Sperm Histone H3 Lysine 4 tri-methylation serves as a metabolic sensor of paternal obesity and is associated with the inheritance of metabolic dysfunction

Authors

Anne-Sophie Pepin[1], Christine Lafleur[2], Romain Lambrot[2], Vanessa Dumeaux[3], Sarah Kimmins[1,2,4*]

Affiliations

[1]Department of Pharmacology and Therapeutics, Faculty of Medicine, McGill University, Montreal, QC H3G 1Y6, Canada

[2]Department of Animal Science, Faculty of Agricultural and Environmental Sciences, McGill University, Montreal, QC H9X 3V9, Canada

[3]Department of Biology, PERFORM Center, Concordia University, Montreal, QC H4B 1R6, Canada

[4]Lead Contact

*Correspondence: sarah.kimmins@mcgill.ca

Running title: Paternal obesity alters sperm H3K4me3 and metabolic outcomes in descendants

Summary sentence: Paternal obesity impacts sperm H3K4me3 and is associated with placenta, embryonic and metabolic outcomes in descendants.

Keywords: obesity, metabolism, chromatin, sperm, epigenetic inheritance

Abstract

Parental environmental exposures can strongly influence descendant risks for adult disease. Metabolic disorders arise from the intersection of environmental and genetic risk factors, with epigenetic inheritance being at the center of the familial cycle of transgenerational disease. How paternal high-fat diet changes the sperm chromatin leading to the acquisition of metabolic disease in offspring remains controversial and ill-defined. Using a genetic model of epigenetic inheritance, we investigated the role of histone H3 lysine 4 methylation (H3K4me3) in the paternal transmission of metabolic dysfunction. We show that obesity-induced alterations in sperm H3K4me3 associated with offspring phenotypes and corresponded to embryonic and placental chromatin profiles and gene expression. Transgenerational susceptibility to metabolic disease was only observed when grandsires had a pre-existing modified sperm epigenome due to overexpression of the histone demethylase KDM1 in the male germline. This non-DNA based knowledge of inheritance has the potential to improve our understanding of how environment shapes heritability and may lead to novel routes for the prevention of disease.

Introduction

The prevalence of obesity and type II diabetes is growing globally at rates indicating that environment rather than genes is the principal driver. Exposures to high-fat diet, toxicants or micronutrient deficiency can impact our health and that of future generations (Gernand *et al.*, 2016; Braun, Messerlian and Hauser, 2017; Donkin and Barrès, 2018; Eberle *et al.*, 2020). Only now are we beginning to identify mechanisms linking these exposures to parental and offspring health. One connection between environment and health is the epigenome. The epigenome refers to the biochemical content associated with DNA that impacts gene expression, chromatin organization, and is transmitted via the gametes to alter phenotypes across generations. Uncovering how genomic information is organized and regulated through epigenetic processes to control gene expression and cell functions in the next generation is still in a nascent stage. We and others have shown that errors in epigenomic profiles in sperm can be induced by environmental exposure to toxicants such as those in insecticides and plastics, obesity, and poor diet (Lambrot *et al.*, 2013; Radford *et al.*, 2014; Donkin *et al.*, 2016; Wu *et al.*, 2017; Lismer, Dumeaux, *et al.*, 2021; Pilsner *et al.*, 2021). We recently demonstrated that these epigenome changes at the level of chromatin can be transmitted via sperm to alter embryonic gene expression, development, and offspring health (Lismer, Dumeaux, *et al.*, 2021). Historically, parental health and fertility have focused predominantly on the mother, although it is clear a father's health and lifestyle can also impact his children's health. How epimutations in sperm functionally impact the embryo urgently require elucidation to prevent transmission of disease from father to offspring.

Metabolic disease including obesity and type II diabetes can in part be attributed to genetic factors with a 5-10% increased risk (Voight *et al.*, 2010). The remaining risk is attributable to environmental-epigenetic interactions including potentially those of our ancestors. This possibility is supported by epidemiological

and animal studies. Transgenerational effects are suggested by studies in humans that linked the food supply of grandfathers to obesity and cardiovascular disease in their grandchildren (Kaati, Bygren and Edvinsson, 2002; Pembrey *et al.*, 2006; Lumey and Poppel, 2013). However, the ability for diet to induce transgenerational effects in animal models remains controversial and requires more in depth studies addressing the underlying molecular mechanisms (Siklenka *et al.*, 2015; Dalgaard *et al.*, 2016; Miska and Ferguson-smith, 2016). To date, studies using mice to assess the impact of diet and obesity in relation to the sperm epigenome, have focused on the DNA methylome and ncRNA as the potential sperm-borne mediators of metabolic disease (Carone *et al.*, 2010; Radford *et al.*, 2014; Grandjean *et al.*, 2015; Chen *et al.*, 2016; Cropley *et al.*, 2016; de Castro Barbosa *et al.*, 2016; Sharma *et al.*, 2016). The role of sperm chromatin in the non-genetic inheritance of metabolic disorders is unknown. In human and mouse sperm, histone H3 lysine 4 trimethylation (H3K4me3) localizes to genes involved in metabolism and development (Lambrot *et al.*, 2019; Lambrot, 2021; Lismer, Dumeaux, *et al.*, 2021). Moreover, sperm H3K4me3 can be altered by folate deficiency and influences embryonic development and gene expression (Siklenka *et al.*, 2015; Lismer *et al.*, 2020). This association of histone modifications in sperm with offspring phenotypes has since been confirmed in other mouse models (Stringer *et al.*, 2018; Lesch *et al.*, 2019). Based on these observations, we hypothesized that sperm H3K4me3 may serve as a metabolic sensor that is implicated in the paternal transmission of obesity-associated disease in offspring.

A focus of this study was to identify whether paternal obesity impacts the F_1-F_2, and if so, to identify potential mechanisms of this non-genetic inheritance. In our transgenic (TG) mouse model of epigenetic inheritance, male mice over-express the histone demethylase KDM1A (TG) specifically in spermatogenesis, resulting in sperm with alterations in H3K4me2 and me3. Of note, only H3K4me3 has been implicated in transgenerational inheritance in this mouse model (Lismer *et al.*, 2020). Therefore, as sperm H3K4me3 is responsive to paternal folate deficiency (Lismer, Dumeaux, *et al.*, 2021), and has been implicated in transgenerational inheritance (Lismer *et al.*, 2020), we targeted this mark in sperm to probe in response to paternal obesity and as a potential mediator of inheritance. In this study, we aimed to: 1) assess the impact of high-fat diet (HFD) induced paternal obesity on sperm H3K4me3 and its association with metabolic dysfunction across generations, and 2) determine if descendants of obese TG sires with a previously altered sperm epigenome would show more severe metabolic dysfunction? To address these aims, we used wildtype (WT), or the germline specific KDM1A-overexpressing TG mice, in combination with a diet-induced obesity. These TG sires have descended from males that have an altered sperm epigenome and whose ancestors had compromised health (see Materials and Methods for details). This TG model is used to represent an at-risk population that may be more susceptible to poor health when challenged with obesity. Here, we demonstrate that a paternal high-fat diet induces F_0 obesity and metabolic dysfunction in the F_1. Remarkably, transgenerational phenotypes were only observed in descendants of obese KDM1A TG males, and this was associated with enhanced alterations in H3K4me3 enrichment in obese TG sperm. This suggests that the risk of transgenerational disease transmission may be greater if an ancestor has had

prior exposures that cause pre-existing damage to the sperm epigenome. Concordant with the metabolic phenotypes observed in offspring, obesity-induced alterations in sperm H3K4me3 occurred at genes involved in development, placenta formation, inflammatory processes, glucose and lipid metabolic pathways. These sperm altered H3K4me3 regions persist in the embryo and placenta, supporting a role for sperm H3K4me3 in paternal origins of adult-onset metabolic disorders.

Results

Paternal obesity induces metabolic phenotypes in a sex-specific manner that are enhanced in KDM1A F_1 and F_2 transgenic descendants

Impact of paternal obesity on offspring bodyweight and fat accruement

Beginning at weaning until 20 weeks, inbred C57BL/6NCrl control mice (WT), or KDM1A heterozygous transgenics (TG) were fed either a calorie-dense high-fat diet (HFD; 60% kcal fat), or a sucrose- and vitamin-matched low fat control diet (CON; 10% kcal fat) (Fig. 1A-C and Table S1). Table S2 provides the animal numbers by sex, generation, and genotype for each metabolic analysis. In the 2 weeks post-weaning, F_0 males on the HFD consumed more calories and gained significantly more weight than CON males irrespective of genotype (Fig. S1A-B). These effects persisted throughout the diet intervention (Fig. S1A-C), with TG HFD males weighing the most at 4 months (Fig. S1C_i). This trend continued in the TG male F_1 and F_2 descendants (fed regular chow), with weights being significantly more than the F_1 and F_2 of TG CON and WT HFD (Fig. S1C_{ii-iii}). Indicating sex-specific responses to paternal obesity, in female descendants the changes in body weight and fat deposition differed from males (Fig. S1C-E). For example, female F_1 HFD descendants showed no changes in body weight, while F_2 TG HFD females weighed more in comparison to F_2 WT HFD females (Fig. S1C_{iv-v}, respectively). Unexpectedly, F_2 WT HFD females weighed less than F_2 WT CON. To assess fat accruement, we measured visceral mesenteric (mWAT) and gonadal white adipose tissue (gWAT). All male (F_0) on the HFD accumulated more mWAT compared to CON males, with no genotype effect (Fig. S1D_i). Male and female F_1 offspring sired by WT HFD or TG HFD had increased mWAT fat mass compared to WT CON and TG CON (Fig. S1D_{ii} and S1D_{iv}, respectively). Strikingly, mWAT stores were greater in TG HFD F_1 and F_2 males and females compared to WT HFD descendants (Fig. S1D_{ii-v}). Gonadal fat depots in F_0 males were not impacted by the HFD (gWAT; Fig. S1E_i), while male WT HFD F_1 showed increased gWAT, and TG HFD F_1 did not (gWAT; Fig. S1E_{ii}). In the male F_2, TG HFD descendants had increased gWAT, but not WT HFD (Fig. S1E_{iii}). Consistent with the unusual decrease in body weight in female F_2 WT HFD vs F_2 WT CON, mWAT and gWAT were also reduced in female F_2 WT HFD (Fig. S1D_v and S1E_v, respectively). Perhaps indicative of a sex-specific over-correction in the female F_2 to grand-sire obesity. Like for body weight and mWAT, female F_2 TG HFD had increased gWAT in comparison to WT HFD (Fig. S1E_v). Overall analysis of body weight and fat accruement revealed sex-specific responses in descendants with transgenerational effects of paternal obesity being detected only in the TG HFD descendants of both males and females.

Impact of paternal obesity on glucose homeostasis

Next, we assessed glucose metabolism and insulin sensitivity by glucose tolerance (GTT), and insulin tolerance tests (ITT). These were conducted following the standard operating procedures of the NIH mouse Metabolic Phenotyping Center (Ayala *et al.*, 2010). First, we assessed the effects of the HFD on baseline fasting blood glucose. Consumption of a HFD resulted in elevated baseline glucose in male (F_0) WT HFD and TG HFD in comparison to WT CON and TG CON, respectively (Fig. S2Ai). Male TG HFD descendants (F_1 and F_2), but not WT HFD descendants had significantly elevated fasting blood glucose (Fig. S2A$_{ii-iii}$). In contrast, the glycemic status of all descendant females (F_1 and F_2) did not differ (Fig. S2A$_{iv-v}$). The same animals used to assess baseline glucose were then given an intraperitoneal glucose challenge and the rate of glucose disposal measured. Analysis of GTT data showed that F_0 WT HFD and TG HFD were glucose intolerant following glucose injection in comparison to F_0 CON males as shown by their elevated blood glucose curves that persisted across the time course of 15-120 min (Fig. 1D$_i$). Note that at T0, before the glucose challenge, F_0 WT HFD males had significantly elevated blood glucose in comparison to F_0 WT CON which is in line with the baseline glucose measures (Fig. 1D$_i$ and Fig. S2A$_i$). In contrast, the T0 blood glucose levels did not differ between male F_0 TG CON and F_0 TG HFD even though baseline measures after fasting were significantly different (Fig. 1D$_i$ and Fig. S2A$_i$). These non-congruent findings are attributable to multiple testing correction to adjust for the high number of simultaneous tests conducted in the analysis of GTT curves. Intergenerational effects of the paternal HFD were apparent as male (F_1) sired by WT HFD and TG HFD were also glucose intolerant (Fig. 1D$_{ii}$). Indicating that there was enhanced metabolic disruption in the male F_1 TG HFD, elevated glucose levels persisted across the GTT time-course for the F_1 TG HFD but not the F_1 WT HFD. Interestingly, glycemic response impairments persisted in the F_2 generation of male descendants of TG HFD only (Fig. 1D$_{iii}$). Although fat measures were impacted in female F_1 and F_2 HFD, they did not exhibit glucose impairment (Fig. 1D$_{iv-v}$). Analysis of the area under the curve (AUC) for the GTT was consistent with the male and female glycemic responses shown in the glucose curves (Fig. S2B$_{i-v}$). In line with the observed glycemic responses, the insulin tolerance test and the corresponding AUC demonstrated that male F_0 WT HFD and TG HFD were insulin insensitive (Fig. 1E$_i$ and S2C$_i$). Analysis of the AUC indicated that F_1 WT HFD and F_1 TG HFD were insulin insensitive (Fig. S2C$_{ii}$). Like the glucose tolerance test, there were more pronounced impairments revealed by the ITT for the F_1 TG HFD in comparison to the F_1 WT HFD and only the F_2 TG HFD showed impaired whole body insulin action (Fig. 1E$_{iii}$ and Fig. S2C$_{ii-iii}$). Like the GTT, there was no indication of insulin impairment in female HFD F_1 nor F_2 (Fig. 1E$_{iv-v,}$ S2C$_{iv-v}$).

To summarize, the effects of paternal high-fat diet on glucose homeostasis were sex-specific; male descendants had impaired glucose homeostasis, whereas females did not. Many factors can influence the outcomes of GTT and ITT measures such as the length of fasting, age and strain of mice (Ayala *et al.*, 2010; Bowe *et al.*, 2012). It is worth considering that even though female descendants of HFD (F_{1-2}) showed alterations in fat accretion but not in glucose homeostasis, there remains the possibility that findings may have been different had measurements been taken after a reduced fasting period, or beyond 4 months of

age. Taken together, the assessments of weight and metabolic testing indicate that the TG descendants had enhanced responses to paternal obesity in comparison to WT descendants.

Paternal obesity was associated with altered liver gene expression in the F_0-F_1 with unique genes being differentially expressed in KDM1A descendants (F_1-F_2)

Obesity contributes to pathophysiological changes in gene expression in liver gene expression (Kim *et al.*, 2004). To determine whether the altered metabolic status of HFD sires and their descendants (F_1-F_2) was associated with differential gene expression in the liver, we performed RNA-sequencing on the left lateral lobe (*lobus hepatis sinister lateralis*) of adult mice (F_0-F_2). Sequencing quality was high with most RNA profiles having a Pearson correlation coefficient > 0.8 (Fig. S3A). Interestingly, principal component analysis of sequencing data revealed distinct hepatic transcriptomic profiles between males and females that was independent of experimental group and genotype (Fig. S3C). We compared hepatic transcriptome profiles by diet, sex, genotype and generation using a gene-level analysis at single-transcript resolution (Yi *et al.*, 2018). As expected, obesity was associated with differential liver gene expression. Liver from obese F_0 WT males showed differential expression of 2,136 genes in comparison to non-obese F_0 WT males (Fig. 2A, Lancaster $p<0.05$). Similarly, when comparing obese F_0 TG to non-obese F_0 TG, 1,476 genes were differentially expressed (Fig. 2B, Lancaster $p<0.05$). Of these differentially expressed genes (DEGs), 448 were commonly altered by obesity in both the F_0 WT and F_0 TG ($p<0.0001$; Fig. 2_i). To identify which genes were altered due to genotype, we compared WT obese to TG obese and identified 524 DEGs, suggesting that obesity had a unique effect in TG mice due to an interaction between diet and genotype (Fig. 2C, Lancaster $p<0.05$).

To determine if the effects of paternal obesity on liver function were intergenerational, we compared the liver transcriptome of male and female F_1. In comparison to F_1 WT CON and TG CON males, livers of F_1 WT HFD and TG HFD, showed differential expression of 1,015 and 794 genes (Fig. 2D and 2E, respectively, Lancaster $p<0.05$). A total of 165 DEGs overlapped between F_1 WT and TG ($p<0.0001$; Fig. 2_{ii}). Of the DEGs between the WT CON and HFD in the F_1, 139 were the same deregulated genes as identified in the F_0 WT CON vs HFD males ($p=0.76$; Fig 2_{iv}). Similarly, there were 103 shared transcripts identified as differentially expressed between the F_1 TG CON vs HFD, that were also altered in the F_0 TG CON vs HFD ($p=0.003$; Fig 2_v). This suggests that a common set of genes maintain dysfunction as a consequence of direct exposures to obesity and these changes are maintained in the non-obese F_1. When comparing genes altered by genotype in the F_1 (WT HFD vs TG HFD), 961 were significantly altered (Fig 2F, Lancaster $p<0.05$), with 78 overlapping DEGs between the F_0 and the F_1 ($p<0.0001$; Fig 2_{vi}). The overlap in deregulated genes between the F_0 and F_1 indicates that the metabolic phenotypes generated by the paternal HFD persist intergenerationally despite the F_1 being fed a regular chow diet. There were also novel genes in the F_1 that were deregulated, suggesting secondary intergenerational effects of the paternal HFD. As expected, there was a reduction in the number of deregulated genes in the F_1 WT and TG males compared to the F_0 which is consistent with the milder metabolic phenotypes observed in the F_1 in

comparison to the directly HFD-exposed F_0. Likely reflecting the enhanced genotype influence of the F_0 TG HFD sires that resulted in more severe weight changes and higher baseline glucose in their descendants compared to the WT HFD descendants, there were more DEG in the male F_0 compared to the male F_1 TG HFD (524 in F_0 versus 961 in F_1; Fig. 2_{vi}).

