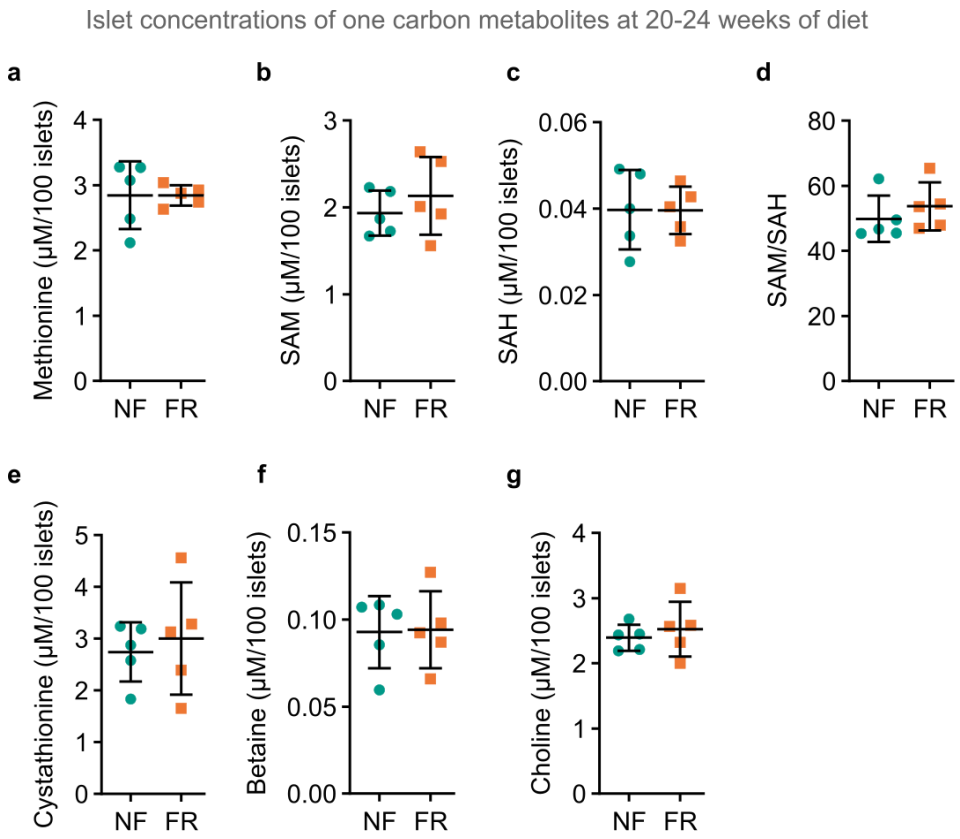

Figure 5.8 | FR diet does not reduce methylation potential in islets. a-g, Islet lysate concentration of methionine (a), SAM (b), SAH (c), SAM/SAH ratio (d), cystathionine (e), betaine (f), and choline (g) in female mice after 20-24 weeks of diet. Showing mean ± SD and values form individual mice; n = 5.

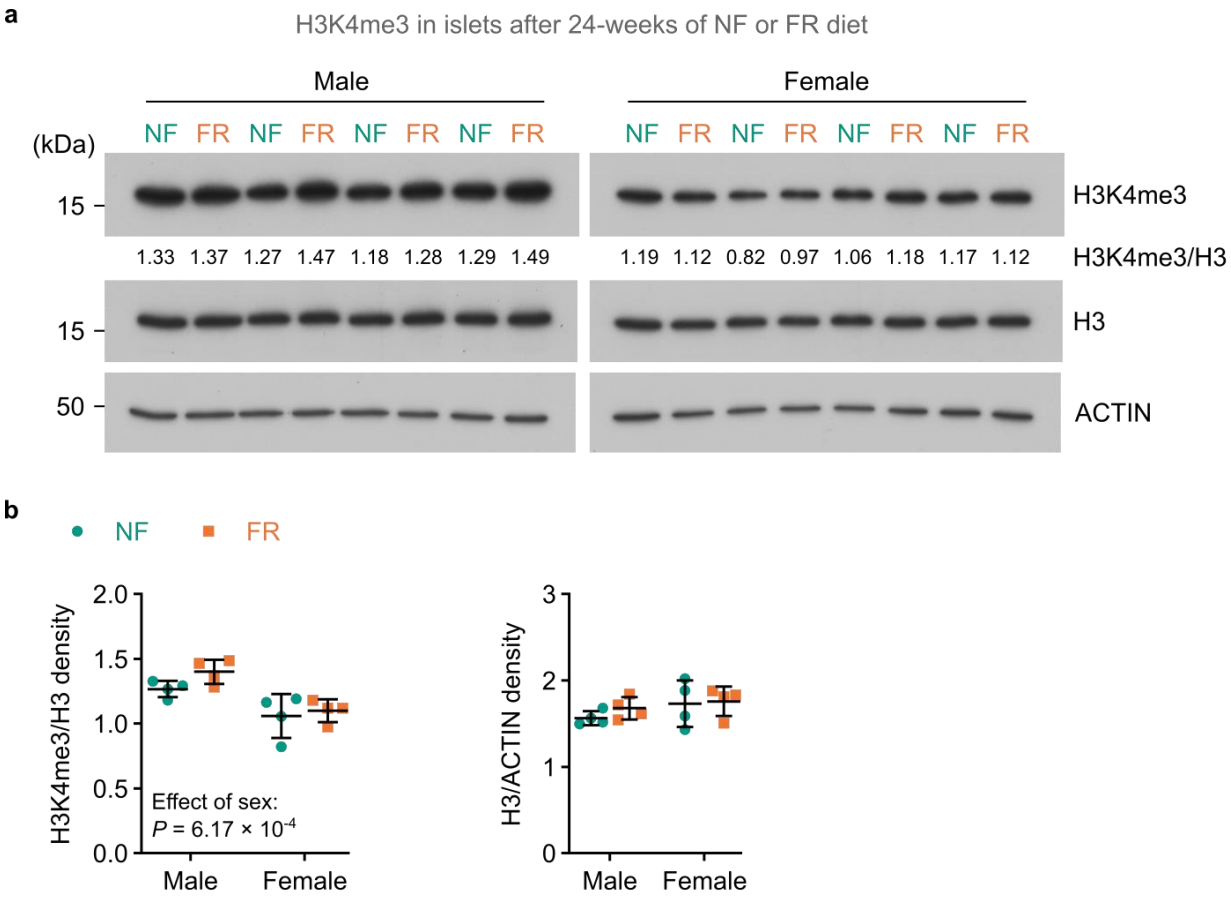

Figure 5.9 | FR diet does not alter H3K4me3 levels in islets. a, Immunoblots of islet lysates collected after 24-weeks of diet showing H3K4me3, total H3, and ACTIN. Male and female samples were blotted on the same membrane but are shown separately for clarity. **b**, Comparison of relative band intensities for data in panel (a). *P*-values calculated using two-way ANOVA with Benjamini-Hochberg correction; only values < 0.05 are shown. Showing mean ± SD with individual values; n = 4.

5.8 Discussion

In this chapter, dietary folic acid restriction was used to reduce methylation potential in mature mice in order to test the relationship between systemic methylation potential and H3K4me3 in islets. Contrary to my hypothesis, the FR diet failed to reduce methylation potential in islets or in circulation.

Compared to the NF diet, the FR diet was formulated with a 10-fold lower concentration of folic acid and led to a 5.3-fold reduction in circulating folate after 24-weeks of feeding. A 1.9-fold increase in circulating homocysteine shows that folic acid restriction impaired the regeneration of methionine from homocysteine, as intended, since this is the reaction that links the folate cycle with the methionine cycle. This was not associated with a decrease in circulating methionine or SAM, however. This finding suggests that dietary methionine was sufficiently robust to maintain normal circulating methionine and SAM levels despite impaired recycling from homocysteine in the FR diet. The rise in homocysteine was also not sufficient to cause a rise in circulating SAH in this model.

The finding that *in vitro* folate restriction impaired insulin secretion from islets suggests that islets use folate. The reduction of circulating folate did not lead to a change in SAM or SAH concentration in islets, however. It is possible that a more extreme restriction of folic acid in the FR diet could have decreased methylation potential—FR diets used in Balaghi *et al.*'s studies lacked any folic acid and led to reduced methylation potential in liver and in whole pancreas of rats [437,440,441]. On the other hand, those reports may reflect species differences or a higher sensitivity to dietary folic acid by those tissues compared to islet or blood cells. Indeed, results

from Zhao *et al.* (2018b) suggest that removing the remaining 0.2 mg/3987 kcal of folic acid in my FR diet may not have changed my results—they reported a 6.3-fold decrease in serum folates and a 1.5-fold increase in serum homocysteine in mice fed a 0 mg/kg folic acid diet for 25-weeks (they did not measure methionine, SAM, or SAH) [330]. These values are not substantially different to the 5.3-fold decrease in plasma folates and 1.9-fold increase in plasma homocysteine at 24-weeks of FR diet shown above. Therefore, future work that aims to reduce methylation potential in islets should take a more direct approach, such as: methionine restriction; islet-specific knockout or knockdown of MAT genes, which convert methionine to SAM; or, islet-specific overexpression of *Gnmt*, which quenches SAM by transferring the methyl moiety to free serine. The first two approaches have been shown to reduce SAM and H3K4me3 in other contexts, but have not been studied in islets, and *Gnmt* overexpression depletes SAM and increases SAH in liver cells but has not been studied with regard to histone methylation [297,317,452,453].

The folate-restricted diet caused excess weight gain and disrupted glucose homeostasis in male mice. This finding is consistent with other studies that have used comparable folate-restricted diet regimes in mice [329,330]. Low serum folate is also associated with insulin and glucose intolerance in humans [454,455] and folic acid supplementation improves glycemic control in patients with T2D [456,457]. The mechanisms by which folate improves glucose and insulin responsiveness are unclear, but several possible explanations exist. First, folate lowers homocysteine, a source of vascular and organ oxidative stress [458,459]. Excess oxidative stress is associated with reduced insulin sensitivity in insulin target tissues and reduced insulin secretion by β-cells [460]. Homocysteine increases glucose output and reduces insulin's effect on glucose output in hepatocytes *in vitro* [461]. In the transformed rat β-cell line RINm5F, excluding folic acid

from culture medium causes accumulation of homocysteine, increases oxidative stress, and lowers insulin production [337]. However, the effects of homocysteine on primary islets have not been studied. Second, folate is an essential cofactor for *de novo* purine and thymidylate synthesis. In replicating mammalian cells in culture, the largest demand for folate is in purine synthesis [286], and, the central role for folate in DNA and RNA subunit synthesis is at least partially responsible for the effectiveness of antifolates—competitive structural analogues of folate derivatives—as anti-cancer therapies [462]. The purine nucleotides ATP and GTP are also messengers for insulin secretion and signaling [463,464]. However, to what extent metabolic outcomes of folate cycle dysregulation are caused by deficits of purine and thymidylate synthesis requires further study. Third, low folate may reduce methylation status of insulin target tissues. While I did not find evidence for a reduction in methylation potential in islets, other tissues may be affected. For example, mice fed a low folate diet displayed lower hepatic SAM/SAH ratios and DNA methylation than mice fed a control diet [465]. Changes to DNA methylation, leading to changes in gene expression, are implicated in the pathologies of folate deficiencies [466]. Liver and kidney are the major sites of homocysteine remethylation [286] so may have the greatest demand for folates. However, the degree to which impairment of methylation reactions contributes to tissue dysfunction during folate deficiency needs more research. Fourth, folic acid may contribute to expansion of β-cell mass via β-cell differentiation [339]. However, relevance of this finding in mature mammals remains to be determined.

In summary (Fig. 5.10), results in this chapter show that folate deficiency reduces the glucose metabolic health of mature male mice but does not reduce methylation potential or H3K4me3 in islets. This conclusion could be strengthened and expanded by future experiments characterizing

β-cell function during folate-deficiency and with homocysteine supplementation, and, by establishing an alternative method to reduce methylation potential in islets to better understand the consequences of that on islet function and gene expression regulation. Whether changes in organismal or cellular methylation potential drives changes in chromatin methylation in β-cells remains to be experimentally tested.

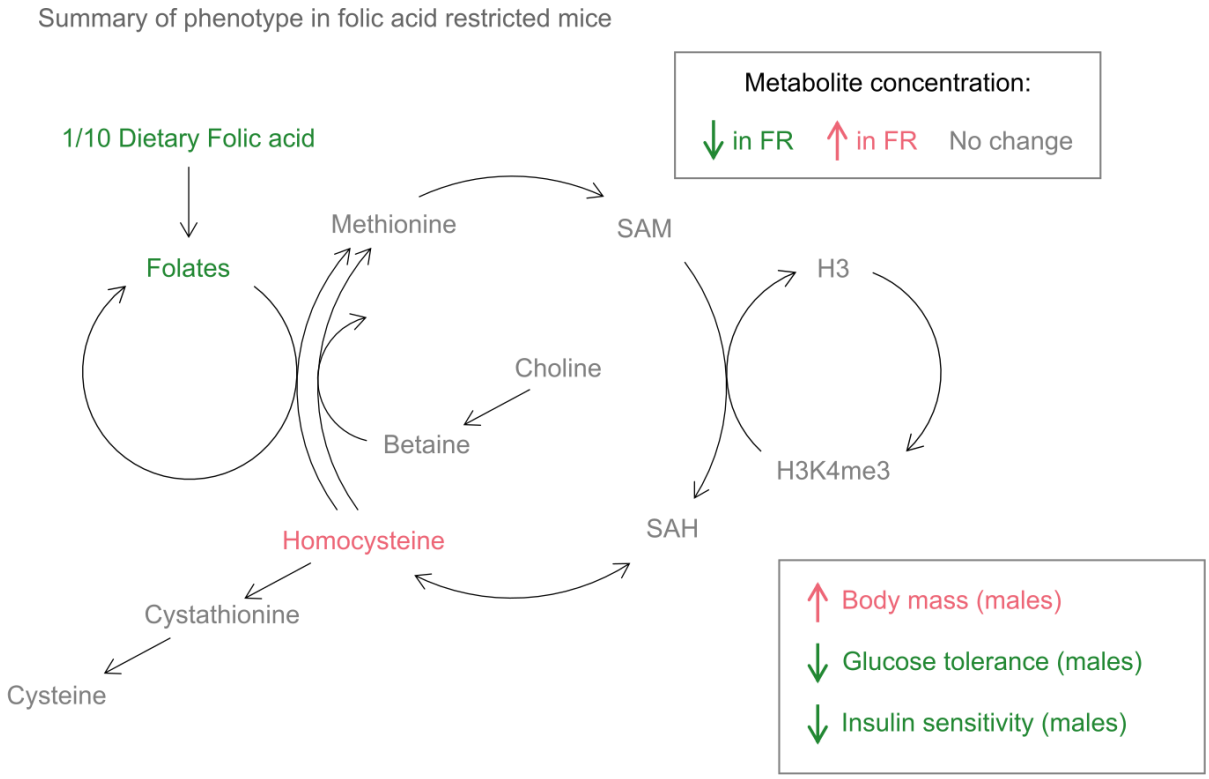

Figure 5.10 | Summary of the phenotype observed in adult mice fed a low folic acid diet from 8 weeks old for 24 weeks. Only molecules that were quantified are shown.

Chapter 6: Conclusions

6.1 Research summary

Pancreatic β-cells, like all terminally differentiated cell types, perform specialized functions that arise from selective expression of specific genes. Expression patterns are established during cell differentiation and maturation, and then maintained during the mature period of that cell's life, which may span decades, cell divisions, and environmental perturbations. The TrxG genes were first described as genes that oppose transcriptional silencing during development in *Drosophila*. Rather than activating gene expression, TrxG proteins are chiefly involved in the stable maintenance and transmission of active gene expression through time and cell divisions, even in absence of the initiating signal. Thus, it was hypothesized that TrxG proteins function in a system of cellular memory [467]. The prototypical TrxG protein Trx and its mammalian homologue MLL1 were later shown to be histone methyltransferases that selectively methylate H3K4 at targeted gene promoters [468,469]. It was immediately recognized that H3K4me3 could be an essential signal in that system of cellular memory [468]. To perform this role, H3K4me3 must regulate gene expression. This has proven challenging to validate [178]. Accumulated evidence suggests that, rather than a fundamental role in transcription, H3K4 methylation in mammals is required for expression of very few genes, primarily during development or in response to stress [244,250,251,351], whereas a requirement during homeostatic conditions in mature tissues is uncertain.

In mature β-cells, H3K4 methylation has been linked with maintenance of metabolic function [275–278], cell identity [279,280], and suppressing replication [282–284], but these links are strictly correlative. The purpose of this thesis was to define the functions performed by H3K4 methylation in regulating β-cell gene expression and function. There were three objectives:

1. Determine whether H3K4me3 regulates gene expression in mature β-cells.
2. Determine if β-cell (dys)function is associated with patterns of H3K4me3 enrichment.
3. Test the sensitivity of H3K4me3 enrichment to external metabolic perturbation.

To address these objectives, I first took advantage of an unexpected quality of inducible *Dpy30*-KO β-cells—that loss of H3K4 methylation shows much slower kinetics than loss of DPY30—to resolve transcriptional defects arising from loss of DPY30 versus loss of H3K4 methylation. Transcriptome remodeling occurred after loss of H3K4 methylation, supporting a function for H3K4 methylation in the maintenance of active gene expression. Global reduction of H3K4me3 in *Dpy30*-KO cells did not lead to global reduction in mRNA expression. However, most active promoters retained detectable enrichment of H3K4me3. Genes which completely lost H3K4me3 were downregulated. Further, the degree to which H3K4 methylation was lost at a given promoter correlates with initial enrichment, with the effect that lowly expressed genes were most likely to completely lose H3K4 methylation and be downregulated. Therefore, even though only ~5% of genes are dysregulated in *Dpy30*-KO cells, these data do not preclude a more general role for H3K4me3 in maintaining active transcription in mature β-cells. This work adds to growing body of evidence that H3K4 methylation is instructive for gene expression, but is not an essential requirement [194,199,201,202,226,254,255,257,259,263,265,267,270,470], and is the first demonstration in a terminally-differentiated mammalian tissue.

I show that global enrichment of H3K27ac in β-cells is partially controlled downstream of H3K4me3. This extends a previous demonstration that H3K27ac at enhancers in mESCs is partially controlled by H3K4me1, and may be caused by reduced recruitment of

acetyltransferases that recognize methylated H3K4 [400–402]. H3K27me3, meanwhile, became enriched in promoters of some genes that were downregulated in *Dpy30*-KO cells. This is consistent with evidence that H3K4me3 serves an anti-repressive role by sterically hindering deposition of H3K27me3 by PcG proteins [267]. Notably, global reduction of active marks H3K4me3, H3K4me1, and H3K27ac did not lead to expansion of the megabase-scale H3K27me3-positive PcG-repressed domains. While TrxG was initially defined as an antagonist of PcG-driven repression [208], this result suggests that H3K4me3 performs this role only at a small number of targeted promoters in β-cells, whereas the boundaries of PcG-driven facultative heterochromatin are delimited by other means. H3K4me3 may also serve an anti-silencing role in β-cells by sterically hindering recruitment of DNA-methyltransferases, but this was not directly tested. More generally, the degree to which H3K4me3 regulates transcription via activation versus anti-repression requires more research.

