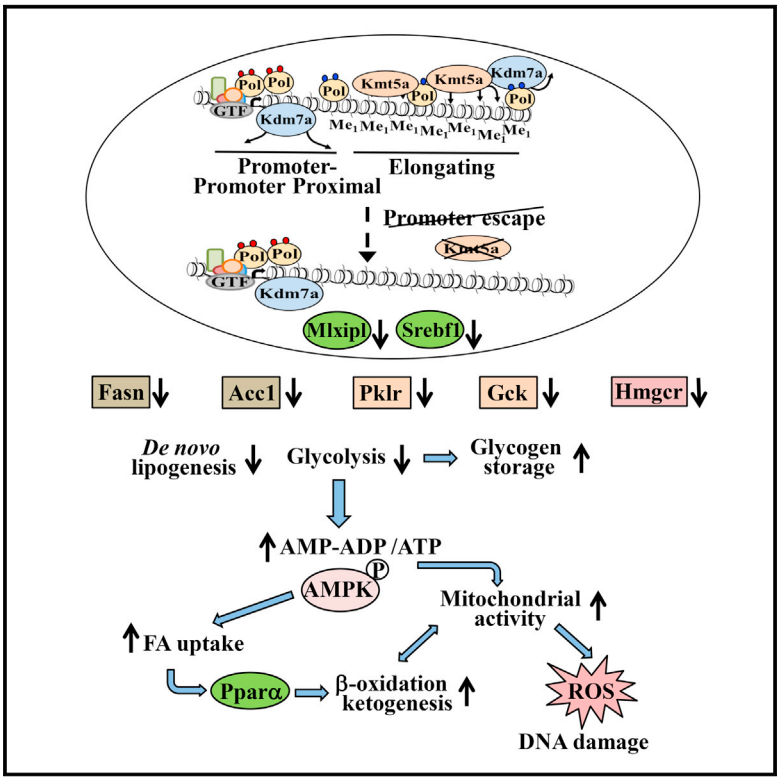
Article

Cell Reports

Kmt5a Controls Hepatic Metabolic Pathways by Facilitating RNA Pol II Release from Promoter-Proximal Regions

Graphical Abstract



Highlights

- H4K20Me$_1$ is dynamically deposited in the gene bodies of active genes
- Kmt5a regulates RNA Pol II release from promoter-proximal pause sites
- Kmt5a regulates metabolic gene transcription
- The transcription regulatory function of Kmt5a is important for genome integrity

Authors

Kostas C. Nikolaou, Panagiotis Moulos, Vangelis Harokopos, George Chalepakis, Iannis Talianidis

Correspondence

talianid@imbb.forth.gr

In Brief

Nikolaou et al. find that Kmt5a regulates the escape of RNA polymerase II from promoter-proximal pause sites and that this step is critical in the regulation of metabolic gene expression. The transcription regulatory function of Kmt5a is important for maintaining genome integrity in non-dividing cells.

Accession Numbers

GSE97338

Nikolaou et al., 2017, Cell Reports *20*, 909–922
July 25, 2017
http://dx.doi.org/10.1016/j.celrep.2017.07.003

CellPress

Cell Reports
Article

Kmt5a Controls Hepatic Metabolic Pathways by Facilitating RNA Pol II Release from Promoter-Proximal Regions

Kostas C. Nikolaou,[1] Panagiotis Moulos,[1] Vangelis Harokopos,[1] George Chalepakis,[2] and Iannis Talianidis[1,3,4,*]
[1]Biomedical Sciences Research Center Alexander Fleming, 16672 Vari, Greece
[2]Department of Biology, University of Crete, 71110 Herakleion, Crete, Greece
[3]Institute of Molecular Biology and Biotechnology, FORTH, 70013 Herakleion, Crete, Greece
[4]Lead Contact
*Correspondence: talianid@imbb.forth.gr
http://dx.doi.org/10.1016/j.celrep.2017.07.003

SUMMARY

H4K20 monomethylation maintains genome integrity by regulating proper mitotic condensation, DNA damage response, and replication licensing. Here, we show that, in non-dividing hepatic cells, H4K20Me1 is specifically enriched in active gene bodies and dynamically regulated by the antagonistic action of Kmt5a methylase and Kdm7b demethylase. In liver-specific Kmt5a-deficient mice, reduced levels of H4K20Me$_1$ correlated with reduced RNA Pol II release from promoter-proximal regions. Genes regulating glucose and fatty acid metabolism were most sensitive to impairment of RNA Pol II release. Downregulation of glycolytic genes resulted in an energy starvation condition partially compensated by AMP-activated protein kinase (AMPK) activation and increased mitochondrial activity. This metabolic reprogramming generated a highly sensitized state that, upon different metabolic stress conditions, quickly aggravated into a senescent phenotype due to ROS overproduction-mediated oxidative DNA damage. The results illustrate how defects in the general process of RNA Pol II transition into a productive elongation phase can trigger specific metabolic changes and genome instability.

INTRODUCTION

Regulation of transcription is a multistep process, initiated by the recruitment of sequence-specific transcription factors that facilitate the assembly of RNA polymerase II (RNA Pol II)-containing preinitiation complexes (PICs) at the regulatory regions of genes (Hochheimer and Tjian, 2003). Although PIC recruitment controls specificity, the steps subsequent to the assembly of RNA Pol II machinery are of pivotal importance. Pausing of RNA Pol II in promoter-proximal regions and its regulated release to the productive phase of transcript elongation are part of the complex mechanisms involved in regulating the transcription of most, if not all, eukaryotic genes (Nechaev et al., 2010; Adelman and Lis, 2012; Kwak et al., 2013; Jonkers and Lis, 2015). Negative elongation factor (NELF) and DRB sensitivity inducing factor (DSIF) stabilize RNA Pol II 30 to 60 nt downstream of transcription start sites (TSSs) (Wada et al., 1998a; Yamaguchi et al., 1999; Narita et al., 2003). Regulated release of RNA Pol II from the pause sites is mediated by the positive transcription elongation factor b (P-TEFb) complex, which phosphorylates the C-terminal domain of RNA Pol II at serine 2 and facilitates the eviction of NELF and DSIF (Marshall and Price, 1995; Wada et al., 1998b).

The role of chromatin structure in the post-recruitment processes is poorly understood. Genome-wide studies have revealed that the densities of H3K79Me$_2$, H2BK120Ub$_1$, and H4K20Me$_1$ at gene-coding regions positively correlate with transcript elongation rates (Barski et al., 2007; Fuchs et al., 2014; Veloso et al., 2014). While these studies point to a regulatory function of specific histone modifications in transcript elongation, their role in the earlier post-recruitment steps, such as pausing or regulated the escape of RNA Pol II, has not yet been investigated.

Among the gene body modifications, H4K20 monomethylation is of special interest because its global levels are highly regulated during the cell cycle. H4K20Me$_1$ concentrations are the highest in the G$_2$/M phase, gradually decline during the G$_1$ phase, and remain very low throughout the S phase (Rice et al., 2002; Beck et al., 2012b; van Nuland and Gozani, 2016). This correlates with the similar cell-cycle-dependent changes in Kmt5a (also known as Setd8 or PR-Set7) protein levels, the sole enzyme catalyzing H4K20Me$_1$. H4K20Me$_1$ mediates chromatin condensation via the recruitment of LMBTL1, which is important for the proper transition to mitosis (Kim et al., 2006; Min et al., 2007; Trojer et al., 2007). H4K20Me$_1$ is used as substrate by Suv4-20h to generate H4K20Me$_2$ and H4K20Me$_3$, which enhance nucleosomal folding and heterochromatin formation (Beck et al., 2012b). Furthermore, H4K20Me$_{2/3}$ is directly bound by ORC1 and ORCA, guiding replication origin selection (Beck et al., 2012a), or by 53BP1 at sites of DNA damage, which initiates DNA double-stranded repair (Oda et al., 2010; Tuzon et al., 2014). These findings established the view that Kmt5a-mediated H4K20 methylation is required for the maintenance of genome integrity (Abbas et al., 2010; Oda et al., 2010). We previously generated adult hepatocyte-specific *Kmt5a* knockout

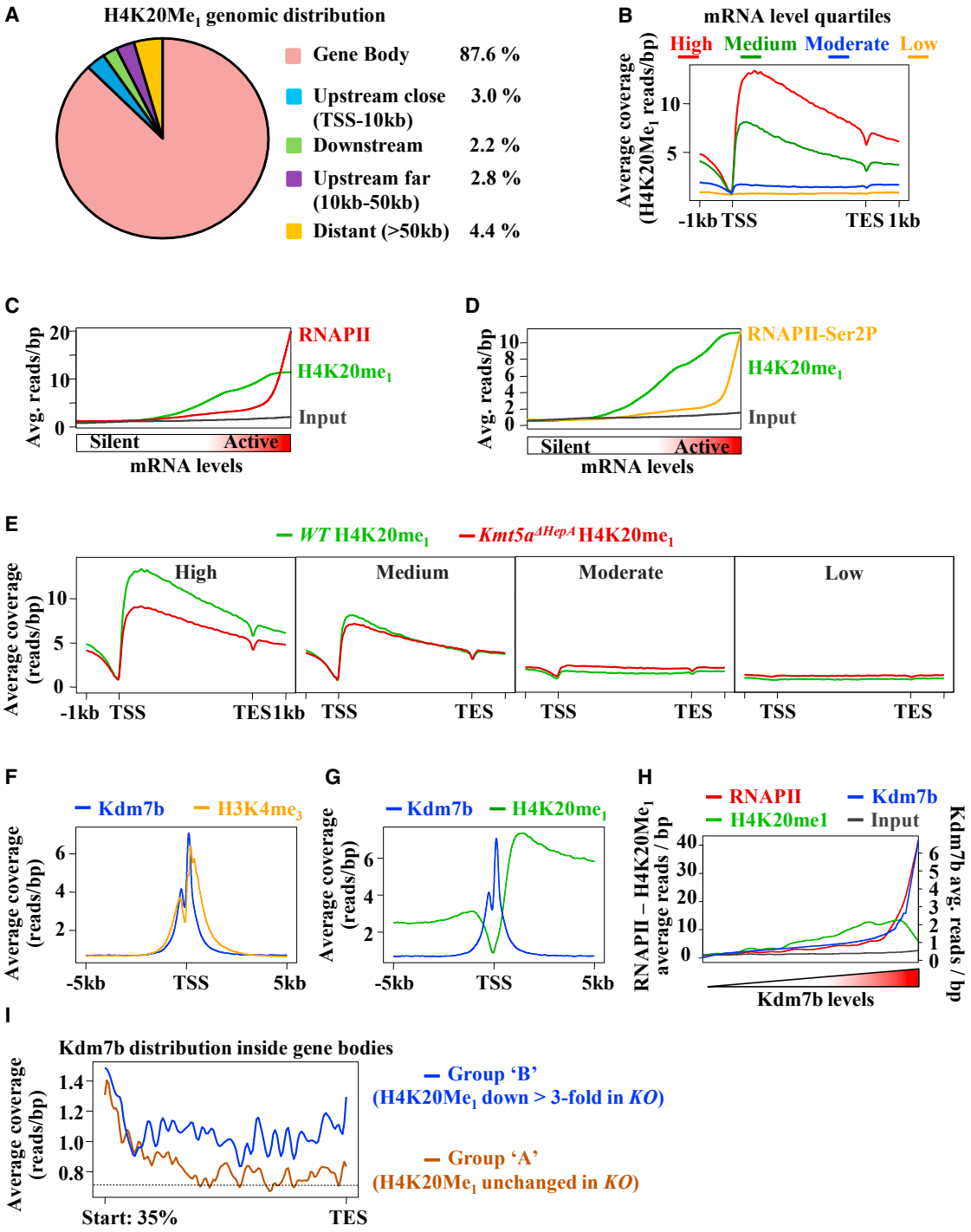


Figure 1. Dynamic H4K20 Monomethylation in the Intragenic Regions of Transcriptionally Active Genes

(A) Genomic distribution of ChIP-seq-enriched signal areas obtained with the H4K20Me$_1$ antibody in wild-type mice. Percentages indicate the proportion of the total (12,156) enriched regions falling in the respective categories.

(B) Average H4K20Me$_1$ coverage profiles at the genes expressed in mouse liver. Annotated genes (18,042 genes after filtering those with <1 kb of length) were divided into equal quartiles (high, medium, moderate, and low) according to their steady-state mRNA levels in wild-type P45 mouse livers as determined by RNA-seq analysis. The graph shows normalized average H4K20Me$_1$ reads/bp at the regions extending from −1 kb of the transcription start site (TSS) to +1 kb of the transcription end site (TES).

(legend continued on next page)

mice (*Kmt5a*$^{\Delta HepA}$) that, consistent with the well-established role of Kmt5a and H4K20Me$_1$ in genome integrity, displayed cell division-dependent DNA damage and hepatocyte necrosis after 3–4 months of age (Nikolaou et al., 2015). At earlier postnatal stages (e.g., postnatal day 45 [P45]), no major morphological or histological alterations could be detected. However, a small but significant increase in serum alanine aminotransferase (ALT) levels in young animals was indicative of hepatocyte dysfunction, which warranted further investigation.

Here, we investigated the early phenotypic changes in liver-specific Kmt5a-deficient mice and the in vivo function of H4K20 methylation in non-dividing hepatocytes. We show that H4K20Me$_1$ turnover in gene bodies positively correlates with gene activity and RNA Pol II release from promoter-proximal regions. Most sensitive to this promoter escape regulation are genes involved in glucose and lipid homeostasis, whose defect resulted in widespread metabolic reprogramming and genome damage. The results suggest that H4K20Me$_1$, in addition to its role in proper replication licensing, mitotic chromatin condensation, and DNA repair, safeguards genome integrity in non-dividing cells by controlling the transcription of metabolic genes at the post-initiation steps.

RESULTS

H4K20 Monomethylation Is Dynamically Deposited over the Gene Bodies of Active Genes

Mapping the genome-wide locations of H4K20Me$_1$-modified nucleosomes in P45 mouse livers revealed that most (87.6%) of them are located in gene body regions between TSSs and transcription end sites (TESs) (Figure 1A). Average coverage plots of H4K20Me$_1$ reads over the coding regions of annotated genes detected high methylation levels in the highly expressed genes, lower levels in the medium mRNA category genes, and low or lack of methylation in the moderately or poorly expressed gene groups (Figure 1B). Ranked coverage plot analyses of RNA Pol II chromatin immunoprecipitation sequencing (ChIP-seq) data obtained with antibody recognizing either all forms of the protein (RNA Pol II) or the form engaged in active elongation (the Ser2 form of RNA polymerase II [RNA Pol II-Ser2]) showed a positive, albeit not perfect, correlation with the density of H4K20Me$_1$ reads (Figures 1C and 1D). Because Kmt5a is the sole enzyme that can catalyze H4K20 monomethylation, the preceding data suggest that the mechanism of methylation of transcriptionally active gene bodies may involve traveling of Kmt5a with RNA Pol II. Despite extensive efforts, we could not obtain reliable ChIP-seq data for Kmt5a distribution, probably due to the short residence time of the protein in the genome (data not shown). However, we could detect in vivo interactions between Kmt5a and all forms of RNA Pol II in co-immunoprecipitation assays (Figure S1A).

As expected, H4K20Me$_1$ read densities were decreased in the livers of P45 *Kmt5a*$^{\Delta HepA}$ mice, which lack Kmt5a in hepatocytes from P20 (Figure 1E) (Nikolaou et al., 2015). However, this drop was mainly observed in the highly expressed gene group. Even in these genes, it was not eliminated, suggesting that H4K20Me$_1$ is a relatively stable modification. This selective gene group-specific decrease of H4K20Me$_1$ suggests that cell duplication-mediated halving cannot account for the observed H4K20Me$_1$ patterns.

To test whether gene-selective enzymatic demethylation may explain the observed patterns, we mapped the binding locations of the known H4K20Me$_1$ demethylase Kdm7b (also known as Phf8). Consistent with its ability to associate with H3K4Me$_3$ marks (Fortschegger et al., 2010), Kdm7b mainly occupied promoter-proximal nucleosomes, highly overlapping with H3K4Me$_3$ ChIP-seq peaks (Figure 1F). This distribution oppositely mirrored that of H4K20Me$_1$, which was sharply increased in gene body nucleosomes located downstream of those modified by H3K4Me$_3$ and occupied by Kdm7b (Figure 1G). The mutually exclusive occupancy pattern indicates that Kdm7b is a main H4K20 demethylase and that its activity keeps promoter-proximal nucleosomes devoid of the H4K20Me$_1$ mark. Average coverage plots ranked by increasing Kdm7b read counts revealed a perfect correlation between Kdm7b and RNA Pol II occupancy (Figure 1H). This is consistent with previous studies in HeLa cells demonstrating that Kdm7b is preferentially recruited to active genes via interaction with the C-terminal domain of RNA Pol II and with nucleosomes trimethylated at H3K4 (Fortschegger et al., 2010). H4K20Me$_1$ was enriched nearly proportionally with increasing Kdm7b occupancy, but only in genes with relatively low Kdm7b levels (Figure 1H). Genes with high levels of Kdm7b contained less H4K20Me$_1$ (Figure 1H). Inspection of individual gene tracks revealed three groups of genes categorized by specific Kdm7b occupancy patterns. The first group, representing about 1.5% of active, RNA Pol II-containing genes (e.g., *Alb* and *Pck1*), had high levels of Kdm7b all over the gene bodies and not concentrated at the promoter-proximal

(C and D) Correlation between gene body RNA Pol II and H4K20Me$_1$ coverage in the livers of P45 wild-type mice. Averaged ChIP-seq read densities of RNA Pol II (red line) and H4K20Me$_1$ (green line) (C) or RNA Pol II-Ser2 (orange line) and H4K20Me$_1$ (green line) (D) were ranked by increasing mRNA levels according to RPGM (reads per gene model) values. The read counts of the input sample are also shown (dark line).