The last comparisons in liver transcriptomes were between the F_1 male and female. Despite the female F_1 having no metabolic phenotype detected by our measures, there was significantly altered gene expression in the livers of F_1 female offspring of WT HFD vs WT CON sires (830; Fig 2G, Lancaster $p<0.05$) Of these, 153 were in common with the F_1 male WT HFD sired offspring ($p<0.0001$; Fig 2_{vii}). Likewise, the F_1 female sired by TG HFD had 1,125 DEGs in comparison to females sired by TG CON (Fig. 2H, Lancaster $p<0.05$) with 148 in common with F_1 male TG HFD sired offspring ($p<0.0001$; Fig. 2_{viii}). Of these altered transcripts, 160 were in common between F_1 female descendants of WT HFD and TG HFD ($p<0.0001$; Fig. 2_{iii}). Like the F_1 male TG HFD offspring, there were unique transcripts altered in F_1 female TG HFD offspring (1,370; Fig. 2I, Lancaster $p<0.05$), with 181 differentially expressed in both F_1 males and females ($p=<0.0001$; Fig. 2_{ix}). These may reflect genes impacted by genotype regardless of sex. An interesting finding from the F_2 phenotyping was those transgenerational metabolic effects of the HFD were only detected in the male descendants of TG. Therefore, we only profiled F_2 male livers by RNA-seq. This analysis revealed differential expression of 2,141 genes between the F_2 WT HFD and TG HFD (Fig. 2J, Lancaster $p<0.05$) with 129 overlapping with the F_1 WT HFD vs TG HFD males ($p=0.06$; Fig 2_x). The number of differentially expressed genes increased every generation in comparisons between the WT HFD and the TG HFD (F_0=524, F_1=961, F_2 = 2,141). This sustained deregulated gene expression in the livers of TG HFD F_2, matches the enhanced metabolic phenotypes observed in only F_2 TG HFD males but not in the F_1 WT HFD. It also indicates a dilution in phenotype by reprogramming in the WT HFD descendants.

Paternal diet-induced obesity disrupts gene expression in functional processes that differ between genotypes, sex and generations

To gain insight into the physiological implications of obesity-induced altered hepatic transcriptomes, we used a gene ontology (GO) approach combined with functional similarity clustering to compare processes in the liver impacted by diet across genotype and sex, and those impacted by genotype across generation (Fig. 3A-C, Supplemental files 1-3 and Table S5-7) (Brionne, Juanchich and Hennequet-Antier, 2019). Interactive heatmaps that facilitate in-depth probing of the gene frequency and the $-\log_{10}$ p-value of enriched GO terms within each cluster are found in Supplemental files 1-3. The non-interactive heatmaps are shown in Fig. 3. Overall, there were similar processes altered by obesity in F_0 WT and TG livers, including lipid, amino acid, and small molecule metabolism (Fig. 3A, Supplemental file 1 and Table S5; clusters 1-5), homeostasis and environmental responses (clusters 8-10), and cellular differentiation and signalling (clusters 11-13). However, the gene frequency (# of genes annotated to that process) within processes differed by genotype. For example, genes involved in lipid metabolic processes (cluster 2), specifically the carboxylic acid metabolic process was more enriched at 19.03% gene frequency in the F_0 WT HFD ($-\log_{10}$

p-value 6.18), than in the F_0 TG HFD at 14.7% gene frequency (-$\log_{10}$ p-value 4.37). In contrast, genes involved in environmental responses (cluster 8) such as to alcohol, were similarly enriched in both the F_0 WT HFD with a gene frequency of 19.19% (-$\log_{10}$ p-value 1.3) and of 19.19% in the F_0 TG HFD (-$\log_{10}$ p-value 2.29). Interestingly, deregulated genes involved in insulin, protein and metal ion transport were only enriched in F_0 TG HFD livers (clusters 6-7).

When the altered functional pathways were compared between F_1 WT CON vs WT HFD males and females, there were clear impacts of obesity on the liver functional pathways of offspring, and these differed by sex (Fig. 3B, Supplemental file 2 and Table S6). Reflecting sex differences, a greater number of GO terms related to inflammation (cluster 4), and cell cycle, differentiation and signalling regulation (clusters 10-11) were significantly enriched in males compared to females. Of note, genes involved in the regulation of proinflammatory cytokines were particularly enriched in males but not females (clusters 4 and 6). This concurs with the more severe phenotypes observed in the males. Conversely, genes involved in DNA/RNA biosynthesis, transcription factors and telomere activity (clusters 1-3), and macromolecule and nitrogen metabolism (cluster 5) were more enriched in females. Interestingly, pathways associated with chromatin and cellular organization and protein metabolism (clusters 8-9) were differentially enriched by paternal obesity in both sexes.

Next, we compared the intergenerational and transgenerational effect of the interaction between the KDM1A transgene with obesity in terms of differences in process enrichment across generations when comparing $F_{0\text{-}2}$ WT HFD with $F_{0\text{-}2}$ TG HFD (Fig. 3C, Supplemental file 3 and Table S7). Reflecting the increasing generational changes in liver gene expression in the TG HFD descendants, there was an increase in the number of significantly enriched GO terms when comparing WT HFD vs TG HFD across generations (F_0 male=79; F_1 male=118; F_1 female=159; F_2 male=206; Supplemental file 4). This finding is concordant with the metabolic phenotypes detected in the TG HFD male F_2 descendants, but not in the F_2 male WT HFD (Fig. 1, Fig. S1, Fig. S2; Fig. 3C and Supplemental file 3). An interesting finding that supports a unique genotype response in the male F_1 TG HFD in comparison to the male F_1 WT HFD, is the over-representation of genes implicated in chromatin remodelling and transcription (clusters 17-19). This may reflect the activation of genetic response to reprogram the KDM1A-diet-induced sperm epimutations. In addition, there was an enrichment in the differential expression of genes with functions related to inflammation and environmental response (clusters 3-5), and metabolic processes (clusters 11-14). Genes involved in lipid catabolic processes (cluster 13) were increasingly enriched across generations, with 3.54% (F_0), 6.69% (F_1 male), 8.66% (F_1 female) and 20.87% (F_2) gene frequency. Like the males, female F_1 WT HFD vs F_1 TG HFD showed differences in enrichment pathways that reflect the interaction between obesity and the KDM1A transgene in the F_0 and the sex differences observed in the phenotypes (Fig. 3C)

Obesity in combination with expression of the KDM1A transgene increases differential enrichment of sperm H3K4me3 at genes involved in metabolism and development

Previously we associated paternal folate deficient diets with altered H3K4me3 in sperm and showed that these alterations are transmitted to the embryo and in line with offspring phenotypes (Lismer et al., 2021). We hypothesized that the sperm epigenome at the level of H3K4me3 would be altered by obesity and that this effect would be enhanced in KDM1A TG males with pre-existing alterations in sperm H3K4me3. To test these hypotheses, we performed native chromatin immunoprecipitation followed by deep sequencing (ChIP-seq) targeting histone H3K4me3, using sperm from individual WT or TG males fed either a CON or HFD (N=5 per experimental group, on average 33.3 million reads per sample with an alignment rate of 97%, Table S8). H3K4me3 localized to 30,745 genomic regions, with a Spearman correlation coefficient of 0.98 between samples (Fig. 4A and S4). Principal component analysis of H3K4me3 profiles revealed a clear separation of samples according to dietary treatment within genotype groups (Fig. 4 B-C). WT samples separated along Principal Component 1 (PC1) with 37.41% of variance attributed to diet (Fig. 4B; PERMANOVA, permutation-based p=0.01). TG samples separated on PC1 with 32.68% of the variability, with diet as the second source of variance (PC2), at 25.56% (Fig. 4C; PERMANOVA, permutation-based p = 0.009).

To focus our analysis on the regions most impacted by diet we selected the top 5% differentially enriched H3K4me3 regions (deH3K4me3, n=1,538) in each genotype (PC1 in WT, PC2 in TG) (Fig. 4Di-iv). The genome distribution analysis for specific annotations showed that obesity-sensitive H3K4me3 regions were predominantly located in CpG islands, promoters, exons, and intergenic regions (Fig. S5). To a lesser extent, deH3K4me3 also occurred at transposable elements (LINE, SINE and LTR), where epigenetic de-repression is associated with the use of alternative promoters and long- and short-range enhancers that are implicated in embryo development and pluripotency (Gerdes *et al.*, 2016) (Fig. S5). Representative genome browser tracks (Fig. S5) showing enrichment gains and losses for H3K4me3 at gene promoters are shown for *Pde1c* (phosphodiesterase 1C; affects the olfactory system), *Bcdin3d* (RNA methyltransferase; highly expressed in embryonic development), *Sh2d4a* (Sh2 domain containing protein 4A; expressed during development and associated with endocrine and liver function), and *Col15a1* (collagen alpha-1; cell differentiation and development, endocrine, cardiovascular system and more) (*Mouse Genome Informatics*, no date).

Next, we compared the regions of H3K4me3 that were altered by obesity, their genomic location, directionality change and functions between diets and genotype (Fig. 4). As a response to obesity, H3K4me3 enrichment gains were more predominant than losses for both F_0 WT HFD and TG HFD (Fig. 4D). In the WT HFD 1,323 regions gained and 215 lost H3K4m3 in comparison to the WT CON (Fig. 4D$_{i-ii}$). Similarly, in the F_0 TG HFD sperm, 1,067 regions gained and 471 lost H3K4me3 in comparison to the F_0 TG CON (Fig 4Diii-iv). Regions with deH3K4me3 in WT HFD had an 15.6% overlap (240/1,538 regions) with those of TG HFD (Fig. 4E). Of those common 240 regions, 162 had the same directionality change in both WT and TG HFD, with 159 regions with a gain and 3 regions with a reduction in H3K4me3 enrichment (Fig. 4 F and G, respectively). The non-overlapping regions of deH3K4me3 in WT HFD and TG HFD sperm could be a consequence of genetic-epigenetic interactions where the TG mice respond uniquely to obesity

as was observed in the phenotypic characterization. The proximity to the TSS of the deH3K4me3 regions in sperm altered by obesity in the F_0 WT HFD and TG HFD were similar (Fig. 4H).

Next, we performed a gene ontology (GO) enrichment analysis on promoters to gain functional insight into the genes disturbed by obesity and how they may relate to offspring phenotypes. This analysis revealed that changes in H3K4me3 occurred at promoters of genes that function in reproduction and development which is expected as these processes are known to be enriched for H3K4me3 in sperm (Fig. 4I and Tables S7-10). Supporting the idea that deH3K4me is specifically impacted by obesity and may be implicated in the transmission of metabolic disease to offspring is the finding that enriched genes were identified in processes related to metabolism and the impact of obesity. These included inflammatory processes, glucose and lipid metabolic pathways, and one-carbon cycle metabolism (Fig. $4I_{i\text{-}iv}$; Tables S7-10). Notably, some of the significantly enriched pathways are concordant with disturbed metabolic phenotypes of the F_0-F_2 including, for example, carbohydrate metabolic processes, glycolysis, growth hormone signaling and insulin signaling (Fig 4I, Tables S7-10).

The metabolic phenotypes of WT HFD and TG HFD descendants were similar, although the $F_{1\text{-}2}$ TG HFD showed enhanced abnormalities. We hypothesized that these differences in offspring metabolic disturbances may relate to the degree of H3K4me3 alteration in F_0 sperm, the directionality of the change (gain versus loss), and the functionality of genes bearing alterations. Interestingly, when comparing WT CON with TG HFD sperm, samples separated along PC2, with 26.69% of variance associated with genotype and diet (Fig. 5A; PERMANOVA, permutation-based p=0.006). Of the top 5% impacted regions selected (n=1,538), a greater proportion showed a gain of enrichment for H3K4me3 in TG HFD sperm in comparison to WT CON (Fig. 5B, n=1,071 regions with gains; Fig. 5C. n=467 regions with losses). We analyzed the detected regions impacted by genotype and diet (n=1,538) for differential enrichment to determine whether obesity in combination with KDM1A overexpression led to greater changes in H3K4me3 enrichment. This analysis identified 264 regions with a significant linear trend, where TG HFD sperm showed a greater degree of change in enrichment, and TG CON and WT HFD showed intermediate changes in comparison to WT CON (Fig. 5D-E, adjusted p<0.2). There were only 9 significant regions with further increase in H3K4me3 in the TG HFD (Fig. 5 D), whereas 255 regions showed a greater loss of H3K4me3 enrichment in the TG HFD (Fig. 5E). Consistent with the stronger metabolic phenotypes observed in the TG HFD $F_{1\text{-}2}$, the functional analysis of the promoters showing significant linear trends (n=104) for H3K4me3 across experimental groups occurred at genes implicated in metabolic and cardiovascular disease progression (Fig. 5F, Table S13).

Paternal obesity impacts sperm H3K4me3 at regions that coincide with open chromatin and gene expression in pre-implantation embryos

We recently demonstrated that sperm H3K4me3 is transmitted to the embryo and associated with gene expression (Lismer, Dumeaux, *et al.*, 2021). We hypothesized that obesity-altered sperm H3K4me3 is transmitted and associated with chromatin accessibility in the early embryo, which in turn could influence

gene expression and offspring phenotypes. To assess this possibility, we investigated the relationship between deH3K4me3 in sperm in relation to H3K4me3 in the embryo, the oocyte and open chromatin (Zhang *et al.*, 2016; Jung *et al.*, 2017; Liu *et al.*, 2019). We used existing datasets from mouse 2-cell embryos (Zhang *et al.*, 2016), generated by crossing males and females of different strains, permitting the assignment of reads to the paternal-specific allele. In line with a preferential paternal contribution of H3K4me3 to the 2-cell embryo, regions enriched for H3K4me3 in sperm, including those altered by obesity are not enriched in the oocyte (Fig 6A).

Next, we examined the relationship between sperm H3K4me3, chromatin accessibility and embryonic gene expression at the 4-cell and morula stages (Fig. 6 and Fig. S6A). We focused on the 4-cell and morula stages as this is when large-scale embryonic transcription directs development (Jukam, Shariati and Skotheim, 2017). There is an association between sperm H3K4me3 and embryonic gene expression (Lismer *et al.*, 2020; Lismer, Dumeaux, *et al.*, 2021) (Fig. S6 A_i). Strikingly, sperm H3K4me3 including obesity-sensitive regions are associated with ATAC-seq signal in pre-implantation embryos (Fig. 6A and 6B). To determine the functional relationship between the H3K4me3 obesity-altered regions and embryonic gene expression, we compared these with 4-cell and morula expressed genes and performed a gene ontology analysis. Of the sperm deH3K4me3 regions overlapping promoters (n=738), 51.8% (n=382) are expressed in the 4-cell embryos, 44.3% (n=327) are expressed in the morula embryos, and 39.7% (292) overlap in both (Fig. S6A_{ii}). To gain insight into what obesity-altered H3K4me3 associated genes in sperm relate to embryonic gene expression, we performed a GO analysis on the regions that are deH3K4me3 in sperm and the corresponding genes expressed in 4-cell and morula embryos (Fig. 6$C_{i\text{-}ii}$). Again, supporting a role for sperm H3K4me3 in paternal transmission of metabolic disease, with both the 4-cell and the morula gene processes significantly enriched specific to metabolism (Fig. 6$C_{i\text{-}ii}$ and Tables S12-13).

Unlike our previous studies on the function of H3K4me3 sperm where we found strong associations between deH3K4me3 and birth defects (Lismer *et al.*, 2020; Lismer, Dumeaux, *et al.*, 2021), in this study the defects in descendants appear to be limited to metabolic dysfunction and not development. To further probe the bias of obesity altered H3K4me3 for genes that function in metabolism and not post-implantation development, we examined the relationship between deH3K4me3 and H3K4me3/H3K27me3 bivalency. Genes that are bivalent for H3K4me3/H3K27me3 in sperm and embryos are termed as poised genes and are expressed later in embryo development (Hammoud *et al.*, 2009; Brykczynska *et al.*, 2010). Some promoters with high H3K4me3 enrichment in sperm are associated with bivalency in the 2-cell embryos (Fig. S6B_i). The obesity-altered H3K4me3 regions are predominantly not associated with H3K4me3/H3K27me3 bivalent genes in the embryo (Fig. S6C_{ii}). Taken together these findings suggest a preferential contribution of H3K4me3 on the paternal chromatin in the early embryo that includes obesity-sensitive regions that may be instructive of metabolic-associated gene expression and a direct route for epigenetic inheritance.

HFD alters the sperm epigenome at regions instructive for placenta development

The placenta is a key extra-embryonic organ that represents the uterine-fetal interface and plays a central role in energy allocation, nutrient exchange, and developmental progression. Placental abnormalities have been linked to late onset cardiometabolic diseases, highlighting the importance of the *in utero* environment for adulthood metabolic health (Perez-Garcia *et al.*, 2018). Our gene ontology analysis on diet-induced deH3K4me3 regions in sperm revealed significant enrichment of genes involved in placenta development (Fig. 4I and Tables S7-10). Given the sperm epigenome influences placental gene expression (Wang *et al.*, 2013), we were interested in the prospect that diet-induced epimutations in sperm affect placenta gene expression that could influence metabolic phenotypes across generations. To investigate this possibility, we used existing H3K4me3 and transcriptomic datasets from mouse trophectoderm – the embryonic precursor of placenta lineage – and placenta (Shen *et al.*, 2012; Wu *et al.*, 2016; Chu *et al.*, 2019). We compared the enrichment profiles of H3K4me3 in sperm, trophectoderm and placenta, at all H3K4me3-enriched regions in sperm (n=30,745) and at those sensitive to diet (n=2,836). Most regions showed presence of H3K4me3 in both the trophectoderm and the placenta (Fig. 7A). Of the 738 deH3K4me3 regions localizing to promoters in sperm, 56.8% (n=418) were expressed in the trophectoderm, 76.8% (n=567) were expressed in the placenta, and 54.6% (n=403) were expressed in both (Fig. S6C_{ii}). Gene ontology analysis of the shared H3K4me3 in sperm with TE and placenta revealed that there was an association with placenta function including at deH3K4me3 regions (Fig. 7Bi and iii, Tables S14 and S16). The GO analysis of the H3K4me3 regions that were not common with TE and placenta were involved in spermatogenesis, fertilization and sperm function (Fig. 7Bii and iv, Tables S15 and S17).

Next, we compared gene enrichment of sperm H3K4me3 with low- and high-expressed genes in the TE and placenta. Suggesting an influential role of sperm H3K4me3, the highly expressed genes and to a lesser extent the lowly expressed genes in placenta were positively correlated with sperm H3K4me3 (Fig. 7Ci-iv). Notably when the same comparisons were made with the deH3K4me3 there was a significant relationship with both lowly- and highly-expressed placenta genes (p=1.2e-11 and p=0.008, respectively; Fig. 7Cv-viii). In addition, the GO analysis of TE- and placenta-expressed genes that overlap with deH3K4me3 promoters are in line with the metabolic phenotypes in offspring (Fig. 7D i-ii, Tables S18-19). Taken together this analysis raises the possibility that obesity-induced alterations in sperm may influence embryonic and placenta gene expression to alter metabolic function of offspring.