There is growing appreciation for the contribution of epigenetic plasticity in the pathogenesis of metabolic diseases [341,471]. Global shifts in H3K27me3 [133] and H3K27ac [416] enrichment, and site-specific changes of DNA methylation [472] and RNA methylation [417] enrichment have been described in islets from donors with T2D or mouse models of T2D. These changes are implicated in driving changes in gene expression and consequent changes to β-cell function. In Chapter 4, I examined the relationship between H3K4me3 peaks, gene expression, and function in *Dpy30*-KO and *Lepr*$^{db/db}$ mice. Whereas H3K27me3 is reduced [133] and H3K27ac is increased [416] in diabetic islets, I did not find evidence for a global shift in H3K4me3 enrichment. Instead of a global shift, H3K4me3 peaks at genes upregulated in *Lepr*$^{db/db}$ islets expand, and peaks at downregulated genes contract. Genes that are up- or downregulated in *Lepr*$^{db/db}$ islets tend to be

downregulated in *Dpy30*-KO islets, indicating that their expression relies on H3K4me3. It is notable that less than 5% of genes are downregulated in *Dpy30*-KO islets but that these genes are significantly overrepresented among those dysregulated in *Lepr*$^{db/db}$ mice. I conclude that H3K4me3 is important for transcription of genes that are dysregulated in a mouse model of T2D. What drives the change in H3K4me3 at these genes was not determined but may stem from differential recruitment of COMPASS complexes and/or histone demethylases by transcription factors. Since I also show that β-cell-enriched genes tend to have broad H3K4me3 domains, determining what gives rise to broad H3K4me3 domains may shed light on why those genes tend to be downregulated in diabetic islets.

Loss of H3K4 methylation in *Dpy30*-KO β-cells led to functional deficits. Mice became hyperglycemic and glucose intolerant due to impaired transcription, maturation, and GSIS. Accumulation of H3K4me3 at cyclin-dependent kinase inhibitor genes has been linked to suppression of replication in mature β-cells [282–284] but I found that replication was unchanged in basal glucose, and inhibited in stimulatory glucose, after loss of H3K4me3. This agrees with previous findings that endocrine cell proliferation is blunted in *Dpy30*-KO mice [351]. Therefore, H3K4me3 may promote transcription of cyclin-dependent kinase inhibitors, but it is not a general "anti-replication" signal in β-cells. Increased rates of replication in islets of MEN1-happloinsufficient mice [283] may be caused by loss of nonenzymatic anti-tumour activities of MEN1 [473] rather than a global reduction of H3K4me3.

Having established that H3K4me3 regulates gene expression and function in mature β-cells, I next tested how resilient islet H3K4me3 enrichment is to availability of an essential methyl

group carrier. H3K4me3 levels were unchanged by dietary restriction of folic acid. While mice became hyperhomocysteinemic, this did not lead to reduction in methylation potential in circulation or islets, suggesting that H3K4me3 is stable because an (undetermined) metabolic compensation preserves systemic SAM pools for methyltransferase reactions. Defects in glucose homeostasis in folic acid-restricted mice may be caused by hyperhomocysteinemia [458,459], reduced methylation potential in other tissues like liver [465,466], or impairments in nucleotide synthesis [286]. H3K4me3 may furthermore have been altered at particular genomic loci in islets but not globally [321,323] but this was not measured. Therefore, I find that global H3K4me3 levels in islets are robust against metabolic perturbation – both by hyperphagia-induced T2D in the *Lepr*$^{db/db}$ model, and by disruption of systemic 1C-metabolism in the folic acid-restriction model.

6.2 Limitations and future directions

This thesis focused on the link between H3K4me3 and gene expression in mature mouse β-cells. Similar strategies could be used to study any of a large number of other chromatin modifications that are found in β-cells. A powerful starting point would compare the genome-wide enrichment patterns of a selection of histone modifications, histone variants, and DNA modifications in β-cells from healthy and diabetic individuals – or, models of β-cell dysfunction, or, incompletely-matured stem cell-derived β-like-cells, or α-cells, according to your interests. Resulting data could then be used to predict epigenetic drivers of gene expression dysregulation that contribute to β-cell dysfunction. Patterns may also emerge that suggest a mechanism driving epigenome dysregulation. For example, a generalized increase of histone acetylation might indicate a shift in the expression or localization acetyltransferases or deacetylases, or perhaps a shift in the nuclear

pool of their essential cofactors acetyl-CoA, Zn^{2+}, and NAD^{+}. Such approaches could also be applied to other tissues for which a causal link between metabolism and chromatin is suspected.

More work is required to determine why some genes are downregulated when H3K4me3 is lost. Ultimately, it is probably due to a change in productive elongation by RNAP2 at those genes. ChIP-seq or CUT&RUN of total and phosphorylated RNAP2 would therefore help determine whether recruitment and/or elongation are altered. Consideration of DNA motifs enriched at downregulated genes may also suggest causative transcription factors that act by recruiting COMPASS complexes, notably MAFA and PAX6 [278,281], or are recruited to sites of H3K4me3, such as TAF3 [265]. Examination of transcriptional bursting kinetics using RNA fluorescence in-situ hybridization would provide greater detail into RNAP2 activity of selected genes. The generalized increase in transcriptional variability I observed in *Dpy30*-KO mice suggests a change in bursting kinetics. PcG proteins repress gene expression by reducing the frequency of transcriptional bursts [197], but it is not known how H3K4me3 affects burst frequency, length, or magnitude, which are fundamental to transcription. Finally, it would be interesting to determine if transcriptional downregulation is reversible by re-introducing DPY30. I predict that most genes which require H3K4me3 to activate transcription could be rescued assuming H3K4me3 acts as an activating mark [202]. Some genes, such as imprinted genes that may be targeted by DNA methyltransferases, would be irreversibly silenced [270], and genes which gained H3K27me3 may be reversibly repressed. Nearby chromatin may therefore influence a gene-specific effect of global H3K4me3 reduction and potential for rescue that could be tested by re-introducing DPY30.

The use of *Dpy30*-KO was useful to separate the effects of COMPASS complex enzymatic and non-enzymatic activity *in vivo* but is imperfect. For one thing, DPY30 may also associate with the nucleosome remodeling NuRF complex [237]. A transcriptional defect was observed several weeks after deletion of *Dpy30*-KO, suggesting defects arise due to loss of H3K4 methylation. However, it is possible that residual DPY30 remained which was below the limit of detection at the early timepoint making my conclusion that DPY30 has negligible role in gene expression inaccurate. The experimental separation of functions performed by DPY30 versus H3K4 methylation would be strengthened by knowledge of why loss of DPY30 leads to loss of H3K4 methylation. In absence of this answer, a complementary model is to introduce an inactivating point mutation to the SET domain of a selection of histone methyltransferases. Using this strategy, H3K4me1-regulated gene expression is currently being defined in stem cells with catalytically dead MLL3/4 [199,226,255,257]. The same should be done to writers of H3K4me3: MLL1/2 and SETD1A/B. Since simple point mutations are not Cre-inducible, studies with these would be limited to stem cells and stem cell differentiations, i.e., developmental contexts. More complex strategies could of course be devised to study mature tissues, perhaps using transduction of a unmodifiable dominant-negative H3 gene [167]. Of particular interest in mature tissues is the question of whether deposition of new H3K4me3 is necessary for activation of transcription in response to external stimuli. This could be an acute signaling event (for example, H3K4me3 rapidly, transiently accumulates in hippocampus during memory formation [474]) or prolonged changes to the external environment, such as starting a high fat diet. Attachment of a degron sequence to DPY30 may help answer this question, since it would prevent addition of new H3K4me3 but already-present H3K4me3 should be unperturbed for days to weeks, at least in β-cells.

Little is known about the actions of one-carbon-metabolism in β-cells. My brief analysis highlights two fundamental questions. First, how is one-carbon metabolism involved in glucose-stimulated insulin secretion? My observation that folic acid potentiates high glucose-stimulated insulin secretion, but is not by itself an insulin secretagogue, suggests that the folate cycle is an amplifying pathway. It should be noted that folic acid is a carrier of one carbon units but does not itself come pre-loaded with one [286]. Glycolysis is a major source of one carbon units for the folate cycle [475], so shuttling of one carbon units derived from glucose through the folate cycle may increase GSIS in some way. Alternatively, red/ox reactions of the folate cycle are a significant source (and consumer) of NAD(P)H in both cytosol and mitochondria [476–478], especially when mitochondrial respiration is impaired [479]. The folate cycle may therefore be an important source of reducing equivalents that amplify GSIS. Second, why do β-cells express both isoforms of methionine adenosyltransferase? *Mat2a* is expressed in fetal liver and extrahepatic cells whereas *Mat1a* expression is widely believed to be restricted to mature liver tissues [480]. These isoforms have different interactomes that also influence their subcellular localization, and unique kinetic properties [481]. For example, MAT1A is minimally sensitive to feedback inhibition, and may be stimulated by SAM, whereas MAT2B is subject to feedback inhibition by physiological levels of SAM [481]. This causes liver cells to have relatively high levels of SAM. Maturing hepatocytes switch from *MAT2A* to *MAT1A* expression, whereas a switch back from *MAT1A* to *MAT2A* may drive liver cell dedifferentiation and proliferation in cancer due to reduced SAM biosynthesis [481]. The consequences of dual expression in β-cells may be investigated using *in vitro* experiments during gene suppression or deletion, or chemical inhibition, of each. Further research could characterize their expression during development and

disease, and co-expression (or not) in individual cells. First, protein expression of both should be confirmed. Promisingly, both MAT1A and MAT2A proteins are detected in pancreatic acini [482]. Pancreatic MAT1A is downregulated and MAT2A is upregulated within two days of starting a choline-deficient diet in mice, concomitant with a 50% drop in SAM levels, suggesting novel regulatory roles [482]. *Mat2a* but not *Mat1a* RNA is downregulated, in *Lepr*$^{db/db}$ islets (data not shown), suggesting differential regulation between each in islets as well.

Overall, the data presented in this thesis demonstrate that H3K4me3 has a surprisingly limited role in regulating gene expression in mature β-cells. Nevertheless, genes that are sensitive to H3K4me3 play major roles in the function of β-cells and are often dysregulated in a model of T2D. Further, global enrichment of H3K4me3 in islets is robust against external insults in mouse models of hyperphagia-induced T2D and dietary restriction of folic acid. Together, these findings further our understanding of the regulation of transcription by chromatin.

References

1. Wild, S., Roglic, G., Green, A., Sicree, R. & King, H. Global Prevalence of Diabetes: Estimates for the year 2000 and projections for 2030. *Diabetes Care* **27**, (2004).
2. International Diabetes Federation. *IDF Diabetes Atlas. 7 ed.* (2015).
3. Saeedi, P. *et al.* Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9th edition. *Diabetes Res. Clin. Pract.* **157**, 1–10 (2019).
4. Diabetes Canada. *Diabetes in Canada: Backgrounder*. (2020).
5. Williams, R. *et al.* Global and regional estimates and projections of diabetes-related health expenditure: Results from the International Diabetes Federation Diabetes Atlas, 9th edition. *Diabetes Res. Clin. Pract.* **162**, (2020).
6. Bommer, C. *et al.* The global economic burden of diabetes in adults aged 20–79 years: a cost-of-illness study. *Lancet Diabetes Endocrinol.* **5**, 423–430 (2017).
7. Bommer, C. *et al.* Global Economic Burden of Diabetes in Adults: Projections From 2015 to 2030. *Diabetes Care* **41**, 963–970 (2018).
8. Pelletier, C. *et al.* Diabetes in Canada: facts and figures from a public health perspective. *Chronic Dis. Inj. Can.* **33**, 53–54 (2012).
9. Calder, S. New Data Shows Diabetes Rates And Economic Burden On Families Continue To Rise In Ontario. *Diabetes Canada* 2–4 (2019).
10. Dinca-Panaitescu, S. *et al.* Diabetes prevalence and income: Results of the Canadian Community Health Survey. *Health Policy (New. York).* **99**, 116–123 (2011).
11. de la Monte, S. M. & Wands, J. R. Alzheimer's Disease Is Type 3 Diabetes—Evidence Reviewed. *J. Diabetes Sci. Technol.* **2**, 1101–1113 (2008).

12. Bădescu, S. V. *et al.* The association between Diabetes mellitus and Depression. *J. Med. Life* **9**, 120–125 (2016).

13. Wingard, D. L. & Barrett-Connor, E. Heart Disease and Diabetes. *Diabetes Am.* **2**, 429–448 (1995).

14. National Kidney Foundation. *Diabetes -- a Major Risk Factor for Kidney Disease*. (2022).

15. Ramirez, M. A. & Borja, N. L. Epalrestat: An aldose reductase inhibitor for the treatment of diabetic neuropathy. *Pharmacotherapy* **28**, 646–655 (2008).

16. Kosiborod, M. *et al.* Vascular complications in patients with type 2 diabetes: Prevalence and associated factors in 38 countries (the DISCOVER study program). *Cardiovasc. Diabetol.* **17**, 1–13 (2018).

17. Cade, W. T. Diabetes-related microvascular and macrovascular diseases in the physical therapy setting. *Physical Therapy* **88**, 1322–1335 (2008).

18. Imam, B., Miller, W. C., Finlayson, H. C., Eng, J. J. & Jarus, T. Incidence of lower limb amputation in Canada. *Canadian Journal of Public Health* **108**, e374–e380 (2017).

19. Beckman, J. A. & Creager, M. A. Vascular complications of diabetes. *Circ. Res.* **118**, 1771–1785 (2016).

20. Moheet, A., Mangia, S. & Seaquist, E. R. Impact of diabetes on cognitive function and brain structure. *Ann. N. Y. Acad. Sci.* **1353**, 60–71 (2015).

21. Barbiellini Amidei, C. *et al.* Association between Age at Diabetes Onset and Subsequent Risk of Dementia. *JAMA - J. Am. Med. Assoc.* **325**, 1640–1649 (2021).

22. Sarwar, N. *et al.* Diabetes mellitus, fasting blood glucose concentration, and risk of vascular disease: A collaborative meta-analysis of 102 prospective studies. *Lancet* **375**, 2215–2222 (2010).

23. Loukine, L., Waters, C., Choi, B. C. K. & Ellison, J. Impact of diabetes mellitus on life expectancy and health-adjusted life expectancy in Canada. *Popul. Health Metr.* **10**, 1–10 (2012).

24. Saeedi, P. *et al.* Mortality attributable to diabetes in 20–79 years old adults, 2019 estimates: Results from the International Diabetes Federation Diabetes Atlas, 9th edition. *Diabetes Research and Clinical Practice* (2020). doi:10.1016/j.diabres.2020.108086

25. Baena-Díez, J. M. *et al.* Risk of cause-specific death in individuals with diabetes: A competing risks analysis. *Diabetes Care* **39**, 1987–1995 (2016).

26. Punthakee, Z., Goldenberg, R. & Katz, P. Definition, Classification and Diagnosis of Diabetes, Prediabetes and Metabolic Syndrome. *Can. J. Diabetes* **42**, S10–S15 (2018).

27. Olefsky, J., Farquhar, J. W. & Reaven, G. Relationship between fasting plasma insulin level and resistance to insulin mediated glucose uptake in normal and diabetic subjects. *Diabetes* **22**, 507–513 (1973).

28. Cohen, A. *et al.* Increasing Incidence of Type 1 and Type 2 Diabetes Among Canadian Children. *Can. J. Diabetes* **46**, 189–195 (2021).

29. Wang, Y. & Lobstein, T. Worldwide trends in childhood overweight and obesity. *Int. J. Pediatr. Obes.* **1**, 11–25 (2006).

30. Daneman, D. Type 1 diabetes. *Lancet* **367**, 847–858 (2006).

31. Aschner, P. New IDF clinical practice recommendations for managing type 2 diabetes in primary care. *Diabetes Res. Clin. Pract.* **132**, 169–170 (2017).

32. Zhang, H., Colclough, K., Gloyn, A. L. & Pollin, T. I. Monogenic diabetes: A gateway to precision medicine in diabetes. *J. Clin. Invest.* **131**, e142244 (2021).

33. Hoffman, L. S., Fox, T. J., Anastasopoulou, C. & Jialal, I. *Maturity Onset Diabetes in the*

Young. (StatPearls Publishing, 2021).

34. Kautzky-willer, A. *et al.* Pronounced Insulin Resistance and Inadequate beta-cell Secretion Characterize Lean Gestational Diabetes During and After Pregnancy. *Diabetes Care* **20**, 1717–1723 (1997).

35. A., R. E. & Enns, L. Role of Gestational Hormones in the Induction of Insulin Resistance *. *J. Clin. Endocrinol. Metab.* **67**, 341–347 (1988).

36. Kampmann, U. *et al.* Gestational diabetes: A clinical update. *World J. Diabetes* **6**, 1065–1072 (2015).

37. Dabelea, D. *et al.* Intrauterine Exposure to Diabetes Conveys Risks for Type 2 Diabetes and Obesity A Study of Discordant Sibships. *Diabetes* **49**, 2208–2211 (2000).

38. Muoio, D. M. & Newgard, C. B. Molecular and metabolic mechanisms of insulin resistance and β-cell failure in type 2 diabetes. *Nat. Rev. Mol. Cell Biol.* **9**, 193–205 (2008).

39. Kwon, H. & Pessin, J. E. Adipokines mediate inflammation and insulin resistance. *Front. Endocrinol. (Lausanne).* **4**, 1–13 (2013).

40. Rabe, K., Lehrke, M., Parhofer, K. G. & Broedl, U. C. Adipokines and insulin resistance. *Mol. Med.* **14**, 741–751 (2008).

41. Borkman, M. *et al.* The Relation between Insulin Sensitivity and the Fatty-Acid Composition of Skeletal-Muscle Phospholipids. *N. Engl. J. Med.* **328**, 238–244 (1993).

42. Griffin, M. E. *et al.* Free fatty acid-induced insulin resistance is associated with activation of protein kinase C θ and alterations in the insulin signaling cascade. *Diabetes* **48**, 1270–1274 (2000).

43. Defronzo, R. A. From the triumvirate to the ominous octet: A new paradigm for the

treatment of type 2 diabetes mellitus. *Diabetes* **58**, 773–795 (2009).

44. Makaronidis, J. M. & Batterham, R. L. Obesity, body weight regulation and the brain: insights from fMRI. *Br. J. Radiol.* **91**, 1–12 (2018).

45. Ding, S. & Lund, P. K. Role of intestinal inflammation as an early event in obesity and insulin resistance. *Curr. Opin. Clin. Nutr. Metab. Care* **14**, 328–333 (2011).

46. Al-Goblan, A. S., Al-Alfi, M. A. & Khan, M. Z. Mechanism linking diabetes mellitus and obesity. *Diabetes, Metab. Syndr. Obes. Targets Ther.* **7**, 587–591 (2014).

47. Hamburg, N. M. *et al.* Physical Inactivity Rapidly Induces Insulin Resistance and Microvascular Dysfunction in Healthy Volunteers. *Arterioscler. Thromb. Vasc. Biol.* **27**, 2650–2656 (2007).

48. Malik, V. S. *et al.* Sugar-sweetened beverages and risk of metabolic syndrome and type 2 diabetes: A meta-analysis. *Diabetes Care* **33**, 2477–2483 (2010).

49. Refaie, M. R. *et al.* Aging is an Inevitable Risk Factor for Insulin Resistance. *J. Taibah Univ. Med. Sci.* **1**, 30–41 (2006).

50. Lawlor, N., Khetan, S., Ucar, D. & Stitzel, M. L. Genomics of Islet (Dys)function and Type 2 Diabetes. *Trends Genet.* **33**, 244–255 (2017).

51. O'Rahilly, S. P. *et al.* Beta-Cell Dysfunction, Rather Than Insulin Insensitivity, Is the Primary Defect in Familial Type 2 Diabetes. *Lancet* **328**, 360–364 (1986).