(E) Average H4K20Me$_1$ coverage profiles of genes expressed at different levels in the livers of P45 wild-type (WT) and P45 *Kmt5a*$^{\Delta HepA}$ mice. High, medium, moderate, and low describe the steady-state mRNA levels of the quartiles as in (B).

(F and G) Overlay of Kdm7b (blue line) with that of H3K4Me$_3$ (orange line) (F) or H4K20Me$_1$ (green line) (G) average ChIP-seq profiles in P45 wild-type mouse livers.

(H) Correlation between Kdm7b occupancy and gene body RNA Pol II or H4K20Me$_1$ read densities in the livers of P45 wild-type mice. The graphs show average coverage profiles of Kdm7b (blue line), RNA Pol II (red line), and H4K20Me$_1$ (green line) in gene bodies ranked by increasing Kdm7b average read densities. The respective average gene coverage of the input sample (dark line) is also indicated.

(I) Distribution of Kdm7b read densities in downstream gene body regions. The graph shows average coverage (reads/bp) of Kdm7b in regions downstream of H3K4Me$_3$-containing nucleosomes in P45 wild-type livers. Group A genes (n = 95) correspond to H4K20Me$_1$-containing active genes, whose overall gene body methylation levels were not significantly changed in *Kmt5a*$^{\Delta HepA}$ mice. Group B genes (n = 95) correspond to H4K20Me$_1$-containing active genes, whose methylation levels decreased more than 3-fold in *Kmt5a*$^{\Delta HepA}$ mice.

See also Figures S1 and S2.

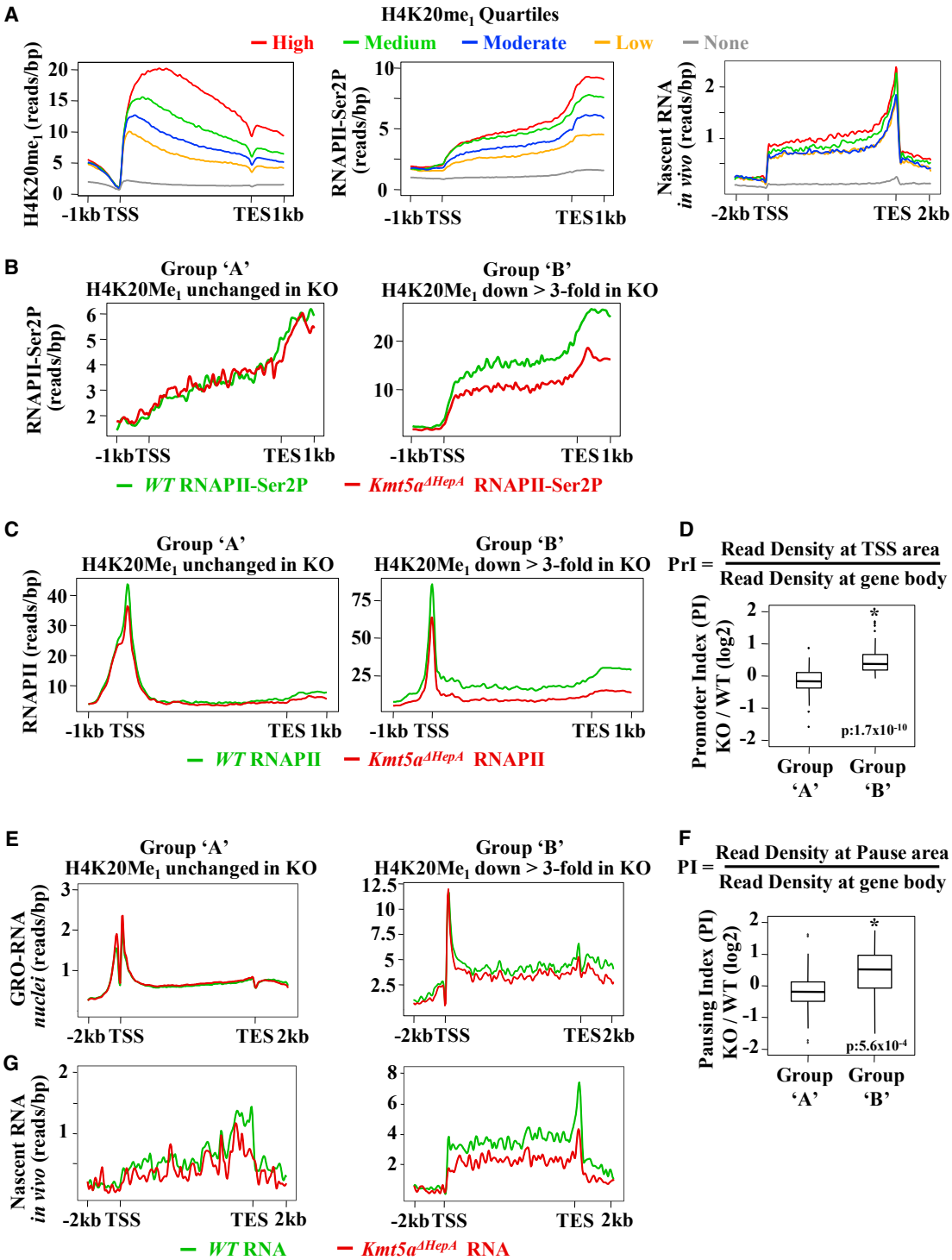


Figure 2. H4K20 Monomethylation Correlates with RNA Pol II Distribution on Active Genes

(A) Average H4K20Me$_1$ (left panel), RNA Pol II-Ser2 (middle panel), and in vivo ethynyl-uridine (EU)-labeled nascent RNA (right panel) coverage profiles at the genes expressed in P45 wild-type mouse livers. Genes containing H4K20Me$_1$ in their gene bodies (n = 6,570 genes with normalized reads per gene length > 0.03 bp) were divided into quartiles (high, red line, n = 1,643; medium, green line, n = 1,642; moderate, blue line, n = 1,642; and low, orange line, n = 1,643) according to H4K20Me$_1$ read densities. The distribution over non-methylated, annotated genes (n = 11,472 genes) is shown by the gray line. The graphs show average coverage of ChIP-seq reads obtained with H4K20Me$_1$ and RNA Pol II-Ser2 antibodies and EU-labeled nascent RNA reads over the gene body regions.

(legend continued on next page)

nucleosomal regions (Figure S1B). These genes are highly active in wild-type livers and lack H4K20Me$_1$, while RNA Pol II ChIP signals were distributed evenly along their promoter and gene bodies in both wild-type and Kmt5a-deficient livers (Figure S1B). The second group (~3.6% of active genes, e.g., *Plekhm2* or *Rfc5*) is characterized by low levels or the absence of Kdm7b, high H4K20Me$_1$ levels, and relatively even distribution of RNA Pol II (Figure S1C). H4K20Me$_1$ levels and RNA Pol II distribution on this group of genes were not changed significantly in Kmt5a-deficient mice. Most (92%) of the active genes belong to the third group, with high levels of Kdm7b and RNA Pol II at promoter-proximal regions and lower densities at gene bodies (Figure S1D). In most of these genes, H4K20 methylation is reduced in *Kmt5a*$^{\Delta HepA}$ proportional to the amounts of Kdm7b in the gene bodies. Genes with a high degree of H4K20Me$_1$ loss in *Kmt5a*$^{\Delta HepA}$ mice had low but detectable amounts of Kdm7b in gene bodies downstream of the promoter-proximal nucleosomes, while those whose H4K20Me$_1$ levels was unchanged had near-baseline Kdm7b levels (Figure 1I; Figures S2A and S2B).

These results suggest that the antagonistic action of Kmt5a methylase and Kdm7b demethylase generates a dynamic H4K20 monomethylation pattern across the gene bodies of transcriptionally active genes, which correlates with specific distribution patterns of RNA Pol II.

Kmt5a Regulates RNA Pol II Escape from the Promoter-Proximal Regions

Analysis of global RNA Pol II distribution profiles revealed that gene body H4K20Me$_1$ positively correlated with the amounts of the elongating, Ser2-phosphorylated form of RNA polymerase II (RNA Pol II-Ser2P) (Figure 2A). H4K20Me$_1$ levels also correlated with nascent RNA reads determined by sequencing of newly synthesized RNA labeled in vivo by ethynyl-uridine (Figure 2A, right). In *Kmt5a*$^{\Delta HepA}$ mice, the amounts of RNA Pol II-Ser2P decreased only in the group of genes (group B) in which significant loss of H4K20Me$_1$ was detected, not in the group (group A) in which H4K20Me$_1$ remained unaffected by *Kmt5a* inactivation (Figure 2B; Figure S2A). Similar differences in gene body locations were observed in ChIP-seq assays performed with an antibody recognizing all forms of RNA Pol II (Figure 2C). In this latter assay, we noticed that promoter and promoter-proximal region-bound RNA Pol II was only marginally reduced in group B genes, while gene body-bound RNA Pol II reads decreased at a larger extent. To obtain a quantitative view, we calculated the normalized RNA Pol II coverage (reads per base pair [bp]) in the 500 bp area centered on the TSSs and divided by the same window length of normalized RNA Pol II coverage in the rest of the gene body (promoter + pausing index [PrI]). Increased PrI values in *Kmt5a*$^{\Delta HepA}$ mice were observed only in genes (group B) whose gene body methylation was highly reduced in the absence of Kmt5a (Figure 2D). As an independent analysis, we grouped active genes according to their PrIs and compared the H4K20Me$_1$ coverage in their gene bodies. The average H4K20Me$_1$ levels were higher in genes with higher PrI values as opposed to those with PrI values near 1 (Figure S2C). These results demonstrate that Kmt5a function is important for the regulation of RNA Pol II escape from the promoter or promoter-proximal regions.

The preceding results were confirmed by the higher-resolution approach of global run-on sequencing (GRO-seq). In these assays, isolated nuclei are incubated with Bromo-Uridine triphosphate (Br-UTP) for a short time in the presence of sarcosyl, which limits the extension of nascent RNAs to already engaged polymerases (Core et al., 2008). In agreement with the RNA Pol II distribution profile, we observed high-level accumulation of short 30–50 nt transcripts that map downstream of the TSS (Figure 2E). As in other systems, we could also detect significant levels of promoter transcription in both group A and group B genes (Figure 2E). The resolution of this analysis also allowed the calculation of pausing indexes (PIs). Similar to the PrI ratios calculated from RNA Pol II distribution data, we observed increased pausing index values in *Kmt5a*$^{\Delta HepA}$ mice only in group B genes (Figure 2F).

The role of H4K20 methylation in nascent RNA synthesis was further demonstrated by the in vivo nascent RNA sequencing (RNA-seq) approach, in which newly synthesized RNA is labeled by ethynyl-uridine treatment of live animals. Although this assay does not detect promoter-proximal short transcripts, it provides valuable information about the in vivo dynamics of new RNA synthesis. Kmt5a inactivation resulted in reduced synthesis of new transcripts only in the group B gene category (Figure 2G).

These results suggest that Kmt5a regulates transcription at the step involving the escape of RNA Pol II from promoter-proximal locations.

Kmt5a-Mediated H4K20 Methylation Regulates Genes Involved in Glucose and Lipid Homeostasis

Gene expression profiling by RNA-seq identified 1,030 downregulated and 760 upregulated genes in the livers of P45 *Kmt5a*$^{\Delta HepA}$ mice compared to wild-type littermates (Figure 3A).

(B and C) Distribution of RNA Pol II-Ser2 (B) and total RNA Pol II (C) in the group of genes displaying low (group A) and high (group B) H4K20Me$_1$ methylation turnover. The graphs show average coverage of ChIP-seq reads.

(D) Quantitative comparison of changes in PrIs in hepatic genes displaying different methylation turnover. The PrI was calculated as the ratio of the normalized RNA Pol II coverage (reads/bp) of a 500 bp area around the TSS of mouse genes divided by the window length (500 bp) to the normalized RNA Pol II coverage over the rest of the gene body. Boxplots depict changes in PrI values between wild-type and *Kmt5a*$^{\Delta HepA}$ mice (knockout [KO]) in group A and group B genes as indicated. Statistical significance was assessed by Welch's t test.

(E) Distribution of nascent RNA reads labeled by Bromo-Uridine (BrU) in isolated nuclei evaluated by GRO-seq (GRO-RNA) in group A and group B genes. Note the different scales of the y axis.

(F) Pausing index (PI) calculated as the ratio of the normalized nascent RNA coverage (reads/bp) in the TSS to the +50 nt area of mouse genes divided by the window length (500 bp) to the normalized nascent RNA coverage over the rest of the gene body. The results are presented in boxplots as in (D).

(G) Distribution of in vivo nascent RNA reads by EU (nascent RNA in vivo) in group A and group B genes. Note the different scales of the y axis.

See also Figure S2.

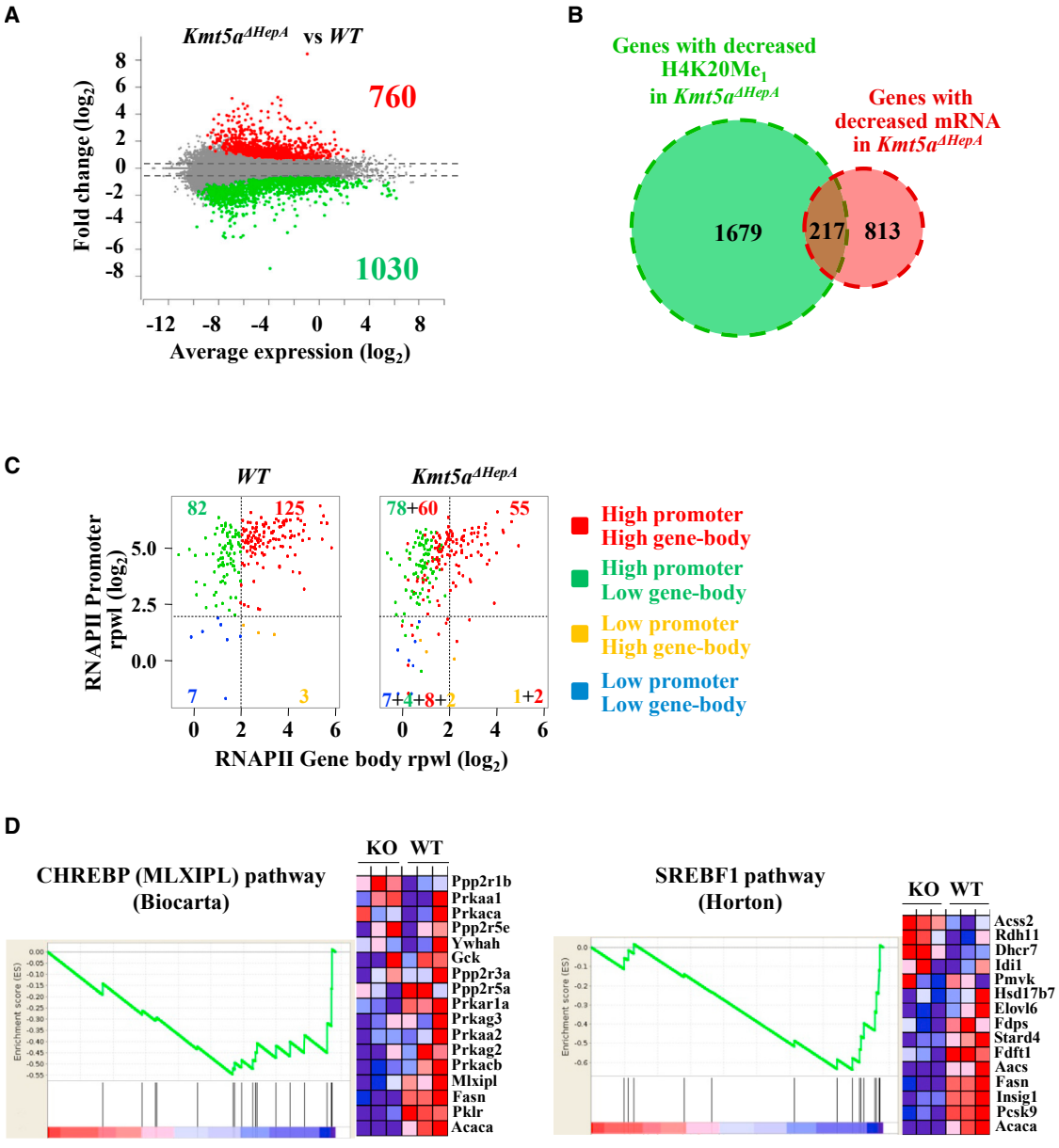


Figure 3. The Steady-State mRNA Levels of Only a Limited Number of Genes Are Affected by H4K20 Monomethylation

(A) Mean-difference plot depicting the gene expression changes in the livers of *Kmt5a*$^{\Delta HepA}$ against wild-type mice. Red dots and green dots correspond to upregulated and downregulated genes with a p value < 0.05, respectively. Gray dots correspond to all other genes.

(B) Venn diagram showing the overlap between genes whose mRNA and H4K20Me$_1$ levels decreased in *Kmt5a*$^{\Delta HepA}$ mice.

(C) Scatterplots showing the relationship between promoter (y axis) and gene body (x axis) occupancy of RNA Pol II in the 217 genes identified in (B) as correlating with changes in H4K20Me$_1$ levels. Genes are categorized, according to their read enrichment in promoters versus gene bodies, into four groups, as indicated at the right. Color codes and the respective numbers indicate the transition of genes among the different categories. For example, 60 of the 125 genes of the high promoter/high gene body category moved to the high promoter/low gene body group in *Kmt5a*$^{\Delta HepA}$ mice.

(D) Gene set enrichment analysis (GSEA) for the CHREBP-regulated (MLXIPL) (left panel) and the SREBF1-regulated (right panel) pathways. The figures depict the enrichment score (ES) as derived by gene set enrichment analysis for the two gene sets.