Obesity-induced sperm epigenomic and hepatic transcriptomic alterations are unrelated

In a recent study, paternal low-protein diet was associated with reduced H3K9me2 at genes in sperm and were suggested to modulate gene expression profiles in the liver (Yoshida *et al.*, 2020). We aimed to assess whether a similar association between obesity-induced deH3K4me3 in sperm would relate to differential expression in the livers of the next generation. We focused on the obesity-associated sperm deH3K4me3 at promoters in F_0 sires and their relationship to differentially expressed genes in the liver (DEGs) of F_1 males. This analysis revealed that genes with differential expression in livers (n=1,644) were by in large unrelated to genes bearing deH3K4me3 in sperm. Only 9.1% (n=67) of promoters with deH3K4me3 in

sperm were differentially expressed in the liver of F_1 males sired by HFD-fed sires (Fig. S7 A-B). We then asked if deH3K4me3 promoters in sperm and liver DEGs had related biological functions. Strikingly, sperm- and liver-altered genes showed few functional similarities (Fig. S7C, Supplemental file 5 and Table S22). Functional pathways specifically enriched in deH3K4me3 promoters involved development and differentiation processes (clusters 12-15). As expected in a paternal obesity model, gene processes altered in offspring livers included: regulation of transcription and RNA splicing (clusters 1-3), protein and histone post-translational modifications (clusters 4-5), and metabolism of lipid, nitrogen and glucose (clusters 6-8). Pathways enriched in both the deH3K4me3 promoters in sperm and the DEGs in liver were involved in cell cycle, transport and signaling (clusters 16-19), and response to stress and inflammation (clusters 20-22). These commonly enriched pathways might reflect obesity-associated systemic inflammation which could affect multiple organs in a similar manner. These findings indicate that paternal obesity alters the sperm epigenome at distinct genes and functional pathways than those differentially expressed in offspring livers and fits with a developmental origin of adult metabolic dysfunction that could be related to alterations in gene expression in the embryo and placenta.

Discussion

Epidemiological studies and animal models point to the world-wide increases in childhood obesity and diabetes being in part attributed to environment-epigenetic interactions in the gametes (Huypens *et al.*, 2016; Isganaitis, Suehiro and Cardona, 2017; Sales, Ferguson-Smith and Patti, 2017). With the prevalence of childhood obesity predicted to reach 70 million by 2025 (Brown *et al.*, 2015), it is urgent to gain a better understanding of the underlying heritable mechanisms. Studies of obesity in mice and men have focussed on the role of sperm DNA methylation and non-coding RNA as mediators of paternal non-genetic inheritance of metabolic dysfunction (Carone *et al.*, 2010; Fullston *et al.*, 2013; Lambrot *et al.*, 2013; Martínez *et al.*, 2014; Chen *et al.*, 2016; Soubry *et al.*, 2016; Youngson *et al.*, 2016; Zhang *et al.*, 2018; Salas-Huetos *et al.*, 2021). At the mechanistic level, how sperm altered DNA methylation and non-coding RNA lead to metabolic phenotypes in offspring remains unresolved. In contrast to the voluminous studies on the response of sperm DNA methylation to paternal obesity/diets, there have been only a handful of studies on the role of chromatin. The few chromatin studies have been limited by either the state of the available technology, methodology used, or gene targeting approaches (Carone *et al.*, 2011; Palmer *et al.*, 2011; Martínez *et al.*, 2014; Donkin *et al.*, 2016).

Diets high in fat alter epigenetic programming (Ng *et al.*, 2010), likely through the alteration of cellular metabolism, which influences the availability of methyl donors and/or the activation or inactivation of chromatin modifying enzymes. In overweight and obese individuals, homocysteine is consistently elevated, and associated with reduced B12 and folate (Sánchez-Margalet *et al.*, 2002; Karatela and Sainani, 2009). In spermatogenesis there are dynamic changes to the sperm epigenome including histone methylation which is susceptible to alterations induced by changes in methyl donor availability (Lambrot *et al.*, 2013; Lismer, Dumeaux, *et al.*, 2021). It follows that the obesity-induced alterations in H3K4me3 we

report here could be a consequence of an altered methyl donor pool. Intriguingly, the effects of obesity on the paternal epigenome were linked with the metabolic dysfunction in the descendants; deH3K4me3 occurred at the promoters of genes involved in fertility, metabolism, and placenta processes. The association with metabolic processes is apparently specific to the HFD and obesity. In our analysis of sperm from folate deficient males, deH3K4me3 was associated more with developmental processes such as skeletal formation which aligned with the developmental defects observed in the offspring (Lismer, Dumeaux, *et al.*, 2021). Indicative of paternal transmission of sperm altered H3K4me3 as a mechanism of metabolic dysfunction was the strong relationship between deH3K4me3, an open chromatin state and gene expression in embryos and placenta.

The enhanced metabolic abnormalities observed in the descendants of obese F_0 TG revealed an increased susceptibility to metabolic disease in the TG line. An explanation for this response is that the F_0 TG were descendants from a lineage with pre-existing alterations in the sperm epigenome due to the genetic modification causing KDM1A overexpression. This genetic stress in combination with the environmental challenge of the HFD resulted in a more severely altered sperm epigenome in comparison to the WT, with consequent enhanced offspring phenotypes. Admittedly speculative, these findings suggest that the higher incidence of poor health in at-risk populations may be attributed to generational exposures to poor diet that leads to an accumulation of sperm epigenome errors that escape reprogramming.

Notably, paternal obesity-induced transgenerational metabolic disturbances in offspring were only observed in descendants of obese TG males. The phenomena of transgenerational inheritance has been most documented in genetic mouse models of epigenetic inheritance and studied in relation to DNA methylation patterns. These include the Avy agouti model (Morgan *et al.*, 1999; Dolinoy *et al.*, 2006; Cropley *et al.*, 2016), the kinky tail model ($Axin^{Fu}$ allele) (Rakyan *et al.*, 2003), and in mice bearing a mutation in the *Mtrr* gene, a folate metabolism enzyme (Padmanabhan *et al.*, 2013). In the context of environmental challenges, paternal transgenerational inheritance has been associated with altered sperm DNA methylation when there has been gestational exposure to toxicants and undernutrition (Anway *et al.*, 2005; Martínez *et al.*, 2014), and in a non-genetic pharmacologically-induced prediabetes model begun at weaning (Wei *et al.*, 2014). Taken together, this growing body of evidence indicates that transgenerational inheritance occurs under genetic influence, or when exposures coincide with developmental programming. The male F_0 mice in this study were exposed to the paternal HFD from weaning and not *in utero*, which may account for why transgenerational effects were not observed in WT HFD descendants. Another possibility is that transgenerational responses in the WT may have become detectable in older mice.

Our analysis indicates that the inherited metabolic disturbances observed in adult descendants originated early in development. In rodent models, paternal obesity and *in utero* undernutrition has been linked to altered gene expression in offspring livers and pancreatic islets with some minor links to concordant DNA methylation changes (Carone *et al.*, 2011; Martínez *et al.*, 2014; Wei *et al.*, 2014). It has been suggested that diet-associated alterations in DNA methylation in sperm are retained through embryogenesis and maintained in adult tissues mediating paternally-induced phenotypes (Martínez *et al.*,

2014; Wei *et al.*, 2014). Consistent with these studies, altered hepatic gene expression occurred in F_{1-2} offspring of obese sires. In contrast, we observed minimal overlap of genes and functional pathways between altered H3K4me3 enrichment in sperm, with those differentially expressed in F_1 livers. Instead, we demonstrate a significant overlap of obese sperm H3K4me3 profiles with the expression of metabolic-related genes in the embryo and placenta. Based on these findings, we suggest that the metabolic phenotypes we observe originate in early embryogenesis and through changes in placental gene expression.

There is a bounty of research linking maternal obesity to adverse metabolic consequences for the offspring that coincide with altered placental gene expression and function (Kerr *et al.*, 2018; Franzago *et al.*, 2019). On the other hand, it is an emerging concept that the paternal environment including factors such as diet and age can influence placental development and function. It is known that paternally expressed genes contribute to placental growth, trophoblast invasion and insulin resistance and adiposity (Moore, 2001; Binder, Hannan and Gardner, 2012; Wang *et al.*, 2013; Rosenfeld, 2015; Naruse *et al.*, 2019; Michelle M Denomme *et al.*, 2020; Michelle M. Denomme *et al.*, 2020). In humans, errors in epigenomic programming have been associated with gestational trophoblast disease and pre-eclampsia, but the role of the obese father in these conditions has been entirely unexplored (Gabory, Attig and Junien, 2011; Nelissen *et al.*, 2011). Previous studies support a connection between paternal diets, obesity, and placental dysfunction as a developmental route to metabolic disease in children. For example, we have shown that a folate deficient paternal diet and altered sperm DNA methylation coincided with deregulated placenta gene expression of *Cav1* and *Txndc16* (Lambrot *et al.*, 2013). Moreover, paternal obesity in mice has been attributed to defective placental development (Binder, Hannan and Gardner, 2012; Binder *et al.*, 2015). In women, altered DNA methylation in the regulation of some genes in preeclampsia has been established. However, many genes with deregulated expression were not associated with DNA methylation raising the possibility of altered chromatin signatures leading to abnormal gene expression in this placental disorder (Leavey *et al.*, 2018). Indeed, upregulated expression of LncRNA by increased H3K4me3 has been observed in preeclampsia placentas (Sun *et al.*, 2020), and the levels of H3K4me3 as detected by immunocytochemistry are decreased (Meister *et al.*, 2021). Until now the connection between sperm chromatin and placenta function has been unexplored. Our analyses revealed that most of the obesity-altered H3K4me3 at promoters occurred at loci involved in placental development and inflammatory processes (56.6% and 76.8% of deH3K4me3 occurred at promoters expressed in the trophectoderm and placenta, respectively). Remarkably, deregulated expression of genes implicated in inflammation have been implicated in hypertensive disorders in pregnancy including pre-eclampsia. This raises the possibility that the paternal sperm epigenome may influence maternal health during pregnancy in addition to that of the developing fetus. Hypertensive disorders in pregnancy have been associated with increased risk for developing cardiovascular disease (Naruse *et al.*, 2019).

As in previous studies we found that paternal obesity resulted in sex-specific differences in metabolism and fat accruement with males being more impacted. The underlying mechanisms that lead to

the greater susceptibility of males may be related to sexually dimorphic placental gene expression (Eriksson *et al.*, 2010). In support of this possibility, paternal environment (diet) influenced placental function in a sex-specific manner (Binder *et al.*, 2015). Alternatively, different metabolic responses in male and female offspring may be due to hormonal responses where estrogen has been shown to protect against altered glucose homeostasis (Gupte, Pownall and Hamilton, 2015; Lainez *et al.*, 2018).

In summary, we provide evidence that paternal obesity induced altered H3K4me3 signatures in sperm contribute to the inheritance of metabolic disease. Our findings indicate further studies on the relationship between sperm chromatin, the placental chromatin and gene expression are warranted and could lead to a better understanding of paternal sperm epigenome transmission of metabolic disease. The translational validation of these findings will be important in developing intervention strategies focused on paternal health that could impact the health of future generations (Barratt *et al.*, 2021).

Materials and Methods

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Sarah Kimmins (sarah.kimmins@mcgill.ca).

Materials availability

This study did not generate new unique reagents.

Data and code availability

The sperm H3K4me3 ChIP-Seq and liver RNA-Seq data generated during this study are available at the following GEO accession number: GSE178096.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Animals

All animal procedures were carried out in accordance with the guidelines of the Faculty Animal Care Committee of McGill University, Montreal. For the wild-type line (WT), C57BL/6NCrl 8-week old males and 6-week old females were purchased from Charles Rivers Laboratory and were allowed one week of acclimation before breeding. For the KDM1A transgenic line (TG), mice were generated as previously described (Siklenka et al., 2015), with the same genetic background as the wild-type line. The F_0 transgenic mice used in this study were from the 11th generation. The earlier generations of mice in this KDM1A TG had severe developmental abnormalities, pre-implantation loss and early post-natal death (Siklenka *et al.*, 2015) (Siklenka et al., 2015). Over time we have selected against the severe phenotype by breeding the mice that survive and are normal. Single males were housed with two females to generate the F_0 generation. All animals were given access to water and food *ad libitum* and were maintained on a controlled light/dark cycle.

METHODS DETAILS

Diet experiments and animal breeding

The low-fat control diet (CON; D12450J) and high-fat diet (HFD; D12492) were obtained from Research Diets, and selected based on the matched amounts of sucrose, vitamin mix and folate. Diets' macronutrients composition are listed in Table S1. Males of the F_0 generation were generated from at least 7 different sires per group. F_0 males were weaned at 3 weeks of age and randomly assigned to either a CON or HFD. The number of animals per group, per sex and per generation, used for all metabolic characterization tests can be found in Table S2. Total body weights were monitored weekly. Cumulative caloric intake was recorded weekly by weighting pellets from the food hopper and calculated as kilocalorie per animal. The diet intervention spanned 10-12 weeks followed by 2 weeks of metabolic testing (at 4 months of age), 1 week of rest and 1-2 weeks of breeding with 7-week old C57BL/6NCrl females. In order to minimize maternal exposure to the intervention diets, females were housed with males overnight and separated during the day, for a maximum of 3 nights per week for 2 weeks, or until a vaginal plug was detected. Litter sizes (number of pups per litter) were recorded, and sex ratios (ratio of male pups over total number of pups) were calculated for all litters generated and can be found in Tables S3 and S4, respectively. The same timeline was used to generate the F_1 and F_2 animals. All females used for breeding and all F_1 and F_2 were fed a regular chow diet (2020X Teklad global soy protein-free extruded rodent diet, Envigo). All animals were sacrificed at 22 weeks (±2 weeks) by carbon dioxide asphyxiation under isoflurane anesthesia.

Metabolic testing

Assessment of metabolic parameters was conducted at 4 months of age within 2 consecutive weeks according to the standard operating procedures of the NIH mouse Metabolic Phenotyping Center (Ayala *et al.*, 2010). For the glucose tolerance test, animals were fasted overnight for 15 hours (± 1 hour) starting at 6:00PM with free access to water. Blood glucose was measured before and 15, 30, 60 and 120 minutes following an intraperitoneal injection of 2 g/kg of a 20% glucose solution (D-glucose, G7021, Sigma Aldrich) with one drop of blood from the tail-tip using a glucometer (Accu-Chek Aviva Nano). For the insulin tolerance test, animals were fasted for 6 hours (1 hour), starting at 9:00AM with free access to water. Blood glucose was measured before and 15, 30, 60 and 120 minutes following an intraperitoneal injection of 1 IU/kg insulin (Insulin solution, I9278, Sigma Aldrich), with one drop of blood from the tail-tip using a glucometer (Accu-Chek Aviva Nano). The area under the curves (AUCs) for the tolerance tests were calculated using the trapezoidal rule (GraphPad Prism version 8). For the baseline blood glucose levels, blood glucose levels were measured after an overnight fasting of 15 hours (± 1 hour) with one drop of blood from the tail-tip using a glucometer (Accu-Chek Aviva Nano).

Tissue collection

At necropsy, mice were dissected to collect adipose tissue (gonadal and mesenteric white adipose depots; gWAT and mWAT, respectively) and a liver lobe (left lateral lobe or *lobus hepatis sinister lateralis* for RNA-sequencing). All tissues were weighed, transferred to a clean tube, snap frozen in liquid nitrogen and stored

at -80°C until subsequent downstream experiments. Cauda epididymides were weighed and immediately used for sperm isolation.

Sperm isolation

Spermatozoa were isolated from paired caudal epididymides (Hisano *et al.*, 2013; Lismer, Lambrot, *et al.*, 2021). Cauda epididymides were cut into 5 mL of freshly-prepared Donners medium (25 mM $NaHCO_3$, 20 mg ml^{-1} BSA, 1 mM sodium pyruvate, 0.53% vol/vol sodium DL-lactate in Donners stock) and gently agitated to allow to swim out for 1 hour at 37°C. The solution was passed through a 40-µm cell strainer (Fisher Scientific, #22363547) and washed three times with phosphate-buffered saline (PBS). The sperm pellet was cryopreserved in freezing medium (Irvine Scientific, cat. #90128) and kept in a -80°C freezer until the chromatin immunoprecipitation experiment.

RNA-Sequencing and library preparation

RNA extraction was performed using the RNeasy Mini Kit (Qiagen, cat. #74104) following the manufacturer's protocol with slight modifications. In brief, 15-20 mg of liver lobes were cut on dry ice using a sterile scalpel and Petri dish. Samples were lysed in 350 µL of a denaturing buffer (*Buffer RLT* with beta-mercaptoethanol) and homogenized with homogenizer pestles. Lysates were centrifuged at maximum speed for 3 minutes and the supernatants transferred to a clear tube. Ethanol (50%) was added to lysates to promote selective binding of RNA molecules to the silica-based membrane when applied to the spin columns. To avoid genomic DNA contamination, an additional DNase digestion was performed. Finally, membranes of the spin columns were washed twice with 500 µL of *Buffer RPE* and total RNA was eluted using 30 µL of RNase-free water. Libraries were prepared and sequenced at the *Génome Québec Innovation Centre* with single-end 50 base-pair reads on the illumina HiSeq 4000 and paired-end 100 base-pair reads on the illumina NovaSeq 6000 S2 sequencing platforms.