52. Wysham, C. & Shubrook, J. Beta-cell failure in type 2 diabetes: mechanisms, markers, and clinical implications. *Postgraduate Medicine* **132**, 676–686 (2020).

53. Gerber, P. A. & Rutter, G. A. The Role of Oxidative Stress and Hypoxia in Pancreatic Beta-Cell Dysfunction in Diabetes Mellitus. *Antioxidants Redox Signal.* **26**, 501–518 (2017).

54. Back, S. H., Kang, S. W., Han, J. & Chung, H. T. Endoplasmic reticulum stress in the β-cell pathogenesis of type 2 diabetes. *Exp. Diabetes Res.* **2012**, 1–11 (2012).

55. Nordmann, T. M. *et al.* The Role of Inflammation in β-cell Dedifferentiation. *Sci. Rep.* **7**, 1–10 (2017).

56. Weir, G. C., Aguayo-Mazzucato, C. & Bonner-Weir, S. β-Cell Dedifferentiation in Diabetes Is Important, But What Is It? *Islets* **5**, 233–237 (2013).

57. Marín-Peñalver, J. J., Martín-Timón, I., Sevillano-Collantes, C. & Cañizo-Gómez, F. J. del. Update on the treatment of type 2 diabetes mellitus. *World J. Diabetes* **7**, 354–395 (2016).

58. Zhou, G. *et al.* Role of AMP-activated protein kinase in mechanism of metformin action. *J. Clin. Invest.* **108**, 1167–1174 (2001).

59. Drucker, D. J. Incretin action in the pancreas: Potential promise, possible perils, and pathological pitfalls. *Diabetes* **62**, 3316–3323 (2013).

60. Proks, P., Reimann, F., Green, N., Gribble, F. & Ashcroft, F. Sulfonylurea stimulation of insulin secretion. *Diabetes* **51**, S368–S376 (2002).

61. Siehler, J., Blöchinger, A. K., Meier, M. & Lickert, H. Engineering islets from stem cells for advanced therapies of diabetes. *Nat. Rev. Drug Discov.* **20**, 920–940 (2021).

62. Shapiro, A. M. J. *et al.* Insulin expression and C-peptide in type 1 diabetes subjects implanted with stem cell-derived pancreatic endoderm cells in an encapsulation device. *Cell Reports Med.* **2**, 1–9 (2021).

63. Ramzy, A. *et al.* Implanted pluripotent stem-cell-derived pancreatic endoderm cells secrete glucose-responsive C-peptide in patients with type 1 diabetes. *Cell Stem Cell* **28**, 2047-2061.e5 (2021).

64. Saito, K., Iwama, N. & Takahashi, T. Morphometrical Analysis on Topographical Difference in Size Distribution, Number and Volume of Islets in the Human Pancreas. *Tohoku J. Exp. Med.* **124**, 177–186 (1978).

65. Da Silva Xavier, G. The cells of the islets of langerhans. *J. Clin. Med.* **7**, 1–17 (2018).

66. Ohneda, K., Ee, H. & German, M. Regulation of insulin gene transcription. *Semin. Cell Dev. Biol.* **11**, 227–233 (2000).

67. Melloul, D., Ben-Neriah, Y. & Cerasi, E. Glucose modulates the binding of an islet-specific factor to a conserved sequence within the rat I and the human insulin promoters. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 3865–3869 (1993).

68. Sharma, A. & Stein, R. Glucose-induced transcription of the insulin gene is mediated by factors required for beta-cell-type-specific expression. *Mol. Cell. Biol.* **14**, 871–879 (1994).

69. Mosley, A. L., Corbett, J. A. & Özcan, S. Glucose Regulation of Insulin Gene Expression Requires the Recruitment of p300 by the β-Cell-Specific Transcription Factor Pdx-1. *Mol. Endocrinol.* **18**, 2279–2290 (2004).

70. Tengholm, A. & Gylfe, E. cAMP signalling in insulin and glucagon secretion. *Diabetes, Obes. Metab.* **19**, 42–53 (2017).

71. Read, M. L., Masson, M. R. & Docherty, K. A RIPE3b1-like factor binds to a novel site in the human insulin promoter in a redox-dependent manner. *FEBS Lett.* **418**, 68–72 (1997).

72. Tillmar, L. & Welsh, N. Glucose-induced binding of the polypyrimidine tract-binding protein (PTB) to the 3′-untranslated region of the insulin mRNA (ins-PRS) is inhibited by rapamycin. *Mol. Cell. Biochem.* **260**, 85–90 (2004).

73. Tillmar, L. & Welsh, N. Hypoxia may increase rat insulin mRNA levels by promoting

binding of the polypyrimidine tract-binding protein (PTB) to the pyrimidine-rich insulin mRNA 3′-untranslated region. *Mol. Med.* **8**, 263–272 (2002).

74. Welsh, M., Nielseng, D. A., Mackrel, A. J. & Steiner, D. F. Control of Insulin Gene Expression in Pancreatic Beta-Cells and in an insulin-producing cell line, RIN-5F cells. *J. Biol. Chem.* **260**, 13590–13594 (1985).

75. Yang, E. *et al.* Decay Rates of Human mRNAs : Correlation With Functional Characteristics and Sequence Attributes. *Genome Res.* **13**, 1863–1872 (2003).

76. Tillmar, L., Carlsson, C. & Welsh, N. Control of Insulin mRNA Stability in Rat Pancreatic Islets. *J. Biol. Chem.* **277**, 1099–1106 (2002).

77. Millership, S. J. *et al.* Neuronatin regulates pancreatic β cell insulin content and secretion. *J. Clin. Invest.* **128**, 3369–3381 (2018).

78. Wicksteed, B. *et al.* Cooperativity between the Preproinsulin mRNA Untranslated Regions Is Necessary for Glucose-stimulated Translation. *J. Biol. Chem.* **276**, 22553–22558 (2001).

79. Haataja, L. *et al.* Disulfide mispairing during proinsulin folding in the endoplasmic reticulum. *Diabetes* **65**, 1050–1060 (2016).

80. Ghiasi, S. M. *et al.* Endoplasmic reticulum chaperone glucose-regulated protein 94 is essential for proinsulin handling. *Diabetes* **68**, 747–760 (2019).

81. Omar-Hmeadi, M. & Idevall-Hagren, O. Insulin granule biogenesis and exocytosis. *Cell. Mol. Life Sci.* **78**, 1957–1970 (2021).

82. Stephens, S. B. *et al.* The Prohormone VGF Regulates β Cell Function via Insulin Secretory Granule Biogenesis. *Cell Rep.* **20**, 2480–2489 (2017).

83. Petersen, M. C. & Shulman, G. I. Mechanisms of insulin action and insulin resistance.

Physiol. Rev. **98**, 2133–2223 (2018).

84. Schmitz, O., Rungby, J., Edge, L. & Juhl, C. B. On high-frequency insulin oscillations. *Ageing Res. Rev.* **7**, 301–305 (2008).

85. Johnson, J. H., Newgard, C. B., Milburn, J. L., Lodish, H. F. & Thorens, B. The high K(m) glucose transporter of islets of Langerhans is functionally similar to the low affinity transporter of liver and has an identical primary sequence. *J. Biol. Chem.* **265**, 6548–6551 (1990).

86. Guillam, M., Dupraz, P. & Thorens, B. Glucose uptake, utilization, and signaling in GLUT2-Null Islets. *Diabetes* **49**, 1485–1491 (2000).

87. McCulloch, L. J. *et al.* GLUT2 (SLC2A2) is not the principal glucose transporter in human pancreatic beta cells: Implications for understanding genetic association signals at this locus. *Mol. Genet. Metab.* **104**, 648–653 (2011).

88. Sternisha, S. M. & Miller, B. G. Molecular and cellular regulation of human glucokinase. *Arch. Biochem. Biophys.* **663**, 199–213 (2019).

89. Fu, Z., R. Gilbert, E. & Liu, D. Regulation of Insulin Synthesis and Secretion and Pancreatic Beta-Cell Dysfunction in Diabetes. *Curr. Diabetes Rev.* **9**, 25–53 (2013).

90. Eto, K. *et al.* Role of NADH shuttle system in glucose-induced activation of mitochondrial metabolism and insulin secretion. *Science (80-.).* **283**, 981–985 (1999).

91. Henquin, J. C., Ravier, M. A., Nenquin, M., Jonas, J. C. & Gilon, P. Hierarchy of the β-cell signals controlling insulin secretion. *Eur. J. Clin. Invest.* **33**, 742–750 (2003).

92. Prentki, M., Matschinsky, F. M. & Madiraju, S. R. M. Metabolic Signaling in Fuel-Induced Insulin Secretion. *Cell Metab.* **18**, 162–185 (2013).

93. Blum, B. *et al.* Functional beta-cell maturation is marked by an increased glucose

threshold and by expression of urocortin 3. *Nat. Biotechnol.* **30**, 261–264 (2012).

94. Henquin, J. C. & Nenquin, M. Immaturity of insulin secretion by pancreatic islets isolated from one human neonate. *J. Diabetes Investig.* **9**, 270–273 (2018).

95. Huang, C. *et al.* Synaptotagmin 4 Regulates Pancreatic β Cell Maturation by Modulating the Ca2+ Sensitivity of Insulin Secretion Vesicles. *Dev. Cell* **45**, 347-361.e5 (2018).

96. Helman, A. *et al.* A Nutrient-Sensing Transition at Birth Triggers Glucose-Responsive Insulin Secretion. *Cell Metab.* **31**, 1004–1016 (2020).

97. Xu, Q. G. *et al.* Insulin as an in vivo growth factor. *Exp. Neurol.* **188**, 43–51 (2004).

98. Jaafar, R. *et al.* MTORC1-to-AMPK switching underlies β cell metabolic plasticity during maturation and diabetes. *J. Clin. Invest.* **129**, 4124–4137 (2019).

99. Stolovich-Rain, M. *et al.* Weaning Triggers a Maturation Step of Pancreatic β Cells. *Dev. Cell* **32**, 535–545 (2015).

100. Puri, S. *et al.* Replication confers β cell immaturity. *Nat. Commun.* **9**, 1–12 (2018).

101. Tan, C., Tuch, B. E., Tu, J. & Brown, S. A. Role of NADH shuttles in glucose-induced insulin secretion from fetal beta-cells. *Diabetes* **51**, 2989–2996 (2002).

102. Pullen, T. J. *et al.* Identification of genes selectively disallowed in the pancreatic islet. *Islets* **2**, 89–95 (2010).

103. Liu, J. S. E. & Hebrok, M. All mixed up: Defining roles for β-cell subtypes in mature islets. *Genes Dev.* **31**, 228–240 (2017).

104. Iynedjian, P. B. Molecular physiology of mammalian glucokinase. *Cell. Mol. Life Sci.* **66**, 27–42 (2009).

105. Sekine, N. *et al.* Low lactate dehydrogenase and high mitochondrial glycerol phosphate dehydrogenase in pancreatic β-cells. Potential role in nutrient sensing. *J. Biol. Chem.* **269**,

4895–4902 (1994).

106. Ishihara, H., Wang, H., Drewes, L. R. & Wollheim, C. B. Overexpression of monocarboxylate transporter and lactate dehydrogenase alters insulin secretory responses to pyruvate and lactate in β cells. *J. Clin. Invest.* **104**, 1621–1629 (1999).

107. Jermendy, A. *et al.* Rat neonatal beta cells lack the specialised metabolic phenotype of mature beta cells. *Diabetologia* **54**, 594–604 (2011).

108. 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

109. Teschendorff, A. E., Sollich, P. & Kuehn, R. Signalling entropy : A novel network-theoretical framework for systems analysis and interpretation of functional omic data. *Methods* **67**, 282–293 (2014).

110. Salinno, C. *et al.* β-Cell Maturation and Identity in Health and Disease. *Int. J. Mol. Sci.* **20**, 1–20 (2019).

111. Gao, T. *et al.* Pdx1 maintains β cell identity and function by repressing an α cell program. *Cell Metab.* **19**, 259–271 (2014).

112. Mitchell, R. K. *et al.* The transcription factor Pax6 is required for pancreatic β cell identity, glucose-regulated ATP synthesis, and Ca2+ dynamics in adult mice. *J. Biol. Chem.* **292**, 8892–8906 (2017).

113. Ediger, B. N. *et al.* Islet-1 is essential for pancreatic β-cell function. *Diabetes* **63**, 4206–4217 (2014).

114. Swisa, A. *et al.* Pax6 maintains pancreatic beta-cell identity by repressing alternative islet cell genes. *J. Clin. Invest.* **127**, 230–243 (2017).

115. Taylor, B., Liu, F. F. & Sander, M. Nkx6.1 Is Essential for Maintaining the Functional State of Pancreatic Beta Cells. *Cell Rep.* **4**, 1262–1275 (2013).

116. Gutiérrez, G. D. *et al.* Pancreatic β cell identity requires continual repression of non – β cell programs. *J. Clin. Invest.* **127**, 244–259 (2017).

117. Gu, C. *et al.* Pancreatic beta cells require neuroD to achieve and maintain functional maturity. *Cell Metab.* **11**, 298–310 (2010).

118. Nishimura, W., Takahashi, S. & Yasuda, K. MafA is critical for maintenance of the mature beta cell phenotype in mice. *Diabetologia* **58**, 566–574 (2015).

119. Vivo, P. I. *et al.* Foxa2 controls Pdx1 Gene Expression in Pancreatic β-Cells In Vivo. *Diabetes* **51**, 2546–2551 (2002).

120. Sund, N. J. *et al.* Tissue-specific deletion of Foxa2 in pancreatic β cells results in hyperinsulinemic hypoglycemia. *Genes Dev.* **15**, 1706–1715 (2001).

121. Gao, N. *et al.* Foxa1 and Foxa2 Maintain the Metabolic and Secretory Features of the Mature β-Cell. *Mol. Endocrinol.* **24**, 1594–1604 (2010).

122. Marchetti, P., Bugliani, M., De Tata, V. D., Suleiman, M. & Marselli, L. Pancreatic beta cell identity in humans and the role of type 2 diabetes. *Front. Cell Dev. Biol.* **5**, 1–8 (2017).

123. Guo, S. *et al.* Inactivation of specific β cell transcription factors in type 2 diabetes. *J. Clin. Invest.* **123**, 3305–3316 (2013).

124. Talchai, C., Xuan, S., Lin, H. V., Sussel, L. & Accili, D. Pancreatic β Cell Dedifferentiation as a Mechanism of Diabetic β Cell Failure. *Cell* **150**, 1223–1234 (2012).

125. Sahebi, L. *et al.* Iranian neonatal diabetes mellitus due to mutation in PDX1 gene : a case report. *J. Med. Case Rep.* **13**, 1–6 (2019).

126. Iacovazzo, D. *et al.* MAFA missense mutation causes familial insulinomatosis and diabetes mellitus. *Proc. Natl. Acad. Sci.* **115**, 1027–1032 (2018).

127. Stekelenburg, C. *et al.* Exome sequencing identifies a de novo FOXA2 variant in a patient with syndromic diabetes. *Pediatr. Diabetes* **20**, 366–369 (2019).

128. Rubio-cabezas, O. *et al.* Homozygous Mutations in NEUROD1 Are Responsible for a Novel Syndome of Permanent Neonatal Diabetes and Neurological Abnormalities. *Diabetes* **59**, 2326–2331 (2010).

129. Aurbach, A. *et al.* NKX2-2 Mutation Causes Congenital Diabetes and Infantile Obesity With Paradoxical Glucose- Induced Ghrelin Secretion. *J. Clin. Endocrinol. Metab.* **105**, 3486–3495 (2020).

130. Solomon, B. D. *et al.* Compound Heterozygosity for Mutations in PAX6 in a Patient With Complex Brain Anomaly , Neonatal Diabetes Mellitus , and Microophthalmia. *Am. J. Med. Genet.* 2543–2546 (2009). doi:10.1002/ajmg.a.33081

131. Bensellam, M., Jonas, J. & Laybutt, D. R. Mechanisms of β-cell dedifferentiation in diabetes: recent findings and future research directions. *J. Endocrinol.* **236**, R109–R143 (2018).

132. Diedisheim, M. *et al.* Modeling human pancreatic beta cell dedifferentiation. *Mol. Metab.* **10**, 74–86 (2018).

133. Lu, T. T. *et al.* The Polycomb-dependent epigenome controls beta cell dysfunction, dedifferentiation and diabetes. *Cell Metab.* **27**, 1–15 (2018).

134. Sun, X., Wang, L., Obayomi, S. M. B. & Wei, Z. Epigenetic Regulation of β Cell Identity and Dysfunction. *Front. Endocrinol. (Lausanne).* **12**, 1–9 (2021).

135. Aguayo-Mazzucato, C. *et al.* Mafa expression enhances glucose-responsive insulin

secretion in neonatal rat beta cells. *Diabetologia* **54**, 583–593 (2011).

136. Bader, E. *et al.* Identification of proliferative and mature beta cells in the islets of Langerhans. *Nature* **535**, 430–434 (2016).

137. Lange, A. *et al.* Fltp T2AiCre: A new knock-in mouse line for conditional gene targeting in distinct mono- and multiciliated tissues. *Differentiation* **83**, S105–S113 (2012).

138. Benninger, R. K. P. & Hodson, D. J. New understanding of β-cell heterogeneity and in situ islet function. *Diabetes* **67**, 537–547 (2018).

139. Salinno, C. *et al.* CD81 marks immature and dedifferentiated pancreatic β-cells. *Mol. Metab.* **49**, 101188 (2021).

140. Modi, A. H. *et al.* Dynamic Ins2 gene activity defines beta cell maturity states. *bioRxiv* (2021).

141. Farack, L. *et al.* Transcriptional Heterogeneity of Beta Cells in the Intact Pancreas. *Dev. Cell* **48**, 115-125.e4 (2019).

142. Johnston, N. R. *et al.* Beta Cell Hubs Dictate Pancreatic Islet Responses to Glucose. *Cell Metab.* **24**, 389–401 (2016).

143. Nasteska, D. *et al.* PDX1LOW MAFALOW β-cells contribute to islet function and insulin release. *Nat. Commun.* **12**, 1–19 (2021).

144. Dominguez-Gutierrez, G., Xin, Y. & Gromada, J. Heterogeneity of human pancreatic β-cells. *Mol. Metab.* **27**, S7–S14 (2019).

145. Morozov, A. V., Havranek, J. J., Baker, D. & Siggia, E. D. Protein-DNA binding specificity predictions with structural models. *Nucleic Acids Res.* **33**, 5781–5798 (2005).

146. Li, Y., Chen, C. yu, Kaye, A. M. & Wasserman, W. W. The identification of cis-regulatory elements: A review from a machine learning perspective. *BioSystems* **138**, 6–17

(2015).

147. Lomvardas, S. *et al.* Interchromosomal Interactions and Olfactory Receptor Choice. *Cell* **126**, 403–413 (2006).

148. Sagai, T., Hosoya, M., Mizushina, Y., Tamura, M. & Shiroishi, T. Elimination of a long-range cis-regulatory module causes complete loss of limb-specific Shh expression and truncation of the mouse limb. *Development* **132**, 797–803 (2005).

149. Bickmore, W. A. The spatial organization of the human genome. *Annu. Rev. Genomics Hum. Genet.* **14**, 67–84 (2013).

150. Ali, T., Renkawitz, R. & Bartkuhn, M. Insulators and domains of gene expression. *Curr. Opin. Genet. Dev.* **37**, 17–26 (2016).

151. Waddington, C. H. *An Introduction to Modern Genetics*. *An Introduction to Modern Genetics* (The MacMillan Company, 1939). doi:10.4324/9781315665412

152. Cavalli, G. & Heard, E. Advances in epigenetics link genetics to the environment and disease. *Nature* **571**, 489–499 (2019).