See also Figure S3.

Among the downregulated genes, only 217 had decreased levels of H4K20 monomethylation (Figure 3B). These 217 genes represent only a fraction of the 1,896 genes whose H4K20Me$_1$ levels decreased in P45 *Kmt5a*$^{\Delta HepA}$ mice (Figure 3B).

The RNA Pol II distribution profile of the 217 genes displayed a clear shift of RNA Pol II from gene bodies to promoter-proximal regions (Figure 3C), which is in agreement with a role of H4K20Me$_1$ in RNA Pol II escape from the promoter-proximal regions.

H4K20Me$_1$-sensitive genes were highly enriched in metabolism-related biological process terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway categories (Figures S3A and S3B). Among them were *Mlxipl* (also called *Chrebp*) and *Srebf1*, two main regulators of glucose and fatty acid synthesis pathways (Foretz et al., 1999; Horton et al., 2002; Iizuka et al., 2004; Uyeda and Repa, 2006). The RNA Pol II distribution, Kdm7b binding, H4K20Me$_1$, nascent RNA synthesis, and GRO-seq profiles of these regulators demonstrate that their expression is directly regulated by Kmt5a-dependent transition of RNA Pol II from the pause sites (Figures S4A and S4B). Gene set enrichment analysis revealed that most Mlxipl and Srebf1 targets were greatly affected by *Kmt5a* inactivation (Figure 3D). The mRNA and protein levels of both *Mlxipl* and *Srebf1*, as well as the mRNAs of key glycolytic, lipogenic, and cholesterol biosynthetic enzymes such as *Pklr*, *Gck*, *Fasn*, *Acc1*, *Scd1*, and *Hmgcr*, were significantly decreased in the livers of P45 *Kmt5a*$^{\Delta HepA}$ mice (Figures 4A–4E). We also detected increased mRNA levels of the genes involved in fatty acid uptake, hydrolysis, and detoxification (*Cd36*, *Acot1*, *Cyp2a4*, *Cyp2a14*, and *Cyp2a10*) and of the genes of the fatty acid β-oxidation and ketogenesis pathways (*Cpt1a*, *Hmgcs2*, and *Fgf21*) (Figures 4F and 4G). These latter genes are known targets of peroxysome proliferator-activated receptor alpha (Pparα), a nuclear receptor, which is activated by endogenous fatty acid ligands (Desvergne et al., 2006; Martinez-Jimenez et al., 2010; Lee et al., 2014). Comparisons with Pparα ChIP-seq data from Lee et al. (2014) revealed that a large part (43%) of the upregulated genes are bona fide functional targets of Pparα (Figure 4H).

The preceding changes resulted in impaired glucose tolerance of *Kmt5a*$^{\Delta HepA}$ mice (Figure 5A), without significant loss in insulin sensitivity (Figures 5B and 5C). Periodic acid Schiff (PAS) staining revealed an excessive accumulation of intracellular glycogen, as expected from the inhibition of the glycolysis pathway (Figure 5D). The shape of the glucose tolerance curve showed higher differences in early time points and full remission 2 hr following glucose challenge (Figure 5A), which points to a mechanism involving delayed glucose uptake due to high cellular glycogen content. Consistent with the changes in the mRNA levels of lipid metabolic enzymes, serum cholesterol and triglyceride levels decreased, while free fatty acids and β-hydroxybutyrate levels were increased in *Kmt5a*$^{\Delta HepA}$ mice (Figure 5E).

These data suggest that Kmt5a-dependent regulation of RNA Pol II escape from promoter-proximal regions affects the steady-state mRNA levels of only a subset of genes. They encode for key regulators and enzymes of glycolysis, de novo lipogenesis, and cholesterol biosynthesis pathways. These are accompanied by significant changes in the levels of serum and in the cellular levels of metabolic intermediates, which trigger additional indirect effects via the activation of Pparα and its target genes.

Kmt5-Dependent Regulation of Metabolic Gene Transcription Is Important for the Maintenance of Genome Integrity

The metabolic changes observed in the livers of *Kmt5a*$^{\Delta HepA}$ mice resemble to a condition of energy starvation, which primarily originates from the defects in the glycolysis pathway. In line with this, AMP-ADP/ATP ratios in the livers of *Kmt5a*$^{\Delta HepA}$ mice were increased significantly (Figure 6A). The overall ATP levels were about 38% higher in *Kmt5a*$^{\Delta HepA}$ mice compared to wild-type littermates, which is likely a result of an increased mitochondrial activity (described later). The increased ADP/ATP ratio arises from a more pronounced (about 125%) increase of ADP levels, which is expected to activate AMP-activated protein kinase (AMPK), the major regulator of cellular energy homeostasis (Cantó and Auwerx, 2010; Hardie, 2011; Burkewitz et al., 2014). We detected constitutively active AMPK in the livers of *Kmt5a*$^{\Delta HepA}$ mice using an antibody specifically recognizing the Thr172-phosphorylated form of the enzyme (Figure 6B). As a consequence of persistent AMPK activation, we detected highly increased mitochondrial activity by staining with the mitochondrial marker carbamoyl phosphate synthetase 1 (CPS1) and by measuring the enzymatic activity of the cytochrome *c* oxidase (COX) complex (Figure 6C). Electron microscopy imaging revealed that in *Kmt5a*$^{\Delta HepA}$ mice, increased mitochondrial activity correlated with an enlargement of most mitochondria (Figure 6D). In addition, consistent with the autophagy-stimulating function of AMPK, we frequently (in five of seven hepatocytes) observed autophagic vesicles by electron microscopy and detected increased levels of the lipidated form of Map1lc3a (LC3-II) in liver extracts from *Kmt5a*$^{\Delta HepA}$ mice (Figures 6D and 6E).

To test whether the gluconeogenesis pathway is affected by *Kmt5a* inactivation, we performed pyruvate tolerance tests. Surprisingly, exogenously added sodium pyruvate was not converted to glucose but instead generated a metabolic imbalance of irreversible hypoglycemia (Figure 7A). In parallel, hepatocytes became abnormally enlarged, resembling to damaged, senescent cells surrounded by invading inflammatory cells (Figure 7B). Similar abnormal enlargement of hepatocytes was observed when *Kmt5a*$^{\Delta HepA}$ mice were exposed to other metabolic stress conditions, such as high-fat diet or fasting for 24 or 48 hr (Figure S5A). As expected from the increased mitochondrial activity, we detected increased amounts of reactive oxygen species (ROS) in the livers of *Kmt5a*$^{\Delta HepA}$ mice (Figures 7C and 7D). These levels probably are not sufficient to induce spontaneous DNA damage, because they were not accompanied by increases in γH2AX-positive cells, at least during the time frame of our analyses (Figure 7E). However, ROS levels were highly increased in response to fasting or sodium pyruvate treatment, resulting in the appearance of a high number of γH2AX-positive hepatocytes (Figures 7C–7E) and an excessive accumulation of the senescence marker SA-β-gal (Figure S5B). The preceding phenotypic changes point to excessive liver damage, which was confirmed by highly increased serum ALT levels (Figure 7F).

Collectively, the preceding results suggest that in P45 *Kmt5a*$^{\Delta HepA}$ hepatocytes, AMPK-dependent pathways are activated to compensate for defects in glycolysis, which results in a labile state highly sensitive to stress conditions. Upon metabolic stress, rapid and irreversible accumulation of DNA damages occurs, leading to excessive cellular senescence and liver function failure.

DISCUSSION

The results of this paper demonstrate that Kmt5a controls the initial phase of transcription elongation and that the genes of

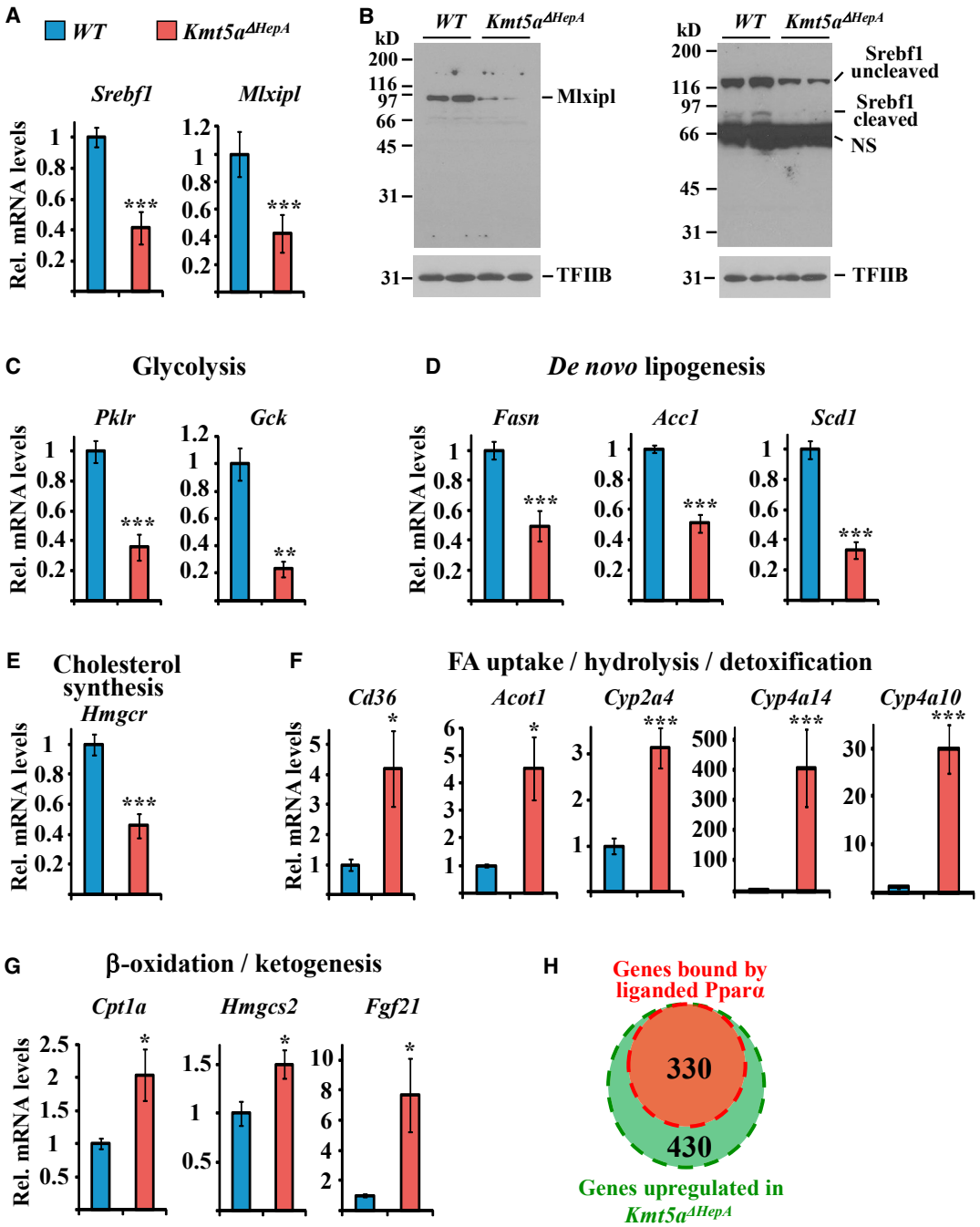


Figure 4. Kmt5a Regulates Genes Involved in Glucose and Lipid Homeostasis

(A) mRNA levels of *Srebf1* and *Mlxipl* in the livers of P45 wild-type and *Kmt5a*$^{\Delta HepA}$ mice. Bars represent mean mRNA levels normalized to Gapdh mRNA and ±SEM from n = 5 individual mice. The data are expressed relative to the values obtained with wild-type mice. ***p < 0.0001.

(B) Western blot analyses of liver extracts were performed with antibodies against Mlxipl (left panel) or Srebf1 (right panel) or control antibodies recognizing TFIIB (bottom panels). NS, non-specific band.

(C–G) mRNA levels of the genes encoding the rate-limiting enzymes of glycolysis (C), de novo lipogenesis (D), cholesterol synthesis (E), Fatty acid metabolism and detoxification (F) and fatty acid oxidation (G) pathways in the livers of P45 wild-type and *Kmt5a*$^{\Delta HepA}$ mice. The genes and the pathways involved are indicated above the graphs. Bars represent mean mRNA levels normalized to Gapdh mRNA and ±SEM from n = 5 individual mice. The data are expressed relative to the values obtained with wild-type mice. *p < 0.05; **p < 0.01; ***p < 0.001.

(H) Venn diagram showing the overlap between genes upregulated in *Kmt5a*$^{\Delta HepA}$ mice and those bound by liganded Pparα. Pparα occupancy data were retrieved from Lee et al. (2014) (GEO: GSE44571).

See also Figure S4 and Table S1.

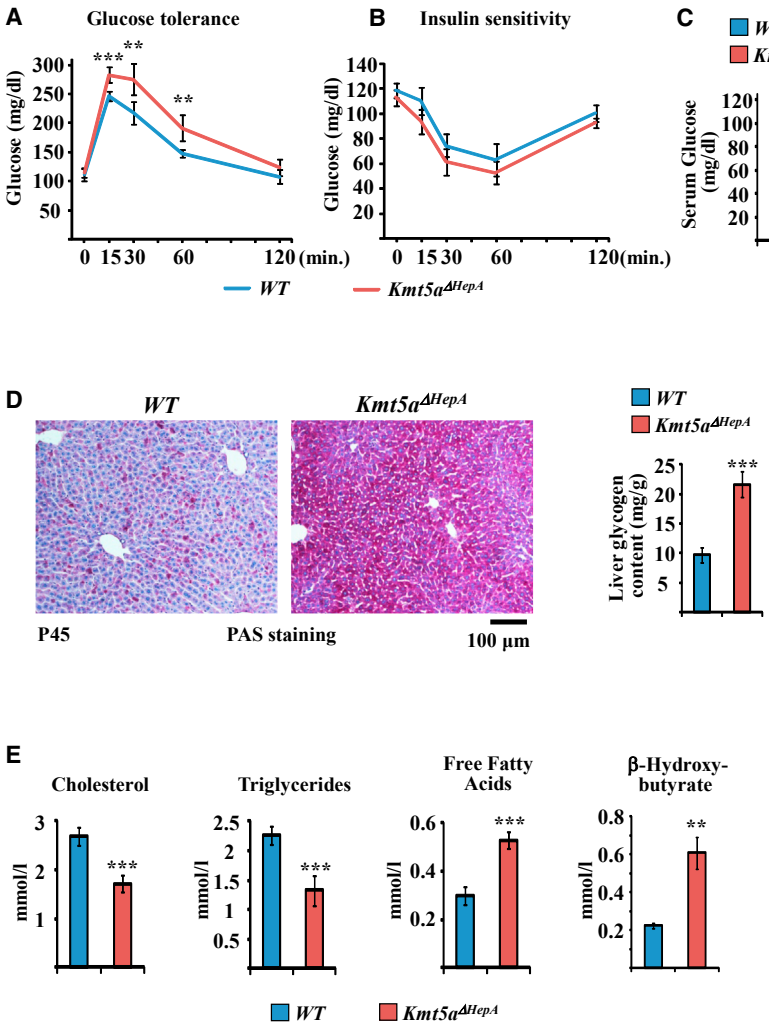


Figure 5. Metabolic Phenotypes in *Kmt5a*$^{\Delta HepA}$ Mice

(A and B) Glucose tolerance and insulin sensitivity tests in P45 wild-type and *Kmt5a*$^{\Delta HepA}$ mice. Plasma glucose levels were measured at the indicated time points after the administration of 1 g/kg glucose (A) or 0.5 units/kg insulin (B) in n = 5 individual mice. $^{**}p < 0.01$; $^{***}p < 0.001$.

(C) Average serum glucose levels in 5 hr fasted P45 mice. n = 5.

(D) Representative images of liver sections stained with periodic acid Schiff (PAS) staining of glycogen. The bar graph at the right shows colorimetric measurement of glycogen in liver extracts from P45 wild-type and *Kmt5a*$^{\Delta HepA}$ mice. n = 5. $^{***}p < 0.001$.

(E) Serum total cholesterol, triglyceride, free fatty acid, and β-hydroxybutyrate levels in P45 wild-type and *Kmt5a*$^{\Delta HepA}$ mice. n = 5. $^{**}p < 0.01$; $^{***}p < 0.001$.

glucose and lipid homeostasis pathways are particularly sensitive to alterations of this regulatory process. This regulatory function at the post-initiation phase of transcription is required for the maintenance of genome stability in non-dividing cells.

A possible role of H4K20Me$_1$ in transcription regulation has been raised by several previous reports, albeit with contradictory conclusions. H4K20 methylation has been linked to transcription activation of estrogen-responsive (Li et al., 2011a), Wnt-inducible (Li et al., 2011b), neuronal (Wang et al., 2015), and PPARγ-activated promoters (Wakabayashi et al., 2009). Other studies demonstrated that H4K20Me$_1$ modification is restricted to gene body regions of active genes in CD4$^+$ T cells (Barski et al., 2007) and in nine Encyclopedia of DNA Elements (ENCODE) cell lines (Beck et al., 2012b) and that gene body H4K20 methylation positively correlates with elongation rates in K562 cells (Veloso et al., 2014). In contrast, other global studies or mRNA analyses of specific genes suggested that H4K20Me$_1$ represses transcription (Congdon et al., 2010; Abbas et al., 2010; Kapoor-Vazirani and Vertino, 2014; Tanaka et al., 2017). Concern about a functional role in transcription also arises from the H4K20Me$_1$ levels in cultured cells being highest in the G$_2$/M phase, when transcription is generally shut down, but highly reduced during the G$_1$ phase and eliminated in the S phase, when genes are actively transcribed (Rice et al., 2002; Abbas et al., 2010; Oda et al., 2010).