ChIP-Sequencing and library preparation

Chromatin immunoprecipitation was performed as we have previously described (Hisano *et al.*, 2013; Lismer, Lambrot, *et al.*, 2021). In brief, spermatozoa samples in freezing media were thawed on ice and washed with 1 mL phosphate-buffered saline. For each sample, two aliquots of 10 µL were used to count spermatozoa in a hemocytometer under microscope, and 10 million spermatozoa were used per experiment. Sperm chromatin was decondensed in 1 M dithiothreitol (DTT; Bio Shop, #3483-12-3) and the reaction quenched with N-ethylmaleimide (NEM). Samples were lysed in lysis buffer (0.3M sucrose, 60mM KCl, 15mM Tris-HCl pH 7.5, 0.5mM DTT, 5mM McGl2, 0.1mM EGTA, 1% deoxycholate and 0.5% NP40). An MNase enzyme (15 units; Roche, #10107921001) was added to aliquots containing 2 million spermatozoa in an MNase buffer (0.3 M sucrose, 85 mM Tris-HCl pH 7.5, 3mM $MgCl_2$ and 2 mM $CaCl_2$), for exactly 5 minutes at 37°C. The digestion was stopped with 5 mM EDTA. Samples were centrifuged at maximum speed for 10 minutes, and the supernatants of aliquots from each sample were pooled back together. Each tube was supplemented with a protease inhibitor to obtain an 1X solution (complete Tablets EASYpack, Roche, #04693116001). Magnetic beads (DynaBeads, Protein A, Thermo Fisher Scientific, #10002D) were pre-blocked in a 0.5% Bovine Serum Albumin (BSA, Sigma Aldrich, #BP1600-100) solution

for 4 hours at 4°C and then used to pre-clear the chromatin for 1 hour at 4°C. Pulling down of the pre-cleared chromatin was performed with the use of magnetic beads that were previously incubated with 5 µg of antibody (Histone H3 Lysine 4 trimethylation; H3K4me3; Cell Signaling Technology, cat. #9751) for 8 hours at 4°C. Immunoprecipitation of the chromatin with the beads-antibody suspension was performed overnight at 4°C. Beads bound to the chromatin were subjected to a 3-step wash, one wash with Washing Buffer A (50 mM Tris-HCl pH 7.5, 10 mM EDTA, 75 mM NaCl) and two washes with Washing Buffer B (50 mM Tris-HCl pH 7.5, 10 mM EDTA, 125 mM NaC). The chromatin was eluted in 250 µL of Elution Buffer (0.1 M HaHCO3, 0.2% SDS, 5 mM DTT) by incubating the beads twice (2 x 125 µL) shaking at 400 rpm for 10 minutes at 65°C, vortexing vigorously and transferring the chromatin elute in a clean tube. The eluted chromatin was finally treated with 5 µL of RNase A (Sigma Aldrich, #10109169001) by shaking in a thermomixer at 400 rpm for 1 hour at 37°C, and then with 5 µL of Proteinase K (Sigma Aldrich, #P2308) overnight at 55°C. DNA was extracted and purified using the *ChIP DNA Clean and Concentrator kit* (Zymo Research, #D5201) using the manufacturer's protocol, eluted with 25 µL of the provided elution buffer. Size selection of the mononucleosomes (147 bp) was performed with the use of Agencourt AMPure XP beads (Beckman Coulter, #A63880). Libraries were prepared in-house using the Ultra-low Input Library kit (Qiagen; #180495). Libraries were sequenced with single-end 50 base-pair reads on the illumina HiSeq 4000 sequencing platform (n=5 samples per experimental group).

Pre-processing

Liver RNA-Sequencing data

All samples were processed with the same parameters with the exception of those sequenced on the NovaSeq platform to adapt for paired-end sequencing and sequencing read length. Reads were trimmed using *Trim Galore* (version 0.5.0, parameters for HiSeq: --phred33 --length 36 -q 5 --stringency 1 -e 0.1; parameters for NovaSeq: --paired --retain_unpaired --phred33 --length 36 -q 5 --stringency 1 -e 0.1) (Krueger, 2015). Trimmed reads were aligned to the *Ensembl* Genome Reference Consortium mouse reference 38 (GRCm38) primary assembly using *hisat2* (version 2.1.0, parameters: -p 8 --dta) (Kim, Langmead and Salzberg, 2015). Aligned files with SAM format were converted to binary SAM format (BAM) and sorted by genomic position using *SAMtools* (version 1.9) (Li *et al.*, 2009). Transcripts were assembled and gene abundances calculated using *Stringtie* (version 2.1.2, parameters: -p 8 -e -B -A) (Pertea *et al.*, 2015).

Sperm ChIP-Sequencing data

Sequencing reads were trimmed using *Trimmomatic* on single-end mode to remove adapters and filter out low-quality reads (version 0.36, parameters: 2:30:15 LEADING:30 TRAILING:30) (Bolger, Lohse and Usadel, 2014). Trimmed reads were aligned to the *Mus Musculus* mm10 genome assembly using *Bowtie2* (version 2.3.4) (Salzberg, 2013). Unmapped reads were removed using *SAMtools* (version 1.9) (Li *et al.*, 2009), and those with 3 mismatches or more were filtered out using *Perlcode*. BAM coverage files (BigWig) were generated using *deeptools2 bamCoverage* function (version 3.2.1, parameters: -of bigwig -bs 25 -p 20 --normalizeUsing RPKM -e 160 --ignoreForNormalization chrX) (Ramírez *et al.*, 2016).

Other publicly available ATAC-Sequencing or ChIP-Sequencing datasets

Raw files for 2-cell H3K4me3 and H3K27me3 ChIP-Seq (Liu *et al.*, 2016) (GEO: GSE73952), MII oocyte H3K4me3 ChIP-Seq (Zhang *et al.*, 2016) (GEO: GSE71434), sperm ATAC-Seq (Jung *et al.*, 2017) (GEO: GSE79230), 4-cell and morula ATAC-Seq (Liu *et al.*, 2019) (NCBI SRA: SRP163205),TE H3K4me3 ChIP-Seq (Liu *et al.*, 2016) (GEO: GSE73952), and placenta H3K4me3 ChIP-Seq (Shen *et al.*, 2012) (GEO: GSE29184) were downloaded from the National Centre for Biotechnology Information (NCBI) using the Sequencing Read Archive (SRA) Toolkit. Files were pre-processed as described above for the sperm H3K4me3 ChIP-Sequencing with slight modifications to adapt for datasets with paired-end reads and for different sequencing read lengths.

Other publicly available RNA-Sequencing data

Raw files for 4-cell and morula (Liu *et al.*, 2019) (NCBI SRA: SRP163205), TE (Liu *et al.*, 2016) (GEO: GSE73952), and placenta (Chu *et al.*, 2019) (NCBI SRA: SRP137723) RNA-Seq were downloaded from the National Centre for Biotechnology Information (NCBI) using the Sequencing Read Archive (SRA) Toolkit. Files were pre-processed as described above for the sperm H3K4me3 ChIP-Sequencing with slight modifications to adapt for datasets with paired-end reads and for different sequencing read lengths.

Paternal allele 2-cell embryo ChIP-Sequencing data

Raw files for 2-cell H3K4me3 ChIP-Seq (Zhang *et al.*, 2016) (GEO: GSE71434) were downloaded from the National Centre for Biotechnology Information (NCBI) using the Sequencing Read Archive (SRA) Toolkit. *SNPsplit* (version 0.3.2) was used to build a reference genome with PWK_PhJ SNPs masked (Krueger and Andrews, 2016). Reads were aligned to the generated PWK_PhJ SNPs N-masked reference genome using *Bowtie2* (parameters: -p 10 -t -q -N 1 -L 25 -X 2000 --no-mixed --no-discordant). Aligned files with SAM format were converted to binary SAM format (BAM) and sorted by genomic position using *SAMtools (*version 1.9) (Li *et al.*, 2009). *SNPsplit* (version 0.3.2) was used to assign reads to either the paternal (PWK_PhJ) or the maternal (C57BL/6) genome based on SNPs origin. BAM coverage files (BigWig) were generated using *deeptools2 bamCoverage* function (parameters: -of bigwig -bs 25 -p 20 --normalizeUsing RPKM -e 160 --ignoreForNormalization chrX).

QUANTIFICATION AND STATISTICAL ANALYSIS

Visualization and statistical analyses for metabolic characterization

Visualization of the metabolic characterization data was performed using Jupyter Notebook (version 6.0.1) with Python (version 3.7.4), with the use of the following packages: *seaborn* (version 0.9.0) (Waskom, 2021), *numpy* (version 1.17.2) (Harris *et al.*, 2020), and *panda* (version 0.25.2) (Mckinney, 2010). The *pyplot* and *patches* modules were loaded from the *matplotlib* library (version 3.4.2) (Hunter, 2007). Statistical analyses were conducted using GraphPad Prism 8. For all tests, a p-value less than 0.05 was considered significant. Significance for individual time points on the blood glucose curves for the glucose and insulin tolerance tests, for cumulative energy intake and for growth trajectories during the diet intervention, was tested using multiple t-test with a Holm-Sidak correction. Significance for total body

weight, mesenteric and gonadal white adipose tissue weight, baseline blood glucose and the area under the curve for the glucose and insulin tolerance tests, was tested using two-way ANOVA followed by Fisher's LSD. Litter sizes and sex ratios were analyzed by 2-way ANOVA to assess genotype and diet effects and interaction, followed by Tukey's multiple comparisons test.

Bioinformatics analysis

All bioinformatics analyses were conducted using R version 4.0.2 (R Core Team, 2018).

Liver RNA-Sequencing data

Transcripts with a mean count below 10 were filtered out, conferring a total of 27,907 and 45,992 detected expressed transcripts in samples sequenced on the illumina HiSeq and NovaSeq platforms, respectively. The samples tended to cluster by RNA Integrity Number (RIN), which was corrected for in the differential analysis (Fig. S3B). Differential expression analysis was conducted using *DESeq2* (version 1.28.1) (Love, Huber and Anders, 2014), by including sample's RIN value and group in the design formula. Independent hypothesis weighting (IHW, version 1.16.0) was used to correct for multiple testing and prioritization of hypothesis testing based on covariate (i.e. the means of normalized counts) (Ignatiadis *et al.*, 2016). IHW calculates weight for each individual p-value and then applies the Benjamini-Hochberg (BH) procedure to adjust weighted p-values (Benjamini and Hochberg, 1995). Finally, we used the Lancaster method to perform a gene-level analysis at single transcript resolution (*aggregation* package, version 1.0.1) (Yi *et al.*, 2018). Lancaster applies aggregation of individual transcripts p-values to obtain differentially expressed genes while capturing changes at the transcript level. Genes with a Lancaster p-value below 0.05 were considered significant.

For data visualization, transcript counts were normalized using variance stabilizing transformation without the use of blind dispersion estimation (i.e. with parameter *blind=FALSE*) (Love, Huber and Anders, 2014). This transformation approach translates data on a $\log_2$ scale, allows correction for library size and removes the dependence of the variance on the mean (heteroscedasticity). Variance-stabilized transcript counts were corrected for RIN values using *limma's removeBatchEffect* function (version 3.44.3) (Ritchie *et al.*, 2015). Pearson correlation heatmaps were generated using the *corrplot* package (version 0.88) (Taiyun, Wei, Simko, 2021), with samples ordered by hierarchical clustering. Principal component analysis was performed using *DEseq's plotPCA* function, with RIN values and sexes labeled. Heatmaps of differentially expressed genes were generated with the *Pheatmap* package (version 1.0.12) (Kolde, 2019), with transcripts ordered by k-means clustering (n kmeans=2) and samples ordered by hierarchical clustering using complete-linkage clustering based on Euclidean distance. Alluvial plots were generated with *ggplot2* (version 3.3.3) (Wickham, 2016), and overlap of differentially expressed genes across genotypes, generations and sexes were determined by the *GeneOverlap* package (version 1.24.0) (Shen, 2014), which uses a Fisher's exact test to compute p-values.

Visualization, Semantic similarity, and Enrichment Analysis of Gene Ontology (ViSEAGO)

Gene ontology (GO) analysis was performed using the *ViSEAGO* package (version 1.2.0) (Brionne, Juanchich and Hennequet-Antier, 2019). Gene symbols and *EntrezGene* IDs from the *org.Mm.eg.db* database were retrieved using the *AnnotationDbi* package. GO annotations were retrieved from *EntrezGene* for the *Mus Musculus* species (ID="10090") using the *ViSEAGO EntrezGene2GO* followed by *annotate* functions. ViSEAGO uses topGO to perform GO terms enrichment tests on the sets of genes of interest (differentially expressed genes). We used the Biological Process (BP) ontology category with Fisher's exact test (classic algorithm), and a p-value below 0.01 was considered significant. Results of enrichment tests for each set of genes of interest were then merged and hierarchical clustering was performed based on Wang's semantic similarity distance and *ward.D2* aggregation criterion. Results are visualized on a heatmap where GO terms are ordered by hierarchical clustering based on their functional similarity and GO terms enrichment significance is shown as a color gradient ($-\log_{10}$ p-value) in each set of differentially expressed genes of interest.

Sperm ChIP-Sequencing data

To detect genomic regions enriched with H3K4me3 in sperm, we used *csaw* (version 1.22.1) (Lun and Smyth, 2016) to scan the genome into windows of 150-bp. Windows with a fold-change enrichment of 4 over bins of 2,000 bp (background) were considered enriched. Enriched regions less than 100 bp apart were merged for a maximum width of 5,000 bp, conferring a total of 30,745 merged enriched regions. Counts in enriched regions were normalized using TMM normalization followed by *ComBat's* correction for batch effects (*sva* package, version 3.36.0) (Leek *et al.*, 2012; Zhang, Parmigiani and Johnson, 2020). Spearman correlation heatmaps and MA-plots were generated using raw and normalized counts at enriched regions using *corrplot* (version 0.88) (Taiyun, Wei, Simko, 2021), and *graphics* packages, respectively.

Principal component analysis was conducted on normalized counts in enriched regions, by comparing WT HFD vs WT CON (effect of diet in WT), TG HFD vs TG CON (effect of diet in TG), and WT CON vs TG HFD (effects of genotype and HFD). Based on visual assessment of the separation of samples according to dietary or genotype groups along Principal Component 1 (PC1; x axis) or 2 (PC2; y axis), the top 5% regions contributing the PC of interest were selected. Permutational multivariate analysis of variance (PERMANOVA) was conducted to determine whether variation is attributed to dietary/genotype group, using the *adonis* function (*vegan* package, version 2.5-7) (Oksanen *et al.*, 2007). Euclidean distances were used as a metric, 999 permutations were performed, and a $p<0.05$ was considered significant. The directionality change in enrichment was identified based on the positive (up-regulated regions) and negative (down-regulated regions) $\log_2$ fold change values of the median of normalized counts using *gtools' foldchange2logratio* function. Regions with increased and decreased enrichment for each comparison of interest were visualized using *Pheatmap* (version 1.0.12) (Kolde, 2019). Regions distance relative to

transcription start site (TSS) were annotated and visualized using the package *chipenrich* (version 2.12.0) (Welch *et al.*, 2014). Gene ontology analysis was performed using *topGO* (version 2.40.0) for genes with increased or decreased H3K4me3 enrichment at the promoter region for each comparison of interest. We used the Biological Process (BP) ontology category with Fisher's exact test *weight01Fisher* algorithm (Alexa, Rahnenführer and Lengauer, 2006), and a p-value less than 0.05 was considered significant. Genomic regions with deH3K4me3 were annotated using *annotatr* (version 1.14.0) (Cavalcante and Sartor, 2017) including CpG annotations and basic genes genomic features. Upset plots were generated using *UpsetR* (version 1.4.0) (Conway, Lex and Gehlenborg, 2017), by ordering each set by frequency and displaying 12 sets. Z-scores were calculated using *regioneR's overlapPermTest* (version 1.20.1) which performs a permutation test (n=1,000 permutations) to assess whether a set of regions is significantly enriched to a specific genomic feature compared to genomic regions from the whole genome (Gel *et al.*, 2016). Genome browser snapshots were generated using *trackplot* (Pohl and Beato, 2014).

To assess linear trends associated with the cumulative exposure of KDM1A overexpression and high-fat feeding in sperm, we ran *DESeq2* (version 1.28.1) on the top 5% regions contributing to Principal Component 2 (PC2; n=1,538 regions) associated with sample separation when comparing WT CON and TG HFD normalized counts. In the design formula, we included sample's batch information, and assigned a numerical value for each sample based on their group category (WT CON=1, WT HFD=2, TG CON=2, TG HFD=3). Independent hypothesis weighting (IHW) was used to correct for multiple testing and prioritization of hypothesis testing based on covariate (i.e. the means of normalized counts) (Ignatiadis *et al.*, 2016). Median of normalized counts were used to depict the increased and decreased trend of significant regions (adjusted p-value less than 0.2) across groups recoded on a numerical scale as defined above.

Competing Interest Statement

The authors declare no competing interests.

Acknowledgements

We thank the team at the McGill University and Génome Québec Innovation Centre for performing the sequencing and Dr. Deborah Sloboda (McMaster University) for advice on metabolic phenotyping methods. This research was funded by the Canadian Institute of Health Research (CIHR) grants to SK (358654 and 350129).

Author Contributions

SK conceived, designed, funded and guided the study. ASP developed the model and conducted the animal studies with assistance from CL and RL with associated laboratory analysis. ASP conducted the

bioinformatic analysis with assistance and oversight from VD. ASP and SK assembled and wrote the manuscript which was edited by VD.

References

Alexa, A., Rahnenführer, J. and Lengauer, T. (2006) 'Improved scoring of functional groups from gene expression data by decorrelating GO graph structure', *Bioinformatics*, 22(13), pp. 1600–1607. doi: 10.1093/bioinformatics/btl140.

Anway, M. D. *et al.* (2005) 'Epigenetic Transgenerational Actions of Endocrine Disruptors and Male Fertility', *Science*, 308(5727), pp. 1466 LP – 1469. doi: 10.1126/science.1108190.

Ayala, J. E. *et al.* (2010) 'Standard operating procedures for describing and performing metabolic tests of glucose homeostasis in mice', *DMM Disease Models and Mechanisms*, 3(9–10), pp. 525–534. doi: 10.1242/dmm.006239.

Barratt, C. L. R. *et al.* (2021) 'A global approach to addressing the policy, research and social challenges of male reproductive health.', *Human reproduction open*, 2021(1), p. hoab009. doi: 10.1093/hropen/hoab009.

Benjamini, Y. and Hochberg, Y. (1995) 'Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing', *Journal of the Royal Statistical Society. Series B (Methodological)*. [Royal Statistical Society, Wiley], 57(1), pp. 289–300. Available at: http://www.jstor.org/stable/2346101.

Binder, N. K. *et al.* (2015) 'Paternal obesity in a rodent model affects placental gene expression in a sex-specific manner', *Reproduction*, 149(5), pp. 435–444. doi: 10.1530/REP-14-0676.

Binder, N. K., Hannan, N. J. and Gardner, D. K. (2012) 'Paternal Diet-Induced Obesity Retards Early Mouse Embryo Development, Mitochondrial Activity and Pregnancy Health', *PLoS ONE*, 7(12). doi: 10.1371/journal.pone.0052304.

Bolger, A. M., Lohse, M. and Usadel, B. (2014) 'Trimmomatic : a flexible trimmer for Illumina sequence data', *Bioinformatics*, 30(15), pp. 2114–2120. doi: 10.1093/bioinformatics/btu170.