153. Berger, S. L., Kouzarides, T., Shiekhattar, R. & Shilatifard, A. An operational definition of epigenetics. *Genes Dev.* **23**, 781–783 (2009).

154. Sharma, U. & Rando, O. J. Metabolic Inputs into the Epigenome. *Cell Metab.* **25**, 544–558 (2017).

155. Subramanian, S. *et al.* Long-term culture-expanded alveolar macrophages restore their full epigenetic identity after transfer in vivo. *Nat. Immunol.* **23**, 458–468 (2022).

156. Richmond, R. K., Sargent, D. F., Richmond, T. J., Luger, K. & Mader, A. W. Crystal structure of the nucleosome resolution core particle at 2 . 8 A. *Nature* **389**, 251–260 (1997).

157. Lorch, Y., LaPointe, J. W. & Kornberg, R. D. Nucleosomes inhibit the initiation of transcription but allow chain elongation with the displacement of histones. *Cell* **49**, 203–210 (1987).

158. Orphanides, G. & Reinberg, D. RNA polymerase II elongation through chromatin. *Nature* **407**, 471–475 (2000).

159. Cairns, B. R. The logic of chromatin architecture and remodelling at promoters. *Nature* **461**, 193–198 (2009).

160. Lee, J. S. & Shilatifard, A. A site to remember: H3K36 methylation a mark for histone deacetylation. *Mutat. Res. - Fundam. Mol. Mech. Mutagen.* **618**, 130–134 (2007).

161. Lee, C. K., Shibata, Y., Rao, B., Strahl, B. D. & Lieb, J. D. Evidence for nucleosome depletion at active regulatory regions genome-wide. *Nat. Genet.* **36**, 900–905 (2004).

162. Allfrey, V. G., Faulkner, R. & Mirsky, A. E. Acetylation and Methylation of Histones and Their Possible Role in the Regulation of RNA Synthesis. *Proc. Natl. Acad. Sci.* **51**, 786–794 (1964).

163. Zhao, Y. & Garcia, B. A. Comprehensive catalog of currently documented histone modifications. *Cold Spring Harb. Perspect. Biol.* **7**, (2015).

164. Bannister, A. J. & Kouzarides, T. Regulation of chromatin by histone modifications. *Cell Res.* **21**, 381–395 (2011).

165. Huang, H. *et al.* Lysine benzoylation is a histone mark regulated by SIRT2. *Nat. Commun.* **9**, 1–11 (2018).

166. Zhang, D. *et al.* Metabolic regulation of gene expression by histone lactylation. *Nature* **574**, 575–580 (2019).

167. Lepack, A. E. *et al.* Dopaminylation of Histone H3 in Ventral Tegmental Area Regulates

Cocaine Seeking. *Science (80-.).* **368**, 197–201 (2020).

168. Zhu, Z. *et al.* Identification of lysine isobutyrylation as a new histone modification mark. *Nucleic Acids Res.* **49**, 177–189 (2021).

169. Maksimovic, I. & David, Y. Non-enzymatic Covalent Modifications as a New Chapter in the Histone Code. *Trends Biochem. Sci.* **46**, 718–730 (2021).

170. Stasevich, T. J. *et al.* Regulation of RNA polymerase II activation by histone acetylation in single living cells. *Nature* **516**, 272–275 (2014).

171. Martin, B. J. E. *et al.* Transcription shapes genome-wide histone acetylation patterns. *Nat. Commun.* **12**, 1–9 (2021).

172. Struhl, K. Histone acetylation and transcriptional regulatory mechanisms. *Genes Dev.* **12**, 599–606 (1998).

173. Cosgrove, M. S. & Wolberger, C. How does the histone code work? *Biochem. Cell Biol.* **83**, 468–476 (2005).

174. Wang, Z. *et al.* Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat. Genet.* **40**, 897–903 (2008).

175. Jamieson, K. *et al.* Loss of HP1 causes depletion of H3K27me3 from facultative heterochromatin and gain of H3K27me2 at constitutive heterochromatin. *Genome Res.* **26**, 97–107 (2016).

176. Hogg, S. J. *et al.* Targeting histone acetylation dynamics and oncogenic transcription by catalytic P300/CBP inhibition. *Mol. Cell* **81**, 2183-2200.e13 (2021).

177. Ferrari, K. J. *et al.* Polycomb-Dependent H3K27me1 and H3K27me2 Regulate Active Transcription and Enhancer Fidelity. *Mol. Cell* **53**, 49–62 (2014).

178. Howe, F. S., Fischl, H., Murray, S. C. & Mellor, J. Is H3K4me3 instructive for

transcription activation? *BioEssays* **39**, 1–12 (2017).

179. Chen, K. *et al.* Broad H3K4me3 is associated with increased transcription elongation and enhancer activity at tumor-suppressor genes. *Nat. Genet.* **47**, 1149–1157 (2015).

180. Benayoun, B. A. *et al.* H3K4me3 breadth is linked to cell identity and transcriptional consistency. *Cell* **158**, 673–688 (2014).

181. Orford, K. *et al.* Differential H3K4 Methylation Identifies Developmentally Poised Hematopoietic Genes. *Dev. Cell* **14**, 798–809 (2008).

182. Bannister, A. J. *et al.* Spatial distribution of di- and tri-methyl lysine 36 of histone H3 at active genes. *J. Biol. Chem.* **280**, 17732–17736 (2005).

183. Xu, Q. *et al.* SETD2 regulates the maternal epigenome, genomic imprinting and embryonic development. *Nat. Genet.* **51**, 844–856 (2019).

184. Shirane, K., Miura, F., Ito, T. & Lorincz, M. C. NSD1-deposited H3K36me2 directs de novo methylation in the mouse male germline and counteracts Polycomb-associated silencing. *Nat. Genet.* **52**, 1088–1098 (2020).

185. Barski, A. *et al.* High-Resolution Profiling of Histone Methylations in the Human Genome. *Cell* **129**, 823–837 (2007).

186. Yokochi, T. *et al.* G9a selectively represses a class of late-replicating genes at the nuclear periphery. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 19363–19368 (2009).

187. Rea, S. *et al.* Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* **406**, 593–599 (2000).

188. Schotta, G. *et al.* A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. *Genes Dev.* **18**, 1251–1262 (2004).

189. Jones, B. *et al.* The histone H3K79 methyltransferase Dot1L is essential for mammalian

development and heterochromatin structure. *PLoS Genet.* **4**, e1000190 (2008).

190. Chantalat, S. *et al.* Histone H3 trimethylation at lysine 36 is associated with constitutive and facultative heterochromatin. *Genome Res.* **21**, 1426–1437 (2011).

191. Ernst, J. & Kellis, M. Chromatin-state discovery and genome annotation with ChromHMM. *Nat. Protoc.* **12**, 2478–2492 (2017).

192. Mikkelsen, T. S. *et al.* Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**, 553–560 (2007).

193. Ernst, J. *et al.* Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature* **473**, 43–49 (2011).

194. Wang, Z. *et al.* Prediction of histone post-translational modification patterns based on nascent transcription data. *Nat. Genet.* **54**, 295–305 (2022).

195. Rando, O. J. Combinatorial complexity in chromatin structure and function: Revisiting the histone code. *Curr. Opin. Genet. Dev.* **22**, 148–155 (2012).

196. Aubert, Y., Egolf, S. & Capell, B. C. The Unexpected Noncatalytic Roles of Histone Modifiers in Development and Disease. *Trends Genet.* **35**, 645–657 (2019).

197. Dobrinić, P., Szczurek, A. T. & Klose, R. J. PRC1 drives Polycomb-mediated gene repression by controlling transcription initiation and burst frequency. *Nat. Struct. Mol. Biol.* **28**, 811–824 (2021).

198. Kim, K. H. *et al.* SWI/SNF-mutant cancers depend on catalytic and non-catalytic activity of EZH2. *Nat. Med.* **21**, 1491–1496 (2015).

199. Yan, J. *et al.* Histone H3 lysine 4 monomethylation modulates long-range chromatin interactions at enhancers. *Cell Res.* **28**, 204–220 (2018).

200. Fursova, N. A. *et al.* Synergy between Variant PRC1 Complexes Defines Polycomb-

Mediated Gene Repression. *Mol. Cell* **74**, 1020-1036.e8 (2019).

201. Cruz, C. *et al.* Tri-methylation of histone h3 lysine 4 facilitates gene expression in ageing cells. *Elife* **7**, 1–24 (2018).

202. Cano-Rodriguez, D. *et al.* Writing of H3K4Me3 overcomes epigenetic silencing in a sustained but context-dependent manner. *Nat. Commun.* **7**, 12284 (2016).

203. Kassis, J. A., Kennison, J. A. & Tamkun, J. W. Polycomb and trithorax group genes in drosophila. *Genetics* **206**, 1699–1725 (2017).

204. Ingham, P. & Whittle, R. Trithorax: A new homoeotic mutation of Drosophila melanogaster causing transformations of abdominal and thoracic imaginal segments. I. Putative role during embryogenesis. *MGG Mol. Gen. Genet.* **179**, 607–614 (1980).

205. Ingham, P. W. trithorax and the regulation of homeotic gene expression in Drosophila: A historical perspective. *Int. J. Dev. Biol.* **42**, 423–429 (1998).

206. Shilatifard, A. The COMPASS family of histone H3K4 methylases: Mechanisms of regulation in development and disease pathogenesis. *Annu. Rev. Biochem.* **81**, 65–95 (2012).

207. Schuettengruber, B., Martinez, A. M., Iovino, N. & Cavalli, G. Trithorax group proteins: Switching genes on and keeping them active. *Nat. Rev. Mol. Cell Biol.* **12**, 799–814 (2011).

208. Schuettengruber, B., Bourbon, H.-M., Di Croce, L. & Cavalli, G. Genome Regulation by Polycomb and Trithorax: 70 Years and Counting. *Cell* **171**, 34–57 (2017).

209. Tamkun, J. W. *et al.* brahma: A regulator of Drosophila homeotic genes structurally related to the yeast transcriptional activator SNF2/SWI2. *Cell* **68**, 561–572 (1992).

210. Filippakopoulos, P. & Knapp, S. Targeting bromodomains: Epigenetic readers of lysine

acetylation. *Nat. Rev. Drug Discov.* **13**, 337–356 (2014).

211. Nislow, C., Ray, E. & Pillus, L. SET1, a yeast member of the Trithorax family, functions in transcriptional silencing and diverse cellular processes. *Mol. Biol. Cell* **8**, 2421–2436 (1997).

212. Miller, T. *et al.* COMPASS: A complex of proteins associated with a trithorax-related SET domain protein. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 12902–12907 (2001).

213. Piunti, A. & Shilatifard, A. Epigenetic balance of gene expression by polycomb and compass families. *Science (80-.).* **352**, (2016).

214. Rao, R. C. & Dou, Y. Hijacked in cancer: The KMT2 (MLL) family of methyltransferases. *Nat. Rev. Cancer* **15**, 334–346 (2015).

215. Bledau, A. S. *et al.* The H3K4 methyltransferase Setd1a is first required at the epiblast stage, whereas Setd1b becomes essential after gastrulation. *Dev.* **141**, 1022–1035 (2014).

216. Hui Ng, H., ois Robert, F., Young, R. A. & Struhl, K. Targeted Recruitment of Set1 Histone Methylase by Elongating Pol II Provides a Localized Mark and Memory of Recent Transcriptional Activity. *Mol. Cell* **11**, 709–719 (2003).

217. Kim, J. *et al.* RAD6-Mediated Transcription-Coupled H2B Ubiquitylation Directly Stimulates H3K4 Methylation in Human Cells. *Cell* **137**, 459–471 (2009).

218. Guenther, M. G. *et al.* Global and Hox-specific roles for the MLL1 methyltransferase. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 8603–8608 (2005).

219. Michurina, A. *et al.* Postnatal expression of the lysine methyltransferase SETD1B is essential for learning and the regulation of neuron-enriched genes. *EMBO J.* **e106459**, 1–19 (2021).

220. Lee, J. H., Tate, C. M., You, J. S. & Skalnik, D. G. Identification and characterization of

the human Set1B histone H3-Lys 4 methyltransferase complex. *J. Biol. Chem.* **282**, 13419–13428 (2007).

221. Bach, C., Mueller, D., Buhl, S., Garcia-Cuellar, M. P. & Slany, R. K. Alterations of the CxxC domain preclude oncogenic activation of mixed-lineage leukemia 2. *Oncogene* **28**, 815–823 (2009).

222. Choudhury, R., Singh, S., Arumugam, S., Roguev, A. & Stewart, A. F. The Set1 complex is dimeric and acts with Jhd2 demethylation to convey symmetrical H3K4 trimethylation. *Genes Dev.* **33**, 1–15 (2019).

223. Hu, D. *et al.* The Mll2 branch of the COMPASS family regulates bivalent promoters in mouse embryonic stem cells. *Nat. Struct. Mol. Biol.* **20**, 1093–1097 (2013).

224. Hu, D. *et al.* The MLL3 / MLL4 Branches of the COMPASS Family Function as Major Histone H3K4 Monomethylases at Enhancers. *Mol. Cell. Biol.* **33**, 4745–4754 (2013).

225. Lee, J. E. *et al.* H3K4 mono- And di-methyltransferase MLL4 is required for enhancer activation during cell differentiation. *Elife* **2013**, 1–25 (2013).

226. Dorighi, K. M. *et al.* Mll3 and Mll4 Facilitate Enhancer RNA Synthesis and Transcription from Promoters Independently of H3K4 Monomethylation. *Mol. Cell* **66**, 568-576.e4 (2017).

227. Ali, A. & Tyagi, S. Diverse roles of WDR5-RbBP5-ASH2L-DPY30 (WRAD) complex in the functions of the SET1 histone methyltransferase family. *J. Biosci.* **42**, 155–159 (2017).

228. Xue, H. *et al.* Structural basis of nucleosome recognition and modification by MLL methyltransferases. *Nature* **573**, 445–449 (2019).

229. Yang, Y. W. *et al.* Essential role of lncRNA binding for WDR5 maintenance of active chromatin and embryonic stem cell pluripotency. *Elife* **3**, 1–19 (2014).

230. Tan, C. C. *et al.* Transcription factor Ap2δ associates with Ash2l and ALR, a trithorax family histone methyltransferase, to activate Hoxc8 transcription. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 7472–7477 (2008).

231. Guarnaccia, A. D. *et al.* Impact of WIN site inhibitor on the WDR5 interactome. *Cell Rep.* **34**, 1-e14 (2021).

232. Bieluszewska, A., Weglewska, M., Bieluszewski, T., Lesniewicz, K. & Poreba, E. PKA-binding domain of AKAP8 is essential for direct interaction with DPY30 protein. *FEBS J.* **285**, 947–964 (2018).

233. Patel, A., Vought, V. E., Dharmarajan, V. & Cosgrove, M. S. A conserved arginine-containing motif crucial for the assembly and enzymatic activity of the mixed lineage leukemia protein-1 core complex. *J. Biol. Chem.* **283**, 32162–32175 (2008).

234. Ali, A., Veeranki, S. N., Chinchole, A. & Tyagi, S. MLL/WDR5 Complex Regulates Kif2A Localization to Ensure Chromosome Congression and Proper Spindle Assembly during Mitosis. *Dev. Cell* **41**, 605-622.e7 (2017).

235. Mittal, A. *et al.* The structure of the RbBP5 β-propeller domain reveals a surface with potential nucleic acid binding sites. *Nucleic Acids Res.* **46**, 3802–3812 (2018).

236. Ee, L.-S. *et al.* An Embryonic Stem Cell-Specific NuRD Complex Functions through Interaction with WDR5. *Stem Cell Reports* **8**, 1488–1496 (2017).

237. Tremblay, V. *et al.* Molecular basis for DPY-30 association to COMPASS-like and NURF complexes. *Structure* **22**, 1821–1830 (2014).

238. Han, J. *et al.* The internal interaction in RBBP5 regulates assembly and activity of MLL1 methyltransferase complex. *Nucleic Acids Res.* **47**, 10426–10438 (2019).

239. Southall, S. M., Wong, P. S., Odho, Z., Roe, S. M. & Wilson, J. R. Structural Basis for the

Requirement of Additional Factors for MLL1 SET Domain Activity and Recognition of Epigenetic Marks. *Mol. Cell* **33**, 181–191 (2009).

240. South, P. F., Fingerman, I. M., Mersman, D. P., Du, H. & Briggs, S. D. A Conserved Interaction between the SDI Domain of Bre2 and the Dpy-30 Domain of Sdc1 Is Required for Histone Methylation and Gene Expression. **285**, 595–607 (2010).

241. Jiang, H. *et al.* Role for Dpy-30 in ES Cell-Fate Specification by Regulation of H3K4 Methylation within Bivalent Domains. *Cell* **144**, 513–525 (2011).

242. Haddad, J. F. *et al.* Structural Analysis of the Ash2L/Dpy-30 Complex Reveals a Heterogeneity in H3K4 Methylation. *Structure* **26**, 1–10 (2018).

243. Murray, S. C. *et al.* H3K4me3 is neither instructive for, nor informed by, transcription. *bioRxiv* (2019). doi:10.1101/709014

244. Hödl, M. & Basler, K. Transcription in the absence of histone H3.2 and H3K4 methylation. *Curr. Biol.* **22**, 2253–2257 (2012).

245. Clouaire, T. *et al.* Cfp1 integrates both CpG content and gene activity for accurate H3K4me3 deposition in embryonic stem cells. *Genes Dev.* **26**, 1714–1728 (2012).

246. Karlić, R., Chung, H. R., Lasserre, J., Vlahoviček, K. & Vingron, M. Histone modification levels are predictive for gene expression. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 2926–2931 (2010).

247. Bonn, S. *et al.* Tissue-specific analysis of chromatin state identifies temporal signatures of enhancer activity during embryonic development. *Nat. Genet.* **44**, 148–156 (2012).

248. Bae, S. & Lesch, B. J. H3K4me1 Distribution Predicts Transcription State and Poising at Promoters. *Front. Cell Dev. Biol.* **8**, 1–11 (2020).

249. Wang, S.-P. *et al.* A UTX-MLL4-p300 Transcriptional Regulatory Network Coordinately

Shapes Active Enhancer Landscapes for Eliciting Transcription. *Mol. Cell* **67**, 308-321.e6 (2017).

250. Rickels, R. *et al.* Histone H3K4 monomethylation catalyzed by Trr and mammalian COMPASS-like proteins at enhancers is dispensable for development and viability. *Nat. Genet.* **49**, 1647–1653 (2017).

251. Mishra, B. P. *et al.* The histone methyltransferase activity of MLL1 is dispensable for hematopoiesis and leukemogenesis. *Cell Rep.* **7**, 1239–1247 (2014).

252. Lee, K. Y., Chen, Z., Jiang, R. & Meneghini, M. D. H3K4 methylation dependent and independent chromatin regulation by JHD2 and SET1 in budding yeast. *G3 Genes, Genomes, Genet.* **8**, 1829–1839 (2018).

253. Calo, E. & Wysocka, J. Modification of Enhancer Chromatin: What, How, and Why? *Mol. Cell* **49**, 825–837 (2013).

254. Kim, J. *et al.* Tudor, MBT and chromo domains gauge the degree of lysine methylation. *EMBO Rep.* **7**, 397–403 (2006).

255. Local, A. *et al.* Identification of H3K4me1-associated proteins at mammalian enhancers. *Nat. Genet.* **50**, (2017).

256. Kang, Y., Kim, Y. W., Kang, J. & Kim, A. R. Histone H3K4me1 and H3K27ac play roles in nucleosome eviction and eRNA transcription, respectively, at enhancers. *FASEB J.* **35**, 1–14 (2021).

