

Figure S1

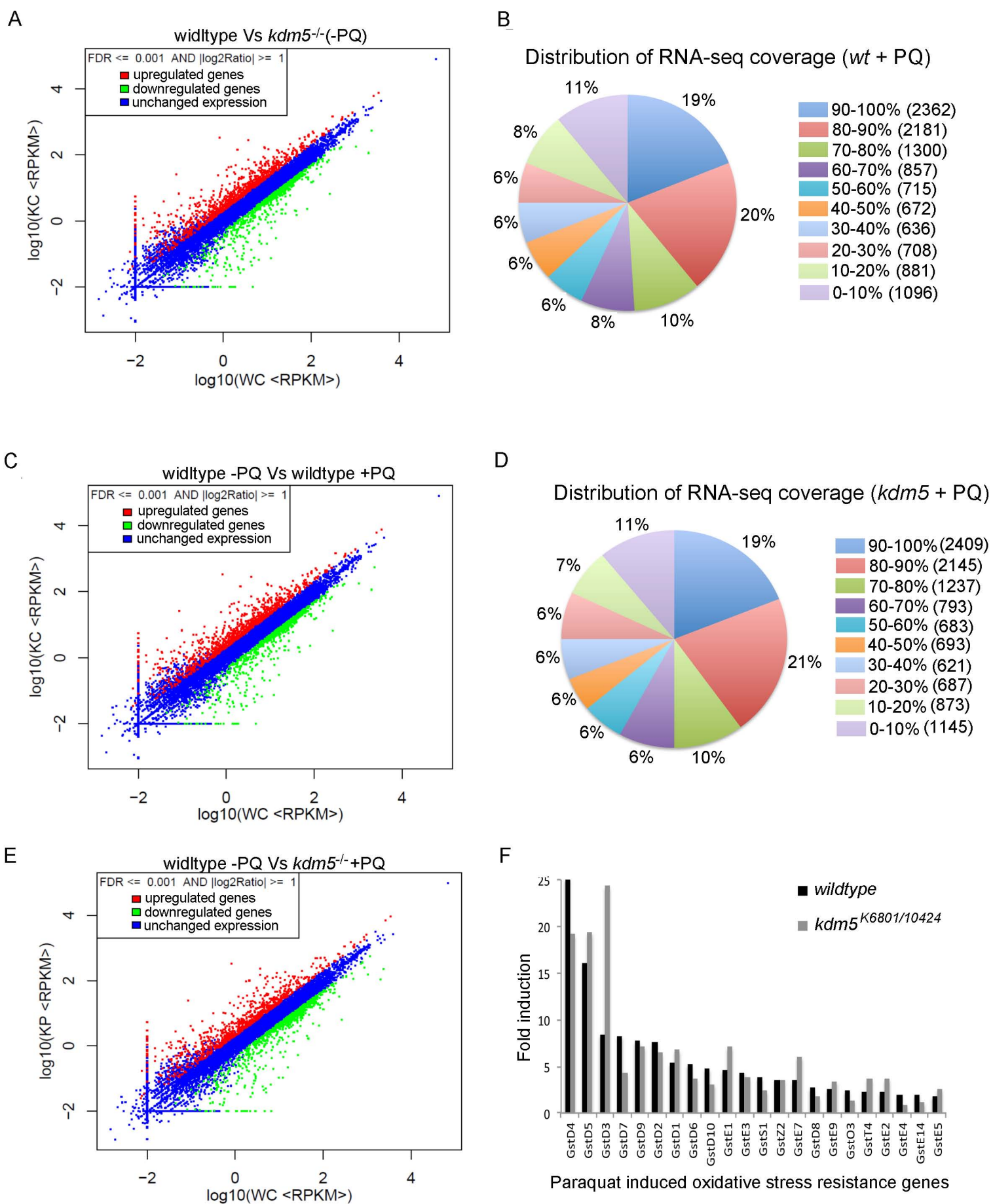


Figure S2

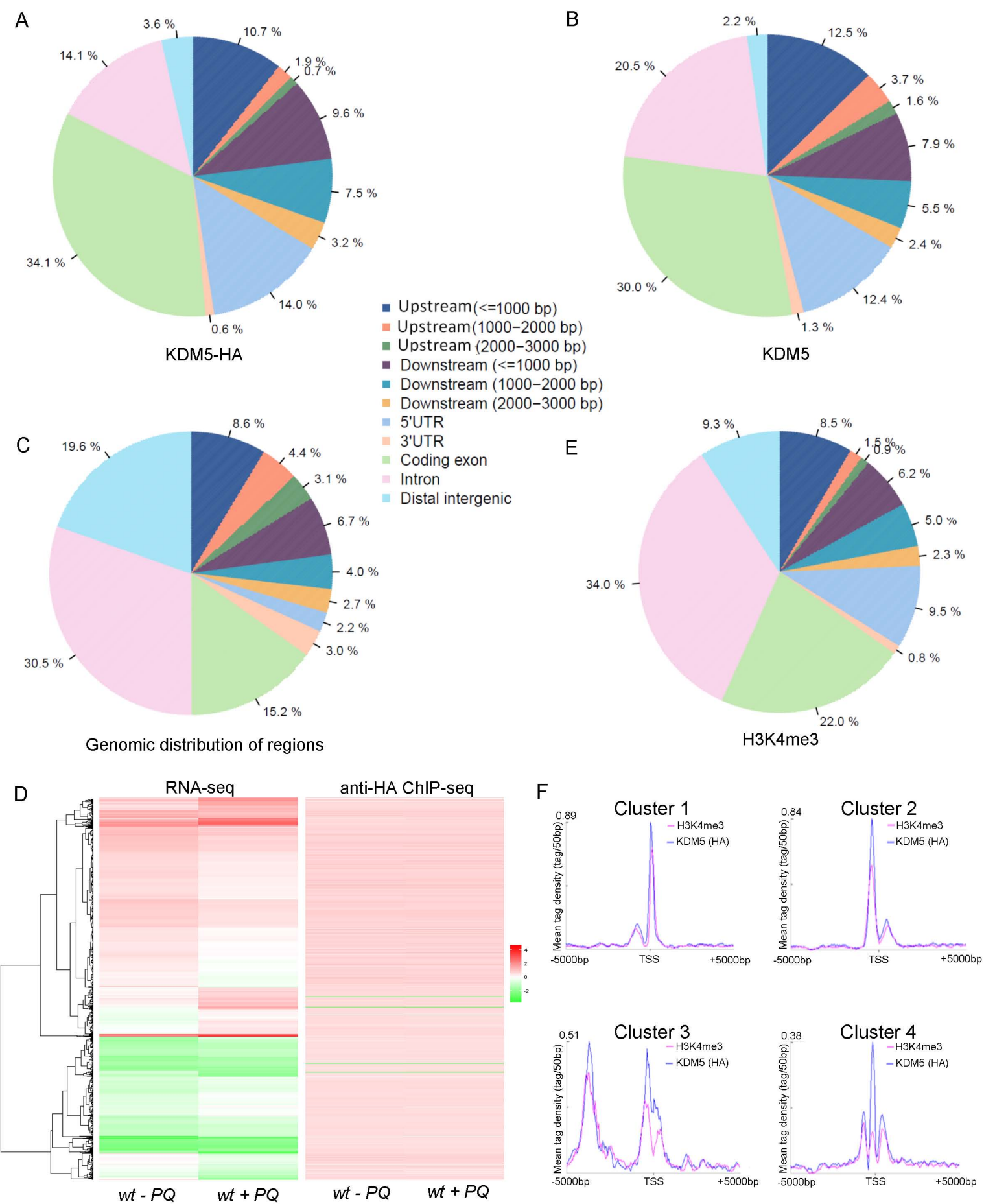


Figure S3

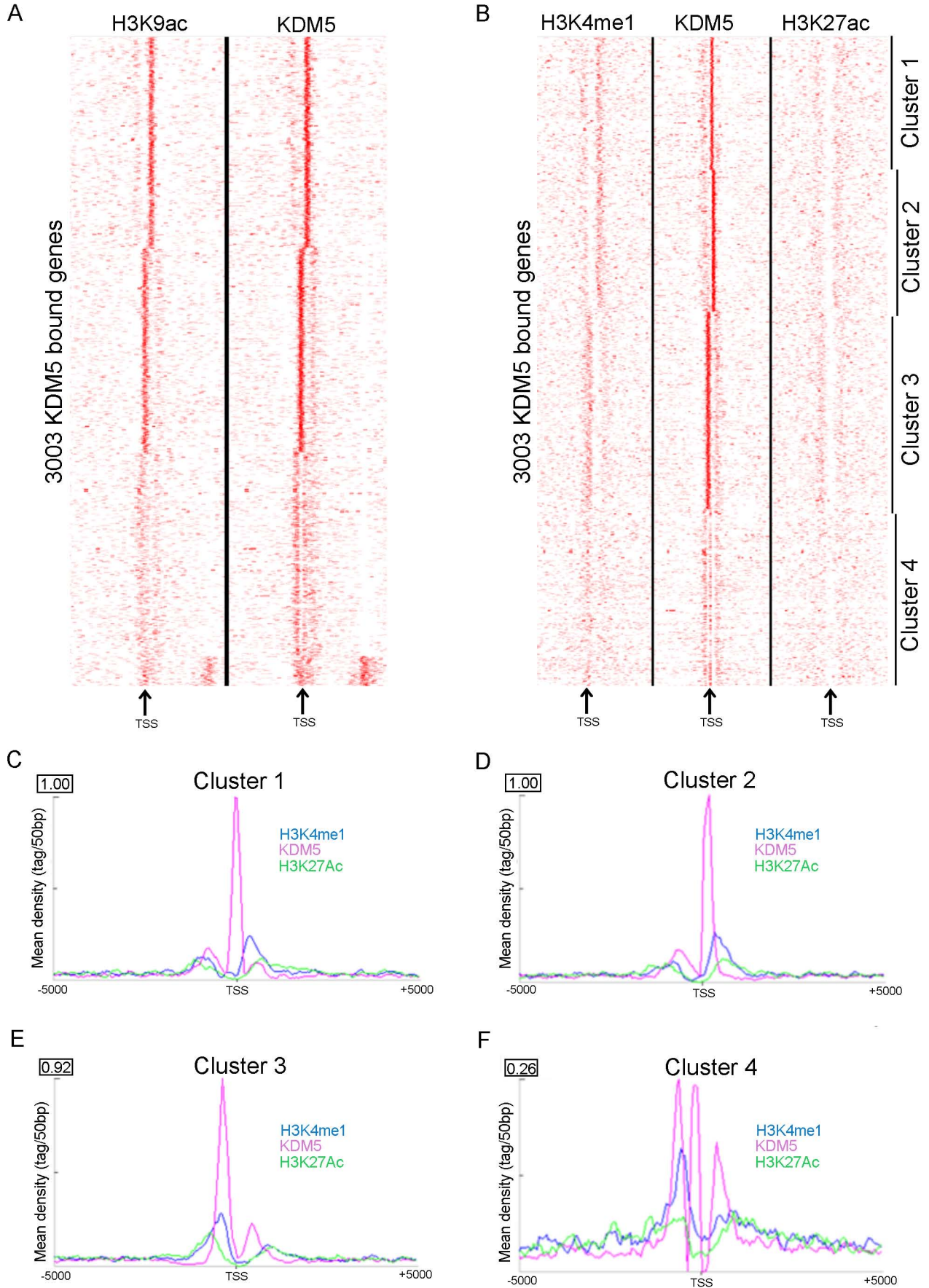


Figure S4

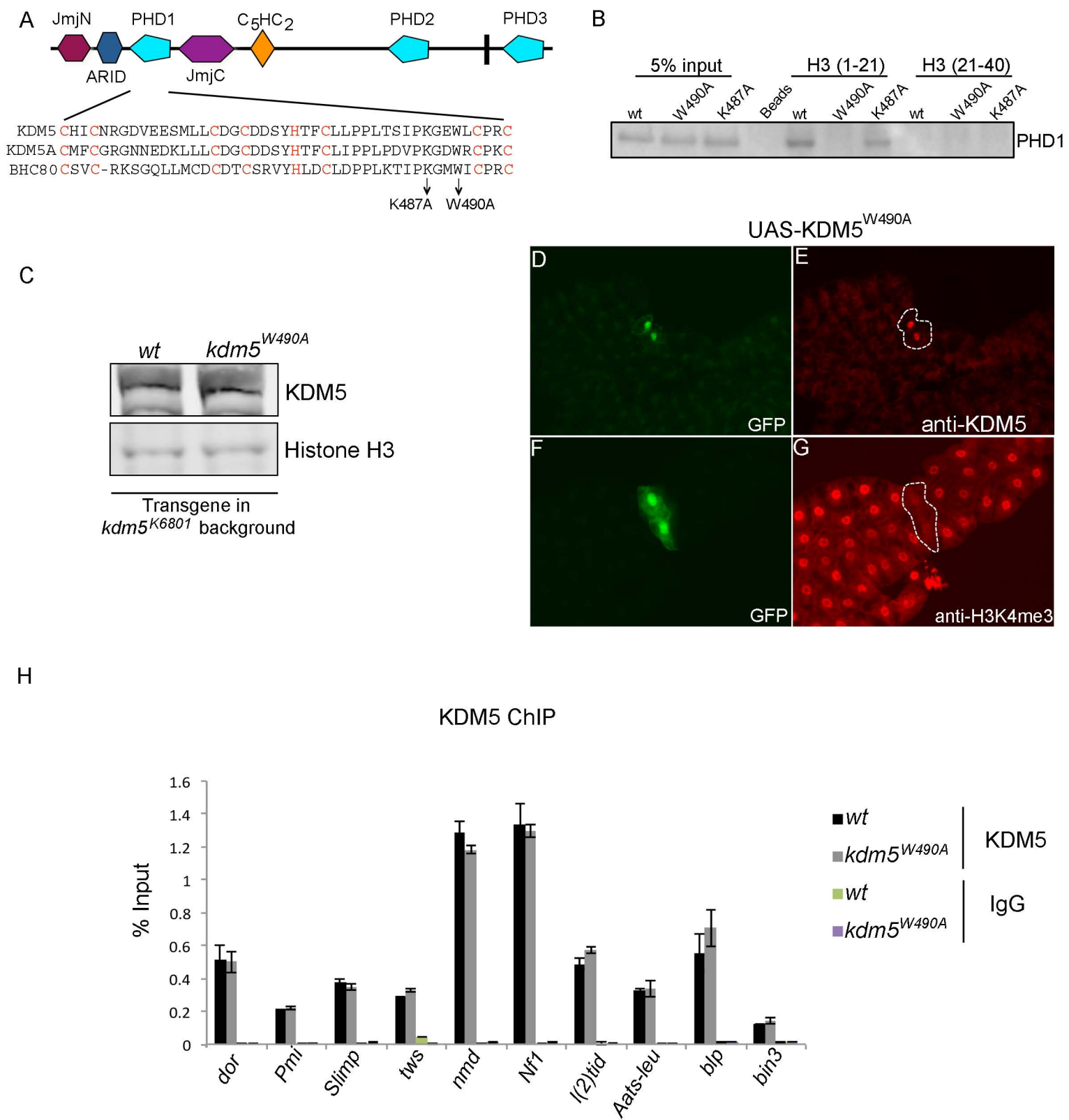


Figure S5

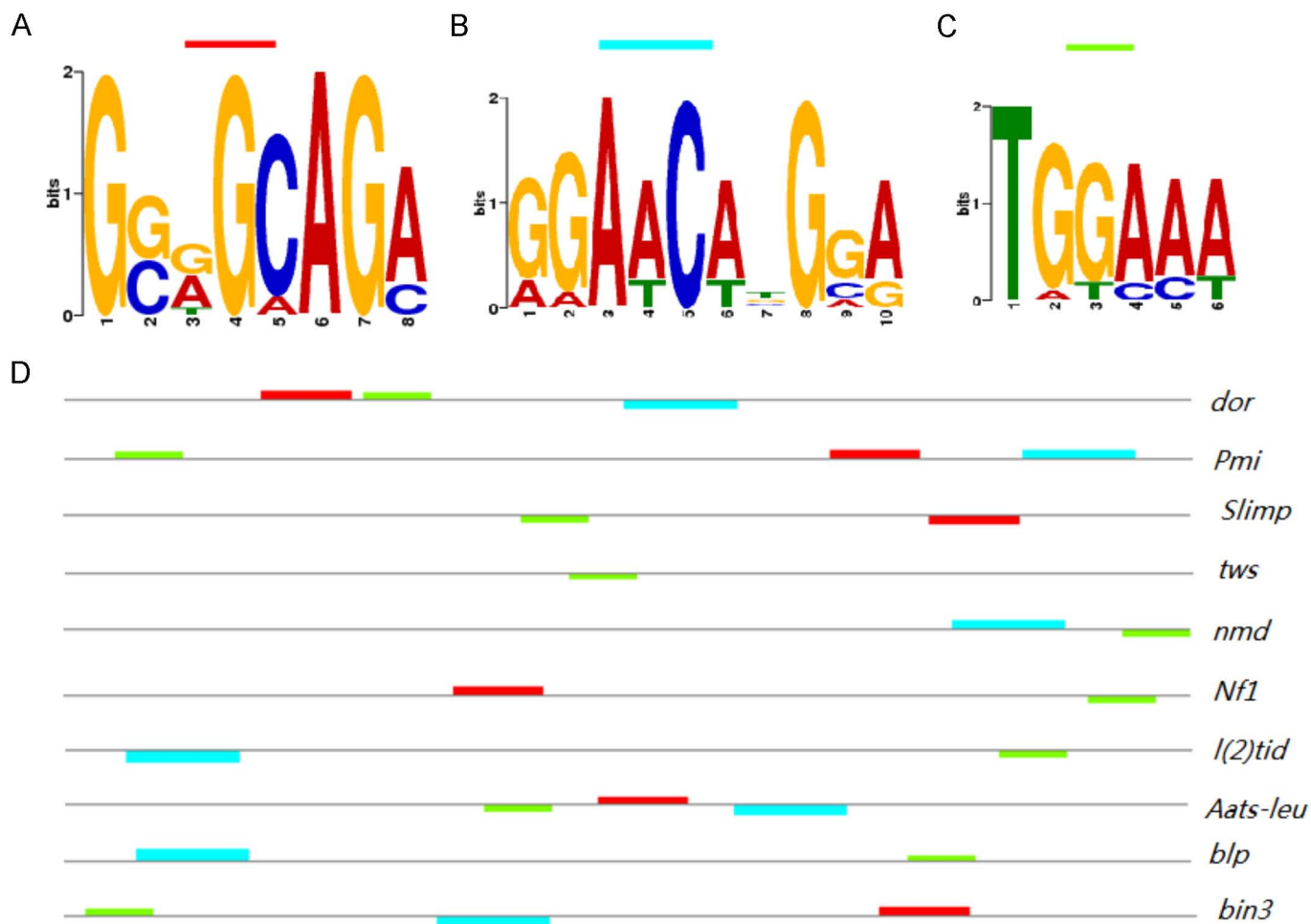


Figure S6

Supplemental Figure Legends

Figure S1: Comparison of gene expression in *kdm5* mutant larvae and adults, related to Figure 1.

(A) Real-time PCR analyses of the genes indicated in *kdm5*^{K6801/10424} mutant 3rd instar