To investigate the role of H4K20Me$_1$ in the in vivo organismal context and to eliminate potential cell-cycle-dependent effects, the present study was conducted in mouse livers, populated mainly by non-dividing G$_0$-phase hepatocytes. In agreement with a potential transcription activation function, we detected high levels of H4K20Me$_1$ in the gene body regions of actively transcribed genes and a strong positive correlation with gene activity. Kmt5a interaction with RNA Pol II in vivo provides a plausible mechanism for Kmt5a recruitment and enzymatic modification of nucleosomes in the actively transcribed genomic regions. H4K20Me$_1$ levels at the different gene body locations are determined by the frequency with which the RNA Pol II/Kmt5a complex encounters the underlying nucleosomes and the localized activity of Kdm7b demethylase.

RNA Pol II distribution in most hepatic genes is bimodal, with a large fraction of RNA Pol II concentrated near the TSS. This distribution is characteristic to genes in which RNA Pol II is paused at promoter-proximal regions (Adelman and Lis, 2012). Reduction of gene body H4K20 monomethylation in *Kmt5a*$^{\Delta HepA}$ mice correlated with decreased levels of elongating RNA Pol II and elevated ratios of promoter/promoter-proximal and gene body RNA Pol II reads (PrI). These data suggest that Kmt5a-mediated H4K20Me$_1$ is involved in the regulation of transition of RNA Pol II from the initiation or the paused state to the active elongation phase. The accumulation of short reads in our GRO-seq experiments indicates that Kmt5a regulates mainly the step of RNA Pol II escape from promoter-proximal pause sites.

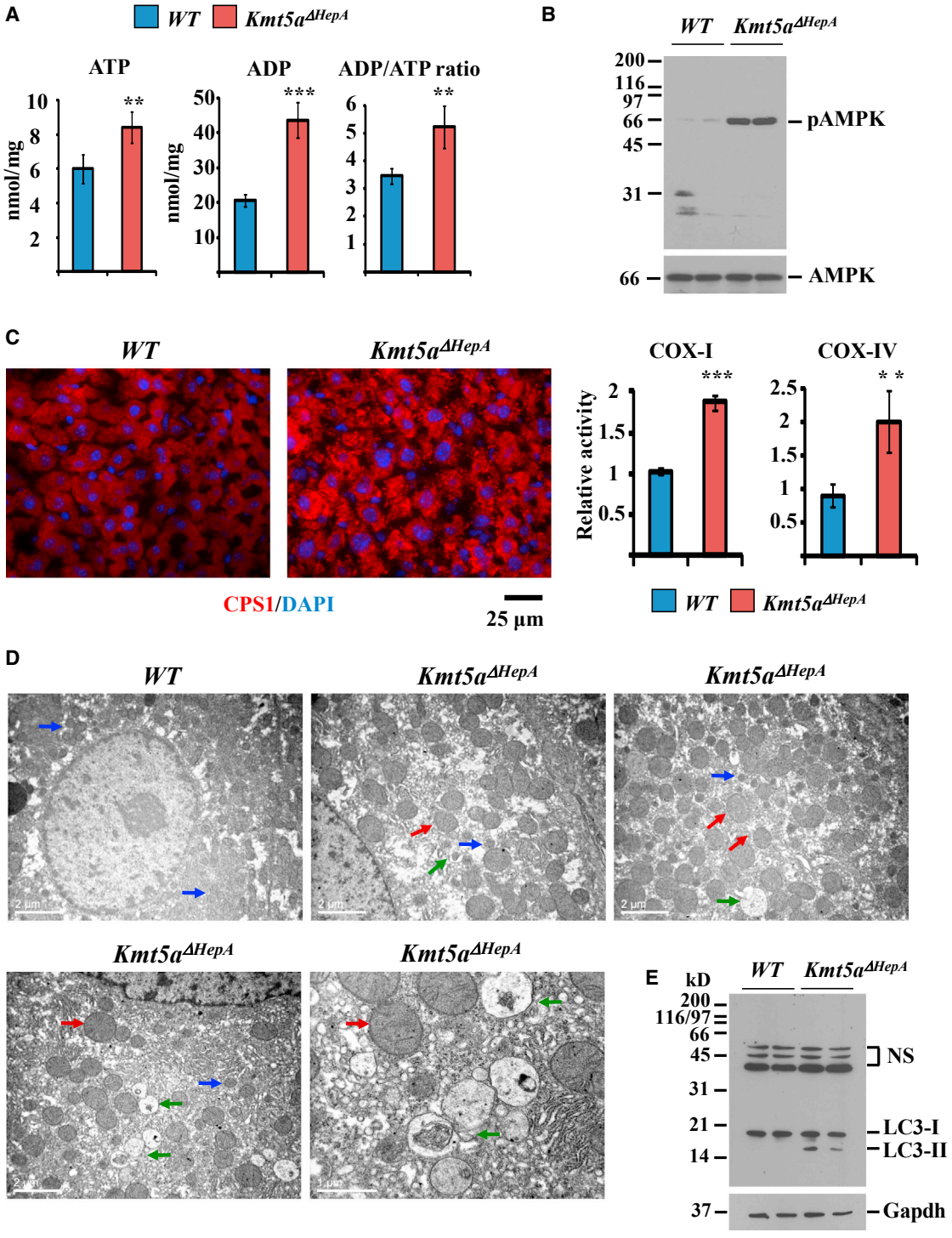


Figure 6. Constitutive AMPK Activation and Increased Mitochondrial Activity in the Livers of *Kmt5a*$^{\Delta HepA}$ Mice

(A) ATP and ADP levels in liver extracts from P45 wild-type and *Kmt5a*$^{\Delta HepA}$ mice. n = 5. **p < 0.01; ***p < 0.001.

(B) Western blot analysis of liver extracts with antibodies recognizing a Thr172-phosphorylated form of AMPK (top panel) or total AMPK protein (bottom panel).

(C) Representative immunohistological staining of liver sections with anti-carbamoyl phosphate synthetase (CPS) antibody recognizing the mitochondrial CPS1 enzyme. Panels at the right show cytochrome oxidase I (COI) and cytochrome oxidase IV (COIV) enzymatic activities in liver extracts from P45 wild-type and *Kmt5a*$^{\Delta HepA}$ mice. n = 5. **p < 0.01; ***p < 0.001.

(legend continued on next page)

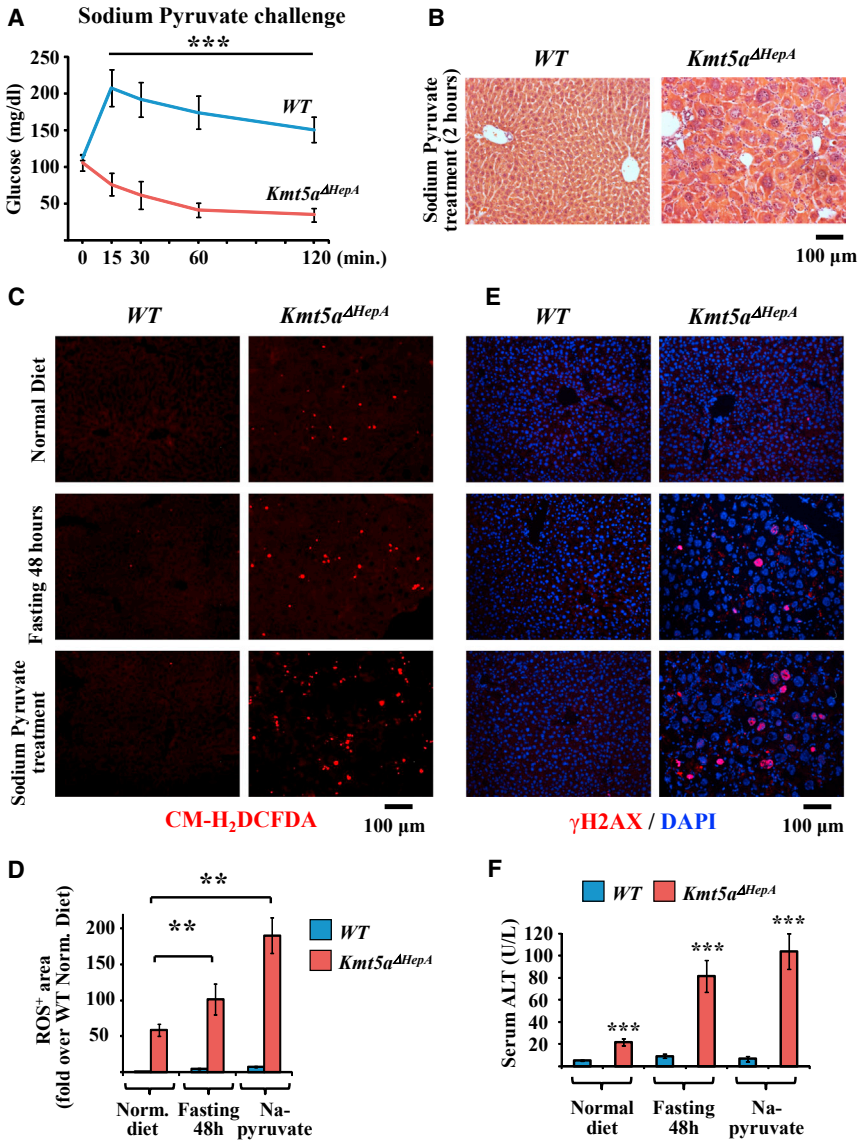


Figure 7. Metabolic Stress Induces Rapid Accumulation of ROS and DNA Damage in the Livers of *Kmt5a*$^{\Delta HepA}$ Mice

(A) Pyruvate tolerance tests in P45 wild-type and *Kmt5a*$^{\Delta HepA}$ mice. Plasma glucose levels were measured at the indicated time points after the administration of 2 g/kg Na-pyruvate in n = 5 individual mice. ***$p < 0.001$.

(B) H&E staining of liver sections from P45 wild-type and *Kmt5a*$^{\Delta HepA}$ mice that were treated with 2 g/kg Na-pyruvate for 2 hr.

(C) Analysis of ROS accumulation. Liver sections from mice fed a normal chow diet, fasted for 48 hr, or treated with 2 g/kg Na-pyruvate for 2 hr were stained with 5-(and-6)-chloromethyl-2',7'-dichlorodihydrofluorescein diacetate, acetyl ester (CM-H_2DCFDA), as indicated.

(D) Quantitative comparison of CM-H_2DCFDA-positive areas in liver sections of (C). Positively stained areas in 20 high-power fields were measured by ImageJ software in liver sections of P45 wild-type or *Kmt5a*$^{\Delta HepA}$ mice. Bars graphs show fold changes of the average positively stained areas from n = 5 mice. **$p < 0.01$.

(E) Analysis of DNA damage. Liver sections from mice treated as in (C) were stained with antibody against γH2AX.

(F) Serum ALT levels were measured in mice treated as in (C). Bars represent mean values of ALT levels and SEM from liver extracts of five individual mice. ***$p < 0.001$.

See also Figures S5 and S6.

Although Kmt5a does not methylate histone residues other than H4K20 (Beck et al., 2012b), similar to other histone-modifying enzymes, it can methylate non-histone substrates such as p53 and proliferating cell nuclear antigen (PCNA) (Shi et al., 2007; Takawa et al., 2012). This raises the possibility that Kmt5a-mediated regulation of transcription elongation may be driven not solely through H4K20 methylation but also through the modification of another, so far unidentified target or targets. Given the strong correlations among H4K20Me$_1$, Kdm7b, and RNA Pol II distribution in different genes, we think that this scenario is unlikely, although it cannot be excluded.

The sharp increase of H4K20Me$_1$ in regions downstream of the promoter-proximal nucleosomes and the opposite distribution of RNA Pol II in most active genes raises the possibility that this modification may act as an obstacle for the movement of RNA Pol II and thus contribute to its retention at pause sites. Consistent with this, full erasure of H4K20Me$_1$ by high levels of Kdm7b in a specific set of genes (e.g., *Alb*) correlated with highly increased transcription and lack of promoter-proximal pausing. However, this simplistic scenario is challenged by the observation of decreased levels of elongating polymerase in genes in which H4K20Me$_1$ was reduced following *Kmt5a* inactivation. Taking the preceding into consideration, we propose that the actual rate of H4K20Me$_1$ turnover, rather than the absolute H4K20Me$_1$ levels, is important for the regulation of RNA Pol II escape from promoter-proximal regions. In other words, rather than acting as a static chromatin modification, H4K20Me$_1$ functions as a transient operational mark whose deposition-removal rate can support RNA Pol II transit into gene bodies. This notion

(D) Electron microscopic images of liver sections from P45 wild-type and *Kmt5a*$^{\Delta HepA}$ mice. Blue arrows indicate normal-sized mitochondria. Red arrows indicate enlarged mitochondria. Green arrows depict autophagosome structures. Note the larger magnification in the bottom-right panel.

(E) Western blot analysis of liver extracts with an antibody recognizing the autophagy marker protein LC3 (top panel) and Gapdh (bottom panel).

See also Figure S6.

is reinforced by the results showing that gene body H4K20 monomethylation is a highly dynamic process and that increased promoter-proximal/gene body ratios of RNA Pol II (PrI) in *Kmt5a*$^{\Delta HepA}$ mice are preferentially detected in genes that are modified in a highly dynamic manner.

Escape of RNA Pol II from promoters and/or promoter-proximal pause sites is a general regulatory step of the transcription mechanism (Adelman and Lis, 2012). Interference with this process is expected to influence the expression of most regulated genes and thus affect a variety of biological functions in any given cell type. Consistent with this notion, genetic or pharmacological inactivation of pTEF-b kinase complex subunits, which play a key role in RNA Pol II release from pause sites, leads to early embryonic lethality in a variety of organisms, including mice, *C. elegans*, and *Drosophila* (Shim et al., 2002; Dahlberg et al., 2015; Oqani et al., 2016).

The results presented here demonstrate that reduced Kmt5a-mediated partial RNA Pol II blockage influences the steady-state mRNA levels of only a fraction of genes. This indicates that compensatory mechanisms, such as mRNA stabilization, are activated to maintain the expression of most hepatic genes. Genes encoding the rate-limiting enzymes of glycolysis and de novo lipogenesis are particularly sensitive to defects of RNA Pol II escape from promoter-proximal sites. Their expression significantly drops in the livers of *Kmt5a*$^{\Delta HepA}$ mice by mechanisms involving reduced gene body H4K20 methylation and the decreased protein levels of the major transactivators Mlxipl and Srebf1. Transcription of *Mlxipl* and *Srebf1* is also a subject of H4K20 methylation-dependent regulation. Thus, metabolic genes are highly regulated at both transcription factor-mediated PIC formation step (Mlxipl and Srebf1 recruitment) and post-recruitment steps involving Kmt5a-dependent RNA Pol II escape. The biological significance of this dual control could be the provision of additional regulatory checkpoints to specific metabolic genes for fast and efficient adjustments of their expression under various conditions.

The defects in metabolic gene transcription in *Kmt5a*$^{\Delta HepA}$ mice result in extensive metabolic reprogramming, which is schematically presented in Figure S6. Impaired glycolytic activity leads to accumulation of glycogen and elevated levels of AMP and ADP, which in turn activate of AMPK. Constitutively active AMPK in *Kmt5a*$^{\Delta HepA}$ hepatocytes increases fatty acid uptake by inducing the expression of CD36 transporter and autophagy. Fatty acids are ligands of PPARα, which stimulates their catabolism via the mitochondrial β-oxidation and ketogenesis pathways, both of which can serve as efficient alternative energy sources. The preceding AMPK-mediated pathways partially compensate for the loss of glucose-dependent energy supply. At the same time, the cells enter a less stable physiological state, which is highly sensitized to metabolic stress conditions. Upon metabolic stress, such as fasting, high-fat diet, or Na-pyruvate challenge, metabolic imbalances cannot be compensated, resulting in excessive increase of intracellular ROS levels and extensive DNA damage. The accumulation of DNA damages quickly aggravates into a senescent phenotype and liver dysfunction.

Altogether, the results of this study highlight a previously unanticipated dimension of Kmt5a function as gatekeeper of genome integrity: They identify a role in the regulation of transcription, through which it controls metabolic homeostasis and protects cells from metabolic stress-mediated DNA damage and cellular senescence.

EXPERIMENTAL PROCEDURES

Mice

Kmt5a$^{\Delta HepA}$ mice have been described previously (Nikolaou et al., 2015). These animals were generated by crossing KMT5A^{loxp} mice carrying the floxed exon 7 allele of KMT5A (Oda et al., 2009) with Alb-Cre mice (Yakar et al., 1999) in a C57BL/6 background. Mice were maintained in grouped cages in a temperature-controlled, pathogen-free facility on a 12 hr light/dark cycle and fed a standard chow diet (19% protein, 5% fat; Altromin 1324) or high-fat diet (34% crude fat, 23% crude protein, 5% crude fiber; Mucedola) and water ad libitum. All animal experiments were approved by the Prefecture of Attica and were performed in accordance with the respective national and European Union regulations. All experiments were performed in randomly chosen age-matched male mice. Typically, each experiment was performed in tissues from at least five individual mice. No blinding was used in this study.

RNA Purification and RT-PCR

Total RNA was prepared by TRIzol extraction as described in Sarris et al. (2016) and Elkouris et al. (2016). For first-strand cDNA synthesis, 1 μg of total RNA was incubated with 200 units of Moloney Murine Leukemia Virus (MMLV) reverse transcriptase in a buffer containing 50 mM Tris-HCl (pH 8.3), 75 mM KCl, 3 mM $MgCl_2$, and 10 mM DTT for 60 min at 37°C. qPCR analyses were carried out in a StepOne real-time PCR detection system using Fast Start Universal SYBR Green Master. Primer sequences are listed in Table S1.