257. Kubo, N., Hu, R., Ye, Z. & Ren, B. MLL3/MLL4 Histone Methyltranferase Activity Dependent Chromatin Organization at Enhancers during Embryonic Stem Cell Differentiation. *bioRxiv* 2021.03.17.435905 (2021).

258. Rada-Iglesias, A. Is H3K4me1 at enhancers correlative or causative? *Nat. Genet.* **50**, 4–5

(2018).

259. Cheng, J. *et al.* A role for H3K4 monomethylation in gene repression and partitioning of chromatin readers. *Mol. Cell* **53**, 979–992 (2014).

260. Dunham, I. *et al.* An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).

261. Sims, R. J. *et al.* Human but not yeast CHD1 binds directly and selectively to histone H3 methylated at lysine 4 via its tandem chromodomains. *J. Biol. Chem.* **280**, 41789–41792 (2005).

262. Hyun, K., Jeon, J., Park, K. & Kim, J. Writing, erasing and reading histone lysine methylations. *Exp. Mol. Med.* **49**, 1–22 (2017).

263. Vermeulen, M. *et al.* Quantitative Interaction Proteomics and Genome-wide Profiling of Epigenetic Histone Marks and Their Readers. *Cell* **142**, 967–980 (2010).

264. Bettridge, J., Na, C. H., Pandey, A. & Desiderio, S. H3K4me3 induces allosteric conformational changes in the DNA-binding and catalytic regions of the V(D)J recombinase. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 1904–1909 (2017).

265. Vermeulen, M. *et al.* Selective Anchoring of TFIID to Nucleosomes by Trimethylation of Histone H3 Lysine 4. *Cell* **131**, 58–69 (2007).

266. Fischer, V. *et al.* The related coactivator complexes SAGA and ATAC control embryonic stem cell self-renewal through acetyltransferase-independent mechanisms. *Cell Rep.* **36**, 109598 (2021).

267. Schmitges, F. W. *et al.* Histone Methylation by PRC2 Is Inhibited by Active Chromatin Marks. *Mol. Cell* **42**, 330–341 (2011).

268. Voigt, P. *et al.* Asymmetrically modified nucleosomes. *Cell* **151**, 181–193 (2012).

269. Ooi, S. K. T. *et al.* DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature* **448**, 714–717 (2007).

270. Kumar, D. & Jothi, R. Bivalent chromatin protects reversibly repressed genes from irreversible silencing. *bioRxiv* (2020).

271. Bhandare, R. *et al.* Genome-wide analysis of histone modifications in human pancreatic islets. *Genome Res.* **20**, 428–433 (2010).

272. Tennant, B. R. *et al.* Identification and analysis of murine pancreatic islet enhancers. *Diabetologia* **56**, 542–552 (2013).

273. Hoffman, B. G. *et al.* Locus co-occupancy, nucleosome positioning, and H3K4me1 regulate the functionality of FOXA2-, HNF4A-, and PDX1-bound loci in islets and liver. *Genome Res.* **20**, 1037–1051 (2010).

274. Pasquali, L. *et al.* Pancreatic islet enhancer clusters enriched in type 2 diabetes risk–associated variants. *Nat. Genet.* **46**, 136–143 (2014).

275. Deering, T. G., Ogihara, T., Trace, A. P., Maier, B. & Mirmira, R. G. Methyltransferase set7/9 maintains transcription and euchromatin structure at islet-enriched genes. *Diabetes* **58**, 185–193 (2009).

276. Maganti, A. V. *et al.* Transcriptional activity of the islet β cell factor Pdx1 Is augmented by lysine methylation catalyzed by the methyltransferase Set7/9. *J. Biol. Chem.* **290**, 9812–9822 (2015).

277. Wortham, M. *et al.* Nutrient regulation of the islet epigenome controls adaptive insulin secretion. *bioRxiv* 1–71 (2019). doi:10.1101/742403

278. Scoville, D. W. *et al.* MLL3 and MLL4 methyltransferases bind to the MafA and MAFB transcription factors to regulate islet β-cell function. *Diabetes* **64**, 3772–3783 (2015).

279. Dhawan, S., Georgia, S., Tschen, S. ing, Fan, G. & Bhushan, A. Pancreatic β Cell Identity Is Maintained by DNA Methylation-Mediated Repression of Arx. *Dev. Cell* **20**, 419–429 (2011).

280. Bramswig, N. C. *et al.* Epigenomic plasticity enables human pancreatic α to β cell reprogramming. *J. Clin. Invest.* **123**, 1275–1284 (2013).

281. Sun, J. *et al.* Pax6 associates with H3K4-specific histone methyltransferases Mll1, Mll2, and Set1a and regulates H3K4 methylation at promoters and enhancers. *Epigenetics Chromatin* **9**, 37 (2016).

282. Dhawan, S., Tschen, S. I. & Bhushan, A. Bmi-1 regulates the Ink4a/Arf locus to control pancreatic β-cell proliferation. *Genes Dev.* **23**, 906–911 (2009).

283. Karnik, S. K. *et al.* Menin regulates pancreatic islet growth by promoting histone methylation and expression of genes encoding p27Kip1 and p18INK4c. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 14659–14664 (2005).

284. Arda, H. E. *et al.* Age-dependent pancreatic gene regulation reveals mechanisms governing human β cell function. *Cell Metab.* **23**, 909–920 (2016).

285. Clare, C. E., Brassington, A. H., Kwong, W. Y. & Sinclair, K. D. One-Carbon Metabolism: Linking Nutritional Biochemistry to Epigenetic Programming of Long-Term Development. *Annu. Rev. Anim. Biosci.* **7**, 263–287 (2019).

286. Ducker, G. S. & Rabinowitz, J. D. One-Carbon Metabolism in Health and Disease. *Cell Metab.* **25**, 27–42 (2017).

287. Mentch, S. J. & Locasale, J. W. One-carbon metabolism and epigenetics: Understanding the specificity. *Ann. N. Y. Acad. Sci.* **1363**, 91–98 (2016).

288. Wright, A. J. A., Dainty, J. R. & Finglas, P. M. Folic acid metabolism in human subjects

revisited: Potential implications for proposed mandatory folic acid fortification in the UK. *Br. J. Nutr.* **98**, 667–675 (2007).

289. Sid, V., Siow, Y. L. & Karmin, O. Role of folate in nonalcoholic fatty liver disease. *Can. J. Physiol. Pharmacol.* **95**, 1141–1148 (2017).

290. Zhao, R., Matherly, L. H. & Goldman, I. D. Membrane transporters and folate homeostasis: Intestinal absorption and transport into systemic compartments and tissues. *Expert Rev. Mol. Med.* **11**, 1–28 (2009).

291. Tibbetts, A. S. & Appling, D. R. Compartmentalization of mammalian folate-mediated one-carbon metabolism. *Annu. Rev. Nutr.* **30**, 57–81 (2010).

292. Ducker, G. S. *et al.* Reversal of Cytosolic One-Carbon Flux Compensates for Loss of the Mitochondrial Folate Pathway. *Cell Metab.* **23**, 1140–1153 (2016).

293. Leung, K. Y., De Castro, S. C. P., Savery, D., Copp, A. J. & Greene, N. D. E. Nucleotide precursors prevent folic acid-resistant neural tube defects in the mouse. *Brain* **136**, 2836–2841 (2013).

294. Wittwer, A. J. & Wagner, C. Identification of the folate-binding proteins of rat liver mitochondria as dimethylglycine dehydrogenase and sarcosine dehydrogenase. Flavoprotein nature and enzymatic properties of the purified proteins. *J. Biol. Chem.* **256**, 4109–4115 (1981).

295. He, H. *et al.* In vivo rate of formaldehyde condensation with tetrahydrofolate. *Metabolites* **10**, 1–15 (2020).

296. Burgos-Barragan, G. *et al.* Mammals divert endogenous genotoxic formaldehyde into one-carbon metabolism. *Nature* **548**, 549–554 (2017).

297. Mentch, S. J. *et al.* Histone Methylation Dynamics and Gene Regulation Occur through

the Sensing of One-Carbon Metabolism. *Cell Metab.* **22**, 861–873 (2015).

298. Lu, S. C. & Mato, J. M. S-adenosylmethionine in liver health, injury, and cancer. *Physiol. Rev.* **92**, 1515–1542 (2012).

299. Bailey, J., Douglas, H., Masino, L., De Carvalho, L. P. S. & Argyrou, A. Human Mat2A Uses an Ordered Kinetic Mechanism and Is Stabilized but Not Regulated by Mat2B. *Biochemistry* **60**, 3621–3632 (2021).

300. Maldonado, L. Y., Arsene, D., Mato, J. M. & Lu, S. C. Methionine adenosyltransferases in cancers: Mechanisms of dysregulation and implications for therapy. *Exp. Biol. Med.* **243**, 107–117 (2018).

301. Lu, S. C. S-Adenosylmethionine. *Int. J. Biochem. Cell Biol.* **32**, 391–395 (2000).

302. Tyagi, N. *et al.* Mechanisms of homocysteine-induced oxidative stress. *Am. J. Physiol. - Hear. Circ. Physiol.* **289**, 2649–2656 (2005).

303. Ganguly, P. & Alam, S. F. Role of homocysteine in the development of cardiovascular disease. *Nutr. J.* **14**, 1–10 (2015).

304. Mosharov, E., Cranford, M. R. & Banerjee, R. The quantitatively important relationship between homocysteine metabolism and glutathione synthesis by the transsulfuration pathway and its regulation by redox changes. *Biochemistry* **39**, 13005–13011 (2000).

305. Weber, R. & Birsoy, K. The Transsulfuration Pathway Makes, the Tumor Takes. *Cell Metab.* **30**, 845–846 (2019).

306. Ye, C., Sutter, B. M., Wang, Y., Kuang, Z. & Tu, B. P. A Metabolic Function for Phospholipid and Histone Methylation. *Mol. Cell* **66**, 180-193.e8 (2017).

307. Katoh, Y. *et al.* Methionine Adenosyltransferase II Serves as a Transcriptional Corepressor of Maf Oncoprotein. *Mol. Cell* **41**, 554–566 (2011).

308. Kera, Y. *et al.* Methionine adenosyltransferase II-dependent histone H3K9 methylation at the COX-2 gene locus. *J. Biol. Chem.* **288**, 13592–13601 (2013).

309. Boukouris, A. E., Zervopoulos, S. D. & Michelakis, E. D. Metabolic Enzymes Moonlighting in the Nucleus: Metabolic Regulation of Gene Transcription. *Trends Biochem. Sci.* **41**, 712–730 (2016).

310. Luka, Z., Moss, F., Loukachevitch, L. V., Bornhop, D. J. & Wagner, C. Histone demethylase LSD1 is a folate-binding protein. *Biochemistry* **50**, 4750–4756 (2011).

311. Ross, S. A. Diet and DNA Methylation Interactions in Cancer Prevention. *Ann. N. Y. Acad. Sci.* **983**, 197–207 (2006).

312. Jill James, S., Melnyk, S., Pogribna, M., Pogribny, I. P. & Caudill, M. A. Elevation in S-Adenosylhomocysteine and DNA hypomethylation: Potential epigenetic mechanism for homocysteine-related pathology. *J. Nutr.* **132**, 2361–2366 (2002).

313. Hoffman, R. M. Is DNA methylation the new guardian of the genome? *Mol. Cytogenet.* **10**, 1–7 (2017).

314. Mandrekar, P. Epigenetic regulation in alcoholic liver disease. *World J. Gastroenterol.* **17**, 2456–2464 (2011).

315. Kharbanda, K. K. Alcoholic liver disease and methionine metabolism. *Semin. Liver Dis.* **29**, 155–165 (2009).

316. Yamamoto, J. *et al.* Histone methylation status of H3K4me3 and H3K9me3 under methionine restriction is unstable in methionine-addicted cancer cells, but stable in normal cells. *Biochem. Biophys. Res. Commun.* **533**, 1034–1038 (2020).

317. Graubard, B. I. *et al.* Influence of Threonine Metabolism on S adenosylmethionine and histone methylation. *Science (80-.).* **339**, 222–226 (2013).

318. Yu, W. *et al.* One-Carbon Metabolism Supports S-Adenosylmethionine and Histone Methylation to Drive Inflammatory Macrophages. *Mol. Cell* **75**, 1147-1160.e5 (2019).

319. Padmanabhan, N. *et al.* Mutation in folate metabolism causes epigenetic instability and transgenerational effects on development. *Cell* **155**, 81–93 (2013).

320. Glier, M. B. *et al.* Tissue-specific relationship of S -adenosylhomocysteine with allele-specific H19 / Igf2 methylation and imprinting in mice with hyperhomocysteinemia. *Epigenetics* **8**, 44–53 (2013).

321. Blake, G. E. T. *et al.* Defective folate metabolism causes germline epigenetic instability and distinguishes Hira as a phenotype inheritance biomarker. *Nat. Commun.* **12**, 1–17 (2021).

322. Padmanabhan, N. *et al.* Abnormal folate metabolism causes age-, sex- and parent-of-origin-specific haematological defects in mice. *J. Physiol.* **596**, 4341–4360 (2018).

323. Lismer, A. *et al.* Histone H3 lysine 4 trimethylation in sperm is transmitted to the embryo and associated with diet-induced phenotypes in the offspring. *Dev. Cell* **56**, 1–16 (2021).

324. Ferreira de Freitas, R., Ivanochko, D. & Schapira, M. Methyltransferase inhibitors: Competing with, or exploiting the bound cofactor. *Molecules* **24**, 1–20 (2019).

325. Rubio-Aliaga, I. *et al.* Alterations in hepatic one-carbon metabolism and related pathways following a high-fat dietary intervention. *Physiol. Genomics* **43**, 408–416 (2011).

326. Tessari, P. *et al.* Effects of insulin on methionine and homocysteine kinetics in type 2 diabetes with nephropathy. *Diabetes* **54**, 2968–2976 (2005).

327. Castaño-Martinez, T. *et al.* Methionine restriction prevents onset of type 2 diabetes in NZO mice. *FASEB J.* **33**, 7092–7102 (2019).

328. Stone, K. P., Wanders, D., Orgeron, M., Cortez, C. C. & Gettys, T. W. Mechanisms of

increased in vivo insulin sensitivity by dietary methionine restriction in mice. *Diabetes* **63**, 3721–3733 (2014).

329. Cheng, C. K. *et al.* A high methionine and low folate diet alters glucose homeostasis and gut microbiome. *Biochem. Biophys. Reports* **25**, 1–7 (2021).

330. Zhao, M. *et al.* Chronic folate deficiency induces glucose and lipid metabolism disorders and subsequent cognitive dysfunction in mice. *PLoS One* **13**, 1–16 (2018).

331. Yarandi, R. B. Effect of folate supplementation on insulin sensitivity and type 2 diabetes: A meta-analysis of randomized controlled trials. *Am. J. Clin. Nutr.* **109**, 1233 (2019).

332. Henderson, A. M. *et al.* Maternal folic acid supplementation with Vitamin B 12 deficiency during pregnancy and lactation affects the metabolic health of adult female offspring but is dependent on offspring diet. *FASEB J.* **32**, 5039–5050 (2018).

333. Achari, A. E. & Jain, S. K. L-cysteine supplementation increases insulin sensitivity mediated by upregulation of GSH and adiponectin in high glucose treated 3T3-L1 adipocytes. *Arch. Biochem. Biophys.* **630**, 54–65 (2017).

334. Zeisel, S. H. Metabolic crosstalk between choline/1-carbon metabolism and energy homeostasis. *Clin. Chem. Lab. Med.* **51**, 467–475 (2013).

335. Patterson, S., Scullion, S. M. J., McCluskey, J. T., Flatt, P. R. & McClenaghan, N. H. Prolonged exposure to homocysteine results in diminished but reversible pancreatic beta-cell responsiveness to insulinotropic agents. *Diabetes. Metab. Res. Rev.* **23**, 324–334 (2007).

336. Patterson, S., Flatt, P. R. & McClenaghan, N. H. Homocysteine-induced impairment of insulin secretion from clonal pancreatic BRIN-BD11 beta-cells is not prevented by catalase. *Pancreas* **34**, 144–151 (2007).

337. Hsu, H. C. *et al.* Folate deficiency triggers an oxidative-nitrosative stress-mediated apoptotic cell death and impedes insulin biosynthesis in RINm5F pancreatic islet β-cells: Relevant to the pathogenesis of diabetes. *PLoS One* **8**, e77931 (2013).

338. Ameri, J. *et al.* Efficient Generation of Glucose-Responsive Beta Cells from Isolated GP2+ Human Pancreatic Progenitors. *Cell Rep.* **19**, 36–49 (2017).

339. Karampelias, C. *et al.* Reinforcing one-carbon metabolism via folic acid/Folr1 promotes β-cell differentiation. *Nat. Commun.* **12**, 1–13 (2021).

340. Hall, E. *et al.* Effects of palmitate on genome-wide mRNA expression and DNA methylation patterns in human pancreatic islets. *BMC Med.* **12**, 1–15 (2014).

341. Tzika, E., Dreker, T. & Imhof, A. Epigenitics and metabolism in health and disease. *Front. Genet.* **9**, 1–8 (2018).

342. Liu, G. *et al.* Single-cell RNA Sequencing Reveals Sexually Dimorphic Transcriptome and Type 2 Diabetes Genes in Mouse Islet β Cells. *Genomics, Proteomics Bioinforma.* **19**, 408–422 (2021).

343. Avrahami, D. *et al.* Single-cell transcriptomics of human islet ontogeny defines the molecular basis of β-cell dedifferentiation in T2D. *Mol. Metab.* **42**, 1–14 (2020).

344. Dayeh, T. *et al.* Genome-Wide DNA Methylation Analysis of Human Pancreatic Islets from Type 2 Diabetic and Non-Diabetic Donors Identifies Candidate Genes That Influence Insulin Secretion. *PLoS Genet.* **10**, e1004160 (2014).

345. Mencucci, M. V., Flores, L. E., Gagliardino, J. J., Abba, M. C. & Maiztegui, B. Integrative transcriptomic analysis of pancreatic islets from patients with prediabetes/type 2 diabetes. *Diabetes. Metab. Res. Rev.* **37**, 1–11 (2021).

346. Xin, Y. *et al.* RNA Sequencing of Single Human Islet Cells Reveals Type 2 Diabetes

Genes. *Cell Metab.* **24**, 608–615 (2016).

347. John, N. A., Ram, R. & Jiang, F.-X. RNA-Seq Analysis of Islets to Characterise the Dedifferentiation in Type 2 Diabetes Model Mice db/db. *Endocr. Pathol.* **29**, 207–221 (2018).

348. Varshney, A. *et al.* Genetic regulatory signatures underlying islet gene expression and type 2 diabetes. *Proc. Natl. Acad. Sci.* **114**, 2301–2306 (2017).

349. Ernst, P. *et al.* Definitive hematopoiesis requires the mixed-lineage leukemia gene. *Dev. Cell* **6**, 437–443 (2004).

350. Glaser, S. *et al.* The histone 3 lysine 4 methyltransferase, Mll2, is only required briefly in development and spermatogenesis. *Epigenetics and Chromatin* **2**, 1–16 (2009).

351. Campbell, S. A., McDonald, C. L., Krentz, N. A. J., Lynn, F. C. & Hoffman, B. G. TrxG Complex Catalytic and Non-catalytic Activity Play Distinct Roles in Pancreas Progenitor Specification and Differentiation. *Cell Rep.* **28**, 1830-1844.e6 (2019).