larvae or adults compared to wildtype animals (*w*¹¹¹⁸) at the same developmental stage. *kdm5*^{K6801/10424} animals were generated by crossing *kdm5*^{K6801}/CyO females to *kdm5*¹⁰⁴²⁴/CyO males. *kdm5*^{K6801/10424} adults eclose at ~50% of the expected Mendelian frequency (Liu et al., 2014; Secombe et al., 2007). All gene expression levels were normalized to the housekeeping gene *rp49*. * indicates genes that were significantly downregulated in *kdm5* mutant larvae (as similar to our published data (Liu et al., 2014)), but unaffected in adults. ** indicates genes that were downregulated in *kdm5* mutant larvae but significantly upregulated in mutant adults ($p < 0.05$). (B-C) Distribution of RNA-seq reads showing that all samples had similar coverage of the reference genome (70-100%). (B) RNA-seq libraries for wildtype without paraquat (C) *kdm5* mutants without paraquat (D) Summary of total number of genes downregulated (green) and upregulated (red) in *kdm5* mutant adults compared to wildtype in non-stressed conditions ($p < 0.05$; FDR < 0.05). Also shown is summary of the total number of genes down (green) and up (red) regulated in response to paraquat treatment ($p < 0.05$; 1.5-fold change or more). (E) Heat map from RNA-seq data showing the gene expression level and clustering of wildtype (*w*¹¹¹⁸) and *kdm5*^{K6801/10424} mutant whole adults in the presence (+PQ) or absence (-PQ) of the oxidative stress agent paraquat. Downregulated genes are shown in green and upregulated in red.