Histology, Metabolite and Enzyme Measurements, and ChIP-Seq and RNA-Seq Assays

Histological examination, metabolic parameter determination, ChIP assays, and steady-state RNA measurements were performed as described in Tataraki et al. (2008) and Nikolaou et al. (2012). Details and modifications of these techniques, including nascent RNA measurements and data analyses, are described in the Supplemental Information.

Statistical Analysis

Statistical significance of the data obtained from different biological replicates was evaluated by two-tailed Student's t test. In the evaluation of pausing indexes or PrIs, Welch's t test was used.

ACCESSION NUMBERS

The accession number for the raw data of ChIP-seq, RNA-seq, GRO-seq, and nascent RNA-seq assays reported in this paper is GEO: GSE97338.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, six figures, and one table and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2017.07.003.

AUTHOR CONTRIBUTIONS

Conceptualization, K.C.N., P.M., and I.T.; Investigation, K.C.N., P.M., V.H., G.C., and I.T.; Supervision, I.T.; Writing, I.T.

ACKNOWLEDGMENTS

We thank Dr. Pantelis Hatzis and Dr. Matthieu Lavigne for comments and discussions. This work was supported by the European Union, ERC Advanced Investigator Grant (ERC-2011-AdG294464) and the AXA Research Fund.

Received: May 22, 2017
Revised: June 23, 2017
Accepted: June 30, 2017
Published: July 25, 2017

REFERENCES

Abbas, T., Shibata, E., Park, J., Jha, S., Karnani, N., and Dutta, A. (2010). CRL4(Cdt2) regulates cell proliferation and histone gene expression by targeting PR-Set7/Set8 for degradation. Mol. Cell *40*, 9–21.

Adelman, K., and Lis, J.T. (2012). Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. Nat. Rev. Genet. *13*, 720–731.

Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. Cell *129*, 823–837.

Beck, D.B., Burton, A., Oda, H., Ziegler-Birling, C., Torres-Padilla, M.E., and Reinberg, D. (2012a). The role of PR-Set7 in replication licensing depends on Suv4-20h. Genes Dev. *26*, 2580–2589.

Beck, D.B., Oda, H., Shen, S.S., and Reinberg, D. (2012b). PR-Set7 and H4K20me1: at the crossroads of genome integrity, cell cycle, chromosome condensation, and transcription. Genes Dev. *26*, 325–337.

Burkewitz, K., Zhang, Y., and Mair, W.B. (2014). AMPK at the nexus of energetics and aging. Cell Metab. *20*, 10–25.

Cantó, C., and Auwerx, J. (2010). AMP-activated protein kinase and its downstream transcriptional pathways. Cell. Mol. Life Sci. *67*, 3407–3423.

Congdon, L.M., Houston, S.I., Veerappan, C.S., Spektor, T.M., and Rice, J.C. (2010). PR-Set7-mediated monomethylation of histone H4 lysine 20 at specific genomic regions induces transcriptional repression. J. Cell. Biochem. *110*, 609–619.

Core, L.J., Waterfall, J.J., and Lis, J.T. (2008). Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. Science *322*, 1845–1848.

Dahlberg, O., Shilkova, O., Tang, M., Holmqvist, P.H., and Mannervik, M. (2015). P-TEFb, the super elongation complex and mediator regulate a subset of non-paused genes during early *Drosophila* embryo development. PLoS Genet. *11*, e1004971.

Desvergne, B., Michalik, L., and Wahli, W. (2006). Transcriptional regulation of metabolism. Physiol. Rev. *86*, 465–514.

Elkouris, M., Kontaki, H., Stavropoulos, A., Antonoglou, A., Nikolaou, K.C., Samiotaki, M., Szantai, E., Saviolaki, D., Brown, P.J., Sideras, P., et al. (2016). SET9-mediated regulation of TGF-β signaling links protein methylation to pulmonary fibrosis. Cell Rep. *15*, 2733–2744.

Foretz, M., Guichard, C., Ferré, P., and Foufelle, F. (1999). Sterol regulatory element binding protein-1c is a major mediator of insulin action on the hepatic expression of glucokinase and lipogenesis-related genes. Proc. Natl. Acad. Sci. USA *96*, 12737–12742.

Fortschegger, K., de Graaf, P., Outchkourov, N.S., van Schaik, F.M., Timmers, H.T., and Shiekhattar, R. (2010). PHF8 targets histone methylation and RNA polymerase II to activate transcription. Mol. Cell. Biol. *30*, 3286–3298.

Fuchs, G., Hollander, D., Voichek, Y., Ast, G., and Oren, M. (2014). Cotranscriptional histone H2B monoubiquitylation is tightly coupled with RNA polymerase II elongation rate. Genome Res. *24*, 1572–1583.

Hardie, D.G. (2011). AMP-activated protein kinase: an energy sensor that regulates all aspects of cell function. Genes Dev. *25*, 1895–1908.

Hochheimer, A., and Tjian, R. (2003). Diversified transcription initiation complexes expand promoter selectivity and tissue-specific gene expression. Genes Dev. *17*, 1309–1320.

Horton, J.D., Goldstein, J.L., and Brown, M.S. (2002). SREBPs: activators of the complete program of cholesterol and fatty acid synthesis in the liver. J. Clin. Invest. *109*, 1125–1131.

Iizuka, K., Bruick, R.K., Liang, G., Horton, J.D., and Uyeda, K. (2004). Deficiency of carbohydrate response element-binding protein (ChREBP) reduces lipogenesis as well as glycolysis. Proc. Natl. Acad. Sci. USA *101*, 7281–7286.

Jonkers, I., and Lis, J.T. (2015). Getting up to speed with transcription elongation by RNA polymerase II. Nat. Rev. Mol. Cell Biol. *16*, 167–177.

Kapoor-Vazirani, P., and Vertino, P.M. (2014). A dual role for the histone methyltransferase PR-SET7/SETD8 and histone H4 lysine 20 monomethylation in the local regulation of RNA polymerase II pausing. J. Biol. Chem. *289*, 7425–7437.

Kim, J., Daniel, J., Espejo, A., Lake, A., Krishna, M., Xia, L., Zhang, Y., and Bedford, M.T. (2006). Tudor, MBT and chromo domains gauge the degree of lysine methylation. EMBO Rep. *7*, 397–403.

Kwak, H., Fuda, N.J., Core, L.J., and Lis, J.T. (2013). Precise maps of RNA polymerase reveal how promoters direct initiation and pausing. Science *339*, 950–953.

Lee, J.M., Wagner, M., Xiao, R., Kim, K.H., Feng, D., Lazar, M.A., and Moore, D.D. (2014). Nutrient-sensing nuclear receptors coordinate autophagy. Nature *516*, 112–115.

Li, Y., Sun, L., Zhang, Y., Wang, D., Wang, F., Liang, J., Gui, B., and Shang, Y. (2011a). The histone modifications governing TFF1 transcription mediated by estrogen receptor. J. Biol. Chem. *286*, 13925–13936.

Li, Z., Nie, F., Wang, S., and Li, L. (2011b). Histone H4 Lys 20 monomethylation by histone methylase SET8 mediates Wnt target gene activation. Proc. Natl. Acad. Sci. USA *108*, 3116–3123.

Marshall, N.F., and Price, D.H. (1995). Purification of P-TEFb, a transcription factor required for the transition into productive elongation. J. Biol. Chem. *270*, 12335–12338.

Martinez-Jimenez, C.P., Kyrmizi, I., Cardot, P., Gonzalez, F.J., and Talianidis, I. (2010). Hepatocyte nuclear factor 4alpha coordinates a transcription factor network regulating hepatic fatty acid metabolism. Mol. Cell. Biol. *30*, 565–577.

Min, J., Allali-Hassani, A., Nady, N., Qi, C., Ouyang, H., Liu, Y., MacKenzie, F., Vedadi, M., and Arrowsmith, C.H. (2007). L3MBTL1 recognition of mono- and dimethylated histones. Nat. Struct. Mol. Biol. *14*, 1229–1230.

Narita, T., Yamaguchi, Y., Yano, K., Sugimoto, S., Chanarat, S., Wada, T., Kim, D.K., Hasegawa, J., Omori, M., Inukai, N., et al. (2003). Human transcription elongation factor NELF: identification of novel subunits and reconstitution of the functionally active complex. Mol. Cell. Biol. *23*, 1863–1873.

Nechaev, S., Fargo, D.C., dos Santos, G., Liu, L., Gao, Y., and Adelman, K. (2010). Global analysis of short RNAs reveals widespread promoter-proximal stalling and arrest of Pol II in *Drosophila*. Science *327*, 335–338.

Nikolaou, K., Tsagaratou, A., Eftychi, C., Kollias, G., Mosialos, G., and Talianidis, I. (2012). Inactivation of the deubiquitinase CYLD in hepatocytes causes apoptosis, inflammation, fibrosis, and cancer. Cancer Cell *21*, 738–750.

Nikolaou, K.C., Moulos, P., Chalepakis, G., Hatzis, P., Oda, H., Reinberg, D., and Talianidis, I. (2015). Spontaneous development of hepatocellular carcinoma with cancer stem cell properties in PR-SET7-deficient livers. EMBO J. *34*, 430–447.

Oda, H., Okamoto, I., Murphy, N., Chu, J., Price, S.M., Shen, M.M., Torres-Padilla, M.E., Heard, E., and Reinberg, D. (2009). Monomethylation of histone H4-lysine 20 is involved in chromosome structure and stability and is essential for mouse development. Mol Cell Biol. *29*, 2278–2295.

Oda, H., Hübner, M.R., Beck, D.B., Vermeulen, M., Hurwitz, J., Spector, D.L., and Reinberg, D. (2010). Regulation of the histone H4 monomethylase PR-Set7 by CRL4(Cdt2)-mediated PCNA-dependent degradation during DNA damage. Mol. Cell *40*, 364–376.

Oqani, R.K., Lin, T., Lee, J.E., Kim, S.Y., Sa, S.J., Woo, J.S., and Jin, D.I. (2016). Inhibition of P-TEFb disrupts global transcription, oocyte maturation, and embryo development in the mouse. Genesis *54*, 470–482.

Rice, J.C., Nishioka, K., Sarma, K., Steward, R., Reinberg, D., and Allis, C.D. (2002). Mitotic-specific methylation of histone H4 Lys 20 follows increased PR-Set7 expression and its localization to mitotic chromosomes. Genes Dev. *16*, 2225–2230.

Sarris, M.E., Moulos, P., Haroniti, A., Giakountis, A., and Talianidis, I. (2016). Smyd3 is a transcriptional potentiator of multiple cancer-promoting genes

and required for liver and colon cancer development. Cancer Cell *29*, 354–366.

Shi, X., Kachirskaia, I., Yamaguchi, H., West, L.E., Wen, H., Wang, E.W., Dutta, S., Appella, E., and Gozani, O. (2007). Modulation of p53 function by SET8-mediated methylation at lysine 382. Mol. Cell *27*, 636–646.

Shim, E.Y., Walker, A.K., Shi, Y., and Blackwell, T.K. (2002). CDK-9/cyclin T (P-TEFb) is required in two postinitiation pathways for transcription in the *C. elegans* embryo. Genes Dev. *16*, 2135–2146.

Takawa, M., Cho, H.S., Hayami, S., Toyokawa, G., Kogure, M., Yamane, Y., Iwai, Y., Maejima, K., Ueda, K., Masuda, A., et al. (2012). Histone lysine methyltransferase SETD8 promotes carcinogenesis by deregulating PCNA expression. Cancer Res. *72*, 3217–3227.

Tanaka, H., Takebayashi, S.I., Sakamoto, A., Igata, T., Nakatsu, Y., Saitoh, N., Hino, S., and Nakao, M. (2017). The SETD8/PR-Set7 methyltransferase functions as a barrier to prevent senescence-associated metabolic remodeling. Cell Rep. *18*, 2148–2161.

Tatarakis, A., Margaritis, T., Martinez-Jimenez, C.P., Kouskouti, A., Mohan, W.S., 2nd, Haroniti, A., Kafetzopoulos, D., Tora, L., and Talianidis, I. (2008). Dominant and redundant functions of TFIID involved in the regulation of hepatic genes. Mol. Cell *31*, 531–543.

Trojer, P., Li, G., Sims, R.J., 3rd, Vaquero, A., Kalakonda, N., Boccuni, P., Lee, D., Erdjument-Bromage, H., Tempst, P., Nimer, S.D., et al. (2007). L3MBTL1, a histone-methylation-dependent chromatin lock. Cell *129*, 915–928.

Tuzon, C.T., Spektor, T., Kong, X., Congdon, L.M., Wu, S., Schotta, G., Yokomori, K., and Rice, J.C. (2014). Concerted activities of distinct H4K20 methyltransferases at DNA double-strand breaks regulate 53BP1 nucleation and NHEJ-directed repair. Cell Rep. *8*, 430–438.

Uyeda, K., and Repa, J.J. (2006). Carbohydrate response element binding protein, ChREBP, a transcription factor coupling hepatic glucose utilization and lipid synthesis. Cell Metab. *4*, 107–110.

van Nuland, R., and Gozani, O. (2016). Histone H4 lysine 20 (H4K20) methylation, expanding the signaling potential of the proteome one methyl moiety at a time. Mol. Cell. Proteomics *15*, 755–764.

Veloso, A., Kirkconnell, K.S., Magnuson, B., Biewen, B., Paulsen, M.T., Wilson, T.E., and Ljungman, M. (2014). Rate of elongation by RNA polymerase II is associated with specific gene features and epigenetic modifications. Genome Res. *24*, 896–905.

Wada, T., Takagi, T., Yamaguchi, Y., Ferdous, A., Imai, T., Hirose, S., Sugimoto, S., Yano, K., Hartzog, G.A., Winston, F., et al. (1998a). DSIF, a novel transcription elongation factor that regulates RNA polymerase II processivity, is composed of human Spt4 and Spt5 homologs. Genes Dev. *12*, 343–356.

Wada, T., Takagi, T., Yamaguchi, Y., Watanabe, D., and Handa, H. (1998b). Evidence that P-TEFb alleviates the negative effect of DSIF on RNA polymerase II-dependent transcription in vitro. EMBO J. *17*, 7395–7403.

Wakabayashi, K., Okamura, M., Tsutsumi, S., Nishikawa, N.S., Tanaka, T., Sakakibara, I., Kitakami, J., Ihara, S., Hashimoto, Y., Hamakubo, T., et al. (2009). The peroxisome proliferator-activated receptor gamma/retinoid X receptor alpha heterodimer targets the histone modification enzyme PR-Set7/Setd8 gene and regulates adipogenesis through a positive feedback loop. Mol. Cell. Biol. *29*, 3544–3555.

Wang, J., Telese, F., Tan, Y., Li, W., Jin, C., He, X., Basnet, H., Ma, Q., Merkurjev, D., Zhu, X., et al. (2015). LSD1n is an H4K20 demethylase regulating memory formation via transcriptional elongation control. Nat. Neurosci. *18*, 1256–1264.

Yakar, S., Liu, J.L., Stannard, B., Butler, A., Accili, D., Sauer, B., and LeRoith, D. (1999). Normal growth and development in the absence of hepatic insulin-like growth factor I. Proc Natl Acad Sci USA *96*, 7324–7329.

Yamaguchi, Y., Takagi, T., Wada, T., Yano, K., Furuya, A., Sugimoto, S., Hasegawa, J., and Handa, H. (1999). NELF, a multisubunit complex containing RD, cooperates with DSIF to repress RNA polymerase II elongation. Cell *97*, 41–51.