Braun, J. M., Messerlian, C. and Hauser, R. (2017) 'Fathers Matter: Why It's Time to Consider the Impact of Paternal Environmental Exposures on Children's Health.', *Current epidemiology reports*, 4(1), pp. 46–55. doi: 10.1007/s40471-017-0098-8.

Brionne, A., Juanchich, A. and Hennequet-Antier, C. (2019) 'ViSEAGO: A Bioconductor package for clustering biological functions using Gene Ontology and semantic similarity', *BioData Mining*. BioData Mining, 12(1), pp. 1–13. doi: 10.1186/s13040-019-0204-1.

Brown, C. L. *et al.* (2015) 'Addressing Childhood Obesity: Opportunities for Prevention.', *Pediatric clinics of North America*, 62(5), pp. 1241–1261. doi: 10.1016/j.pcl.2015.05.013.

Brykczynska, U. *et al.* (2010) 'Repressive and active histone methylation mark distinct promoters in human and mouse spermatozoa', *Nature Structural and Molecular Biology*. Nature Publishing Group, 17(6), pp. 679–687. doi: 10.1038/nsmb.1821.

Carone, B. R. *et al.* (2010) 'Paternally induced transgenerational environmental reprogramming of

metabolic gene expression in mammals', *Cell*. Elsevier Inc., 143(7), pp. 1084–1096. doi: 10.1016/j.cell.2010.12.008.

Carone, B. R. *et al.* (2011) 'Paternally-induced transgenerational environmental reprogramming of metabolic gene expression in mammals Benjamin', 143(7), pp. 1084–1096. doi: 10.1016/j.cell.2010.12.008.Paternally-induced.

de Castro Barbosa, T. *et al.* (2016) 'High-fat diet reprograms the epigenome of rat spermatozoa and transgenerationally affects metabolism of the offspring', *Molecular Metabolism*. Elsevier GmbH, 5(3), pp. 184–197. doi: 10.1016/j.molmet.2015.12.002.

Cavalcante, R. G. and Sartor, M. A. (2017) 'annotatr: genomic regions in context.', *Bioinformatics (Oxford, England)*, 33(15), pp. 2381–2383. doi: 10.1093/bioinformatics/btx183.

Chen, Q. *et al.* (2016) 'Sperm tsRNAs contribute to intergenerational inheritance of an acquired metabolic disorder', *Science*, 351(6271), pp. 397–400. doi: 10.1126/science.aad7977.

Chu, A. *et al.* (2019) 'The Placental Transcriptome in Late Gestational Hypoxia Resulting in Murine Intrauterine Growth Restriction Parallels Increased Risk of Adult Cardiometabolic Disease.', *Scientific reports*, 9(1), p. 1243. doi: 10.1038/s41598-018-37627-y.

Conway, J. R., Lex, A. and Gehlenborg, N. (2017) 'UpSetR: an R package for the visualization of intersecting sets and their properties', *Bioinformatics*, 33(18), pp. 2938–2940. doi: 10.1093/bioinformatics/btx364.

Cropley, J. E. *et al.* (2016) 'Male-lineage transmission of an acquired metabolic phenotype induced by grand-paternal obesity', *Molecular Metabolism*. Elsevier GmbH, 5(8), pp. 699–708. doi: 10.1016/j.molmet.2016.06.008.

Dalgaard, K. *et al.* (2016) 'Trim28 Haploinsufficiency Triggers Bi-stable Epigenetic Obesity', *Cell*. Elsevier, 164(3), pp. 353–364. doi: 10.1016/j.cell.2015.12.025.

Denomme, Michelle M. *et al.* (2020) 'Advanced paternal age directly impacts mouse embryonic placental imprinting', *PLoS ONE*, 15(3), pp. 1–13. doi: 10.1371/journal.pone.0229904.

Denomme, Michelle M *et al.* (2020) 'The inherited methylome landscape is directly altered with paternal aging and associated with offspring neurodevelopmental disorders', *Aging Cell*. John Wiley & Sons, Ltd, 19(8), p. e13178. doi: https://doi.org/10.1111/acel.13178.

Dolinoy, D. C. *et al.* (2006) 'Maternal genistein alters coat color and protects Avy mouse offspring from obesity by modifying the fetal epigenome', *Environmental Health Perspectives*, 114(4), pp. 567–572. doi: 10.1289/ehp.8700.

Donkin, I. *et al.* (2016) 'Obesity and bariatric surgery drive epigenetic variation of spermatozoa in humans', *Cell Metabolism*, 23(2), pp. 369–378. doi: 10.1016/j.cmet.2015.11.004.

Donkin, I. and Barrès, R. (2018) 'Sperm epigenetics and influence of environmental factors', *Molecular Metabolism*, 14, pp. 1–11. doi: https://doi.org/10.1016/j.molmet.2018.02.006.

Eberle, C. *et al.* (2020) 'Paternal metabolic and cardiovascular programming of their offspring: A systematic scoping review.', *PloS one*, 15(12), p. e0244826. doi: 10.1371/journal.pone.0244826.

Eriksson, J. G. *et al.* (2010) 'Boys live dangerously in the womb.', *American journal of human biology : the official journal of the Human Biology Council*, 22(3), pp. 330–335. doi: 10.1002/ajhb.20995.

Franzago, M. *et al.* (2019) 'Nutrigenetics , epigenetics and gestational diabetes : consequences in mother and child', *Epigenetics*. Taylor & Francis, 14(3), pp. 215–235. doi: 10.1080/15592294.2019.1582277.

Fullston, T. *et al.* (2013) 'Paternal obesity initiates metabolic disturbances in two generations of mice with incomplete penetrance to the F2 generation and alters the transcriptional profile of testis and sperm microRNA content', *FASEB Journal*, 27(10), pp. 4226–4243. doi: 10.1096/fj.12-224048.

Gabory, A., Attig, L. and Junien, C. (2011) 'Developmental programming and epigenetics', *The American Journal of Clinical Nutrition*, 94(suppl_6), pp. 1943S-1952S. doi: 10.3945/ajcn.110.000927.

Gel, B. *et al.* (2016) 'regioneR: an R/Bioconductor package for the association analysis of genomic regions based on permutation tests', *Bioinformatics*, 32(2), pp. 289–291. doi: 10.1093/bioinformatics/btv562.

Gerdes, P. *et al.* (2016) 'Transposable elements in the mammalian embryo: pioneers surviving through stealth and service.', *Genome biology*, 17, p. 100. doi: 10.1186/s13059-016-0965-5.

Gernand, A. D. *et al.* (2016) 'Micronutrient deficiencies in pregnancy worldwide: health effects and prevention.', *Nature reviews. Endocrinology*, 12(5), pp. 274–289. doi: 10.1038/nrendo.2016.37.

Grandjean, V. *et al.* (2015) 'RNA-mediated paternal heredity of diet-induced obesity and metabolic disorders', *Scientific Reports*. Nature Publishing Group, 5(June), pp. 1–9. doi: 10.1038/srep18193.

Gupte, A. A., Pownall, H. J. and Hamilton, D. J. (2015) 'Estrogen : An Emerging Regulator of Insulin Action and Mitochondrial Function', 2015.

Hammoud, S. S. *et al.* (2009) 'Distinctive chromatin in human sperm packages genes for embryo development', *Nature*. Nature Publishing Group, 460(7254), pp. 473–478. doi: 10.1038/nature08162.

Harris, C. R. *et al.* (2020) 'Array programming with NumPy', *Nature*, 585(7825), pp. 357–362. doi: 10.1038/s41586-020-2649-2.

Hisano, M. *et al.* (2013) 'Genome-wide chromatin analysis in mature mouse and human spermatozoa', *Nature Protocols*, 8(12), pp. 2449–2470. doi: 10.1038/nprot.2013.145.

Hunter, J. D. (2007) 'Matplotlib: A 2D Graphics Environment', *Computing in Science & Engineering*, 9(3), pp. 90–95. doi: 10.1109/MCSE.2007.55.

Huypens, P. *et al.* (2016) 'Epigenetic germline inheritance of diet-induced obesity and insulin resistance', *Nature Genetics*, 48(5), pp. 497–499. doi: 10.1038/ng.3527.

Ignatiadis, N. *et al.* (2016) 'Data-driven hypothesis weighting increases detection power in genome-scale multiple testing', *Nature Methods*, 13(7), pp. 577–580. doi: 10.1038/nmeth.3885.

Isganaitis, E., Suehiro, H. and Cardona, C. (2017) 'Who's your daddy?: paternal inheritance of metabolic disease risk.', *Current opinion in endocrinology, diabetes, and obesity*. England, 24(1), pp. 47–55. doi: 10.1097/MED.0000000000000307.

Jukam, D., Shariati, S. A. M. and Skotheim, J. M. (2017) 'Zygotic Genome Activation in Vertebrates.', *Developmental cell*, 42(4), pp. 316–332. doi: 10.1016/j.devcel.2017.07.026.

Jung, Y. H. *et al.* (2017) 'Chromatin States in Mouse Sperm Correlate with Embryonic and Adult Regulatory Landscapes', *Cell Reports*. ElsevierCompany., 18(6), pp. 1366–1382. doi: 10.1016/j.celrep.2017.01.034.

Kaati, G., Bygren, L. O. and Edvinsson, S. (2002) 'Cardiovascular and diabetes mortality determined by nutrition during parents' and grandparents' slow growth period', *European Journal of Human Genetics*, 10(11), pp. 682–688. doi: 10.1038/sj.ejhg.5200859.

Karatela, R. A. and Sainani, G. S. (2009) 'Plasma homocysteine in obese, overweight and normal weight hypertensives and normotensives.', *Indian heart journal*. India, 61(2), pp. 156–159.

Kerr, B. *et al.* (2018) 'Foetoplacental epigenetic changes associated with maternal metabolic dysfunction.', *Placenta*. Netherlands, 69, pp. 146–152. doi: 10.1016/j.placenta.2018.04.006.

Kim, D., Langmead, B. and Salzberg, S. L. (2015) 'HISAT: a fast spliced aligner with low memory requirements', *Nature Methods*, 12(4), pp. 357–360. doi: 10.1038/nmeth.3317.

Kim, S. *et al.* (2004) 'Hepatic gene expression profiles in a long-term high-fat diet-induced obesity mouse model.', *Gene*. Netherlands, 340(1), pp. 99–109. doi: 10.1016/j.gene.2004.06.015.

Kolde, R. (2019) 'pheatmap: Pretty Heatmaps. R package'. Available at: https://cran.r-project.org/package=pheatmap.

Krueger, F. (2015) 'Trim galore', *A wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files*, 516, p. 517.

Krueger, F. and Andrews, S. R. (2016) 'SNPsplit: Allele-specific splitting of alignments between genomes with known SNP genotypes [version 1; referees: 3 approved]', *F1000Research*, 5, pp. 1–15. doi: 10.12688/F1000RESEARCH.9037.1.

Lainez, N. M. *et al.* (2018) 'Diet-induced obesity elicits macrophage infiltration and reduction in spine density in the hypothalami of male but not female mice', *Frontiers in Immunology*, 9(SEP), pp. 1–16. doi: 10.3389/fimmu.2018.01992.

Lambrot, R. *et al.* (2013) 'Low paternal dietary folate alters the mouse sperm epigenome and is associated with negative pregnancy outcomes', *Nature Communications*, 4. doi: 10.1038/ncomms3889.

Lambrot, R. *et al.* (2019) 'The genomic distribution of histone H3K4me2 in spermatogonia is highly conserved in sperm†.', *Biology of reproduction*. United States, 100(6), pp. 1661–1672. doi: 10.1093/biolre/ioz055.

Lambrot, R. (2021) 'Whole genome sequencing of H3 lysine 4 tri-methylation and DNA methylation in human sperm reveals regions of overlap and exclusion linked to fertility, development and epigenetic inheritance', *Cell Reports*, CELL-REPOR.

Leavey, K. *et al.* (2018) 'Epigenetic regulation of placental gene expression in transcriptional subtypes of preeclampsia', *Clinical Epigenetics*, 10(1), p. 28. doi: 10.1186/s13148-018-0463-6.

Leek, J. T. *et al.* (2012) 'The sva package for removing batch effects and other unwanted variation in high-throughput experiments.', *Bioinformatics (Oxford, England)*, 28(6), pp. 882–883. doi: 10.1093/bioinformatics/bts034.

Lesch, B. J. *et al.* (2019) 'Intergenerational epigenetic inheritance of cancer susceptibility in mammals', *eLife*, 8, pp. 1–29. doi: 10.7554/elife.39380.

Li, H. *et al.* (2009) 'The Sequence Alignment/Map format and SAMtools.', *Bioinformatics (Oxford, England)*, 25(16), pp. 2078–2079. doi: 10.1093/bioinformatics/btp352.

Lismer, A. *et al.* (2020) 'Sperm histone H3 lysine 4 trimethylation is altered in a genetic mouse model of transgenerational epigenetic inheritance', *Nucleic Acids Research*. doi: 10.1093/nar/gkaa712.

Lismer, A., Lambrot, R., *et al.* (2021) 'ChIP-seq protocol for sperm cells and embryos to assess environmental impacts and epigenetic inheritance', *STAR Protocols*, 2(2), p. 100602. doi: https://doi.org/10.1016/j.xpro.2021.100602.

Lismer, A., Dumeaux, V., *et al.* (2021) 'Histone H3 lysine 4 trimethylation in sperm is transmitted to the embryo and associated with diet- induced phenotypes in the offspring Article Histone H3 lysine 4 trimethylation in sperm is transmitted to the embryo and associated with diet-induce', *Developmental Cell*. Elsevier Inc., pp. 1–16. doi: 10.1016/j.devcel.2021.01.014.

Liu, L. *et al.* (2019) 'An integrated chromatin accessibility and transcriptome landscape of human pre-implantation embryos', *Nature communications*. Nature Publishing Group UK, 10(1), p. 364. doi: 10.1038/s41467-018-08244-0.

Liu, X. *et al.* (2016) 'Distinct features of H3K4me3 and H3K27me3 chromatin domains in pre-implantation embryos', *Nature*. Nature Publishing Group, 537(7621), pp. 558–562. doi: 10.1038/nature19362.

Love, M. I., Huber, W. and Anders, S. (2014) 'Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2', *Genome Biology*, 15(12), pp. 1–21. doi: 10.1186/s13059-014-0550-8.

Lumey, L. H. and Poppel, F. W. A. Van (2013) 'The Dutch Famine of 1944-45 as a Human Laboratory : Changes in the Early Life Environment and Adult Health', in Lumey, L. H. and Vaiserman, A. (ed.) *Early Life Nutrition and Adult Health and Development*. Nova Science Publishers, Inc., pp. 59–76.

Lun, A. T. L. and Smyth, G. K. (2016) 'csaw: a Bioconductor package for differential binding analysis of ChIP-seq data using sliding windows.', *Nucleic acids research*, 44(5), p. e45. doi: 10.1093/nar/gkv1191.

Martínez, D. *et al.* (2014) 'In utero undernutrition in male mice programs liver lipid metabolism in the second-generation offspring involving altered Lxra DNA methylation.', *Cell metabolism*. United States, 19(6), pp. 941–951. doi: 10.1016/j.cmet.2014.03.026.

Mckinney, W. (2010) 'Data Structures for Statistical Computing in Python', in *Proceedings of the 9th Python in Science Conference*, pp. 56–61.

Meister, S. *et al.* (2021) 'Epigenetic modification via H3K4me3 and H3K9ac in human placenta is reduced in preeclampsia', *Journal of Reproductive Immunology*, 145, p. 103287. doi: https://doi.org/10.1016/j.jri.2021.103287.

Miska, E. A. and Ferguson-smith, A. C. (2016) 'Transgenerational inheritance: Models and mechanisms of non – DNA sequence – based inheritance', 354(6308), pp. 778–782.

Moore, T. (2001) 'Genetic conflict, genomic imprinting and establishment of the epigenotype in relation to growth.', *Reproduction (Cambridge, England)*. England, 122(2), pp. 185–193. doi:

10.1530/rep.0.1220185.
Morgan, H. D. *et al.* (1999) 'Epigenetic inheritance at the agouti locus in the mouse', *Nature Genetics*, 23(3), pp. 314–318. doi: 10.1038/15490.
Mouse Genome Informatics (no date).
Naruse, K. *et al.* (2019) 'Preliminary evidence of a paternal-maternal genetic conflict on the placenta: Link between imprinting disorder and multi-generational hypertensive disorders', *Placenta*, 84, pp. 69–73. doi: https://doi.org/10.1016/j.placenta.2019.02.009.
Nelissen, E. C. M. *et al.* (2011) 'Epigenetics and the placenta', *Human Reproduction Update*, 17(3), pp. 397–417. doi: 10.1093/humupd/dmq052.
Ng, S.-F. *et al.* (2010) 'Chronic high-fat diet in fathers programs β-cell dysfunction in female rat offspring', *Nature*, 467(7318), pp. 963–966. doi: 10.1038/nature09491.
Oksanen, J. *et al.* (2007) 'The vegan package', *Community ecology package*, 10(631–637), p. 719.
Padmanabhan, N. *et al.* (2013) 'Mutation in Folate Metabolism Causes Epigenetic Instability and Transgenerational Effects on Development', *Cell*. Elsevier Inc., 155(1), pp. 81–93. doi: 10.1016/j.cell.2013.09.002.
Palmer, N. O. *et al.* (2011) 'SIRT6 in mouse spermatogenesis is modulated by diet-induced obesity.', *Reproduction, fertility, and development*. Australia, 23(7), pp. 929–939. doi: 10.1071/RD10326.
Pembrey, M. E. *et al.* (2006) 'Sex-specific, male-line transgenerational responses in humans', *European Journal of Human Genetics*, 14(2), pp. 159–166. doi: 10.1038/sj.ejhg.5201538.
Perez-Garcia, V. *et al.* (2018) 'Placentation defects are highly prevalent in embryonic lethal mouse mutants', *Nature*, 555(7697), pp. 463–468. doi: 10.1038/nature26002.
Pertea, M. *et al.* (2015) 'StringTie enables improved reconstruction of a transcriptome from RNA-seq reads', *Nature Biotechnology*, 33(3), pp. 290–295. doi: 10.1038/nbt.3122.
Pilsner, J. R. *et al.* (2021) 'Aging-induced changes in sperm DNA methylation are modified by low dose of perinatal flame retardants', *Epigenomics*. Future Medicine, 13(4), pp. 285–297. doi: 10.2217/epi-2020-0404.
Pohl, A. and Beato, M. (2014) 'bwtool: a tool for bigWig files', *Bioinformatics*, 30(11), pp. 1618–1619. doi: 10.1093/bioinformatics/btu056.
Radford, E. J. *et al.* (2014) 'In utero undernourishment perturbs the adult sperm methylome and intergenerational metabolism', *Science*, 345(6198), p. 1255903. doi: 10.1126/science.1255903.
Rakyan, V. K. *et al.* (2003) 'Transgenerational inheritance of epigenetic states at the murine AxinFu allele occurs after maternal and paternal transmission', *Proceedings of the National Academy of Sciences of the United States of America*, 100(5), pp. 2538–2543. doi: 10.1073/pnas.0436776100.
Ramírez, F. *et al.* (2016) 'deepTools2: a next generation web server for deep-sequencing data analysis.', *Nucleic acids research*, 44(W1), pp. W160-5. doi: 10.1093/nar/gkw257.
Ritchie, M. E. *et al.* (2015) 'limma powers differential expression analyses for RNA-sequencing and microarray studies.', *Nucleic acids research*, 43(7), p. e47. doi: 10.1093/nar/gkv007.