Figure S2: Analyses of gene expression levels and KDM5 binding in normal and oxidative stress conditions, related to Figure 1.

(A) Scatter plot showing differentially expressed genes for the pair-wise comparisons of wildtype and *kdm5* mutants (B) Distribution of RNA-seq reads in wildtype with paraquat (C) Scatter plot showing differentially expressed genes for the pair-wise comparison of wildtype with and without paraquat treatment. (D) Distribution of RNA-seq reads in *kdm5* mutants both treated with paraquat. (E) Scatter plot showing differentially expressed genes for the pair-wise comparison of *kdm5* mutants with and without paraquat treatment (F) Paraquat-mediated activation of oxidative stress resistance genes of the glutathione S transferase D and E family from RNA-seq data. Shown is the level of activation observed in wildtype

(black) flies and *kdm5*^{K6801/10424} flies (grey). Levels of paraquat-mediated activation in *wt* and *kdm5*^{K6801/10424} animals is shown relative to their expression level in the absence of oxidative stress conditions.

Figure S3: Quality control analyses for ChIP-seq data, related to Figure 1.

(A) Distribution of ChIP-seq peaks from anti-HA ChIP-seq (from KDM5:HA dataset #1) (B) Distribution of ChIP-seq peaks from anti-KDM5 ChIP-seq (C) whole genome distribution of the genomic features. (D) Heat map showing genes induced and repressed by the addition of paraquat (left) and KDM5 (using anti-HA dataset) binding in normal and oxidative stress conditions (right). While there are many changes to gene expression upon paraquat treatment, KDM5 binding does not significantly change. (E) Distribution of ChIP-seq peaks from anti-H3K4me3 (F) Distribution of KDM5 binding and the H3K4me3 chromatin mark within the four clusters identified in Figure 1J.

Figure S4: KDM5 binding is strongly correlated with H3K9 acetylation (H3K9ac) and weakly with H3K27 acetylation (H3K27ac) and H3K4 monomethylation (H3K4me1). Related to Figure 1.

(A-B) Heat maps showing distribution of KDM5, H3K9ac, H3K4me1 and H3K27ac relative to TSS. H3K9ac, H3K4me1 and H3K27ac datasets were generated as part of the modENCODE project (Negre et al., 2011). Heat maps only show 3003 KDM5-bound genes. (A) Heat map showing distribution of KDM5 and H3K9ac relative to TSS (indicated by arrow) and clustered using k-means. (B) Heat map showing distribution of H3K4me1, KDM5 and H3K9ac relative to TSS clustered using k-means. (C-F) Distribution of KDM5 binding and the H3K4me27ac and H3K4me1 chromatin marks within clusters 1 to 4 identified in part B.

FigureS5: Preventing KDM5-mediated binding to H3K4me0 does not affect promoter recruitment, related to Figure 6.

(A) Schematic representation of the N-terminal PHD motif of KDM5 (PHD1) and alignment with human KDM5A and BHC80. Mutations in KDM5 that potential abolish H3 binding (K487 and K490) are based on the crystal structure of BHC80 (Lan et al., 2007). (B) *in vitro* binding between biotinylated histone H3

peptides that include amino acids 1-21 or 21-40 and GST-PHD1. The interaction between PHD1 and H3 (1-21) is abolished by GST-PHD1^{W490A} and reduced by GST-PHD1^{W487AA}. (C) Western blot analyses from five female adult heads showing levels of KDM5 and the loading control histone H3. Genotypes are as follows: *wt* is *kdm5*^{K680I} ; gKDM5^{WT} (only source of KDM5 is from genomic rescue transgene) and *kdm5*^{W490A} is *kdm5*^{K680I} ; gKDM5^{W490A}. (D-G) Clones of cells marked by the presence of GFP expressing UAS-KDM5^{W490A} transgenes in fat body cells. Clones were generated by crossing *hs*-FLP; UAS-KDM5^{W490A} females to *actin*>CD2>Gal4, UAS-GFP males. Cells expressing a transgene are shown by the co-expression of GFP (D, F) and are outlined in the adjacent panels showing KDM5 levels (E) and H3K4me3 (G). (H) ChIP-PCR analyses of KDM5 levels at the target genes indicated in wildtype flies (*kdm5*^{K680I} mutant flies rescued by a wildtype genomic rescue construct (*wt*)) and *kdm5*^{W490A} mutant flies. No significant changes were observed.

Figure S6: KDM5 binding regions of mitochondrial function genes show motif enrichment, related to Figure 6.

(A-C) Motifs one through three identified using MEME-ChIP. The height of each nucleotide represents its preference at that position. (D) Schematic representation of KDM5-bound regions of 10 mitochondrial function genes. The position of each motif is shown in different colors. Motifs indicated by a box above the line are in the same orientation shown in parts A-C. Motifs indicated under the line were on the reverse strand.