Cell Reports, Volume *20*

Supplemental Information

Kmt5a Controls Hepatic Metabolic Pathways by Facilitating RNA Pol II Release from Promoter-Proximal Regions

Kostas C. Nikolaou, Panagiotis Moulos, Vangelis Harokopos, George Chalepakis, and Iannis Talianidis

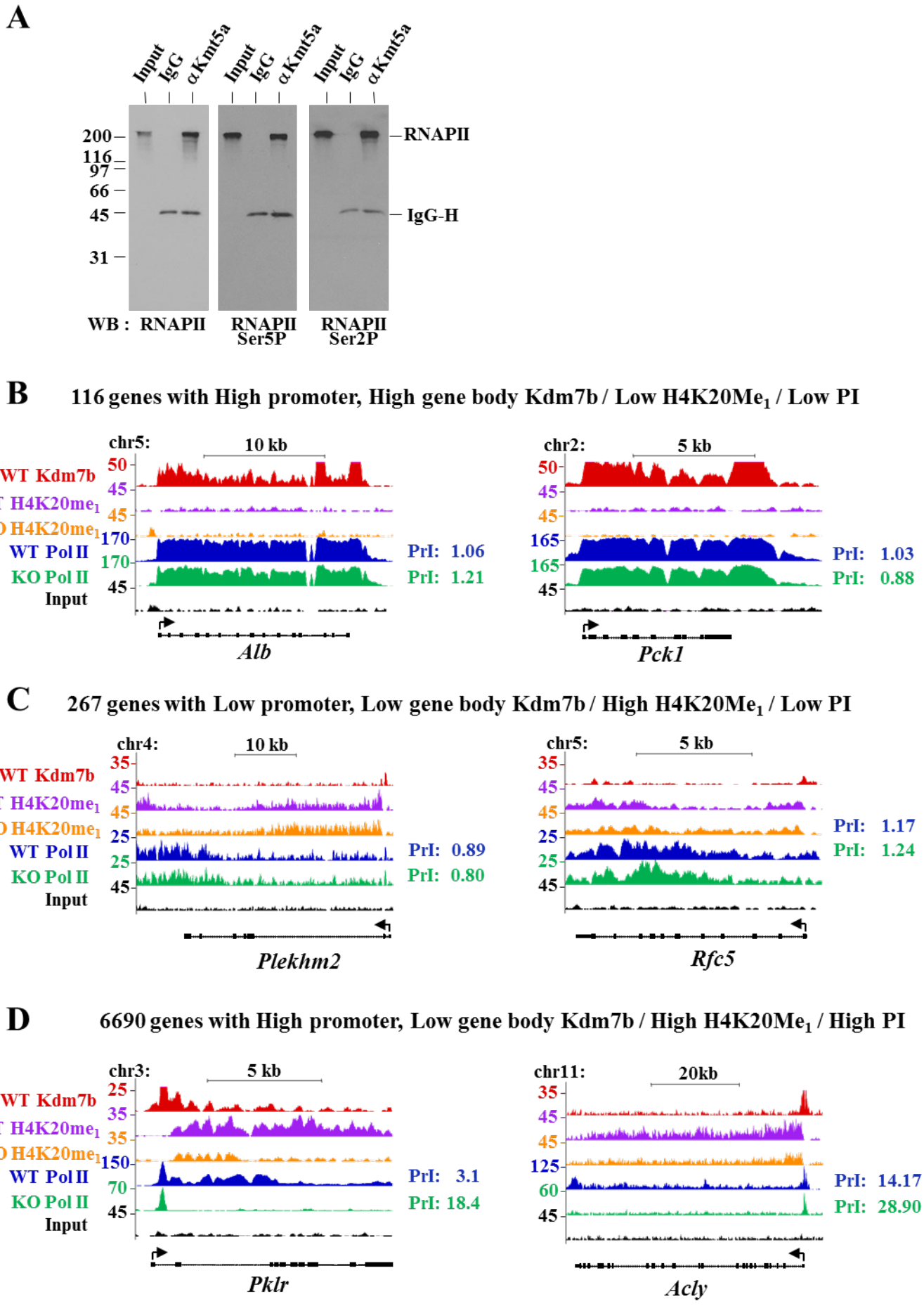
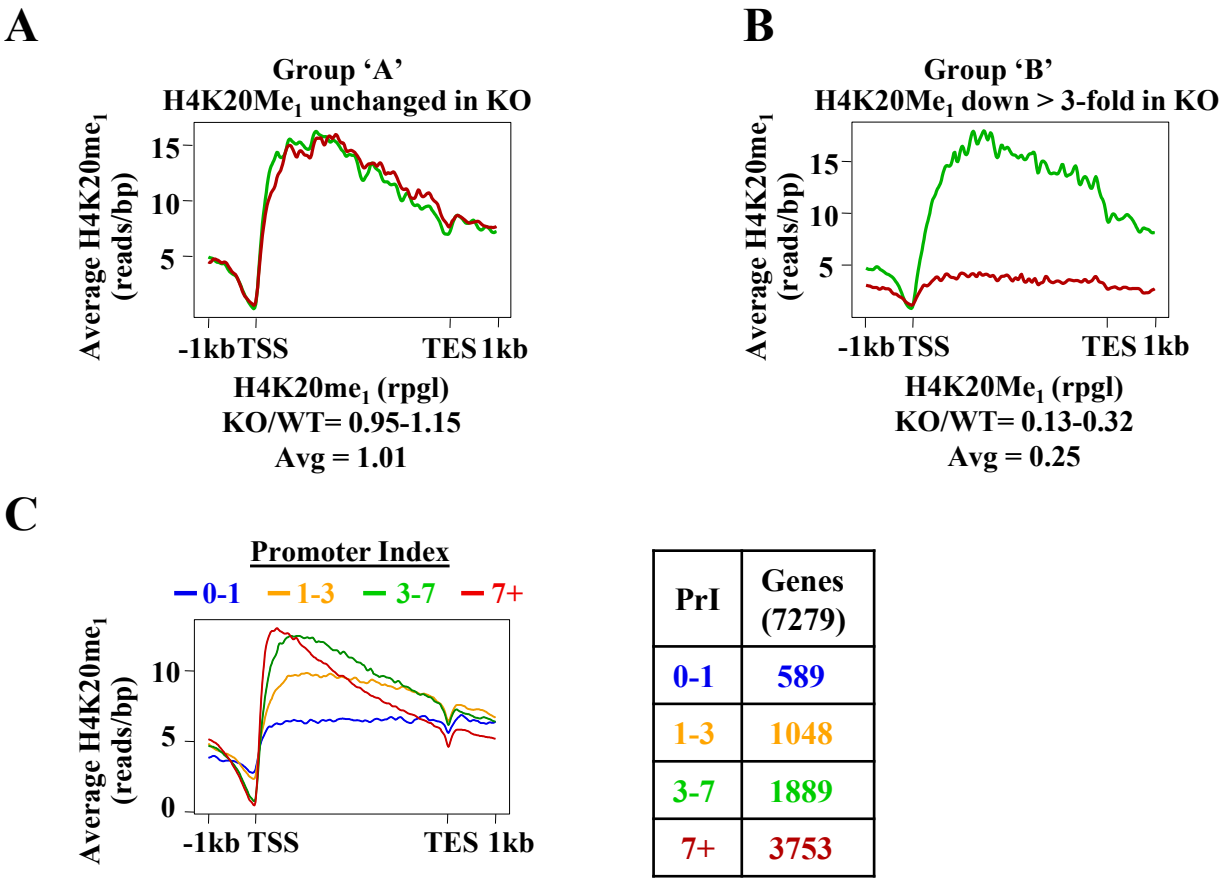

A
Input
IgG
αKmt5a
200
116
97
66
45
31
RNAPII
IgG-H
WB : RNAPII
RNAPII Ser5P
RNAPII Ser2P
B
116 genes with High promoter, High gene body Kdm7b / Low H4K20Me1 / Low PI
chr5:
10 kb
WT Kdm7b
WT H4K20me1
KO H4K20me1
WT Pol II
KO Pol II
Input
PrI: 1.06
PrI: 1.21
Alb
chr2:
5 kb
PrI: 1.03
PrI: 0.88
Pck1
C
267 genes with Low promoter, Low gene body Kdm7b / High H4K20Me1 / Low PI
chr4:
10 kb
PrI: 0.89
PrI: 0.80
Plekhm2
chr5:
5 kb
PrI: 1.17
PrI: 1.24
Rfc5
D
6690 genes with High promoter, Low gene body Kdm7b / High H4K20Me1 / High PI
chr3:
5 kb
PrI: 3.1
PrI: 18.4
Pklr
chr11:
20kb
PrI: 14.17
PrI: 28.90
Acly

Figure S1. Related to Figure 1.

(A) Interaction between Kmt5a and RNA Polymerase-II (RNAPII) in mouse liver. Liver nuclear extracts from P45 wild type mice were prepared and used for immunoprecipitation with anti-Kmt5a antibody. The presence of different forms of RNAPII in the immunoprecipitates was evaluated by western blots assays using anti-RNAPII, recognizing all forms or anti-RNAPII-Ser5P, recognizing the Ser5 phosphorylated form of the C-terminal domain or anti-RNAPII-Ser2P recognizing the Ser2 phosphorylated form of the C-terminal domain of the enzyme.

(B, C and E) UCSC genome browser tracks showing Kdm7b occupancy, H4K20Me$_1$ and RNAPII distribution on representative genes from different categories in the livers of P45 wild type (WT) and *Kmt5a*$^{\Delta HepA}$ (KO) mice. Active, RNAPII-occupied genes (7279) were categorized according to Phf8 occupancy, H4K20Me$_1$ levels and Promoter Index (Read densities at TSS area divided by read densities at gene bodies; PrI). The number of genes fulfilling the indicated triple criteria is indicated at the lines above the graphs. 226 genes did not belong to any of the indicated group. The Promoter Index values of the individual genes are indicated at right.



PrI	Genes (7279)
0-1	589
1-3	1048
3-7	1889
7+	3753

Figure S2. Related to Figures 1 and 2.

(A and B) The distribution of $H4K20Me_1$ in the group of genes displaying low (Group A) and high (Group B) $H4K20Me_1$ methylation turnover. The graphs show average coverage of ChIP-seq reads obtained with an antibody specifically recognizing $H4K20Me_1$.

(C) Active, RNAPII-occupied genes (7279) were categorized according to their Promoter Index (PrI). Average coverage of $H4K20Me_1$ ChIP-seq reads is shown for each category. The number of genes in each category is indicated in the table at right.

A

Biological Process	Enrichment	FDR
Metabolic process	0.0222	1.41E-09
Oxidation-reduction process	0.0275	3.34E-05
Phosphorylation	0.0261	4.68E-05
Lipid metabolic process	0.0440	0.0001
Response to nutrient	0.2105	0.0005
Epithelial cell development	0.3750	0.0012
Positive regulation of glycolysis	0.3333	0.0013
Cholesterol biosynthetic process	0.1538	0.0014
Regulation of transcription from RNA polymerase II promoter	0.0343	0.0019
Lipid biosynthetic process	0.0555	0.0029
Transport	0.0141	0.0046
Cellular response to insulin stimulus	0.1000	0.0049
Ion transport	0.0217	0.0051
Isoprenoid biosynthetic process	0.1875	0.0055
Male gonad development	0.0869	0.0069
Negative regulation of signal transduction	0.0833	0.0072
Mammary gland alveolus development	0.1578	0.0076

B

KEGG Pathway description	Enrichment	FDR
Insulin signaling pathway	0.0592	7.23E-05
Drug metabolism - other enzymes	0.1071	9.91E-05
Pantothenate and CoA biosynthesis	0.1875	0.0037
Type II diabetes mellitus	0.0801	0.0045
Circadian rhythm - mammal	0.1363	0.0050
Valine, leucine and isoleucine degradation	0.0816	0.0052
Fatty acid biosynthesis	0.3333	0.0060
Maturity onset diabetes of the young	0.1153	0.0062
Purine metabolism	0.0363	0.0066
Bile secretion	0.0563	0.0072
Butanoate metabolism	0.1001	0.0076
Propanoate metabolism	0.0967	0.0076
Pyruvate metabolism	0.0731	0.0147
Fatty acid metabolism	0.0666	0.0179
Terpenoid backbone biosynthesis	0.1428	0.0214
Steroid hormone biosynthesis	0.0566	0.0250
Glycolysis / Gluconeogenesis	0.0508	0.0319
Oxidative phosphorylation	0.0305	0.0444

Figure S3. Related to Figure 3.

(A and B) Top Biological Process terms as derived from the respective Gene Ontology enrichment analysis (A) and the top list of biochemical pathways as derived from the respective KEGG Pathway enrichment analysis (B) of the 217 genes showing decreased H4K20Me$_1$ coverage and decreased mRNA levels in the livers of P45 *Kmt5a*$^{\Delta HepA}$ mice. The Biological Processes are ranked by the False Discovery Rate (FDR) values and the number of the support genes that are connected to the respective process according to the Gene Ontology hierarchy.

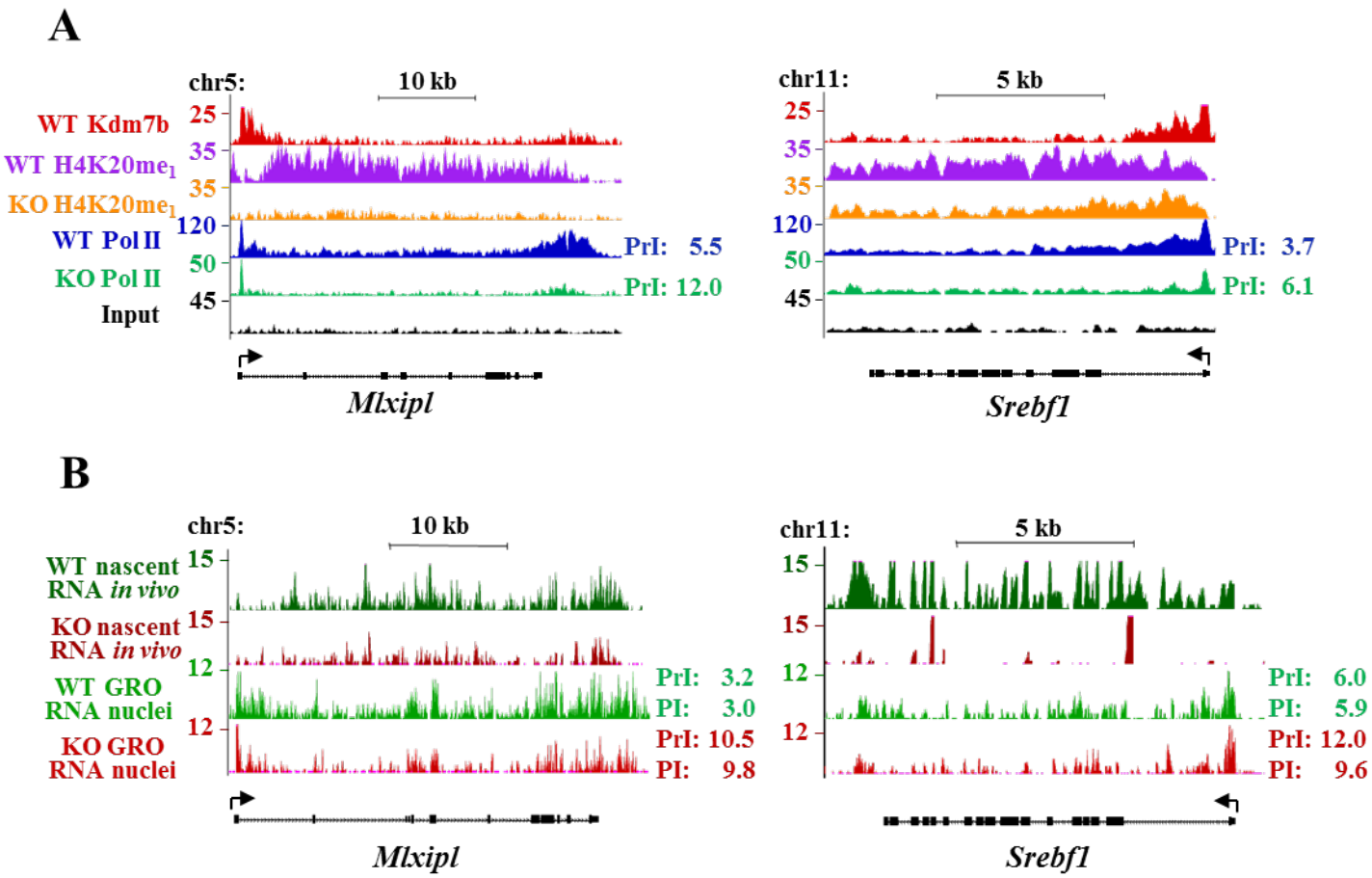


Figure S4. Related to Figure 4.

(A) UCSC genome browser tracks showing Kdm7b occupancy, H4K20Me$_1$ and RNAPII distribution on genes encoding Mlxipl and Srebf1 transcription factors in the livers of P45 wild type (WT) and *Kmt5a*$^{\Delta HepA}$ (KO) mice. The Promoter Index (PrI) value of the individual genes is indicated at right.

(B) UCSC genome browser tracks showing the distribution on Mlxipl and Srebf1 genes of in vivo EU-labeled nascent RNA reads and reads obtained from Global Run-On sequencing in isolated nuclei (GRO-seq) in the livers of P45 wild type (WT) and *Kmt5a*$^{\Delta HepA}$ (KO) mice. The Promoter+Pausing Index (PrI) value (-100 to +100 relative to TSS region reads divided by the reads in downstream gene body region) and the Pausing Index (PI) value (TSS to +50 nt reads per downstream gene body reads) are indicated at right.

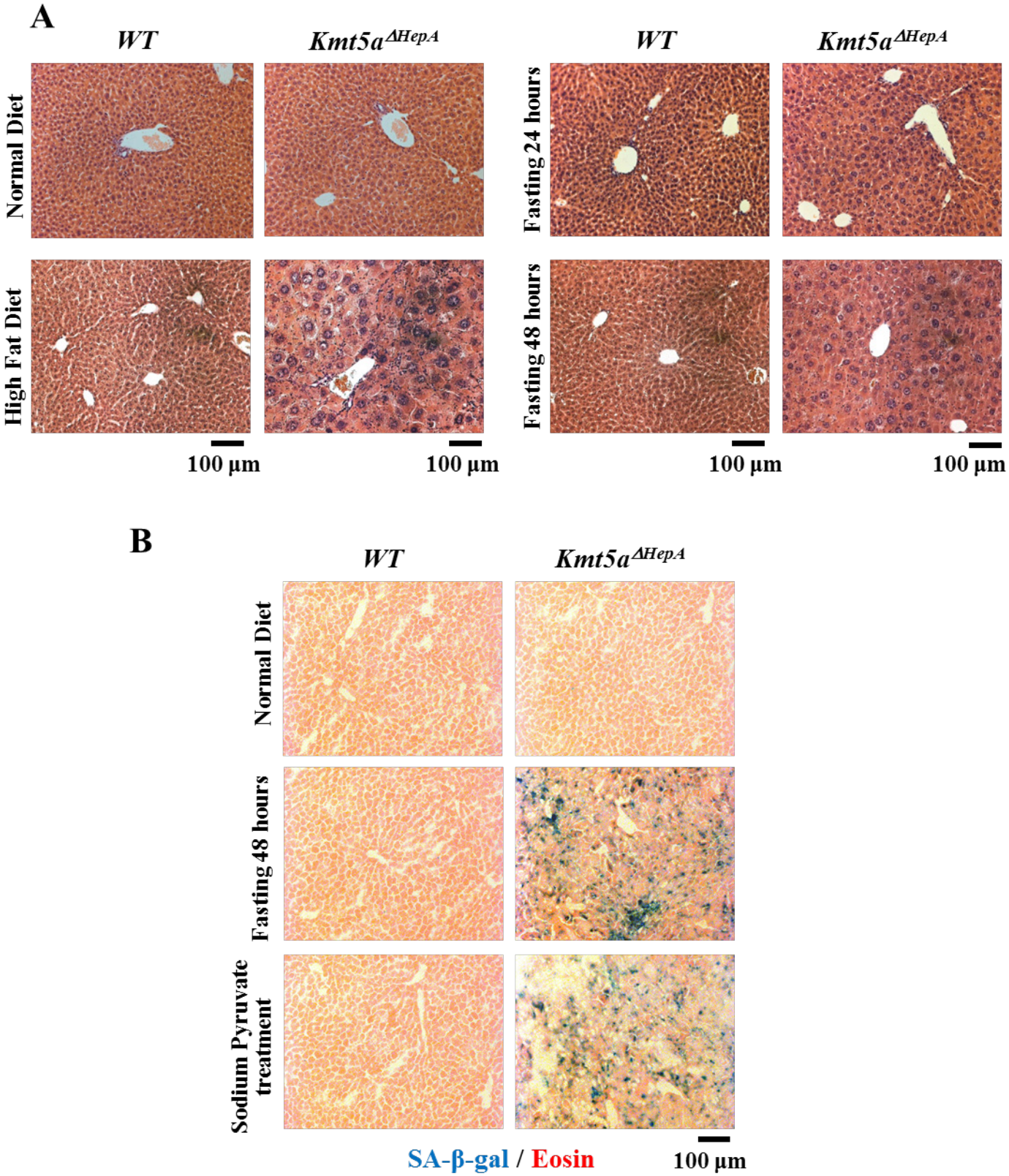


Figure S5. Related to Figure 7.