Rosenfeld, C. S. (2015) ‘Sex-Specific Placental Responses in Fetal Development’, *Endocrinology*, 156(10), pp. 3422–3434. doi: 10.1210/en.2015-1227.

Salas-Huetos, A. *et al.* (2021) ‘Sperm DNA methylation changes after short-term nut supplementation in healthy men consuming a Western-style diet.’, *Andrology*. England, 9(1), pp. 260–268. doi: 10.1111/andr.12911.

Sales, V. M., Ferguson-Smith, A. C. and Patti, M. E. (2017) ‘Epigenetic Mechanisms of Transmission of Metabolic Disease across Generations’, *Cell Metabolism*. Elsevier Inc., 25(3), pp. 559–571. doi: 10.1016/j.cmet.2017.02.016.

Salzberg, B. L. and S. L. (2013) ‘Fast gapped-read alignment with Bowtie 2’, *Nature Methods*, 9(4), pp. 357–359. doi: 10.1038/nmeth.1923.Fast.

Sánchez-Margalet, V. *et al.* (2002) ‘Elevated plasma total homocysteine levels in hyperinsulinemic obese subjects’, *The Journal of Nutritional Biochemistry*, 13(2), pp. 75–79. doi: https://doi.org/10.1016/S0955-2863(01)00197-8.

Sharma, U. *et al.* (2016) ‘Biogenesis and function of tRNA fragments during sperm maturation and fertilization in mammals’, 351(6271), pp. 391–396. doi: 10.1126/science.aad6780.Biogenesis.

Shen, L. (2014) ‘GeneOverlap : An R package to test and visualize gene overlaps’, pp. 1–10.

Shen, Y. *et al.* (2012) ‘A map of the cis-regulatory sequences in the mouse genome.’, *Nature*, 488(7409), pp. 116–120. doi: 10.1038/nature11243.

Siklenka, K. *et al.* (2015) ‘Disruption of histone methylation in developing sperm impairs offspring health transgenerationally’, *Science*, 350(6261). doi: 10.1126/science.aab2006.

Soubry, A. *et al.* (2016) ‘Obesity-related DNA methylation at imprinted genes in human sperm: Results from the TIEGER study’, *Clinical Epigenetics*, 8(1), p. 51. doi: 10.1186/s13148-016-0217-2.

Stringer, J. M. *et al.* (2018) ‘Reduced PRC2 function alters male germline epigenetic programming and paternal inheritance’, *BMC Biology*. BMC Biology, 16(1), pp. 1–20. doi: 10.1186/s12915-018-0569-5.

Sun, N. *et al.* (2020) ‘H3K4me3-Mediated Upregulation of LncRNA-HEIPP in Preeclampsia Placenta Affects Invasion of Trophoblast Cells’, 11(December), pp. 1–11. doi: 10.3389/fgene.2020.559478.

Taiyun, Wei, Simko, V. (2021) ‘R package “corrplot”: Visualization of a Correlation Matrix.’ Available at: https://github.com/taiyun/corrplot.

Team, R. C. (2018) ‘R: A language and environment for statistical computing.’ Vienna, Austria: R Foundation for Statistical Computing.

Voight, B. F. *et al.* (2010) ‘Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis’, *Nature Genetics*, 42(7), pp. 579–589. doi: 10.1038/ng.609.

Wang, X. *et al.* (2013) ‘Paternally expressed genes predominate in the placenta’, *Proceedings of the National Academy of Sciences*, 110(26), pp. 10705–10710. doi: 10.1073/pnas.1308998110.

Waskom, M. L. (2021) ‘seaborn : statistical data visualization Statement of need’, *The Journal of Open Source Software*, 6, pp. 1–4. doi: 10.21105/joss.03021.

Wei, Y. *et al.* (2014) ‘Paternally induced transgenerational inheritance of susceptibility to diabetes in

mammals', *Proceedings of the National Academy of Sciences*, 111(5), pp. 1873–1878. doi: 10.1073/pnas.1321195111.

Welch, R. P. *et al.* (2014) 'ChIP-Enrich: gene set enrichment testing for ChIP-seq data', *Nucleic Acids Research*, 42(13), pp. e105–e105. doi: 10.1093/nar/gku463.

Wickham, H. (2016) *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag. Available at: https://ggplot2.tidyverse.org.

Wu, H. *et al.* (2017) 'Preconception urinary phthalate concentrations and sperm DNA methylation profiles among men undergoing IVF treatment: a cross-sectional study.', *Human reproduction (Oxford, England)*, 32(11), pp. 2159–2169. doi: 10.1093/humrep/dex283.

Wu, J. *et al.* (2016) 'The landscape of accessible chromatin in mammalian preimplantation embryos', *Nature*. Nature Publishing Group, 534(7609), pp. 652–657. doi: 10.1038/nature18606.

Yi, L. *et al.* (2018) 'Gene-level differential analysis at transcript-level resolution', *Genome Biology*. Genome Biology, 19(1), pp. 1–11. doi: 10.1186/s13059-018-1419-z.

Yoshida, K. *et al.* (2020) 'ATF7-Dependent Epigenetic Changes Are Required for the Intergenerational Effect of a Paternal Low-Protein Diet', *Molecular Cell*. Elsevier Inc., 78(3), pp. 445-458.e6. doi: 10.1016/j.molcel.2020.02.028.

Youngson, N. *et al.* (2016) 'Obesity-induced sperm DNA methylation changes at satellite repeats are reprogrammed in rat offspring YR - 2016/12/1', *Asian Journal of Andrology*, (6 UL- https://www.ajandrology.com/article.asp?issn=1008-682X;year=2016;volume=18;issue=6;spage=930;epage=936;aulast=Youngson;t=5), pp. 930 OP-936 VO–18. doi: 10.4103/1008-682X.163190.

Zhang, B. *et al.* (2016) 'Allelic reprogramming of the histone modification H3K4me3 in early mammalian development', *Nature*. Nature Publishing Group, 537(7621), pp. 553–557. doi: 10.1038/nature19361.

Zhang, Y., Parmigiani, G. and Johnson, W. E. (2020) 'ComBat-seq: batch effect adjustment for RNA-seq count data', *NAR Genomics and Bioinformatics*, 2(3). doi: 10.1093/nargab/lqaa078.

Zhang, Yunfang *et al.* (2018) 'Dnmt2 mediates intergenerational transmission of paternally acquired metabolic disorders through sperm small non-coding RNAs', *Nature Cell Biology*. Springer US, 20(May). doi: 10.1038/s41556-018-0087-2.

Pepin_Fig1

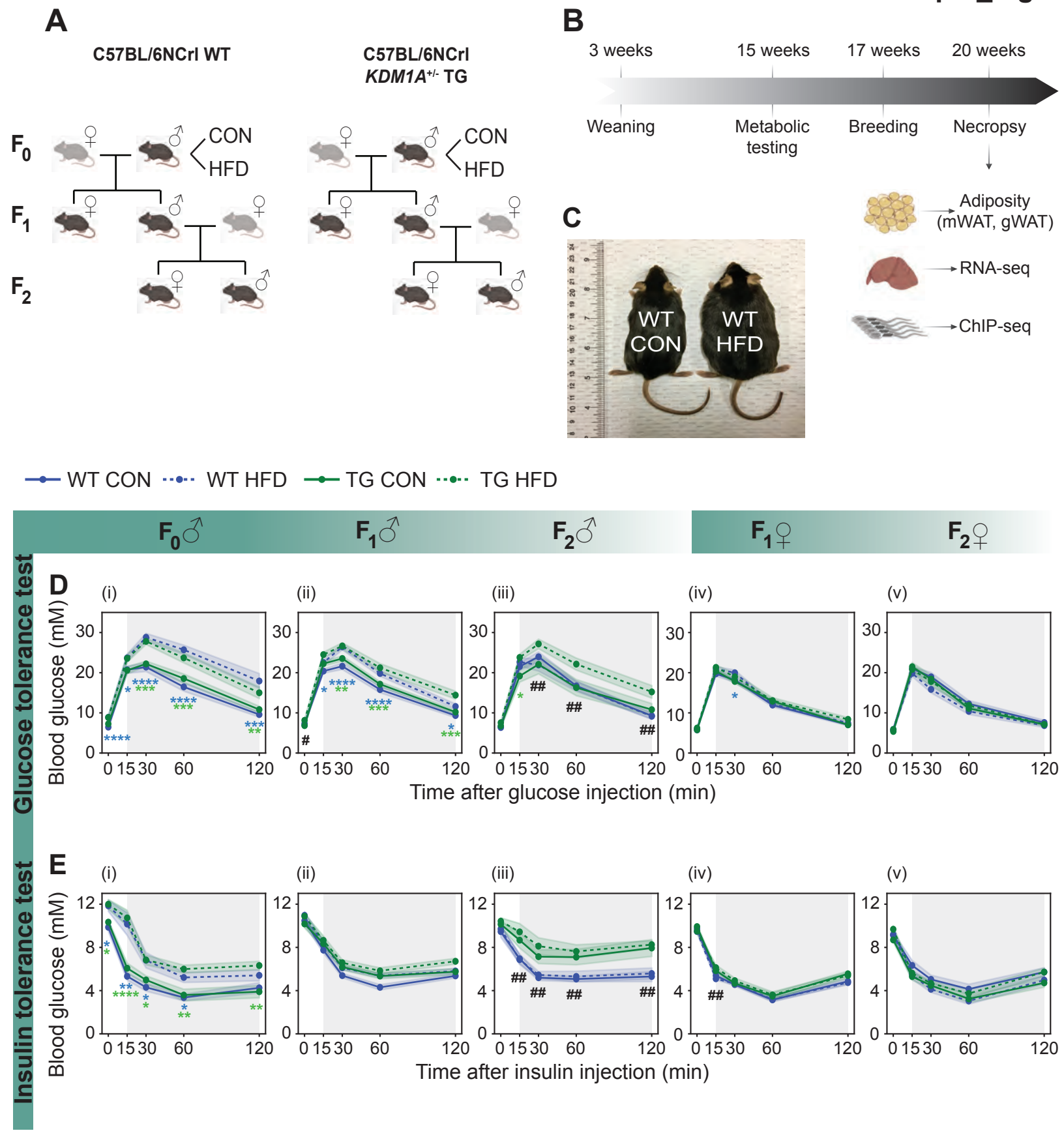
A
C57BL/6NCrl WT
C57BL/6NCrl
KDM1A+/- TG
F0
F1
F2
CON
HFD
B
3 weeks
15 weeks
17 weeks
20 weeks
Weaning
Metabolic testing
Breeding
Necropsy
Adiposity (mWAT, gWAT)
RNA-seq
ChIP-seq
C
WT CON
WT HFD
WT CON
WT HFD
TG CON
TG HFD
F0♂
F1♂
F2♂
F1♀
F2♀
Glucose tolerance test
D
(i)
(ii)
(iii)
(iv)
(v)
Blood glucose (mM)
Time after glucose injection (min)
Insulin tolerance test
E
Time after insulin injection (min)

Figure 1: Paternal obesity induces transgenerational metabolic phenotypes in a sex-specific manner that are enhanced in KDM1A descendants A) Experimental mouse model depicting breeding scheme and generations studied. Male C57BL6NCrl (WT) and KDM1A$^{+/-}$ transgenics (TG, C57BL6NCrl) were fed either a control diet (CON) or high-fat diet (HFD) from weaning for 10-12 weeks, then mated to 8-week-old C57BL6NCrl females fed a regular chow diet (CD). Animals studied per experimental group: F_0 (n=15-25 males), F_1 (n=28-49 per sex) and F_2 (n=8-21 per sex) B) Experimental timeline for metabolic testing and downstream experiments performed for each generation (F_{0-2}). Metabolic profiles were measured after the diet intervention at 15 weeks of age and included: baseline blood glucose, and intraperitoneal glucose and insulin tolerance tests (ipGTT and ipITT, respectively). Visceral adipose depots were weighed (mWAT: mesenteric white adipose tissue and gWAT: gonadal white adipose tissue) and the left lateral lobe of the liver used for RNA-sequencing (RNA-seq). Sperm from cauda epididymides were used for chromatin immunoprecipitation followed by sequencing (ChIP-seq), targeting histone H3 lysine 4 tri-methylation (H3K4me3). C) Age-matched male mice fed either a control diet (left) or a high-fat diet (right) for 12 weeks. D) Glucose tolerance test. Blood glucose levels before and after (shaded in grey) an intraperitoneal glucose injection, after overnight fasting (15 ±1 hour) at 4 months of age in F_0 males (i), F_1 males (ii), F_2 males (iii), F_1 females (iv) and F_2 females (v). E) Insulin tolerance test. Blood glucose levels before and after (shaded in grey) an intraperitoneal insulin injection, after a 6-hour (±1 hour) fasting at 4 months of age in F_0 males (i), F_1 males (ii), F_2 males (iii), F_1 females (iv) and F_2 females (v). Results are shown as mean ± SEM. Statistical analyses were performed using multiple t-test with Holm-Sidak correction. *P<0.05, **P<0.01, ***P<0.001, ****P<0.0001 (in blue; WT CON vs WT HFD, in green; TG CON vs TG HFD) and $^{\#}$P<0.05, $^{\#\#}$P<0.01 (WT HFD vs TG HFD).

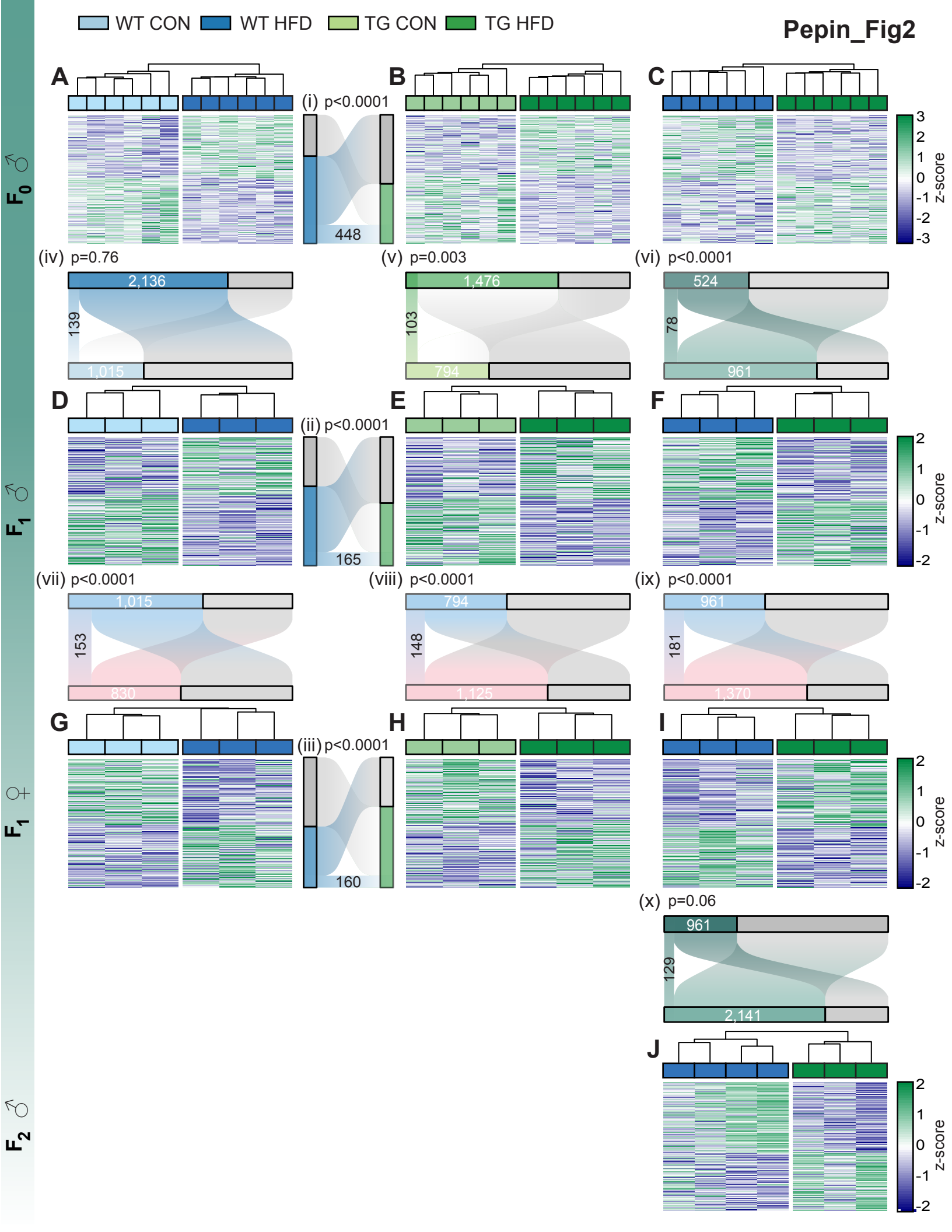
WT CON
WT HFD
TG CON
TG HFD
Pepin_Fig2
F0 ♂
F1 ♂
F1 ♀
F2 ♂
A
B
C
D
E
F
G
H
I
J
(i) p<0.0001
448
(iv) p=0.76
2,136
139
1,015
(v) p=0.003
1,476
103
794
(vi) p<0.0001
524
78
961
(ii) p<0.0001
165
(vii) p<0.0001
1,015
153
830
(viii) p<0.0001
794
148
1,125
(ix) p<0.0001
961
181
1,370
(iii) p<0.0001
160
(x) p=0.06
961
129
2,141
z-score
3
2
1
0
-1
-2
-3

Figure 2: Paternal obesity is associated with altered gene expression in the livers of the F_0-F_2

A-J) Heatmaps of normalized expression values scaled by row (z-score) for transcripts that code for differentially expressed hepatic genes (Lancaster p-value<0.05) for each comparison assessed across sex and generation. Individual transcripts (rows) are ordered by k-means clustering and samples (columns) are arranged by hierarchical clustering, using complete-linkage clustering based on Euclidean distance. F_0 WT CON vs WT HFD males (A), F_0 TG CON vs TG HFD males (B), F_0 WT HFD vs TG HFD males (C), F_1 WT CON vs WT HFD males (D), F_1 TG CON vs TG HFD males (E), F_1 WT HFD vs TG HFD males (F), F_1 WT CON vs WT HFD females (G), F_1 TG CON vs TG HFD females (H), F_1 WT HFD vs TG HFD females (I), and F_2 WT HFD vs TG HFD males (J). i-x) Alluvial plots depicting frequency distributions of significant (colored boxes) and non-significant (grey boxes) genes for each comparison and their overlap across genotype (i-iii), across F_0 and F_1 males (iv-vi), across F_1 males and females (vii-ix) and across F_1 and F_2 males (x). Significance of overlap between differentially expressed genes lists was calculated by Fisher's exact test. P-values are included for each comparison above the respective alluvial plot.