Table S1: Summary of RNA-seq mapping data, related to Figure 1.

Sample	Total reads	Total Base Pairs	Total mapped reads	Perfect match	<=2bp mismatch	Unique match	Multi-position match	Total unmapped reads
wt - PQ	35,238,779	1,726,700,171	32,324,139 (91%)	26,085,851 (74%)	6,238,188 (18%)	27,938,478 (79%)	4,385,661 (12%)	2,914,640 (8%)
wt +PQ	36,233,226	1,775,428,074	33,260,091 (92%)	26,880,300 (74%)	6,379,791 (17%)	27,880,985 (77%)	5,379,106 (15%)	2,973,135 (8%)
kdm5 - PQ	35,742,402	1,751,377,698	31,659,948 (89%)	25,630,786 (72%)	6,029,162 (17%)	27,921,672 (78%)	3,738,276 (10%)	4,082,454 (11%)
kdm5 - PQ	37,699,108	1,847,256,292	33,235,414 (88%)	26,962,497 (72%)	6,272,917 (17%)	27,685,041 (74%)	5,550,373 (15%)	4,463,694 (12%)

A total of 35742402 and 35238779 reads for wildtype (wt) and *kdm5*^{K6801/10424} adults were sequenced respectively. Approximately 88% for wildtype and 91% for *kdm5*^{K6801/10424} were able to be mapped a single location in the genome. There was also a similar read distribution of all coverage genes in the four RNA-seq libraries (see also Figure S1), with over 50% of the reference genes having 70%-100% coverage. The reads therefore could cover the reference genome for both libraries sequenced. For wildtype and *kdm5* mutant paraquat treated flies, a total of 37699108 and 36233226 reads were sequenced, respectively.

Table S2: Gene expression levels of TCA cycle enzymes from RNA-seq data, related to Figure 3.

TCA cycle enzyme	Gene name	Relative expression in <i>kdm5</i> mutants	KDM5 ChIP signal?	p value
citrate synthase	knockdown (kdn)	2.3	+/-	0.00E+00
aconitase	Acon	2.5	-	0
isocitrate dehydrogenase	Idh	1.8	-	0
alpha-ketoglutarate dehydrogenase/oxoglutarate dehydrogenase	E1 - oxoglutarate dehydrogenase (OGDH)	nd	nd	nd
	E2- dihydrolipoyl succinyltransferase (DLST; CG5214)	1.8	+	0
	E3 - dihydrolipoyl dehydrogenase (DLP; CG7430)	1.9	+	0
succinyl-CoA synthetase	scs- α	1.6	-	5.80E-198
	sucb	1.07	+	0.08
succinate dehydrogenase	SdhA	2.5	-	0
	SdhB	2	-	0
Fumarase	I(1)G0255	1.6	+/-	7.08E-189
Malate dehydrogenase	Mdh2	2.1	-	0

Supplementary Table S3: Metabolite analyses of *kdm5* mutant flies, related to Figure 3.

Table is larger than 3 pages and is provided as a separate file.

Table S4: Primers used for RT-PCR and CHIP analyses, related to Figures 2, 4, 5 and 6.

RT-PCR primer	
Aats-leu fw	GCTCAGTGGAGGACAGTTCA
Aats-leu rv	GGATACGGGAACATGGAGAG
blp fw	GCG GCT CCT TCT ATA TCC AG
blp rv	CTG CTT CCG TAT TTG ACG AC
ttm50 fw	ACT GCT GGA CCT TAT CGC TT
ttm50 rv	TGG CGA TAG TAG TGG AGC AC
sesb fw	CAA GGA TTT CGA TGC TGT TG
sesb rv	GGG CTG ATT TGT TTC GAG AT
L(2)tid fw	CTC CTA CCG AAT CCT CTC CA
l(2)tid rv	AGC GTG GCG TAG TAG TCC TT
ifc fw	CTA TGC ACA ACC GGA TCT TG
ifc rv	GTG TCA ATT GCC TCA TCA CC
Porin fw	CTG ACC ACC AAC AAC TTT GC
porin rv	TGT ACA CCC ACA TCC AGC TT
dor fw	TCC TGA ACA CCT TCA GCT TG
dor rv	GAC TTG TCC GTC TCA GCG TA
khc fw	CAG AGT GCG AGC GTC TCT AC
khc rv	TGG CTC TGC TGA TTG ATC TC
tws fw	ACG TTT ATG CGA CAG TCG AG
tws rv	ACG CCT CTC CAT TAC TTG CT
nf1 fw	CAA ACA GGT GAC CGA ATC TG
nf1 rv	ATT GGA GTC GGT GTT GTT GA
slimp fw	TGG TCG GAG CTA CAA CAG AG
slimp rv	GTA CCG CAT TCG TTT GTG TC
tspo fw	CCT TAA GTT CCC GTC CTT CA
tspo rv	ACA CCA GGT AGG AGC CGT AG
fis1fw	CCG AAG TCG GTA CAC AAA TG
fis1rv	GAG CAT TAC CAA ACG CAA GA
drp1 fw	CAC GAA ACT GGA TCT CAT GG
drp1 rv	TTC TGC GAG CGA TTC ATA AC
bin3 fw	TCT TTG TGC GTA TTC CCA AG
bin 3rv	GAG CTG CTC CAG ATG TCT CA
marf fw	AAC ATC AAA CAA CTG GCG AA
marf rv	ACA CCA GGT GAG TCC ACA AA
Tpc1 fw	CCT TCT GGA AGG GAC ACA AT
Tpc1 rv	GCT CGT ACG TCC AGA ACT GA
Aats-leu fw	GAT ACA GCG CTG ACC AAG AA
Aats-leu rv	GGT TTA TAA AGC GGG CGT AG
Nmd fw	CCA GCA CAA GGA TCT CTT CA
nmd rv	GCT ATC AGC GTC TTT CCA CA
pmi fw	CAT GAA AGA TTC GAG GCT GA
pmi rv	AAT CCA ACG ACC AGT TTC GT
Chip-qPCR primer	
cg2794 pfw	AGG CCT GGT AAT TGG TGT TC
cg2794 prv	CTT GAT GGC CTC CTC AAT TT