352. Muzumdar, M. D., Tasic, B., Miyamichi, K., Li, L. & Luo, L. A Global Double-Fluorescent Cre Reporter Mouse. *Genesis* **45**, 593–605 (2007).

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

354. Thorens, B. *et al.* Ins1 Cre knock-in mice for beta cell-specific gene recombination. *Diabetologia* **58**, 558–565 (2015).

355. Brind'Amour, J. *et al.* An ultra-low-input native ChIP-seq protocol for genome-wide profiling of rare cell populations. *Nat. Commun.* **6**, 1–8 (2015).

356. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: A flexible trimmer for Illumina

sequence data. *Bioinformatics* **30**, 2114–2120 (2014).

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

358. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).

359. Quinlan, A. R. & Hall, I. M. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).

360. Kurtenbach, S. & William Harbour, J. SparK: A Publication-quality NGS Visualization Tool. *bioRxiv* 1–4 (2019). doi:10.1101/845529

361. Ramírez, F. *et al.* deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165 (2016).

362. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, (2008).

363. Amemiya, H. M., Kundaje, A. & Boyle, A. P. The ENCODE Blacklist: Identification of Problematic Regions of the Genome. *Sci. Rep.* **9**, 1–5 (2019).

364. Dobin, A. *et al.* STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).

365. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **14**, 417–419 (2017).

366. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 1–21 (2014).

367. Zhu, A., Ibrahim, J. G. & Love, M. I. Heavy-Tailed prior distributions for sequence count data: Removing the noise and preserving large differences. *Bioinformatics* **35**, 2084–2092 (2019).

368. Srivastava, A., Malik, L., Smith, T., Sudbery, I. & Patro, R. Alevin efficiently estimates accurate gene abundances from dscRNA-seq data. *Genome Biol.* **20**, (2019).

369. Sarkar, H., Zakeri, M., Malik, L. & Patro, R. Towards selective-alignment: Bridging the accuracy gap between alignment-based and alignment-free transcript quantification. *bioRxiv* 1–7 (2017). doi:10.1101/138800

370. Stuart, T. *et al.* Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902.e21 (2019).

371. 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).

372. 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

373. Street, K. *et al.* Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics* **19**, 1–16 (2018).

374. Campbell, S. A. *et al.* H3K4 trimethylation is required for postnatal pancreatic endocrine cell functional maturation. *bioRxiv* (2020). doi:https://doi.org/10.1101/2020.11.29.402990

375. Okita, N. *et al.* Modified Western blotting for insulin and other diabetes-associated peptide hormones. *Sci. Rep.* **7**, 1–11 (2017).

376. Jenuwein, T. & Allis, C. D. Translating the Histone Code. *Science (80-.).* **293**, 1074–1080 (2001).

377. Bernstein, B. E. *et al.* Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell* **120**, 169–181 (2005).

378. Chih, L. L. *et al.* Single-nucleosome mapping of histone modifications in S. cerevisiae. *PLoS Biol.* **3**, e328 (2005).

379. Zhang, X., Bernatavichute, Y. V., Cokus, S., Pellegrini, M. & Jacobsen, S. E. Genome-wide analysis of mono-, di- and trimethylation of histone H3 lysine 4 in Arabidopsis thaliana. *Genome Biol.* **10**, 1–14 (2009).

380. Hu, D. *et al.* Not All H3K4 Methylations Are Created Equal: Mll2/COMPASS Dependency in Primordial Germ Cell Specification. *Mol. Cell* **65**, 460-475.e6 (2017).

381. Wang, W. *et al.* SET-9 and SET-26 are H3K4me3 readers and play critical roles in germline development and longevity. *Elife* **7**, 1–33 (2018).

382. Lauberth, S. M. *et al.* H3K4me3 interactions with TAF3 regulate preinitiation complex assembly and selective gene activation. *Cell* **152**, 1021–1036 (2013).

383. Ruthenburg, A. J., Allis, C. D. & Wysocka, J. Methylation of Lysine 4 on Histone H3: Intricacy of Writing and Reading a Single Epigenetic Mark. *Mol. Cell* **25**, 15–30 (2007).

384. Soares, L. M. *et al.* Determinants of Histone H3K4 Methylation Patterns. *Mol. Cell* **68**, 773-785.e6 (2017).

385. Chen, K. *et al.* The Overlooked Fact: Fundamental Need for Spike-In Control for Virtually All Genome-Wide Analyses. *Mol. Cell. Biol.* **36**, 662–667 (2016).

386. Orlando, D. A. *et al.* Quantitative ChIP-Seq Normalization Reveals Global Modulation of the Epigenome Resource Quantitative ChIP-Seq Normalization Reveals Global Modulation of the Epigenome. *Cell Rep.* **9**, 1163–1170 (2014).

387. Zhang, J. *et al.* An integrative ENCODE resource for cancer genomics. *Nat. Commun.* **11**, 3696 (2020).

388. Taruttis, F. *et al.* External calibration with Drosophila whole cell spike ins delivers

absolute mRNA fold changes from human RNA seq and qPCR data. *Biotechniques* **62**, 53–61 (2017).

389. Heintzman, N. D. *et al.* Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat. Genet.* **39**, 311–8 (2007).

390. 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).

391. Kennison, J. A. The Polycomb and trithorax group proteins of Drosophila: Trans-Regulators of Homeotic Gene Function. *Annu. Rev. Genet.* **29**, 289–303 (1995).

392. Collins, B. E., Sweatt, J. D. & Greer, C. B. Broad domains of histone 3 lysine 4 trimethylation are associated with transcriptional activation in CA1 neurons of the hippocampus during memory formation. *Neurobiol. Learn. Mem.* **161**, 149–157 (2019).

393. Bramswig, N. C. *et al.* Epigenomic plasticity enables human pancreatic a to b cell reprogramming. *J. Clin. Invest.* **123**, 1275–1284 (2013).

394. Camunas-Soler, J. *et al.* Patch-Seq Links Single-Cell Transcriptomes to Human Islet Dysfunction in Diabetes. *Cell Metab.* **31**, 1017-1031.e4 (2020).

395. Oppenländer, L. *et al.* Vertical sleeve gastrectomy triggers fast β-cell recovery upon overt diabetes. *Mol. Metab.* **54**, 101330 (2021).

396. Orlando, D. A. *et al.* Quantitative ChIP-Seq normalization reveals global modulation of the epigenome. *Cell Rep.* **9**, 1163–1170 (2014).

397. Taruttis, F. *et al.* External calibration with Drosophila whole-cell spike-ins delivers absolute mRNA fold changes from human RNA-Seq and qPCR data. *Biotechniques* **62**, 53–61 (2017).

398. Chagoyen, M. & Poyatos, J. F. Complex genetic and epigenetic regulation deviates gene expression from a unifying global transcriptional program. *PLoS Comput. Biol.* **15**, e1007353 (2019).

399. Wang, C. *et al.* Enhancer priming by H3K4 methyltransferase MLL4 controls cell fate transition. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 11871–11876 (2016).

400. Martin, D. G. E. *et al.* The Yng1p Plant Homeodomain Finger Is a Methyl-Histone Binding Module That Recognizes Lysine 4-Methylated Histone H3. *Mol. Cell. Biol.* **26**, 7871–7879 (2006).

401. Taverna, S. D. *et al.* Yng1 PHD Finger Binding to H3 Trimethylated at K4 Promotes NuA3 HAT Activity at K14 of H3 and Transcription at a Subset of Targeted ORFs. *Mol. Cell* **24**, 785–796 (2006).

402. Bian, C. *et al.* Sgf29 binds histone H3K4me2/3 and is required for SAGA complex recruitment and histone H3 acetylation. *EMBO J.* **30**, 2829–2842 (2011).

403. Li, J., Liu, Y., Kim, T. H., Min, R. & Zhang, Z. Gene expression variability within and between human populations and implications toward disease susceptibility. *PLoS Comput. Biol.* **6**, (2010).

404. Ecker, S., Pancaldi, V., Valencia, A., Beck, S. & Paul, D. S. Epigenetic and Transcriptional Variability Shape Phenotypic Plasticity. *BioEssays* **40**, 1–11 (2018).

405. Rangaraju, S. *et al.* Suppression of transcriptional drift extends C. elegans lifespan by postponing the onset of mortality. *Elife* **4**, 1–39 (2015).

406. Benninger, R. K. P. & Kravets, V. The physiological role of β-cell heterogeneity in pancreatic islet function. *Nat. Rev. Endocrinol.* **0123456789**, (2021).

407. Fukaya, T., Lim, B. & Levine, M. Enhancer Control of Transcriptional Bursting. *Cell* **166**,

358–368 (2016).

408. Larsson, A. J. M. *et al.* Genomic encoding of transcriptional burst kinetics. *Nature* **565**, 251–254 (2019).

409. Chatterjee, S., Khunti, K. & Davies, M. J. Type 2 diabetes. *Lancet* **389**, 2239–2251 (2017).

410. Fadista, J. *et al.* Global genomic and transcriptomic analysis of human pancreatic islets reveals novel genes influencing glucose metabolism. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 13924–13929 (2014).

411. Ling, C. & Groop, L. Epigenetics: A molecular link between environmental factors and type 2 diabetes. *Diabetes* **58**, 2718–2725 (2009).

412. Vanderkruk, B. & Hoffman, B. G. Metabolism as a central regulator of β-cell chromatin state. *FEBS J.* 1–11 (2020). doi:10.1111/febs.15562

413. Kim, H. *et al.* PRMT1 is required for the maintenance of mature β-cell identity. *Diabetes* **69**, 355–368 (2020).

414. Lu, T. T. H. *et al.* The Polycomb-Dependent Epigenome Controls β Cell Dysfunction, Dedifferentiation, and Diabetes. *Cell Metab.* **27**, 1294-1308.e7 (2018).

415. Zhang, L. *et al.* CBP/p300 HAT maintains the gene network critical for β cell identity and functional maturity. *Cell Death Dis.* **12**, 1–12 (2021).

416. Kuo, T. *et al.* Identification of C2CD4A as a human diabetes susceptibility gene with a role in β cell insulin secretion. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 20033–20042 (2019).

417. De Jesus, D. F. *et al.* m6A mRNA methylation regulates human β-cell biology in physiological states and in type 2 diabetes. *Nat. Metab.* **1**, 765–774 (2019).

418. Dai, Z., Mentch, S. J., Gao, X., Nichenametla, S. N. & Locasale, J. W. Methionine

metabolism influences genomic architecture and gene expression through H3K4me3 peak width. *Nat. Commun.* **9**, 1–12 (2018).

419. Neelankal John, A., Ram, R. & Jiang, F.-X. RNA-Seq Analysis of Islets to Characterise the Dedifferentiation in Type 2 Diabetes Model Mice db/db. *Endocr. Pathol.* **29**, 207–221 (2018).

420. Hsiao, L. L. *et al.* A compendium of gene expression in normal human tissues. *Physiol. Genomics* **2002**, 97–104 (2002).

421. Cahill, K. M., Huo, Z., Tseng, G. C., Logan, R. W. & Seney, M. L. Improved identification of concordant and discordant gene expression signatures using an updated rank-rank hypergeometric overlap approach. *Sci. Rep.* **8**, 1–11 (2018).

422. Zhang, J. *et al.* The diabetes gene Hhex maintains δ -cell differentiation and islet function. *Genes Dev.* **28**, 829–834 (2014).

423. Sabatini, P. V. *et al.* Neuronal PAS Domain Protein 4 Suppression of Oxygen Sensing Optimizes Metabolism during Excitation of Neuroendocrine Cells. *Cell Rep.* **22**, 163–174 (2018).

424. 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).

425. Dalgaard, K. *et al.* Trim28 Haploinsufficiency Triggers Bi-stable Epigenetic Obesity. *Cell* **164**, 353–364 (2016).

426. Campbell, S. A. & Hoffman, B. G. Chromatin Regulators in Pancreas Development and Diabetes. *Trends Endocrinol. Metab.* **27**, 142–152 (2016).

427. Fan, J., Krautkramer, K. A., Feldman, J. L. & Denu, J. M. Metabolic regulation of histone post-translational modifications. *ACS Chem. Biol.* **10**, 95–108 (2015).

428. Alonso, L. C. *et al.* Glucose infusion in mice: A new model to induce β-cell replication. *Diabetes* **56**, 1792–1801 (2007).

429. Yang, Z. *et al.* The DPY30 subunit in SET1 / MLL complexes regulates the proliferation and differentiation of hematopoietic progenitor cells. *Blood* **124**, 2025–2034 (2016).

430. Simboeck, E. *et al.* DPY30 regulates pathways in cellular senescence through ID protein expression. *EMBO J.* **32**, 2217–30 (2013).

431. Jiang, H. *et al.* Role for Dpy-30 in ES cell-fate specification by regulation of H3K4 methylation within bivalent domains. *Cell* **144**, 513–525 (2011).

432. Mato, J. M., Martínez-Chantar, M. L. & Lu, S. C. Methionine metabolism and liver disease. *Annu. Rev. Nutr.* **28**, 273–293 (2008).

433. Poirier, L. A. *et al.* Blood S-adenosylmethionine concentrations and lymphocyte methylenetetrahydrofolate reductase activity in diabetes mellitus and diabetic nephropathy. *Metabolism.* **50**, 1014–1018 (2001).

434. Rees, W. D. Manipulating the sulfur amino acid content of the early diet and its implications for long-term health. *Proc. Nutr. Soc.* **61**, 71–77 (2002).

435. Varela-Rey, M. *et al.* S-adenosylmethionine levels regulate the schwann cell DNA methylome. *Neuron* **81**, 1024–1039 (2014).

436. Okun, J. G. *et al.* Liver alanine catabolism promotes skeletal muscle atrophy and hyperglycaemia in type 2 diabetes. *Nat. Metab.* **3**, 394–409 (2021).

437. Balaghi, M., Horne, D. W. & Wagner, C. Hepatic one-carbon metabolism in early folate deficiency in rats. *Biochem. J.* **291**, 145–149 (1993).

438. Tang, X. *et al.* Comprehensive Profiling of Amino Acid Response Uncovers Unique Methionine-Deprived Response Dependent on Intact Creatine Biosynthesis. *PLoS Genet.*

11, 1–28 (2015).

439. Li, Z. *et al.* Methionine metabolism in chronic liver diseases: an update on molecular mechanism and therapeutic implication. *Signal Transduct. Target. Ther.* **5**, 1–9 (2020).

440. Balaghi, M. & Wagner, C. Biochemical and Molecular Roles of Nutrients Methyl Group Metabolism in the Pancreas of Folate-Deficient Rats. *J. Nutr.* **122**, 1391–1396 (1992).

441. Balaghi, M., Horne, D. W., Woodward, S. C. & Wagner, C. Pancreatic one-carbon metabolism in early folate deficiency in rats. *Am. J. Clin. Nutr.* **58**, 198–203 (1993).

442. Mahajan, A. *et al.* Effect of imbalance in folate and vitamin B12 in maternal/parental diet on global methylation and regulatory miRNAs. *Sci. Rep.* **9**, 1–21 (2019).

443. Nilsson, E. *et al.* Epigenetic alterations in human liver from subjects with type 2 diabetes in parallel with reduced folate levels. *J. Clin. Endocrinol. Metab.* **100**, E1491–E1501 (2015).

444. Pogribny, I. P., Tryndyak, V. P., Muskhelishvili, L., Rusyn, I. & Ross, S. A. Methyl deficiency, alterations in global histone modifications, and carcinogenesis. *J. Nutr.* **137**, 216S–222S (2007).

445. Bistulfi, G., VanDette, E., Matsui, S. I. & Smiraglia, D. J. Mild folate deficiency induces genetic and epigenetic instability and phenotype changes in prostate cancer cells. *BMC Biol.* **8**, 1–12 (2010).

446. Ulanovskaya, O. A., Zuhl, A. M. & Cravatt, B. F. NNMT promotes epigenetic remodeling in cancer by creating a metabolic methylation sink. *Nat. Chem. Biol.* **9**, 300–306 (2013).

447. Arts, R. J. W. *et al.* Glutaminolysis and Fumarate Accumulation Integrate Immunometabolic and Epigenetic Programs in Trained Immunity. *Cell Metab.* **24**, 807–819 (2016).

448. Sperber, H. *et al.* The metabolome regulates the epigenetic landscape during naive-to-primed human embryonic stem cell transition. *Nat. Cell Biol.* **17**, 1523–1535 (2015).

449. Reeves, P. G. & Suppl, M. Symposium : Animal Diets for Nutritional and Toxicological Research Components of the AIN-93 Diets as Improvements in the AIN-76A Diet 1 , 2. *Exp. Biol.* **127**, 838–841 (1997).

450. Berwick, D. C., Heaton, G. R., Azeggagh, S. & Harvey, K. LRRK2 Biology from structure to dysfunction: Research progresses, but the themes remain the same. *Mol. Neurodegener.* **14**, 1–22 (2019).

451. Glier, M. B., Green, T. J. & Devlin, A. M. Methyl nutrients, DNA methylation, and cardiovascular disease. *Mol. Nutr. Food Res.* **58**, 172–182 (2014).

452. Tang, S. *et al.* Methionine metabolism is essential for SIRT 1-regulated mouse embryonic stem cell maintenance and embryonic development . *EMBO J.* **36**, 3175–3193 (2017).

453. Luka, Z., Mudd, S. H. & Wagner, C. Glycine N-methyltransferase and regulation of S-adenosylmethionine levels. *J. Biol. Chem.* **284**, 22507–22511 (2009).

454. Li, J. *et al.* Association between serum folate and insulin resistance among U.S. nondiabetic adults. *Sci. Rep.* **7**, 1–7 (2017).

455. Asbaghi, O. *et al.* Folic acid supplementation improves glycemic control for diabetes prevention and management: A systematic review and dose-response meta-analysis of randomized controlled trials. *Nutrients* **13**, 2355 (2021).

456. Sudchada, P. *et al.* Effect of folic acid supplementation on plasma total homocysteine levels and glycemic control in patients with type 2 diabetes: A systematic review and meta-analysis. *Diabetes Res. Clin. Pract.* **98**, 151–158 (2012).

457. Zhao, J. V., Schooling, C. M. & Zhao, J. X. The effects of folate supplementation on

glucose metabolism and risk of type 2 diabetes: a systematic review and meta-analysis of randomized controlled trials. *Ann. Epidemiol.* **28**, 249-257.e1 (2018).

458. Cagnacci, A. *et al.* Folate administration decreases oxidative status and blood pressure in postmenopausal women. *Eur. J. Nutr.* **54**, 429–435 (2015).

459. Mursleen, M. T. & Riaz, S. Implication of homocysteine in diabetes and impact of folate and vitamin B12 in diabetic population. *Diabetes Metab. Syndr. Clin. Res. Rev.* **11**, S141–S146 (2017).

460. Newsholme, P., Keane, K. N., Carlessi, R. & Cruzat, V. Oxidative stress pathways in pancreatic β-cells and insulin-sensitive cells and tissues: Importance to cell metabolism, function, and dysfunction. *Am. J. Physiol. - Cell Physiol.* **317**, C420–C433 (2019).

461. Yu, X., Huang, Y., Hu, Q. & Ma, L. Hyperhomocysteinemia stimulates hepatic glucose output and PEPCK expression. *Acta Biochim. Biophys. Sin. (Shanghai).* **41**, 1027–1032 (2009).

462. Gonen, N. & Assaraf, Y. G. Antifolates in cancer therapy: Structure, activity and mechanisms of drug resistance. *Drug Resist. Updat.* **15**, 183–210 (2012).

463. Wollheim, C. B. & Maechler, P. Beta-Cell Mitochondria and Insulin Secretion: Messenger role of nucleotides and metabolites. *Diabetes* **51**, 37–42 (2002).