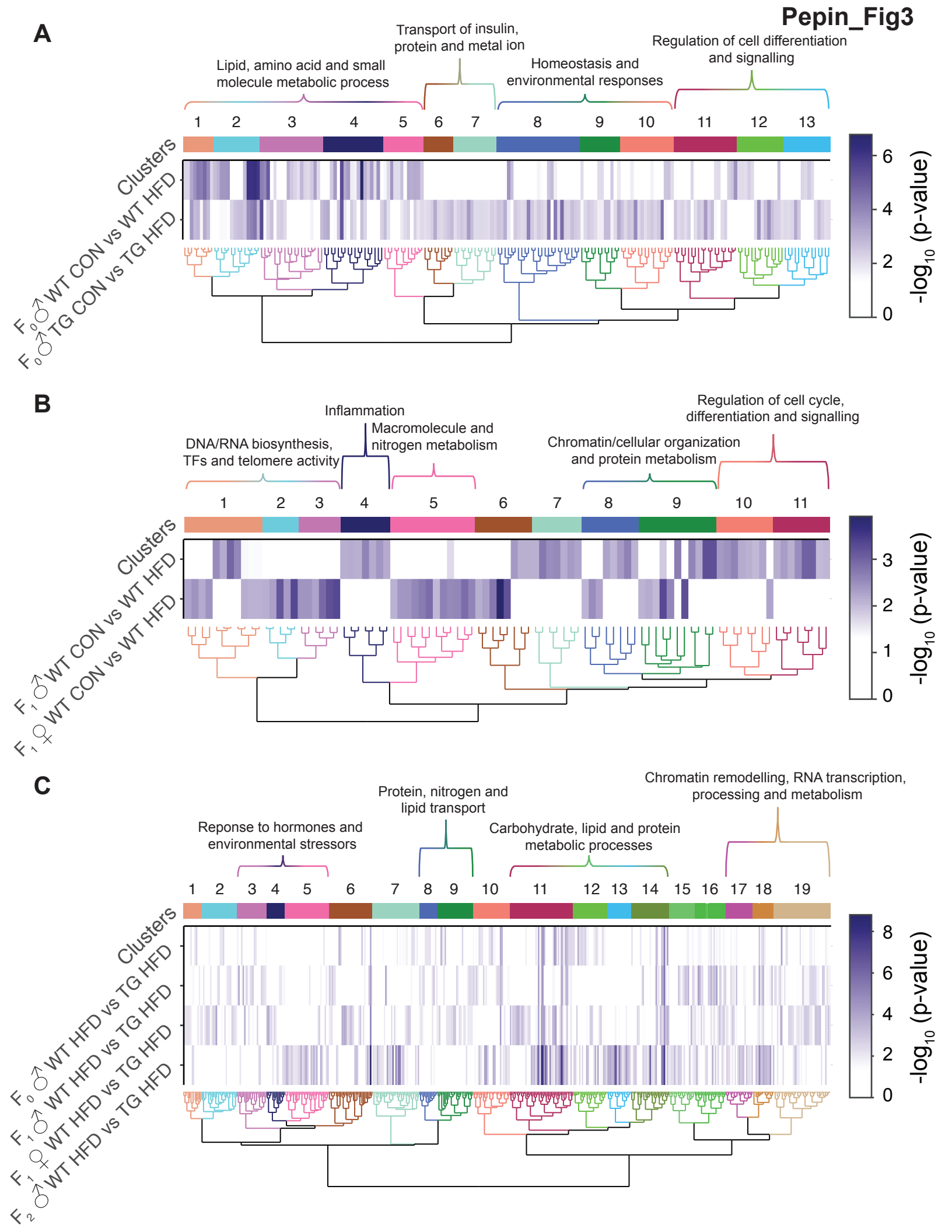
Pepin_Fig3
A
Lipid, amino acid and small molecule metabolic process
Transport of insulin, protein and metal ion
Homeostasis and environmental responses
Regulation of cell differentiation and signalling
1 2 3 4 5 6 7 8 9 10 11 12 13
Clusters
F0 ♂ WT CON vs WT HFD
F0 ♀ TG CON vs TG HFD
-log10 (p-value)
0 2 4 6
B
DNA/RNA biosynthesis, TFs and telomere activity
Inflammation
Macromolecule and nitrogen metabolism
Chromatin/cellular organization and protein metabolism
Regulation of cell cycle, differentiation and signalling
1 2 3 4 5 6 7 8 9 10 11
Clusters
F1 ♂ WT CON vs WT HFD
F1 ♀ WT CON vs WT HFD
-log10 (p-value)
0 1 2 3
C
Reponse to hormones and environmental stressors
Protein, nitrogen and lipid transport
Carbohydrate, lipid and protein metabolic processes
Chromatin remodelling, RNA transcription, processing and metabolism
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
Clusters
F0 ♂ WT HFD vs TG HFD
F1 ♂ WT HFD vs TG HFD
F1 ♀ WT HFD vs TG HFD
F2 ♂ WT HFD vs TG HFD
-log10 (p-value)
0 2 4 6 8

Figure 3: Obesity-induced hepatic transcriptome disturbances show functional similarities across genotype, sex and generation A-C) Heatmaps of significant gene ontology (GO) terms clustered by functional similarity, comparing enriched biological functions for each comparison of interest across genotype (A), sex (B) and generation (C). Columns represent enriched GO terms which are ordered by hierarchical clustering based on Wang's semantic similarity distance and *ward.D2* aggregation criterion. Each row represents a comparison of interest for which enriched GO terms were annotated based on the list of significant genes. The color gradient depicts the GO term enrichment significance ($-\log_{10}$ p-value). Interactive versions of these figures can be found in Supplemental files 1-3 and the complete lists of significantly enriched GO terms can be found in Tables S5-7.

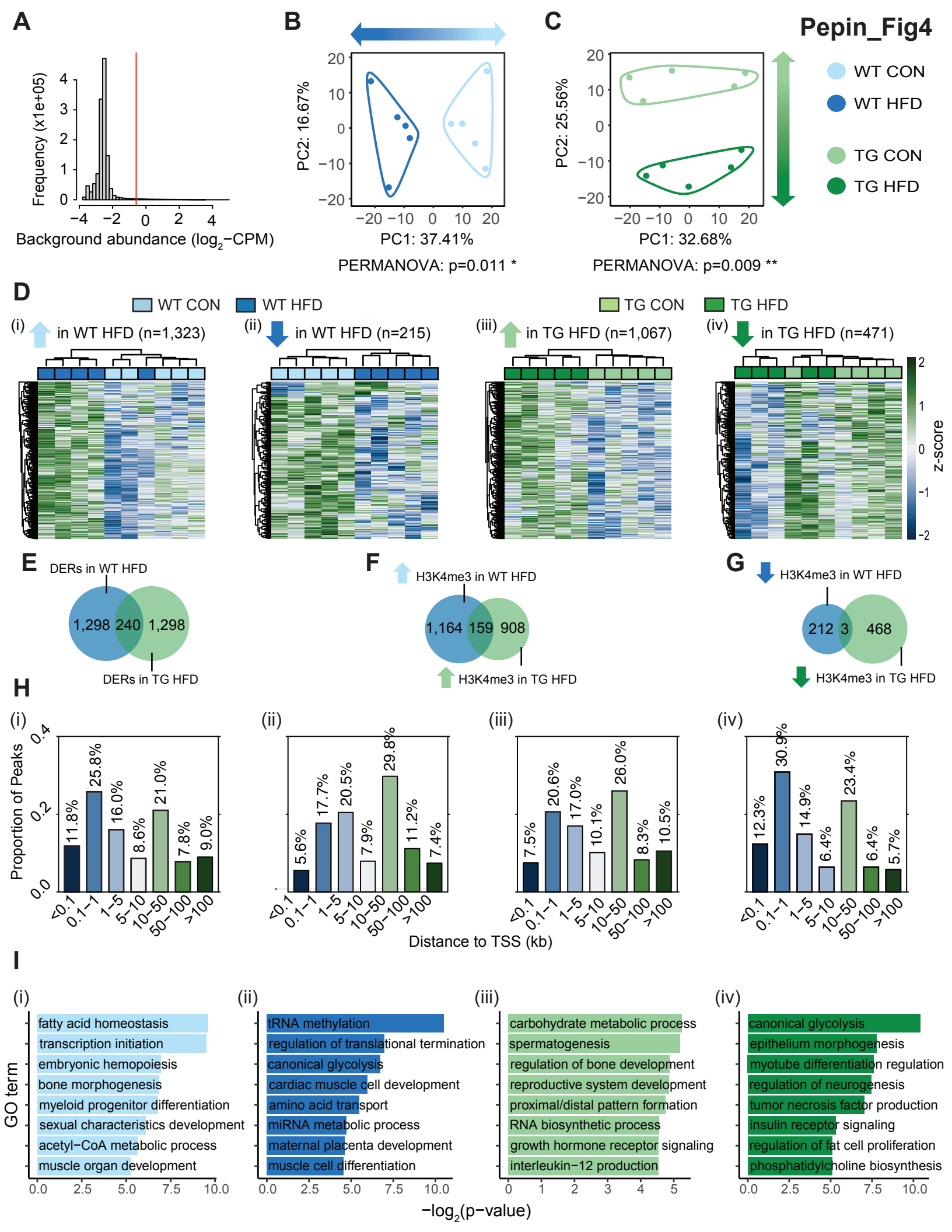
Pepin_Fig4
A
Frequency (x1e+05)
Background abundance (log2−CPM)
B
PC2: 16.67%
PC1: 37.41%
PERMANOVA: p=0.011 *
C
PC2: 25.56%
PC1: 32.68%
PERMANOVA: p=0.009 **
WT CON
WT HFD
TG CON
TG HFD
D
(i) in WT HFD (n=1,323)
(ii) in WT HFD (n=215)
(iii) in TG HFD (n=1,067)
(iv) in TG HFD (n=471)
z-score
E
DERs in WT HFD
1,298 240 1,298
DERs in TG HFD
F
H3K4me3 in WT HFD
1,164 159 908
H3K4me3 in TG HFD
G
H3K4me3 in WT HFD
212 3 468
H3K4me3 in TG HFD
H
Proportion of Peaks
(i) 11.8% 25.8% 16.0% 8.6% 21.0% 7.8% 9.0%
(ii) 5.6% 17.7% 20.5% 7.9% 29.8% 11.2% 7.4%
(iii) 7.5% 20.6% 17.0% 10.1% 26.0% 8.3% 10.5%
(iv) 12.3% 30.9% 14.9% 6.4% 23.4% 6.4% 5.7%
<0.1 0.1–1 1–5 5–10 10–50 50–100 >100
Distance to TSS (kb)
I
GO term
(i) fatty acid homeostasis; transcription initiation; embryonic hemopoiesis; bone morphogenesis; myeloid progenitor differentiation; sexual characteristics development; acetyl−CoA metabolic process; muscle organ development
(ii) tRNA methylation; regulation of translational termination; canonical glycolysis; cardiac muscle cell development; amino acid transport; miRNA metabolic process; maternal placenta development; muscle cell differentiation
(iii) carbohydrate metabolic process; spermatogenesis; regulation of bone development; reproductive system development; proximal/distal pattern formation; RNA biosynthetic process; growth hormone receptor signaling; interleukin−12 production
(iv) canonical glycolysis; epithelium morphogenesis; myotube differentiation regulation; regulation of neurogenesis; tumor necrosis factor production; insulin receptor signaling; regulation of fat cell proliferation; phosphatidylcholine biosynthesis
−log2(p−value)

Figure 4: Genomic location, directionality change and functions of regions with altered H3K4me3 enrichment by obesity A) Histogram showing frequency distributions of read abundances in 150 bp windows throughout the genome. Windows with an abundance below $log_2(4)$ fold over background bins of 2,000 bp were filtered out as indicated by the vertical red line. Enriched regions less than 100 bp apart were merged for a maximum width of 5,000 bp, conferring a total of 30,745 merged enriched regions. Reads were counted in merged enriched regions and normalized counts were used for downstream analyses. (see Material and Methods) B-C) Principal component analysis on normalized counts at merged enriched regions comparing WT CON vs WT HFD (B) and TG CON vs TG HFD (C). The top 5% regions contributing to separation of samples along Principal Component 1 (in B; PC1; x axis) or PC2 (in C; y axis) were selected. The PERMANOVA p-values indicating significance associated with dietary treatment are included under each PCA plot. D) Heatmaps of log2 normalized counts of deH3K4me3 regions in sperm with increased enrichment in WT HFD (i; n=1,323), decreased enrichment in WT HFD (ii; n=215), increased enrichment in TG HFD (iii; n=1,067) and decreased enrichment in TG HFD (iv; n=471) in each group. Samples (columns) and regions (rows) are arranged by hierarchical clustering using complete-linkage clustering based on Euclidean distance. Colored boxes indicate sample groups (light blue=WT CON, dark blue=WT HFD, light green=TG CON, dark green=TG HFD). E-G). Venn diagrams showing the overlap of deH3K4me3 in sperm of WT HFD (blue) and in TG HFD (green), for all detected regions (E), those gaining H3K4me3 (F) and those losing H3K4me3 (G). H) Barplots showing the distribution of altered regions based on the distance from the TSS of the nearest gene, for regions with increased enrichment in WT HFD (i; n=1,323), decreased enrichment in WT HFD (ii; n=215), increased enrichment in TG HFD (iii; n=1,067), and decreased enrichment in TG HFD (iv; n=471). The color gradient represents the distance of the regions to TSS in kilobase. I) Gene ontology analysis of diet-induced deH3K4me3 regions at promoters with increased enrichment in WT HFD (i; n=381), decreased enrichment in WT HFD (ii; n=34), increased enrichment in TG HFD (iii; n=230) and decreased enrichment in TG HFD (iv; n=150). Barplots show 8 selected significant GO terms with their respective $-log_2$(p-value). Tables S9-12 include the complete lists of significantly enriched GO terms.

Pepin_Fig5

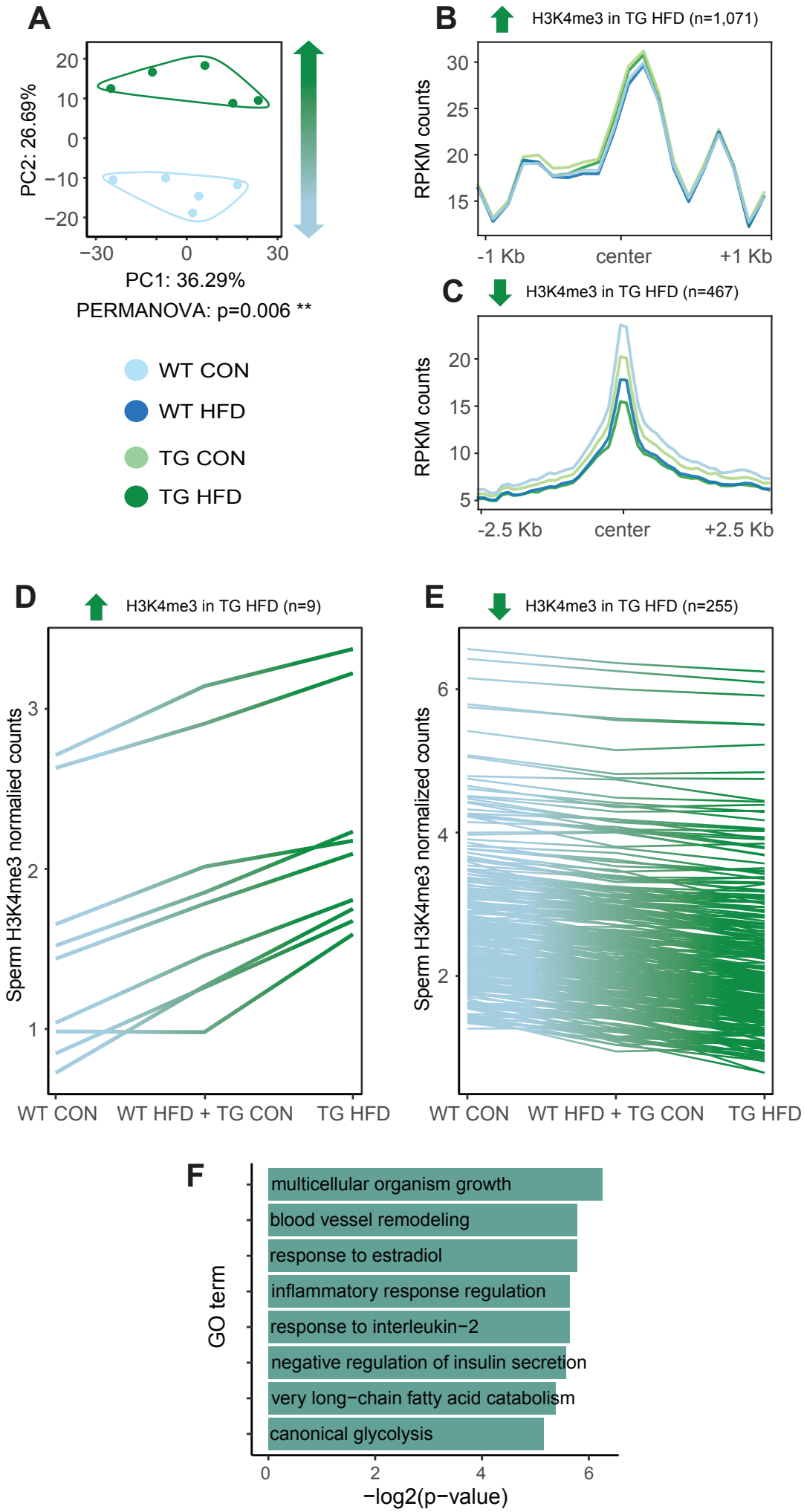
A
PC2: 26.69%
PC1: 36.29%
PERMANOVA: p=0.006 **
WT CON
WT HFD
TG CON
TG HFD
B
H3K4me3 in TG HFD (n=1,071)
RPKM counts
-1 Kb
center
+1 Kb
C
H3K4me3 in TG HFD (n=467)
RPKM counts
-2.5 Kb
center
+2.5 Kb
D
H3K4me3 in TG HFD (n=9)
Sperm H3K4me3 normalied counts
WT CON
WT HFD + TG CON
TG HFD
E
H3K4me3 in TG HFD (n=255)
Sperm H3K4me3 normalized counts
WT CON
WT HFD + TG CON
TG HFD
F
GO term
multicellular organism growth
blood vessel remodeling
response to estradiol
inflammatory response regulation
response to interleukin−2
negative regulation of insulin secretion
very long−chain fatty acid catabolism
canonical glycolysis
−log2(p−value)

Figure 5: Additive effects of KDM1A overexpression and diet-induced obesity in the sperm epigenome at the level of H3K4me3

A) Principal component analysis on normalized counts at merged enriched regions comparing WT CON vs TG HFD. The top 5% regions contributing to separation of samples along Principal Component 2 (PC2; y axis) were selected. The PERMANOVA p-value under the plot indicates significance. B-C) Profile plots of RPKM H3K4me3 counts +/- 1 kilobase around the center of regions with increased H3K4me3 (B) and +/- 2.5 kilobase around the center of regions with decreased H3K4me3 enrichment in TG HFD (C). D-E) Line plots showing the median of normalized sperm H3K4me3 counts for each experimental group at regions showing a significant trend (n=264, adjusted p-value<0.2) with a linear increase in H3K4me3 enrichment (D; n=9) or a linear decrease in H3K4me3 enrichment (E; n=255) from WT CON, WT HFD, TG CON to TG HFD groups. F) Gene ontology analysis on the regions associated with a significant linear trend at promoters (n=104). Barplots show 8 selected significant GO terms with their respective $-\log_2$(p-value). Table S13 includes the complete list of significantly enriched GO terms.