ent1 pfw	TCG AAT ACT TCT CCA CAA TTA CAG A
ent1 prv	GTA CGC AGC CCT TGG TAT TC
blp pfw	GCA TCT CAG GAA GCA GCA C
blp prv	ATT CAG GAT CTG CTT GGC TT
tws pfw	AAA GGA GAA GGC CAA AGG AG
tws prv	ATC GGC AGC TAC AAC AAC AA
l(2)tid pfw	GAC ATC AAG AAA GCC TAC TAC CAG
l(2)tid prv	ATC CGG ATC CTC CTT GTT C
slimp pfw	CAC ATC ACG CTC AGC TGT C
slimp pRV	GCT GAA CCG CAC TAC TAC CA
dor pfw	TAT GCC TAA CCA GCC CAA GT
dor prv	ATT TCA TCC TCA TCG TCC GT
nmd pfw	ACA TGG ACA ACT TCG GAC TG
nmd prv	AGA CAG GCG AAC CAG TAC CT
bin3 pfw	CGG CAG CTA TCG AGT TCT TA
bin3 prv	GGT GTT CAC TTG GCA CAT TC
Aats-leu pfw	TGA TCT CCG GTA TGT CGA TG
Aats-leu prv	GGA AGA ACC CAC TGC ACT CT
nf1 pfw	TTG TAA AGC CAA ACC ATC CA
nf1 prv	GGA GTG GCA ATT GGT TGA G
pmi pfw2	ATC CTG CAG GGT CAC TTC TC
pmi prv2	GGA AGG AAG ATC CAT GTA GCA

Supplemental experimental procedures.

RNA extraction and cDNA library preparation: RNA was extracted from groups of five adult flies (1-3 days old) using TRIZOL (Invitrogen) and cleaned up using DNA-free (Ambion). Control genotype flies were w^{1118} and *kdm5* mutant flies were $w^{1118}; kdm5^{K6801/10424}$. Flies were subjected to control conditions (5% sucrose for 6 hours) or conditions of oxidative stress (paraquat; six hours of 20mM paraquat in 5% sucrose) as previously described (Liu et al., 2014). RNA concentrations were quantified using a NanoDrop Spectrophotometer and sample integrity shown using an Agilent 2100 Bioanalyzer (Agilent Technologies Inc.; only samples with values above 8.0 were used for experiments). cDNA libraries were prepared using an Illumina TruSeq RNA sample prep kit using three micrograms of total RNA. The average size of the library cDNAs was 150 bp (excluding the adapters). Prior to sequencing the integrity and quality of cDNA libraries were assessed using an Agilent 2100 Bioanalyzer and an ABI StepOnePlus real-time PCR system.

RNA-seq analyses: After the sequencing platform generated the sequencing images, the pixel-level raw data collection, image analysis, and base calling were performed by Illumina's Real Time Analysis (RTA) software on a Dell PC attached to the HiSeq2000 sequencer. The base call files (*.BCL) were converted to qseq files by the Illumina's BCL Converter, and the qseq files were subsequently converted to FASTQ files for downstream analysis. The RNA-Seq reads from the FASTQ files were mapped to the *Drosophila* reference genome (dm3) using SOAP (Short Oligonucleotide Analysis Package) (Li et al., 2009). The output files in BAM (binary alignment/map) format were analyzed by Cufflinks to estimate the transcript abundance and the presence of putative novel mRNA isoforms. The transcript structure predictions generated by Cufflinks carried out by comparing reads with the reference annotation, Ensembl GTF version 65 using Cuffcompare. RNA-seq sequencing and analysis was performed by BGI company. RNA-Seq have been submitted to NCBI GEO (Accession number GSE70591).

Real-time PCR: 1ug of total RNA was reverse transcribed at 50°C for 40 minutes using Verso cDNA kit (Thermo Scientific) with oligo (dt) primer to generate cDNA. qRT-PCR reactions were performed in triplicate in total volumes of 10 µl containing Fast SYBR Green Master Mix, 0.25 µl of each gene-specific

primer, 0.5 µl of first strand cDNA template, and nuclease free water. All qRT-PCR reactions were performed with the following conditions: 50°C for 2 min, 95°C for 5 min followed by 40 cycles of denaturation at 95°C for 3 s, annealing at 60°C for 30s. Primers used for real-time PCR analyses are shown in Table S2.