464. Boucher, J., Kleinridders, A. & Kahn, C. R. Insulin Receptor Signaling in Normal and Insulin-Resistant States. *Cold Spring Harb. Perspect. Biol.* **6**, a009191 (2014).

465. Devlin, A. M. *et al.* Effect of Mthfr genotype on diet-induced hyperhomocysteinemia and vascular function in mice. *Blood* **103**, 2624–2629 (2004).

466. Crider, K. S., Yang, T. P., Berry, R. J. & Bailey, L. B. Folate and DNA methylation: A review of molecular mechanisms and the evidence for Folate's role. *Adv. Nutr.* **3**, 21–38

(2012).

467. Steffen, P. A. & Ringrose, L. What are memories made of? How polycomb and trithorax proteins mediate epigenetic memory. *Nat. Rev. Mol. Cell Biol.* **15**, 340–356 (2014).

468. Nakamura, T. *et al.* ALL-1 is a histone methyltransferase that assembles a supercomplex of proteins involved in transcriptional regulation. *Mol. Cell* **10**, 1119–1128 (2002).

469. Klymenko, T. & Jürg, M. The histone methyltransferases Trithorax and Ash1 prevent transcriptional silencing by Polycomb group proteins. *EMBO Rep.* **5**, 373–377 (2004).

470. Kaufmann, K. *et al.* Balancing of Histone H3K4 Methylation States by the Kdm5c / SMCX Histone Demethylase Modulates Promoter and Enhancer Function. *Cell Rep.* **3**, 1071–1079 (2013).

471. van Dijk, S. J., Tellam, R. L., Morrison, J. L., Muhlhausler, B. S. & Molloy, P. L. Recent developments on the role of epigenetics in obesity and metabolic disease. *Clin. Epigenetics* **7**, 66 (2015).

472. Dayeh, T. *et al.* Genome-Wide DNA Methylation Analysis of Human Pancreatic Islets from Type 2 Diabetic and Non-Diabetic Donors Identifies Candidate Genes That Influence Insulin Secretion. *PLoS Genet.* **10**, (2014).

473. Abou Ziki, R., Luo, Y., Vlaeminck-Guillem, V., Le Romancer, M. & Zhang, C. X. Involvement of the MEN1 Gene in Hormone-Related Cancers: Clues from Molecular Studies, Mouse Models, and Patient Investigations. *Endocrines* **1**, 58–81 (2020).

474. Gupta, S. *et al.* Histone Methylation Regulates Memory Formation. *J. Neurosci.* **30**, 3589–3599 (2010).

475. El-Hattab, A. W. Serine biosynthesis and transport defects. *Mol. Genet. Metab.* **118**, 153–159 (2016).

476. Zheng, Y. *et al.* Mitochondrial One-Carbon Pathway Supports Cytosolic Folate Integrity in Cancer Cells. *Cell* **175**, 1546-1560.e17 (2018).

477. Fan, J. *et al.* Quantitative flux analysis reveals folate-dependent NADPH production. *Nature* **510**, 298–302 (2014).

478. Barua, S., Kuizon, S., Brown, W. T. & Junaid, M. A. DNA methylation profiling at single-base resolution reveals gestational folic acid supplementation influences the epigenome of mouse offspring cerebellum. *Front. Neurosci.* **10**, 168 (2016).

479. Yang, L. *et al.* Serine Catabolism Feeds NADH when Respiration Is Impaired. *Cell Metab.* **31**, 809-821.e6 (2020).

480. TORRES, L. *et al.* Liver-specific methionine adenosyltransferase MAT1A gene expression is associated with a specific pattern of promoter methylation and histone acetylation: implications for MAT1A silencing during transformation . *FASEB J.* **14**, 95–102 (2000).

481. Ramani, K. & Lu, S. C. Methionine adenosyltransferases in liver health and diseases. *Liver Res.* **1**, 103–111 (2017).

482. Lu, S. C. *et al.* Role of S-adenosylmethionine in two experimental models of pancreatitis. *FASEB J.* **17**, 56–58 (2003).

Appendices

Appendix A | Diet formulations used in folic acid-restriction studies.

Product #	A20071401	A20071402
	Normal folic acid (NF)	Folic acid-restricted (FR)
Ingredient	**gm**	**gm**
L-Cystine	4.2	4.2
L-Isoleucine	7.6	7.6
L-Leucine	15.8	15.8
L-Lysine	13.2	13.2
L-Methionine	5.1	5.1
L-Phenylalanine	8.4	8.4
L-Threonine	7.2	7.2
L-Tryptophan	2.1	2.1
L-Valine	9.3	9.3
L-Histidine	4.6	4.6
L-Alanine	5.1	5.1
L-Arginine	6.0	6.0
L-Aspartic Acid	12.1	12.1
L-Glutamic Acid	38.2	38.2
Glycine	3.0	3.0
L-Proline	17.8	17.8
L-Serine	10.0	10.0
L-Tyrosine	9.2	9.2
Total L-Amino Acids	*178.9*	*178.9*
Corn Starch	381	381
Maltodextrin 10	110	110
Dextrose	150	150
Cellulose, BW200	75	75
Inulin	25	25
Soybean Oil	70	70
Mineral Mix S10026	10	10
DiCalcium	13	13

Phosphate		
Calcium Carbonate	5.5	5.5
Potassium Citrate, 1 H2O	16.5	16.5
Sodium Bicarbonate	**7.5**	**7.5**
Vitamin Mix V10001 (w/ Folic Acid)	**10**	0
Vitamin Mix V14901 (w/o Folic Acid)	0	**10**
Choline Bitartrate	2	2
Folic Acid	0	**0.0002**
Succinylsulfathiazole	**10.5**	**10.5**
FD&C Yellow Dye #5	0.05	0
FD&C Red Dye #40	0	0.05
FD&C Blue Dye #1	0	0
Total	**1064.95**	**1064.95**
Diet #	**A20071401**	**A20071402**
gm		
Protein	178.9	178.9
Carbohydrate	660.4	660.4
Fat	70.0	70.0
Fiber	100.0	100.0
gm %		
Protein	16.8	16.8
Carbohydrate	62.0	62.0
Fat	6.6	6.6
Fiber	9.4	9.4
kcal		
Protein	716	716
Carbohydrate	2642	2642
Fat	630	630
Total	3987	3987

kcal %		
Protein	18	18
Carbohydrate	66	66
Fat	16	16
Total	100	100
kcal / gm	3.7	3.7
Succinylsulfathiazole (%)	**1.0**	**1.0**
Folic Acid (mg/3987 kcal)	**2**	**0.2**

Appendix B | Dpy30-KO/WT β-cell differentially expressed genes

B.1 15-days post-tamoxifen

Downregulated $P_{adj} < 0.01$, fold-change > 2

geneSymbol	baseMean	padj
Dpy30	375.58	1.37E-19
Edn3	270.10	0.000151
C3	127.74	0.003238

B.2 45-days post-tamoxifen

Downregulated $P_{adj} < 0.01$, fold-change > 2

geneSymbol	baseMean	padj
Jag2	340.40	9.66E-71
Col6a3	276.56	9.13E-70
Tmsb10	512.81	9.13E-70
Spred3	223.85	4.69E-66
Lrrc24	254.37	8.06E-62
Tceal8	1448.46	2.06E-61
Mlph	133.14	2.48E-54
As3mt	1034.84	2.84E-48
Kcnj11	2432.59	1.98E-46
Thbd	178.32	1.70E-45
Syt3	161.56	2.38E-45
Cldn3	249.32	1.40E-43
Ptges	193.09	4.20E-41
Cyp4f39	409.87	1.47E-40
Fgd2	534.22	3.90E-39
Rab34	275.15	3.26E-38
Slc12a4	98.90	4.39E-38
Agtrap	834.42	5.28E-38
Ffar1	2632.46	5.28E-38
Gm281	155.60	2.72E-37
Trp53i11	422.53	1.93E-35
Cldn7	768.93	4.47E-35
Oxtr	308.35	9.01E-35
Rab37	2579.99	9.48E-35
Col9a2	66.05	1.38E-33
Nat14	166.74	2.13E-33
Tmem59l	823.91	3.02E-33
Eef1akmt3	188.52	9.12E-33
Npr1	386.29	7.98E-32
Cpm	378.51	3.94E-31
Atf5	3013.02	5.67E-31
Dpy30	308.64	6.50E-31
Emilin1	2356.60	7.50E-31
Adh1	2168.53	1.49E-29
Hspa2	424.38	4.26E-29
Tmem121b	118.02	3.58E-27
Scrt1	350.09	5.28E-27
Snhg18	164.90	8.64E-27
Cd44	726.74	4.01E-26
Zfp516	3166.63	8.25E-26
Dnhd1	131.74	1.53E-25
Tmem215	4764.61	5.25E-25
Dlg4	1075.20	6.36E-25
Kcnab3	295.65	5.75E-24
Igfbp5	232.47	1.17E-23
Gpr119	310.48	1.29E-22
Slc17a9	373.61	3.18E-22
Gpr162	98.41	9.94E-22
Nupr1	891.74	1.50E-21
Cspg5	104.82	2.41E-21
Nfic	1429.08	2.63E-21
Dusp26	373.88	6.25E-21
Tmem38a	924.48	7.93E-21
Arhgef25	219.99	1.02E-20
Slc2a2	12428.44	2.72E-20
Gys1	234.73	3.00E-20
Gstm1	720.36	3.51E-20
Kdelr3	114.55	4.07E-20
Zmynd8	2257.38	5.25E-20
Arpc1b	316.08	5.48E-20
Fam83g	697.51	5.51E-20
Vwa1	70.51	1.25E-19
Nynrin	435.80	2.12E-19
Mpz	151.61	2.19E-19
Myt1	1251.85	4.49E-19
Tubb4a	200.96	5.09E-19
Zfp239	63.69	5.88E-19

Lgi1	892.06	8.10E-19
Nnat	402.84	9.87E-19
Rnf208	108.58	1.02E-18
Dgkg	112.36	1.15E-18
Kif21b	43.85	1.55E-18
Arhgef17	2084.51	1.65E-18
Kdm2b	1415.35	1.89E-18
Sfrp5	973.41	1.91E-18
Ceacam1	166.46	5.32E-18
Mxra8	157.86	5.73E-18
Stat6	433.63	1.37E-17
Il11ra1	338.75	2.14E-17
Stx16	872.52	3.92E-17
H2bu2	116.03	7.25E-17
Crip1	168.48	7.54E-17
Mxra7	121.05	8.65E-17
Sult4a1	165.21	1.22E-16
Rassf2	57.85	1.68E-16
Rnf112	98.97	1.69E-16
Resp18	14861.93	1.82E-16
Entpd2	44.77	3.14E-16
Derl3	270.66	3.47E-16
Kcnh3	69.93	1.30E-15
Tmem151b	121.78	1.88E-15
B3gnt9	297.55	2.00E-15
Alox8	64.98	2.16E-15
Gm12397	79.97	3.11E-15
Rhobtb1	2234.93	3.11E-15
Cadm4	199.46	3.34E-15
Tnfsf12	301.22	3.86E-15
Nol3	212.67	7.61E-15
Tmem30b	694.30	7.74E-15
Serhl	171.75	9.69E-15
Bco1	134.35	1.63E-14
Usp35	1219.12	1.63E-14
Gm11769	117.39	1.80E-14
Elf4	102.59	2.70E-14
Tspyl3	98.94	4.55E-14
Serinc2	93.30	4.64E-14
Mfsd9	47.35	5.47E-14
Cpne4	73.19	6.39E-14
Efna2	78.11	6.85E-14
Prrt3	70.71	6.85E-14

Syngr3	48.87	7.05E-14
Gm15651	101.71	7.22E-14
Tceal3	1720.13	7.36E-14
Fam189b	249.34	1.03E-13
Clip3	603.74	1.06E-13
Dbn1	197.60	1.35E-13
Hps1	348.17	2.17E-13
Trip10	102.25	2.52E-13
Arhgef10	132.13	2.62E-13
Wnt5b	47.05	2.82E-13
P3h3	161.56	3.15E-13
Degs2	165.31	3.35E-13
Dlec1	75.26	3.42E-13
Wdr17	52.50	3.91E-13
Galnt4	843.14	4.24E-13
Agap2	123.21	6.14E-13
Scnn1a	49.81	6.14E-13
Arhgef19	210.34	6.29E-13
Slc38a5	96.82	6.90E-13
Adcy3	92.25	7.90E-13
Bik	50.05	1.12E-12
Samd14	201.51	1.59E-12
Rab3d	3927.74	1.81E-12
B4galt2	162.14	1.98E-12
Arg1	683.38	2.06E-12
Col4a1	850.96	2.21E-12
Unc93b1	337.96	2.29E-12
Tspoap1	209.19	2.34E-12
Col6a6	997.35	2.43E-12
Nxph3	38.95	3.55E-12
Plut	270.50	4.42E-12
Gm45137	33.22	5.28E-12
Dhdh	443.43	5.60E-12
Muc4	886.71	6.13E-12
Trim46	230.91	6.97E-12
Pex11g	267.60	7.15E-12
Iglon5	45.87	1.16E-11
Ggnbp1	69.04	1.23E-11
Ccdc8	194.04	1.36E-11
Gcgr	519.37	1.59E-11
Ache	184.53	1.88E-11
Plekha4	120.18	1.98E-11
Ajm1	153.29	2.13E-11

Hykk	77.63	2.16E-11
Cldn6	174.67	2.31E-11
Ubxn10	993.18	2.83E-11
Fndc5	135.00	2.99E-11
Gm17767	351.46	3.12E-11
Ass1	96.38	3.35E-11
Gm45847	359.00	3.64E-11
Ttc28	2125.95	3.80E-11
2700081O15Rik	393.31	4.17E-11
Upk3a	30.98	4.53E-11
Rassf4	716.57	7.11E-11
AI463170	35.90	8.00E-11
Glp1r	12929.15	1.02E-10
Prr7	56.38	1.10E-10
Usp11	429.60	1.21E-10
Hsf4	88.03	1.93E-10
St3gal4	83.34	2.69E-10
Serpini1	992.82	2.99E-10
Arap1	1062.68	3.23E-10
2610528J11Rik	26.12	3.34E-10
4930578M01Rik	151.21	3.68E-10
Cd164l2	125.95	4.48E-10
Ddx25	503.40	5.88E-10
Cfap74	198.10	6.23E-10
Cdkn1c	33.26	7.16E-10
Pde4a	112.78	7.41E-10
Renbp	40.77	7.49E-10
Il17re	235.00	7.85E-10
Slc30a8	14436.74	9.07E-10
Emp2	70.41	1.01E-09
Gprc5b	169.60	1.10E-09
Cox6a2	637.98	1.11E-09
Snhg11	365.09	1.52E-09
Kctd12b	721.45	1.55E-09
Ptk7	58.28	1.80E-09
Nbl1	368.90	1.97E-09
Rom1	249.08	2.10E-09
Dtnbos	195.24	2.20E-09
Dll4	189.17	2.40E-09
Cd81	1645.94	2.60E-09
Baiap3	2389.57	2.86E-09
Kcnj12	203.67	3.51E-09
Clip4	124.83	3.59E-09

Tcf7l1	84.61	3.72E-09
Sema6b	137.27	3.99E-09
Prss41	25.99	4.22E-09
Gm45606	354.89	4.45E-09
Cchcr1	136.82	4.69E-09
Pnma2	47.80	4.73E-09
Myom1	63.97	5.37E-09
1810010H24Rik	104.53	5.75E-09
Dmtn	49.74	5.79E-09
Egf	447.00	5.86E-09
Oit1	33.34	5.94E-09
Ush1g	24.38	6.61E-09
1810034E14Rik	69.98	6.62E-09
Caly	65.83	7.00E-09
Rcor2	134.89	7.02E-09
Gstm2	271.05	7.04E-09
L1cam	89.37	7.97E-09
Rapgef3	58.02	8.21E-09
Edn3	174.50	8.69E-09
Ttc12	29.15	8.74E-09
Nrxn2	35.07	8.80E-09
Hhatl	37.99	1.14E-08
Ifitm2	252.00	1.19E-08
St6galnac2	31.40	1.45E-08
Tuba1c	246.88	1.51E-08
Actn2	37.64	1.71E-08
Rcn3	143.50	1.76E-08
Kcnd1	118.97	1.86E-08
Jph4	117.76	1.91E-08
Gm14261	41.83	1.91E-08
Gnat2	549.63	1.94E-08
Itga9	36.86	2.04E-08
Npr2	390.51	2.27E-08
Syt17	30.46	2.38E-08
Tmc6	269.27	3.47E-08
Ctsz	838.77	3.67E-08
Grin1	931.90	3.70E-08
Inava	114.62	3.70E-08
Pxdn	99.83	4.05E-08
Fbxo27	82.12	4.14E-08
Dync1i1	99.87	4.16E-08
Ffar3	206.86	5.47E-08
Pycr1	553.15	5.87E-08

Foxred2	509.49	6.16E-08
Neu2	39.11	6.33E-08
Gm49492	169.88	6.37E-08
Trpm5	1266.34	6.59E-08
Stap2	37.69	7.73E-08
Agrn	182.78	9.87E-08
Zfr2	309.61	1.18E-07
Rin2	287.00	1.20E-07
Cxcl16	27.13	1.23E-07
Mfsd4b1	70.04	1.24E-07
Gm30173	142.05	1.35E-07
Acsf2	120.84	1.36E-07
Sall4	145.82	1.41E-07
Medag	28.76	1.58E-07
Hspb1	336.32	1.60E-07
Psd	729.93	1.67E-07
Hmgb3	125.19	1.79E-07
Sox12	372.01	1.79E-07
Rhd	93.79	1.88E-07
C130021I20Rik	95.43	2.76E-07
Ikzf3	114.41	2.85E-07
Mif4gd	245.66	2.85E-07
Zcchc12	293.94	2.91E-07
Aatk	819.38	3.03E-07
Ier3	81.61	3.37E-07
Unkl	83.99	3.38E-07
Ctf1	106.50	3.52E-07
Flywch2	41.68	3.82E-07
Cilp	82.36	4.12E-07
Dnajc24	688.94	4.46E-07
Kndc1	49.69	4.66E-07
Frrs1l	102.69	4.72E-07
Gm50353	77.62	4.85E-07
Gm37459	100.13	4.97E-07
Tekt2	83.84	5.02E-07
Chp2	97.49	5.70E-07
Zfp583	44.33	5.92E-07
Minar2	27.90	6.47E-07
3-Sep	179.61	7.07E-07
Scn1b	71.96	9.05E-07
Doc2b	37.37	9.44E-07
Lin28a	47.73	9.59E-07
AW011738	46.20	9.61E-07

Rtkn	99.30	1.04E-06
Clmn	98.75	1.05E-06
Nectin4	44.89	1.08E-06
Neil2	336.69	1.22E-06
Per2	553.81	1.40E-06
C1qtnf4	209.51	1.40E-06
Gm13421	100.34	1.58E-06
Bok	35.93	1.65E-06
Il3ra	109.87	1.65E-06
Tcim	839.51	1.70E-06
Bdnf	19.58	1.92E-06
Hspa12b	128.66	2.20E-06
1810044D09Rik	51.30	2.22E-06
Plekhg6	28.71	2.25E-06
Pnck	125.99	2.32E-06
Ntm	88.33	2.37E-06
Krt80	65.25	2.41E-06
Pyroxd2	94.73	2.53E-06
Ddah2	25.40	2.63E-06
Dnaaf4	47.45	2.67E-06
Map3k21	295.30	2.80E-06
Igdcc4	88.87	2.86E-06
Lrrc27	128.07	3.03E-06
Vamp5	19.69	3.10E-06
Vgf	700.08	3.20E-06
Cst6	69.23	3.50E-06
Carmil2	165.89	3.71E-06
Ppp1r3b	119.27	3.82E-06
Syt2	307.79	4.12E-06
Necab2	703.34	4.30E-06
Zim1	1065.44	4.43E-06
Cnr1	65.15	4.46E-06
Shfl	1146.48	4.80E-06
Gm44850	109.88	5.06E-06
Sytl1	617.89	5.68E-06
Wwc2	26.46	5.75E-06
Ankrd24	899.00	5.82E-06
Scarf2	91.09	5.90E-06
Slc8a2	140.28	6.08E-06
Gm43909	35.27	6.24E-06
Hapln4	2124.56	7.28E-06
Kcnip2	141.78	7.42E-06
Mapk13	67.11	8.82E-06