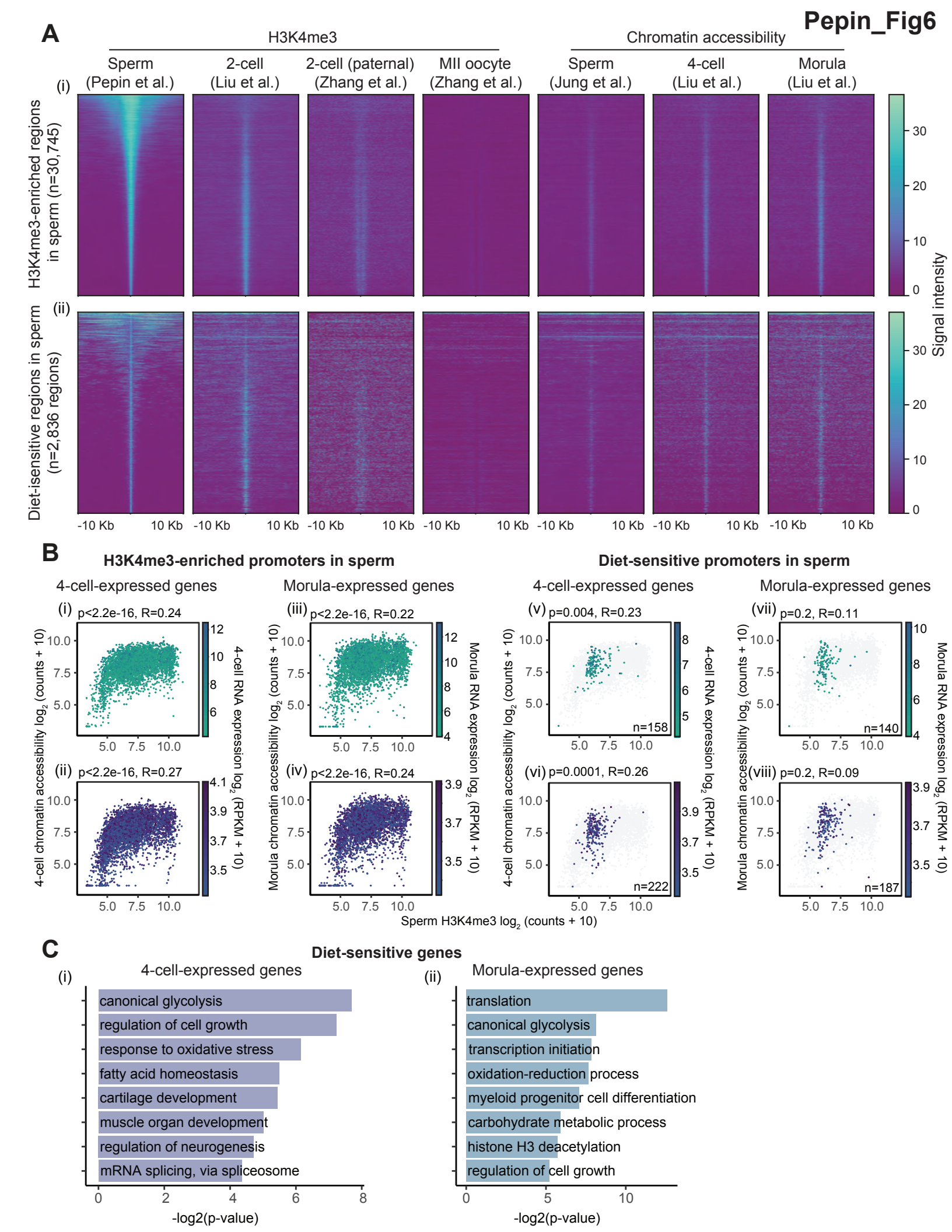
Pepin_Fig6
A
H3K4me3
Chromatin accessibility
Sperm (Pepin et al.)
2-cell (Liu et al.)
2-cell (paternal) (Zhang et al.)
MII oocyte (Zhang et al.)
Sperm (Jung et al.)
4-cell (Liu et al.)
Morula (Liu et al.)
(i)
H3K4me3-enriched regions in sperm (n=30,745)
(ii)
Diet-isensitive regions in sperm (n=2,836 regions)
Signal intensity
-10 Kb
10 Kb
B
H3K4me3-enriched promoters in sperm
Diet-sensitive promoters in sperm
4-cell-expressed genes
Morula-expressed genes
(i) p<2.2e-16, R=0.24
(ii) p<2.2e-16, R=0.27
(iii) p<2.2e-16, R=0.22
(iv) p<2.2e-16, R=0.24
(v) p=0.004, R=0.23
(vi) p=0.0001, R=0.26
(vii) p=0.2, R=0.11
(viii) p=0.2, R=0.09
n=158
n=222
n=140
n=187
4-cell chromatin accessibility log2 (counts + 10)
Morula chromatin accessibility log2 (counts + 10)
4-cell RNA expression log2 (RPKM + 10)
Morula RNA expression log2 (RPKM + 10)
Sperm H3K4me3 log2 (counts + 10)
C
Diet-sensitive genes
(i) 4-cell-expressed genes
canonical glycolysis
regulation of cell growth
response to oxidative stress
fatty acid homeostasis
cartilage development
muscle organ development
regulation of neurogenesis
mRNA splicing, via spliceosome
-log2(p-value)
(ii) Morula-expressed genes
translation
canonical glycolysis
transcription initiation
oxidation-reduction process
myeloid progenitor cell differentiation
carbohydrate metabolic process
histone H3 deacetylation
regulation of cell growth
-log2(p-value)

Figure 6: Sperm H3K4me3 regions sensitive to obesity occur at genes with an open chromatin state and expressed in the pre-implantation embryo A) Heatmaps of RPKM counts signal +/- 10 kilobase around the center of regions enriched with H3K4me3 in sperm (i; n=30,745) and regions with obesity-induced deH3K4me3 in sperm (ii; n=2,836) for H3K4me3 enrichment levels in sperm (this study), 2-cell embryo (Liu *et al.*, 2016), 2-cell embryo on the paternal allele and MII oocyte (Zhang *et al.*, 2016), and for chromatin accessibility signal in sperm (Jung *et al.*, 2017), 4-cell embryo and morula embryo (Liu *et al.*, 2019). B) Scatterplots showing H3K4me3 enrichment in sperm (x axis; log2 counts + 10), chromatin accessibility signal (y axis; log2 counts + 10; (Jung *et al.*, 2017)) and gene expression levels (color gradient; log2 FPKM + 10; (Liu *et al.*, 2019)) in 4-cell (i,ii,v,vi) or in morula (iii,iv,vii,viii) embryos, at either all genes with promoters enriched with H3K4me3 in sperm (i-iv) or at diet-sensitive genes (v-viii). The top row of scatterplots includes lowly-expressed genes (bottom 50%) in 4-cell (i and v) or morula (iii or vii) embryos. The bottom row of scatterplots includes highly-expressed genes (top 50%) in 4-cell (ii and iv) or morula (vi and viii) embryos. Pearson's correlation coefficients and their associated p-values are indicated above each scatterplot, comparing H3K4me3 enrichment in sperm versus H3K4me3 enrichment in 4-cell or morula embryos. C) Gene ontology analysis of genes expressed in the 4-cell (i) or the morula (ii) embryos, overlapping with diet-sensitive promoters in sperm. Barplots show 8 selected significant GO terms with their respective $-\log_2$(p-value). Tables S14-15 include the complete lists of significantly enriched GO terms.

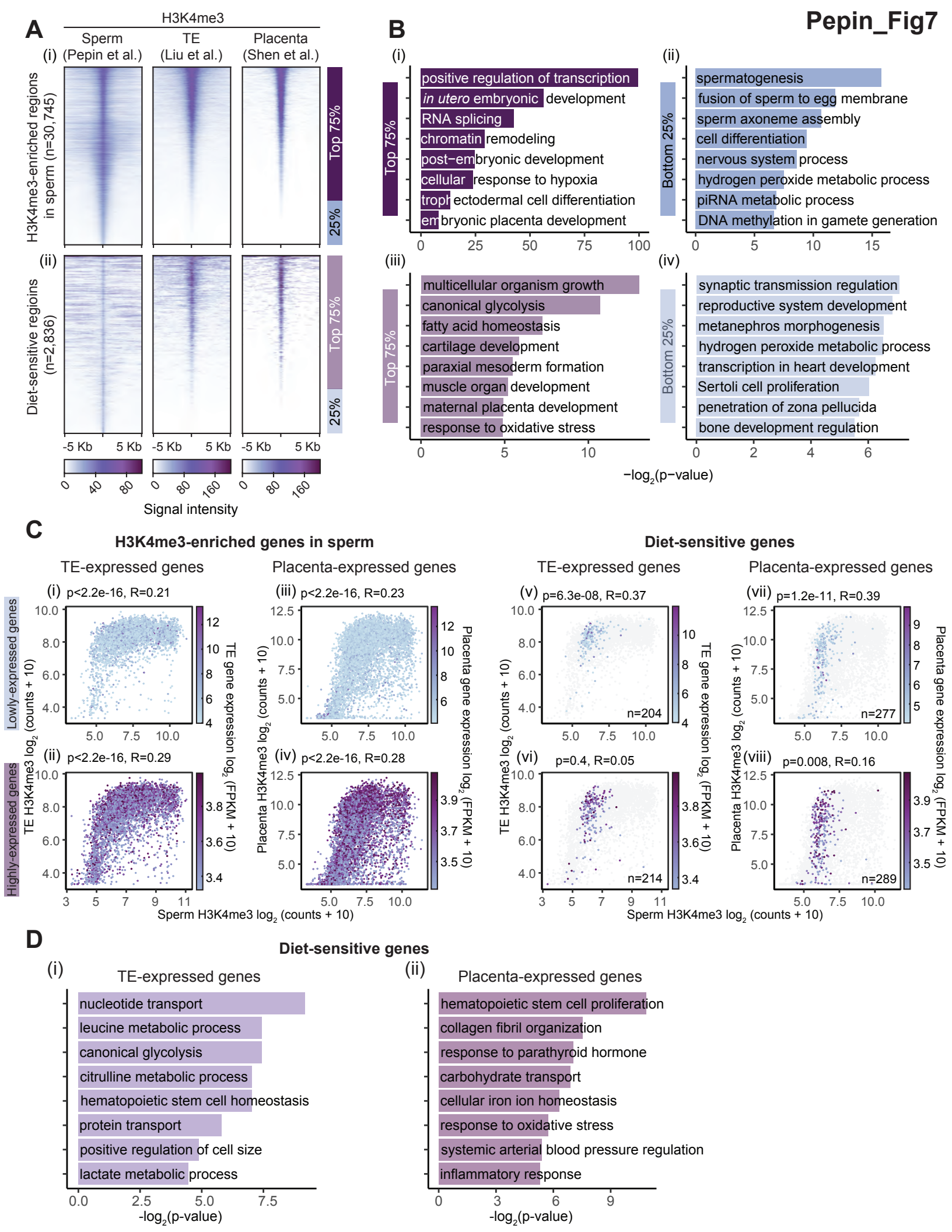
Pepin_Fig7
A
H3K4me3
Sperm (Pepin et al.)
TE (Liu et al.)
Placenta (Shen et al.)
(i)
H3K4me3-enriched regions in sperm (n=30,745)
Top 75%
25%
(ii)
Diet-sensitive regioins (n=2,836)
Top 75%
25%
-5 Kb
5 Kb
Signal intensity
B
(i)
Top 75%
positive regulation of transcription
in utero embryonic development
RNA splicing
chromatin remodeling
post−embryonic development
cellular response to hypoxia
tropho ectodermal cell differentiation
embryonic placenta development
(ii)
Bottom 25%
spermatogenesis
fusion of sperm to egg membrane
sperm axoneme assembly
cell differentiation
nervous system process
hydrogen peroxide metabolic process
piRNA metabolic process
DNA methylation in gamete generation
(iii)
Top 75%
multicellular organism growth
canonical glycolysis
fatty acid homeostasis
cartilage development
paraxial mesoderm formation
muscle organ development
maternal placenta development
response to oxidative stress
(iv)
Bottom 25%
synaptic transmission regulation
reproductive system development
metanephros morphogenesis
hydrogen peroxide metabolic process
transcription in heart development
Sertoli cell proliferation
penetration of zona pellucida
bone development regulation
−log2(p−value)
C
H3K4me3-enriched genes in sperm
Diet-sensitive genes
TE-expressed genes
Placenta-expressed genes
TE-expressed genes
Placenta-expressed genes
(i) p<2.2e-16, R=0.21
(ii) p<2.2e-16, R=0.29
(iii) p<2.2e-16, R=0.23
(iv) p<2.2e-16, R=0.28
(v) p=6.3e-08, R=0.37
(vi) p=0.4, R=0.05
(vii) p=1.2e-11, R=0.39
(viii) p=0.008, R=0.16
Lowly-expressed genes
Highly-expressed genes
TE H3K4me3 log2 (counts + 10)
Placenta H3K4me3 log2 (counts + 10)
TE gene expression log2 (FPKM + 10)
Placenta gene expression log2 (FPKM + 10)
n=204
n=214
n=277
n=289
Sperm H3K4me3 log2 (counts + 10)
D
Diet-sensitive genes
(i)
TE-expressed genes
nucleotide transport
leucine metabolic process
canonical glycolysis
citrulline metabolic process
hematopoietic stem cell homeostasis
protein transport
positive regulation of cell size
lactate metabolic process
-log2(p-value)
(ii)
Placenta-expressed genes
hematopoietic stem cell proliferation
collagen fibril organization
response to parathyroid hormone
carbohydrate transport
cellular iron ion homeostasis
response to oxidative stress
systemic arterial blood pressure regulation
inflammatory response
-log2(p-value)

Figure 7: Obesity-induced deH3K4me3 regions overlap with genes marked by H3K4me3 and expressed in the trophectoderm and placenta A) Heatmaps of RPKM counts signal +/- 5 kilobase around the center of regions enriched with H3K4me3 in sperm (i; n=30,745) and at regions with diet-induced deH3K4me3 in sperm (n=2,836) for H3K4me3 enrichment levels in sperm (this study), trophectoderm (TE) (Liu *et al.*, 2016) and placenta (Shen *et al.*, 2012). B) Gene ontology analysis of regions enriched with H3K4me3 in sperm, TE and placenta (top 75% from A i) (i), regions enriched with H3K4me3 in sperm only (bottom 25% from A i) (ii), diet-sensitive regions enriched with H3K4me3 in sperm, TE and placenta (top 75% from A ii) (iii), and diet-sensitive regions enriched with H3K4me3 in sperm only (bottom 25% from A ii) (iv). Barplots show 8 selected significant GO terms with their respective $-\log_2$(p-value). Tables S16-19 include the complete lists of significantly enriched GO terms. C) Scatterplots showing H3K4me3 enrichment at promoters in sperm (x axis; log2 counts + 10), H3K4me3 enrichment (y axis; log2 counts + 10) and gene expression levels (color gradient; log2 FPKM + 10) in the trophectoderm (i,ii,v,vi; (Liu *et al.*, 2016)) or in the placenta (iii,iv,vii,viii; (Shen *et al.*, 2012; Chu *et al.*, 2019)), at either all genes with promoters enriched with H3K4me3 in sperm (i-iv) or at diet-sensitive genes (v-viii). The top row of scatterplots includes lowly-expressed genes (bottom 50%) in trophectoderm (i and v) or placenta (iii or vii). The bottom row includes highly-expressed genes (top 50%) in trophectoderm (ii and iv) or placenta (vi and viii). Pearson's correlation coefficients and associated p-values are indicated above each scatterplot, comparing H3K4me3 enrichment in sperm versus H3K4me3 enrichment in the trophectoderm or placenta. D) Gene ontology analysis of genes expressed in the trophectoderm (i) or the placenta (ii), overlapping with diet-sensitive promoters in sperm. Barplots show 8 selected significant GO terms with their respective $-\log_2$(p-value). Tables S20-21 include the complete lists of significantly enriched GO terms.