Chromatin immunoprecipitation: To identify KDM5 binding sites ChIP-seq was carried out using HA-tagged KDM5 genomic rescue transgene flies (Liu et al., 2014) and anti-HA (with or without paraquat treatment for 6 hours). Wildtype *w¹¹¹⁸* adults were used for anti-KDM5 and anti-H3K4me3 ChIP-seq. ChIP was performed according to a previously published method with the following modifications (Bai et al., 2013). 200–250 10 day old adult females (~200 mg) were pooled for each ChIP sample. Two biological replicates were prepared for KDM5 (anti-HA) ChIP and one sample for H3K4me3 and KDM5. Immunoprecipitation was performed using Protein A beads (Invitrogen), anti-HA (Abcam), anti-KDM5 (Secombe et al., 2007), or anti-H3K4me3 (Active Motif) following the wash with LiCl and TE. DNA-protein complexes were eluted from the Protein A beads and reverse cross-linked. After Proteinase K digestion, bound DNA fragments were purified and diluted in Tris-HCl buffer. About 20ng of ChIP DNA and input DNA (DNA sample before the immunoprecipitation) were used for library preparation (carried out by BGI for anti-HA and the epigenetics core facility of Albert Einstein College of Medicine for H3K4me3 and KDM5). Libraries were prepared using Tru-Seq adaptors and then size-selected (150 bp–350 bp) and purified by agarose gel. Sequencing was performed using 50 bp (HA-Chip) and 100 bp (H3K4me3 and KDM5 chip) single-end reads from an Illumina HiSeq 2000 in Rapid Run mode. Reads were mapped to the Dm3 reference genome.

To map KDM5 and H3K4me3 reads, we pooled the raw reads (about 20 million reads per sample) from one data file and aligned it to *Drosophila* reference genome using SOAP (Li et al., 2009) for KDM5-HA ChIP and Bowtie short read aligner (Langmead et al., 2009) for KDM5 and H3K4me3 ChIP. About 70% of raw reads have at least one alignment. Enrichment of KDM5 and H3K4me3 binding between ChIP DNA and input DNA was determined using peak calling package MACS2 (version 2.1.0.20140616.0) (Feng et al., 2012). We used CEAS (Shin et al., 2009) to determine the genomic distribution of peaks using dm3

refGene track and PAVIS (Huang et al., 2013) to map peaks to the nearest gene and using maximum distance of 5kb from the TSS. ChIP-Seq data is has been submitted to NCBI GEO (Accession number GSE70591). The binding signal of H3K4me3 ChIP-seq and KDM5 ChIP-seq in +/- 5kb of the TSS and gene body was analyzed using seqMINER (Ye et al., 2011) to generate a heatmap. The H3K9Ac (GSM439459), H3K4me1 (GSM439465) and H3K27Ac (GSM439459) data are publically available (Negre et al., 2011)and were analyzed in a similar manner to the KDM5/H3K4me3 data.

For ChIP-PCR analysis, the binding signals were calculated as a percentage of input DNA (signal obtained from ChIP-PCR is divided by the signal obtained from the input DNA). IgG was used as a negative control. Primers used for ChIP-PCR are shown in Table S2.

References:

- Bai, H., Kang, P., Hernandez, A.M., and Tatar, M. (2013). Activin signaling targeted by insulin/dFOXO regulates aging and muscle proteostasis in *Drosophila*. *PLoS Genet* 9, e1003941.
- Feng, J., Liu, T., Qin, B., Zhang, Y., and Liu, X.S. (2012). Identifying ChIP-seq enrichment using MACS. *Nature protocols* 7, 1728-1740.
- Huang, W., Loganantharaj, R., Schroeder, B., Fargo, D., and Li, L. (2013). PAVIS: a tool for Peak Annotation and Visualization. *Bioinformatics* 29, 3097-3099.
- Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome biology* 10, R25.
- Li, R., Yu, C., Li, Y., Lam, T.W., Yiu, S.M., Kristiansen, K., and Wang, J. (2009). SOAP2: an improved ultrafast tool for short read alignment. *Bioinformatics* 25, 1966-1967.
- Liu, X., Greer, C., and Secombe, J. (2014). KDM5 interacts with Foxo to modulate cellular levels of oxidative stress. *PLoS Genet* 10, e1004676.
- Negre, N., Brown, C.D., Ma, L., Bristow, C.A., Miller, S.W., Wagner, U., Kheradpour, P., Eaton, M.L., Loriaux, P., Sealfon, R., *et al.* (2011). A cis-regulatory map of the *Drosophila* genome. *Nature* 471, 527-531.
- Secombe, J., Li, L., Carlos, L.S., and Eisenman, R.N. (2007). The Trithorax group protein Lid is a trimethyl histone H3K4 demethylase required for dMyc-induced cell growth. *Gene Dev* 21, 537-551.
- Shin, H., Liu, T., Manrai, A.K., and Liu, X.S. (2009). CEAS: cis-regulatory element annotation system. *Bioinformatics* 25, 2605-2606.
- Ye, T., Krebs, A.R., Choukrallah, M.A., Keime, C., Plewniak, F., Davidson, I., and Tora, L. (2011). seqMINER: an integrated ChIP-seq data interpretation platform. *Nucleic Acids Res* 39, e35.