(A) Hematoxylin and eosin staining of liver sections from P45 wild type (*WT*) and in *Kmt5a*$^{\Delta HepA}$ mice under different feeding conditions. The mice were fed with Normal chow diet (Altromin 1324; 19% protein, 5% fat) or High Fat Diet (34% Fat, 23% Protein, 5% fiber by Mucedola s.r.l.) or fasted for 24 hours or 48 hours, as indicated.
(B) Representative SA-β-gal stainings of liver sections from mice fed with normal chow diet or fasted for 48 hours or treated with 2g/kg Na-pyruvate for 2 hours.

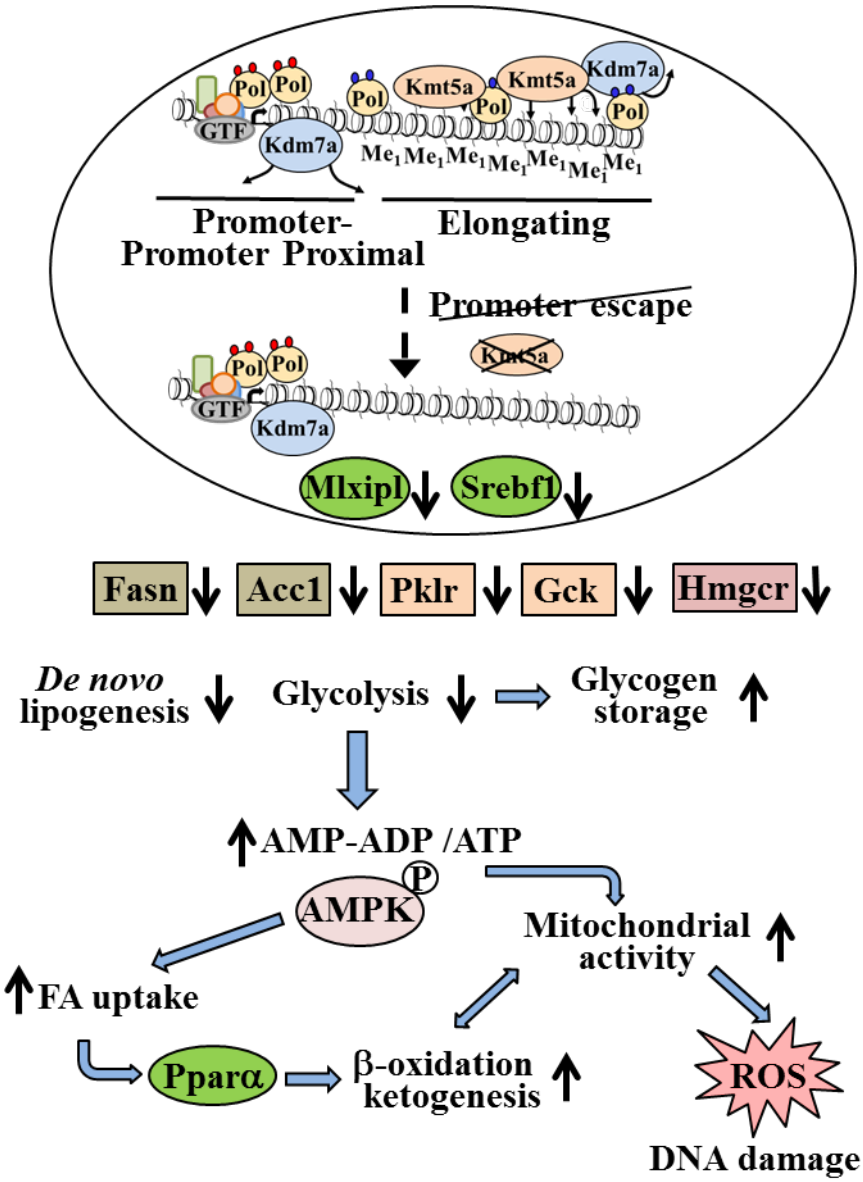


Figure S6. Related to Figures 6 and 7.

Schematic presentation of how alterations of H4K20Me$_1$ in gene bodies lead to ROS-mediated DNA damage in *Kmt5a*$^{\Delta HepA}$ mice.

Table S1 Related to Figure 4. Sequences of PCR primers used in this study.

CD36	NM_001159558.1	5' CAA AGT TGC CAT AAT TGA GTC C 3' 5' CGA ACA CAG CGT AGA TAG ACC T 3'
CYP7B1	NM_007825.4	5' GGG CTA AAT CAC AGT CGG ATG T 3' 5' GAG TGC ACC AAA TCA CGG AG 3'
CYP8B1	NM_010012.3	5' GTA CGC TTC CTC TAT CGC CTG A 3' 5' TGG AGG GAT GGC GTC TTA TG 3'
SCD1	NM_009127.4	5' GGG CAG TTC TGA GGT GAT TAG AG 3' 5' TGG CAG TGG GTA GGT AGT CTT G 3'
ACACA	NM_133360.2	5' GGG GAC AGG TTG TCA CTA TGG G 3' 5' AAA CAC CAG CAC CGA GAC TGA A 3'
ACLY	NM_001199296.1	5' GGC TTC ATT GGG CAC TAC CTT G 3' 5' GTA GGG CTC CTG GCT CAG TTA CA 3'
ACOT1	NM_012006.2	5' TTG GGT AGT GAG TGT TTG CAT GT 3' 5' CAG CCC TTG AAT CAG CAC TAT TT 3'
MLXIPL	NM_021455.4	5' TTG GGG ATG GAG AAA AAC CTA GT 3' 5' GAA GTT GTG AGG AAA CAG ACC G 3'
CPT1a	NM_013495.2	5' GGA AGT GAG GGG CTT GGT AGT C 3' 5' TGG CTG CCT GTG TCA GTA TGC 3'
FASN	NM_007988.3	5' GGA CTT GGG TGC TGA CTA CAA C 3' 5' CCT CCC GTA CAC TCA CTC GT 3'
FGF21	NM_020013.4	5' GAG CAT GGT AGA GCC TTT ACA 3' 5' CGT AAT AAA TAA GTG GAA ACC CAA 3'
HMGCS2	NM_008256.4	5' GAT TCA ACT AGG GGA TGA TCT C 3' 5' CAC TTT CAG GAA TGG GTT ATC T 3'
HMGCR	NM_008255.2	5' TAA TTC ATT TGG TGC TGG TCT AT 3' 5' GCC TCT CTG AAC AAA GAC TCA C 3'
PKLR	NM_013631.2	5' GCG ACC TGG CTC TGG CTA TAC 3' 5' GGC AGA TGT GGG ACT ATG GG 3'
PPARα	NM_011144.6	5' CCT TCT ACG CTC CCG ACC CA 3' 5' CCA TGT CCA TAA ATC GGC ACC A 3'
SREBF1	NM_011480.4	5' GGA CTT CCC ATC TGT TGT AAG G 3' 5' TGT GTT TCT GGA TAT ACG GGG TTA 3'
GAPDH	NM_008084	5' CCA ATG TGT CCG TCG TGG ATC T 3' 5' GTT GAA GTC GCA GGA GAC AAC C 3'
GCK	NM_010292.5	5' GGC CAC AAA CAT TCC AGA 3' 5' GCT GCC ATC CTG CTA ACA 3'
CYP7A1	NM_007824.2	5' CCG TTT CTG TGC AGC CTA AGT 3' 5' TGG AAT TGT AGA CAT AAC GCA AAG A 3'
Kmt5a	NM_030241.3	5' AAA GGC AGG GGT GTG ATC GCT 3' 5' AGT CTC CCT AGG CGG TTC GTT 3'

Supplemental Methods

Antibodies

Antibodies used for this study were from:
Cell Signaling Technology: anti-phospho histone H2A.X #9718; anti-phospho Ser2-Rpb1 CTD #13499; anti-phospho Ser5 Rpb1 CTD #13523; anti-AMPKα #2532; anti-phospho AMPKα Cat#4188.
Santa Cruz Biotrechnology: anti-ChREBP #SC-33764; anti-SREBP1 #SC-8984; anti-GAPDH #SC-32233; anti-TFIIB #SC-225.
Euromedex: anti-RNA Pol II #PB-7G5.
Abcam: anti-CPS1 #ab45956; anti-phospho Ser2-RNA Pol-II #ab5095; anti-PHF8 #ab36068; anti-histone H4 mono methyl K20 #ab9051; anti-LC3B #ab48394

Chromatin Immunoprecipitation (ChIP)

Liver tissue was minced to small pieces in PBS and after addition of formaldehyde to a 1% final concentration immediately was subjected to 10 strokes of dounce homogenization. Cross-linking was continued for 10 min and stopped by the addition of glycine at 0.125 M final concentration. Cross-linked cells were treated with a buffer containing 0.25% Triton-X100, 10 mM EDTA, 0.5 mM EGTA, 20 mM HEPES pH 7.9 and then with a buffer containing 0.15 M NaCl, 1 mM EDTA, 0.5 mM EGTA and 20 mM HEPES pH 7.9 for 10 minutes each. The nuclei were resuspended in 10 volumes of sonication buffer containing 50 mM Hepes pH 7.9, 140 mM NaCl, 1 mM EDTA, 0.1% Na-deoxycholate and 0.5% Sarkosyl and sonicated for 10 minutes in Covaris Sonicator instrument at maximal settings. After centrifugation at 14000 rpm for 15 minutes, the extracted chromatin was supplemented with 0.5 volume of 50 mM Hepes pH 7.9, 140 mM NaCl, 1 mM EDTA, 3% Triton X-100, 0.1% Na-deoxycholate, and incubated overnight with Dynabeads Protein G, that were prebound by 10 μg of the respective antibodies. The beads were washed sequentially first with a buffer containing 50 mM Hepes pH 7.9, 140 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% Na-deoxycholate and 0.1% SDS, then with a buffer containing 50 mM Hepes pH 7.9, 500 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% Na-deoxycholate and 0.1% SDS, followed by a buffer containing 20 mM Tris pH 8.0, 1 mM EDTA, 250 mM LiCl, 0.5 % NP-40 and 0.5 % Na-deoxycholate and finally with TE buffer (10 mM Tris pH 8.0, 1 mM EDTA). Immunoprecipitated chromatin was eluted from the beads by two sequential incubations with a buffer containing 50 mM Tris, pH 8.0, 1 mM EDTA, 1 % SDS, 50 mM $NaHCO_3$ at 65 ^{0}C for 15 minutes. After de-crosslinking by incubation at 65^{0}C for 16 hours, RNAs and proteins were removed by incubation with 10 μg/ml RNAse-I for 1 hour at 37 ^{0}C, followed by incubation with 10 μg/ml Proteinase-K for 2 hours at 42 ^{0}C. DNA was extracted by phenol/chlorophorm and precipitated with ethanol.
About 10 ng of the immunoprecipitated DNA and input DNA was used for library preparation following the optimized workflow of Ion Proton system (Life Technologies).

RNA purification and RT-PCR

Total RNA was prepared by homogenizing liver pieces in 10 volumes of Trizol reagent followed by the addition of 2 volumes of chlorophorm and centrifugation at 12000 g for 15 minutes. The aqueous phases were precipitated with ethanol. The RNA samples were digested with 10 units of DNase I for 10 min at 37 ^{0}C, followed by purification with phenol/chlorophorm extraction and ethanol precipitation. For first strand cDNA synthesis 1 μg of total RNA was incubated with 200 units of MMLV reverse transcriptase in a buffer containing 50 mM Tris-HCl pH 8.3, 75mM KCl, 3mM MgCl2, 10mM DTT for 60 minutes at 37 ^{0}C. Quantitative PCR analyses were carried out in STEP-ONE Real time PCR detection system using Fast Start Universal SYBR Green Master.

RNA sequencing

RNA-seq was performed on an Ion Proton^{TM} System as described previously (Sarris et al., 2016). Briefly, approximately 20 μg of total RNAs was used for mRNA isolation with the Dynabeads mRNA DIRECT^{TM} Micro Kit (Life Technologies, Carlsbad, CA, USA). The isolated mRNA was digested with RNase III, hybridized and ligated to Ion Adaptors, reverse transcribed, barcoded and amplified, using the Ion Total RNA-Seq Kit v2 (Life Technologies). Samples were processed on an OneTouch 2 instrument and enriched on a One Touch ES station. Templating was performed using the Ion PITM Template OT2 200 Kit (Life,Technologies) and sequencing with the Ion PITM Sequencing 200 Kit on Ion Proton PITM chips (Life Technologies) according to the manufacturer's protocols.

GRO-seq assays in isolated nuclei

Nuclei from mouse livers were isolated by dounce homogenization of liver pieces in a buffer containing 10 mM Tris pH 7.5, 2 mM $MgCl_2$, 3 mM $CaCl_2$, 2 U/ml RNaseOUT. After filtering through a cell strainer (100 μm, BD Biosciences), the lysates were centrifuged at 400xg for 10 min, and resuspended in the above buffer supplemented with 10% glycerol and 1% NP40. After incubation in ice for 5 min and two successive washes with the same buffer, the nuclei were resuspended in 50 mM Tris pH 8.3, 40% glycerol, 5 mM $MgCl_2$, 0.1 mM EDTA.
$2x10^7$ nuclei were mixed with an equal volume of run-on buffer (10 mM Tris pH 8.0, 5 mM $MgCl_2$, 1 mM DTT, 300 mM KCl, 200 U/ml RNaseOUT, 1% Sarkosyl, 500 μM ATP, GTP, CTP and Br-UTP) and incubated for 5 min at 30°C. RNA was extracted with Trizol reagent. Purified RNA was treated with 10 Units DNaseI (Invitrogen) for 10 min at 37

^{0}C, extracted with acid-phenol and precipitated with ethanol. The RNA was fragmented by hydrolysis in 0.2N NaOH in ice for 10 min. Fragmented RNA was extracted with acid-phenol and precipitated with ethanol. Next, 10 µg anti-BrdU antibody was conjugated to 50 µl Dynabeads by incubation in blocking buffer, containing 0.25x SSPE, 38 mM NaCl, 1 mM EDTA, 0.05% Tween-20, 0.1% PVP, and 0.1% BSA for 1h and the nuclear RNA was incubated with beads for 1 hour, followed by two washes in binding buffer (0.25x SSPE, 38 mM NaCl, 1 mM EDTA, 0.05% Tween-20), two washes in low salt buffer (0.2x SSPE, 1 mM EDTA, 0.05% Tween-20), one wash in high salt buffer (0.2x SSPE, 1 mM EDTA, 0.05% Tween-20, 138 mM NaCl), and two washes in TET buffer (TE pH 7.4, 0.05% Tween-20). BrU-labeled RNA was extracted from the beads using Trizol Reagent and treated with 5 Units RNA 5' Pyrophosphohydrolase (Rpph) for 1h at 37 ^{0}C. The 5' end repaired RNA was extracted using acid phenol followed by ethanol precipitation and treated with 10 units T4 polynucleotide kinase without ATP for 30 min and then with ATP for an additional 30 min at 37 ^{0}C.
The RNAs were then directly hybridized with Ion adaptor Mix V2, by successive incubations at 65°C for 10 min and at 30°C for 5 min, followed by ligation reaction at 30°C 70 min. The samples were then reverse transcribed in a mixture containing Ion RT Primer V2 and Superscript III enzyme. The cDNAs were then amplified by Platinum High Fidelity PCR Supermix, Ion Xpress RNA 3' Barcode primer and Ion Xpress RNAseq BC Barcode primer 1-24.
Due to the T4 polynucleotide kinase step above "strandedness" detection provided by the adaptor primers is not possible in these libraries.

***In vivo* nascent RNA-seq assay (EU-seq)**
This assay is based on short *in vivo* labeling of nascent RNA with the nucleoside analogue 5-ethynyl uridine. EU-labeled nuclear RNA is subjected to fragmentation, followed by biotinylation and purification using streptavidin beads. This procedure, similar to the recently described TT-seq method (Schwalb et al., 2016), uniformly maps nascent transcripts along genes and also maps transient RNA downstream of the polyadenylation sites.
Mice were injected intraperitoneally with 25 mg/kg 5-ethynyl uridine (EU). Mice were sacrificed 30 minutes following EU treatment and liver nuclei were isolated as above. Nuclear RNA was isolated by Trizol extraction and subjected to fragmentation by hydrolysis in 0.2N NaOH in ice for 10 min, followed by biotinylation using Click-iT Nascent RNA Capture Kit according to manufacturer's instructions. Biotinylated RNA was incubated with Dynabeads MyOne Streptavidin T1 magnetic beads for 30 min at room temperature. The beads were washed 5 times with Click-iT reaction wash buffer and biotinylated RNA was isolated by extraction with Trizol. The RNAs were directly processed to hybridizations with Ion adaptor Mix V2, ligations and cDNA synthesis and library preparation as above, without prior removal of 5'Cap structures by RppH treatment.