Aldh1l2	976.79	8.82E-06
Pabpc1l	53.42	8.99E-06
5330434G04Rik	148.64	9.30E-06
Tmem179	95.68	9.65E-06
Ush2a	784.25	1.06E-05
Gm19426	15.39	1.06E-05
Spock3	144.24	1.12E-05
Adnp2	475.76	1.16E-05
Slc28a2	43.56	1.19E-05
Gm20400	53.16	1.19E-05
Cpne2	21.48	1.19E-05
Mfap2	42.72	1.19E-05
Mthfr	834.96	1.19E-05
Rell2	121.83	1.23E-05
Msi1	215.32	1.39E-05
Gm26737	20.78	1.43E-05
Gm28523	16.27	1.45E-05
Gstt2	132.01	1.49E-05
Atp6v1g2	69.86	1.51E-05
Slc7a2	5977.16	1.58E-05
Grin2c	256.83	1.59E-05
Nrep	499.60	1.66E-05
Hspa1a	20177.01	1.66E-05
Selenbp1	39.87	1.73E-05
Adamts10	20.72	1.98E-05
Nek5	101.70	1.99E-05
Gpr173	58.48	2.03E-05
Fmo1	99.27	2.08E-05
Dlk1	97.85	2.23E-05
1810021B22Rik	32.01	2.37E-05
Fgfr4	21.87	2.50E-05
Gstk1	107.26	2.63E-05
Aqp4	255.05	2.65E-05
Epha8	34.00	2.82E-05
Crybg2	31.13	2.83E-05
Asb16	20.49	2.88E-05
Tas1r3	19.84	2.88E-05
Lmtk3	610.60	3.92E-05
Gm15417	131.27	3.97E-05
Rbp4	2655.50	3.99E-05
Cyp2d22	68.60	3.99E-05
Gm10941	92.00	4.04E-05
Cthrc1	63.06	4.22E-05

Acer2	58.86	4.28E-05
Fkbp11	822.02	4.29E-05
Fcgbp	22.15	4.30E-05
Cpt1c	230.52	4.34E-05
Ednra	31.57	4.45E-05
Gm46123	61.60	4.70E-05
Fgf11	22.15	4.90E-05
4-Sep	47.12	4.93E-05
Pdzrn3	36.61	4.95E-05
Clec11a	57.87	4.99E-05
Gm30648	542.89	5.92E-05
B3gat2	19.39	6.00E-05
Acta1	20.46	6.08E-05
E530011L22Rik	47.23	6.28E-05
Gprasp2	1839.44	7.28E-05
2610035D17Rik	120.24	7.39E-05
Chst14	37.49	7.73E-05
Coro2a	64.52	7.86E-05
Cbs	1460.59	8.04E-05
Aqp7	28.92	8.22E-05
Cacna1b	203.55	8.35E-05
AC149090.1	4117.28	8.64E-05
Amt	221.94	8.71E-05
Smco4	45.30	8.87E-05
Plin3	94.00	9.00E-05
Ppp4r1l-ps	1299.00	9.45E-05
Kazald1	77.05	1.08E-04
Myo3a	26.56	1.08E-04
Dixdc1	713.66	1.10E-04
P2ry12	60.26	1.13E-04
Klhl33	26.35	1.17E-04
Ttll10	71.27	1.20E-04
6430562O15Rik	33.25	1.20E-04
Syt16	89.53	1.24E-04
Ak5	116.63	1.29E-04
Kdm6b	858.53	1.30E-04
Gm41349	29.85	1.32E-04
Togaram2	43.55	1.37E-04
Gm46603	82.73	1.38E-04
Tmlhe	95.80	1.48E-04
Ankrd45	32.87	1.49E-04
Gfra3	52.80	1.53E-04
Tvp23a	78.62	1.56E-04

Palm	161.23	1.57E-04
Zfp951	66.26	1.57E-04
Sema5a	15.54	1.58E-04
Tmed6	255.67	1.59E-04
Dok1	80.25	1.59E-04
BC051142	72.04	1.63E-04
Rab15	49.49	1.66E-04
Oasl1	16.22	1.66E-04
Sytl3	60.41	1.67E-04
Speg	43.40	1.79E-04
Pcsk2os1	375.40	1.80E-04
Nr1h3	161.37	1.81E-04
Cacna2d2	113.79	1.86E-04
Eppk1	90.21	1.92E-04
Lmx1b	163.46	1.99E-04
Zmiz1	1276.53	1.99E-04
Nckap5l	64.00	2.00E-04
Mapk15	372.42	2.03E-04
Ikbke	246.01	2.14E-04
Nxpe4	38.16	2.18E-04
Slc7a3	19.50	2.18E-04
AC109619.1	53.80	2.19E-04
Slc12a5	256.21	2.20E-04
Nav2	2180.84	2.26E-04
Gm37859	36.89	2.29E-04
Angptl7	159.81	2.31E-04
Lrrc3	39.88	2.38E-04
Pyy	18636.08	2.42E-04
Rab26	30.95	2.50E-04
Srrm4	166.83	2.50E-04
Scd1	64.79	2.54E-04
Guca1a	21.32	2.64E-04
Shroom1	27.80	2.73E-04
Crocc	1106.94	2.95E-04
Tppp	17.70	3.09E-04
Egfl7	24.38	3.12E-04
Gm30015	24.13	3.28E-04
Adam32	53.37	3.31E-04
Ada	29.32	3.41E-04
Gm9403	31.51	3.58E-04
Cd24a	1542.38	3.66E-04
Fam129a	272.18	3.72E-04
Nol4l	345.50	4.22E-04
Gm10602	23.26	4.22E-04
Cacng4	58.14	4.26E-04
Gna14	21.23	4.31E-04
Fes	18.79	4.35E-04
Hyal1	48.74	4.35E-04
1500004A13Rik	41.98	4.62E-04
Zswim6	72.33	4.73E-04
Tagln3	34.79	4.73E-04
Chd5	611.64	5.51E-04
Zfp385a	34.57	5.72E-04
Mest	162.15	5.78E-04
Mfng	18.04	5.78E-04
Aox1	254.20	5.99E-04
Pisd-ps1	154.90	6.02E-04
A2ml1	54.91	6.16E-04
Dcdc5	42.64	6.33E-04
Kcnc3	567.26	6.63E-04
Fut4	89.66	6.95E-04
Scn3a	535.50	7.01E-04
Neat1	5602.04	7.05E-04
Gm37834	18.23	7.06E-04
Nyap1	136.55	7.09E-04
4931431B13Rik	109.35	7.13E-04
Zfp872	47.12	7.25E-04
Slc2a10	31.32	7.29E-04
Susd2	23.01	7.33E-04
Iqsec1	180.13	7.43E-04
Syngap1	736.44	7.67E-04
Gm26691	26.04	7.71E-04
Xrcc3	240.95	7.73E-04
Top1mt	25.79	8.29E-04
Gprin2	25.97	8.32E-04
Unc5c	257.61	8.41E-04
Ssc4d	83.87	8.41E-04
Catip	14.56	8.53E-04
Caps2	49.62	9.21E-04
Gm16035	17.23	9.31E-04
Rnf138rt1	424.82	9.55E-04
D430040D24Rik	17.76	9.62E-04
Gm4791	137.81	9.75E-04
Hspa1b	23664.15	9.88E-04
Rnasel	59.11	9.93E-04
9130230L23Rik	31.75	9.95E-04

Sec14l4	79.11	0.001011
9430021M05Rik	17.26	0.001054
Zfp518a	450.67	0.001072
Xrra1	31.02	0.001103
Pcdh15	558.68	0.001173
Gm50039	14.74	0.001175
Acvr1c	157.22	0.00123
Tesc	31.16	0.001234
Ldb3	21.39	0.001255
Jhy	94.12	0.001295
Mycbpap	77.83	0.001299
Pik3r5	44.33	0.001398
Tspan8	971.23	0.001444
Gm20627	19.05	0.001501
Tcf24	1395.55	0.001578
1700108N11Rik	13.74	0.001611
Shank2	76.85	0.001669
Ppy	18091.19	0.0017
Rem2	113.29	0.001755
Sh3bgrl2	90.19	0.00182
Nrip3	66.72	0.001825
Atp1a3	26.90	0.001909
Dok7	67.78	0.001999
Barhl1	20.89	0.002005
Rab29	20.34	0.002026
Astn2	13.25	0.00205
Nova2	215.75	0.002063
4632411P08Rik	28.15	0.002108
Gm45546	20.63	0.002108
Gm46565	52.94	0.002138
Uba7	204.72	0.002145
4930539E08Rik	59.98	0.002255
Gpr62	33.32	0.002273
Nsg1	44.62	0.002368
Rhov	31.76	0.002399
Crb1	165.50	0.002457
F5	145.98	0.002469
Gcnt1	35.86	0.002473
Sema4c	169.58	0.002507
Slc38a11	18.27	0.002514
Rad51ap2	116.63	0.002523
Gm3924	39.88	0.002642
Chrm4	57.91	0.002687

Wdr49	64.33	0.002788
Rbms3	162.85	0.002846
Ksr1	23.29	0.002888
Gm31518	29.83	0.002927
Gm50462	17.67	0.003184
Fev	113.04	0.003226
Afap1l1	19.18	0.003249
Wipf3	36.12	0.003249
B3galt2	306.49	0.003314
Caskin1	148.34	0.003321
Gm16283	21.49	0.003503
Nav3	95.83	0.003528
Tcirg1	144.31	0.003529
Dtna	31.56	0.003818
Esyt3	31.59	0.004164
H2-Q1	10.72	0.004219
Slc8b1	49.13	0.004324
Efnb2	26.40	0.004411
Rab3il1	19.81	0.00449
Pappa2	5015.56	0.00453
Ano9	36.46	0.004594
Dnajc22	28.17	0.004608
B3gat1	30.39	0.004613
Cfap73	15.72	0.004623
Yap1	17.25	0.004882
Efcab12	23.05	0.004941
Hhex	395.23	0.00495
Zan	25.59	0.004963
Pdk4	221.06	0.004981
C1qtnf1	25.07	0.005311
Radil	49.38	0.005615
Fam13a	16.15	0.005655
Tro	105.75	0.005703
Serpinh1	197.65	0.005874
Zc3h12a	19.79	0.005904
Syt6	20.55	0.005952
Rasd1	2225.66	0.00605
Smim5	21.06	0.006199
AC107792.1	18.40	0.00637
Masp1	25.39	0.006674
Sorcs1	18.96	0.006895
Sdc1	33.35	0.006921
Gm10366	13.90	0.007237

Col11a2	102.38	0.007809
Rhbdf1	48.85	0.007811
Ccdc60	11.10	0.00821
Zfp36l1	22.73	0.008386
Gm10282	215.07	0.008622
Ankef1	29.60	0.008721
Ehf	377.65	0.00891
Gm37090	53.17	0.00892
Dock6	250.90	0.009089
Adora3	18.25	0.009238
A330023F24Rik	320.50	0.009263
Pcdhgc4	817.95	0.00987

Upregulated $P_{adj} < 0.01$, fold-change > 2

geneSymbol	baseMean	padj
Serpina7	1424.83	2.55E-44
Sptb	557.98	5.38E-37
Lipa	365.67	9.86E-36
Tacr3	123.14	6.67E-33
Grtp1	1236.10	4.20E-32
Dner	1142.24	3.19E-29
Sparcl1	105.97	1.84E-23
Galns	1228.28	4.61E-23
Ckb	6851.15	8.21E-22
Epdr1	307.55	8.65E-21
B4galt4	307.59	1.52E-19
Mapt	1808.46	2.87E-19
Pcp4	545.95	6.06E-19
S100a4	244.62	2.79E-18
Gsto1	438.69	2.81E-18
Me3	254.04	4.84E-18
P3h2	206.02	5.76E-18
Pex5l	3053.19	1.37E-17
Tnfrsf12a	253.30	1.82E-16
Wasf1	143.77	4.92E-16
Wfdc16	240.93	1.18E-14
Txnrd3	184.71	1.38E-13
Bzw2	497.54	1.45E-13
Gucy2c	1337.39	1.92E-13
Mmachc	557.87	4.58E-13
Cldn2	130.42	5.20E-13
Pwp2	487.74	1.63E-12
Coq7	674.57	2.03E-12
Ptrh1	133.20	3.11E-12
Pfkp	1130.26	3.31E-12
Gpx2	1018.85	4.42E-12
Vegfc	98.89	9.11E-12
Cdkn1a	2053.60	1.37E-11
Gnb1l	184.26	1.63E-11
Pcp4l1	600.83	2.94E-11
Sept6	114.32	4.17E-11
Ndufaf2	249.79	7.85E-11
Tnfrsf23	728.00	1.08E-10
Fam189a2	143.92	2.09E-10
Trim9	1122.85	3.48E-10
Rabepk	277.38	3.77E-10
Srl	92.65	7.87E-10
Ybx3	2309.47	1.05E-09
Ndufb6	1422.94	1.55E-09
Exosc8	211.04	2.50E-09
Rps2-ps10	89.05	2.82E-09
Dach2	192.99	4.33E-09
Hsd17b10	741.79	5.61E-09
Mad2l2	325.14	5.66E-09
Spsb4	109.25	7.56E-09
Rps6ka6	188.63	1.62E-08
Nrsn1	702.66	1.66E-08
Vcl	447.62	1.68E-08
Ehd4	464.63	1.71E-08
Dapl1	1522.34	2.04E-08
2200002J24Rik	56.38	3.63E-08
Hsd17b12	4365.05	3.70E-08
Dctpp1	160.74	3.97E-08
Thyn1	518.00	5.35E-08
Abcb11	78.86	5.61E-08
Ccdc85a	111.11	7.01E-08
Atp5b	22658.88	7.43E-08
Kntc1	161.71	8.11E-08
Aass	241.26	8.27E-08
Sergef	239.25	8.77E-08
Nhp2	337.70	1.27E-07
Gtse1	45.19	1.30E-07
Nt5c	399.13	1.48E-07
Prr29	32.38	1.83E-07

Timm50	628.55	1.94E-07
Nrtn	364.18	2.00E-07
Il13ra1	919.05	2.26E-07
Nln	404.07	2.82E-07
Cd68	23.30	2.90E-07
Apex1	508.62	4.74E-07
Mrm3	152.64	7.48E-07
Timm10	190.79	8.08E-07
Gc	7383.39	8.78E-07
Vav3	52.69	1.27E-06
Rab32	27.36	1.35E-06
Ccdc33	26.17	1.85E-06
Srd5a1	39.99	1.96E-06
Msra	348.64	1.99E-06
Gm49477	100.42	1.99E-06
Taf4b	105.77	2.05E-06
Insyn1	41.32	4.44E-06
Zfp365	450.40	4.77E-06
Cyc1	2008.44	7.35E-06
Adgra3	428.20	7.43E-06
Nit2	159.24	7.59E-06
Eef1akmt4	126.50	7.69E-06
Gldc	72.39	7.74E-06
Ccdc87	23.54	8.11E-06
Tst	51.36	8.20E-06
BC051226	132.11	9.64E-06
Cd320	384.05	1.01E-05
Hectd2os	96.23	1.02E-05
Nop16	409.17	1.15E-05
Clybl	242.06	1.22E-05
Dnaic1	290.23	1.23E-05
Agt	720.65	1.35E-05
Pinx1	111.67	1.38E-05
Ppargc1a	204.81	1.43E-05
Pim2	2187.24	1.56E-05
Stom	131.67	1.59E-05
H2aj	964.90	1.63E-05
Hsd17b7	276.14	1.77E-05
Iqgap2	262.05	1.79E-05
Cklf	99.47	1.79E-05
Dhrs13	155.29	1.84E-05
Ttc36	23.36	1.94E-05
Ascl1	47.43	2.11E-05

2600014E21Rik	38.18	2.26E-05
Svopl	40.34	2.32E-05
Phactr3	267.07	2.58E-05
Gm13127	45.58	2.76E-05
Cntn4	31.45	3.15E-05
Mapkapk3	425.66	4.22E-05
Mrpl41	744.85	4.39E-05
Zp2	92.74	4.61E-05
Grwd1	297.86	4.68E-05
Rasl10b	447.43	5.19E-05
Cenpn	40.35	5.20E-05
Stpg1	22.06	5.69E-05
Comtd1	184.15	6.52E-05
Noc4l	280.88	7.25E-05
Gcat	810.29	8.48E-05
Pdss1	82.23	9.60E-05
Uckl1os	109.90	9.67E-05
Tubb5	3183.26	1.22E-04
Asf1b	68.41	1.26E-04
Dipk1a	78.36	1.45E-04
Itpka	87.91	1.59E-04
Ttyh1	620.96	1.66E-04
Gm16006	36.35	1.71E-04
Psmc3ip	77.37	1.87E-04
Ppm1e	605.83	2.27E-04
Steap1	81.80	2.30E-04
Inka1	14.72	2.31E-04
Gm10658	17.87	3.13E-04
Dlg2	96.21	3.53E-04
Mrpl19	523.08	3.59E-04
Gm13461	87.09	3.83E-04
Lmo2	25.84	4.19E-04
Alb	21.76	4.84E-04
Phlda3	206.19	6.16E-04
Plb1	24.96	6.40E-04
Phb	1659.12	6.58E-04
Gm20754	22.34	7.64E-04
Espl1	29.89	7.89E-04
Gm4798	28.37	8.14E-04
Col16a1	165.27	8.22E-04
Gsto2	139.12	0.001042
Gdap1l1	1834.03	0.001241
Gm40477	31.80	0.001254

Dynll1	1468.59	0.001268
Polr1e	194.12	0.001282
Inka2	359.29	0.001301
Sftpd	117.73	0.001399
Pabpn1l	17.07	0.001468
Tmem74b	16.83	0.001566
Ube2l6	315.83	0.001588
Slc14a2	149.48	0.001808
Panx1	139.36	0.001931
Barx2	63.95	0.002138
Pthlh	133.22	0.002507
Baz1a	198.05	0.002507
Gm42031	41.66	0.003022
Tyms	236.30	0.003258
Neb	101.33	0.003323
Gm44717	44.07	0.003529
Gm34680	22.46	0.00378
Immp2l	103.30	0.003819
A330009N23Rik	42.22	0.003972
Ccdc141	139.62	0.003976
Ntn1	21.86	0.004002
Srpk3	47.65	0.004401
Nudc-ps1	55.69	0.004509
Pecam1	21.92	0.004752
Oip5	17.56	0.005655
Gm26712	17.87	0.005864
4833419F23Rik	14.26	0.006644
C78197	31.51	0.006698
Gm8960	20.77	0.007161
Gm5869	19.35	0.007317
Shroom4	71.37	0.007736
Erich6	81.70	0.00778
Gm2115	373.39	0.008048
Sh3bp1	17.51	0.008376
Gm9797	26.23	0.00843
Ttk	19.62	0.008493
Pdlim3	12.87	0.008826