Short read mapping
The RNA-Seq FASTQ files obtained from the Ion Proton sequencing procedure were mapped on the UCSC mm9 reference genome using a two phase mapping procedure. Firstly, the short reads were mapped using tophat2 (Kim et al., 2013), with default settings and using additional transcript annotation data for the mm9 genome from Illumina iGenomes (http://cufflinks.cbcb.umd.edu/igenomes.html). Next, the reads which remained unmapped were submitted to a second round of mapping using Bowtie2 (Langmead and Salzberg, 2012) against the mm9 transcriptome with the *--local* and *--very-sensitive* local switches turned on. The ChIP-Seq FASTQ files were mapped to the mm9 genome using Bowtie2 with standard parameters. The resulting ChIP-Seq BAM files were processed so as all Ion Torrent Proton reads would have a length of 180bp. The GRO-Seq FASTQ files were mapped using Bowtie2 with standard parameters in two rounds: in the first mapping round, GRO-Seq samples were aligned against the Mus musculus clone RP23-225M6 45S pre-ribosomal RNA gene (https://www.ncbi.nlm.nih.gov/nuccore/22297525) to remove ribosomal RNA contamination which led to the removal of 3 to 10% of short reads. The remaining were aligned against the mm9 genome. All the resulting BAM files were visualized in the UCSC Genome Browser using bedtools and tools provided by the UCSC Genome Browser toolkit.

Reference mouse gene annotation
Prior to any subsequent analyses (gene differential expression, examination of binding genomic profiles etc.) and in order to avoid ambiguities when for example determining the genomic location distribution of H4K20me1 profiles due to potential alternative isoform expression in the mouse strains used, the following data-driven procedure was followed: firstly, a combined mouse gene (UCSC mm9/GRCm37) annotation was constructed from several sources (UCSC, Ensembl, NCBI Entrez) using the Bioconductor package *OrganismDbi*. At the same time, a combined BAM file of RNA PolII short reads in WT was constructed from the two RNA PollII sequenced biological replicates. Then, genomic areas +/- 1kb from the Transcription Start Site (TSS, referred as *promoters*) and +/- 0.5kb from the Transcription End Site (TES, referred as *downstream regions*) were retrieved from the combined annotation described above and the RNA PolII occupancy inside these regions was measured. Based on these occupancy data, a final unambiguous set of 19583 dominant transcripts was constructed by resolving ambiguities (e.g. different isoforms of the same gene) by keeping those whose promoters showed the greatest RNA PolII occupancy as well as the downstream region.

ChIP-seq analyses

H4K20me1 enriched regions calling

The two biological replicates of H4K20Me$_1$ WT ChIP were analyzed with SICER (Xu et al., 2014) in order to detect genomic regions enriched in H4K20Me$_1$ signal. SICER was applied for each replicate with the calling procedure which uses control files (Input DNA for the H4K20me1 samples) with the following parameters: *redundancy threshold=1*, *window size=400*, *fragment size=180*, *effective genome fraction=0.67*, *gap size=1200*, *fdr=0.05*. Prior to applying SICER and in order to compensate for differences in library sizes between ChIP and Input DNA sequencing samples, the total number of reads of each sample was equalized by uniformly downsampling reads relatively to the sample with the lower number of reads. The latter were used for calling H4K20Me$_1$ enriched regions. At the same time, combined H4K20Me$_1$ tracks were created for WT, Kmt5a$^{\Delta HepA}$ and Input DNA by merging the respective duplicates and normalizing as above. These tracks were used for further quality control of the SICER enriched region calling as well as quantifications of the difference between H4K20Me$_1$ signals in WT and Kmt5a$^{\Delta HepA}$. The final set of H4K20Me$_1$ enriched regions was determined by merging the regions called by SICER for each individual replicate and retaining only those demonstrating a fold enrichment greater than 0.58 in log2 scale or 1.5 in natural scale (where fold enrichment is the ratio of reads under an enriched region in the H4K20Me$_1$ ChIP sample to the respective number of reads in the Input DNA sample).

H4K20Me$_1$, H3K4Me$_3$, RNA PolII and Kdm7b data processing

Prior to any subsequent analyses (differential expression, examination of binding genomic profiles etc.), all genes/transcripts derived from the procedure described in the section 'Reference mouse gene annotation' with length less than 1kb were excluded, resulting in a set of 18042 genes/transcripts.

In order to determine the location distribution of the identified H4K20Me$_1$-enriched regions in relation to their distance from the closest genes, facilities from the R/Bioconductor platform were used in a custom script. In addition, the absolute quantifications of H4K20Me$_1$ signals over annotated genes in WT and Kmt5a$^{\Delta HepA}$, were determined by calculating the number of H4K20Me$_1$ Reads Per Window Length (RPWL) normalized by the gene length (taken to be the window length) in H4K20Me$_1$ regions characterized as enriched by SICER. The library sizes were also accounted for the normalization process described in the 'ChIP-Seq analysis' section. A gene was considered to present H4K20Me$_1$ signal if its RPWL was above 0.03 and to present differential H4K20Me$_1$ signal if the RPWL ratio between Kmt5a$^{\Delta HepA}$ and WT was greater than 0.58 in log2 scale or 1.5 in natural scale.

Absolute quantifications for RNA PolII and Ser2 were also determined by calculating the number of sequenced ChIP Reads Per Window Length (RPWL) normalized by the gene length (taken to be the window length) in gene bodies. The Pausing Index (PI) was calculated as the ratio of average normalized coverage (read density) in a window of +/-250bp from the TSS to the average normalized coverage in the rest of the gene body. For the PI calculations, only RNA PolII active genes were used, which were those presenting an absolute quantification greater than 0.01 in WT. H3K4Me$_3$ data were retrieved from Gene Expression Omnibus (GSE44571) and the four sets of H3K4Me$_3$ data in GSE44571 were merged into one.

GRO-seq and Nascent RNA-seq analyses

The two biological replicates of WT and Kmt5a$^{\Delta HepA}$ for Br-UTP-labeled or EU-labeled nascent RNA-seq were merged into one, to create one sample for each condition. Normalization between the two conditions was performed as follows: firstly, the reads overlapping exons were counted and assembled to create a read count value for each gene. Subsequently, the standard Reads Per Kilobase per Million reads (RPKM) score was calculated for each gene, using the sum of exon lengths as the actual gene length in the calculations, resulting in a matrix of RPKM scores, where each row represented a gene while each column represented a sample. Then, genes presenting an RPKM value less than the median RPKM value of all genes and genes non-expressed at all (RPKM=0) in one of WT or Kmt5a$^{\Delta HepA}$ conditions were filtered out. The remaining genes were used to for a simple regression model: $y = ax + b$, where y corresponds to Kmt5a$^{\Delta HepA}$ gene expression values and x to WT gene expression values. The fitted slope a was used for normalization of the coverage profiles, which were calculated after removing duplicate reads from each merged sample and down-sampling the largest library to the size of the lowest to adjust for differences in library sizes.

Coverage profiles

All average coverage profiles were calculated and visualized with the Bioconductor package *recoup* (https://bioconductor.org/packages/release/bioc/html/recoup.html). Briefly, for coverage profiles across gene bodies, as the latter have different lengths, the gene bodies were split in 200 bins of dynamic length each (so that all genes are split in the same number of bins). Then, the ChIP-seq signal was calculated as normalized (from normalized BAM files, see 'ChIP-seq analysis' above) coverage per base-pair for each bin and averaged per bin, resulting in 200 data points for each gene body, comprising the binding pattern for each gene. For coverage profiles across the TSS, the ChIP-seq signal was calculated as normalized coverage per base-pair.

RNA-seq data analyses

The resulting RNA-Seq BAM files were analyzed with the Bioconductor package *metaseqR* (Moulos and Hatzis, 2015). Briefly, the raw BAM files, one for each RNA-Seq sample, were summarized to an exon read counts table, using the

Bioconductor package GenomicRanges (Lawrence et al., 2013) and the Ensembl version 67 (NCBIM37 release, corresponding to mm9). In the resulting read counts table, each row represented one exon, each column one RNA-Seq sample and each cell, the corresponding read counts associated with each row and column. The exon read counts were filtered for artifacts that could affect the subsequent normalization and statistical testing procedures as follows: if an annotated gene had up to 5 exons, read presence was required in at least 2 of the exons, else if an annotated gene had more than 5 exons, then read presence was required in at least $0.2x\lceil E \rceil$ exons, where $\lceil . \rceil$ is the ceiling mathematical function. The final read counts for each gene model were calculated as the sums of their exon reads, creating a gene counts table where each row corresponded to an Ensembl gene model and each column corresponded to an RNA-Seq sample. The gene counts table was normalized for inherent systematic or experimental biases using the Bioconductor package *DESeq* after removing genes that had zero counts over all the RNA-Seq samples. Prior to the statistical testing procedure, the gene read counts were filtered for possible artifacts that could affect the subsequent statistical testing procedures. Genes presenting any of the following were excluded from further analysis: i) genes/transcripts with length less than 500, ii) genes/transcripts with read counts below the median read counts of the total normalized count distribution. Similar expression thresholds (e.g. the median of the count distribution) have been previously used in the literature, where the authors use the median RPKM value instead of normalized counts). The resulting gene counts table was subjected to differential expression analysis for the contrasts Kmt5a-KO versus WT using the statistical test from the Bioconductor package *DESeq*. Genes presenting a *DESeq* p-value less than 0.05 and fold change (for each contrast) greater than 0.58 or less than -0.58 in log2 scale corresponding to 1.5 times up and down in natural scale respectively were considered differentially expressed.

GSEA, GO and pathway analysis

Gene set enrichment analysis was performed with GSEA (Broad Institute) using default gene sets as well as mouse specific gene sets from the Bioconductor package *gskb* (https://bioconductor.org/packages/release/data/experiment/html/gskb.html). Gene Ontology (GO) enrichment and biochemical pathway analysis was performed using GeneCodis (Tabas et al., 2012). The results of GeneCodis for GO and pathway analysis were further filtered in order to remove very general (e.g. 'metabolic_process') or very specific functions (e.g. functions representing single-gene leaves in the GO hierarchy) based on the enrichment score reported by GeneCodis.

H&E staining and immunohistochemistry

Liver tissue was fixed in 4% Formaldehyde/PBS overnight, followed by tissue dehydration and paraffin infiltration. Preparation of tissue sections and staining with hematoxylin and eosin was performed as described in (Tatarakis et al. 2008; Elkouris et al. 2016).

For immunofluorescence staining, freshly isolated liver tissues were embedded in Optimal Cutting Temperature (OCT) embedding medium and were frozen in liquid nitrogen. Frozen sections (5μm thick) were fixed in 4% formaldehyde, blocked in 1% BSA/0,1% Triton X-100 for 1 hr and then incubated with the primary antibodies at room temperature for 2 hr or at 4^0C overnight. After incubation with AlexaFluor 568, (Molecular Probes) secondary antibodies for 1 hr at room temperature and counterstaining with DAPI, fluorescence images were observed using a Zeiss Axioscope 2 Plus microscope.

Accumulation of Reactive Oxygen Species (ROS) was measured as described (Nikolaou et al. 2012), by staining with 5-(and-6)-chloromethyl-20,70-dichlorodihydrofluoresceindiacetate, acetyl ester (CM-H_2DCFDA).

Periodic acid Schiff (PAS) staining was performed in formalin fixed paraffin-embedded liver sections. The sections were treated with 0.5% Periodic acid for 5min and stained in Schiff reagent (Sigma) for 10 min. After washings with PBS, the sections were counterstained with hematoxylin. Liver glycogen levels were measured by using Glycogen Assay II kit (Abcam).

Senescence-associated beta-galactosidase (SA-β-gal) activity was performed in frozen liver sections. After fixation in 0.5% glutaraldehyde for 20 min, the sections were washed with PBS and stained for 3 hours with 5mM of $K_3Fe(CN)_6$ and $K_4Fe(CN)_6$ supplemented with 1mg/ml X-Gal, pH=6. Sections were counterstained with eosin and observed with light microscopy.

Electron microscopy

Liver tissues were fixed for 2 h at room temperature in 0.08 M sodium cacodylate buffer, containing 2% of each glutaraldehyde and paraformaldehyde, followed by 1 h post-fixation with 1% osmium tetroxide. Following 1% uranyl acetate treatment for 20 min, samples were dehydrated with serial ethanol treatment and subsequently embedded in LR White resin/propylene oxide (Polysciences). Approximately 100 nm thin sections on copper grids were observed at 80 kV with a JEOL JM2100 transmission electron microscope.

Glucose tolerance (GTT), Insulin sensitivity (IST) and Pyruvate tolerance (PTT) test

Baseline blood glucose levels were measured from mice fasted for 5 hours (0 min time point). Then, 1g/kg glucose (for GTT) or 0.5 units/kg insulin (for IST) or 2g/kg sodium pyruvate (for PTT) was administered via intraperitoneal injection and blood glucose levels were measured 15, 30, 60, 120 minutes following the treatments. Blood samples

were collected from the tail vein. Glucose concentration was measured using AlphaTRAK glucose monitoring system (Abbott Laboratories).

Serum metabolite and enzyme measurements
Total serum cholesterol, triglyceride, free fatty acid and β-hydroxybutyrate levels were determined by using Cholesterol-cholesteryl Ester quantification kit (Abcam), the Triglyceride Quantification Assay kit (Abcam), the Free Fatty Acid Quantification Assay kit (Abcam) and the beta-Hydroxybutyrate Assay kit (Abcam), respectively. Serum alanine aminotransferase (ALT) activity was determined by using the ALT assay kit from Diasys.

Colorimetric ADP and ATP assays and Cox-I and Cox-IV activity
ADP/ATP levels were determined with ADP/ATP colorimetric assay kits (Biovision). Complex I and IV enzyme activities were measured by using enzyme activity dipstick assay kits (Abcam).

Whole Cell extract preparation
Freshly dissected liver pieces were rinsed in PBS and supplemented with 10 volumes of modified RIPA Buffer containing 50 mM Tris pH7,5, 1% NP40, 0,25% Na-Deoxycholate, 150 mM NaCl, 1mM EDTA pH8, 10% Glycerol, 1mM PMSF and 2 ug/ml Aprotinin. After homogenization with Polytron tissumizer for 10 seconds, the extracted tissue was incubated at 4^0C for 20 min with constant agitation. Extracted proteins were recovered after centrifugation at 14000 rpm for 20 minutes and stored at -80^0C

Nuclear Extracts preparation
Nuclear extracts from liver tissues were prepared as described in (Tatarakis et al., 2008). Briefly, liver tissue was minced to small pieces in a buffer containing 0.32 M sucrose, 15 mM Hepes pH 7.9, 60 mM KCl, 2 mM EDTA, 0.5 mM EGTA, 0.5 % BSA, 0.5 mM spermidine, 0.15 mM spermine and 0.5 mM DTT, supplemented with protease inhibitor coctail (Roche). After 10 strokes of homogenization with Teflon pestle the nuclei were layered over equal volume of a buffer containing 30% sucrose, 15 mM Hepes pH 7.9, 60 mM KCL, 2 mM EDTA, 0.5 mM EGTA, 0.5 mM spermidine, 0.15 mM spermine and 0.5 mM DTT and centrifuged for 15 min at 3000 rpm. Pelleted nuclei were washed with PBS, resuspended in 10 volumes of modified RIPA buffer containing 50 mM Tris pH 7.5, 1% NP40, 0.25% deoxycholate, 150 mM mM NaCl, 1 mM EDTA, 10% glycerol and protease inhibitor coctail (Roche) and incubated at 4 ^{0}C for 20 min with constant agitation. After centrifugation at 14000 rpm, the extracts were collected.

Immunoprecipitation and Western blot assays
Immunoprecipitation or western blot analysis as described (Elkouris et al., 2016). Briefly, nuclear extracts containing ~ 500 μg total protein were mixed with primary antibodies and incubated with 50 μl bed volume Protein-G magnetic Dynabeads for 5 hours at 4 ^{0}C with constant shaking. The beads were washed 4 times with excess RIPA buffer.
For western blot analysis extracted or immunoprecipitated proteins were separated by SDS-PAGE and transferred to nitrocellulose membranes (BIORAD). The membranes were blocked by incubation in TBST buffer containing 20 mM Tris, pH 7.5, 150 mM NaCl and 0.1% Tween 20, supplemented with 5% non-fat dry milk for 1 hour at room temperature. Incubations with primary and HRP-conjugated secondary antibodies were performed in TBST buffer for 1 hour each. After 4 washing steps with excess TBST buffer, the membranes were developed using ECL chemiluminescent kit from Thermo Fisher and exposed to X-ray films.

Supplemental References

Schwalb, B., Michel, M., Zacher, B., Frühauf, K., Demel, C., Tresch, A., Gagneur, J., and Cramer, P. (2016). TT-seq maps the human transient transcriptome. Science *352*, 1225-1228.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. *14,* R36

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

Moulos, P., and Hatzis, P. (2015). Systematic integration of RNA-Seq statistical algorithms for accurate detection of differential gene expression patterns. Nucleic Acids Res *43*, e25

Tabas-Madrid, D., Nogales-Cadenas, R., Pascual-Montano, A. (2012). GeneCodis3: a non-redundant and modular enrichment analysis tool for functional genomics. Nucleic Acids Res. *40,* W478-83.

Xu, S., Grullon, S., Peng, W. (2014). Spatial Clustering for Identification of ChIP-Enriched Regions (SICER) to Map Regions of Histone Methylation Patterns in Embryonic Stem Cells. Methods Mol. Biol. *1150,* 97-111.