Supplemental Information

Active N^6-Methyladenine demethylation by DMAD regulates gene expression by coordinating with Polycomb protein in neurons

Bing Yao[1, *, §], Yujing Li[1, *], Zhiqin Wang[1], Li Chen[1,2,7], Mickael Poidevin[1], Can Zhang[3], Li Lin[1], Feng Wang[1], Han Bao[1], Bin Jiao[1], Junghwa Lim[1], Ying Cheng[1], Luoxiu Huang[1], Brittany Lynn Phillips[4], Tianlei Xu[2], Ranhui Duan[5], Kenneth H. Moberg[3], Hao Wu[2], Peng Jin[1,6, §]

[1]Department of Human Genetics,

[2]Department of Biostatistics and Bioinformatics,

[3]Department of Cell Biology,

[4]Department of Pharmacology, Emory University School of Medicine, Atlanta, GA, 30322, USA;

[5]State Key Laboratory of Medical Genetics, School of Life Science, Central South University, Changsha, Hunan, 410078, China.

[6]Lead Contact

[7]Current address: Department: Health Outcomes Research and Policy, Auburn University Harrison School of Pharmacy, Auburn, AL, 36849, USA.

*These authors contributed equally to this work.

[§]Correspondence should be addressed to Bing Yao (bing.yao@emory.edu) or Peng Jin (peng.jin@emory.edu).

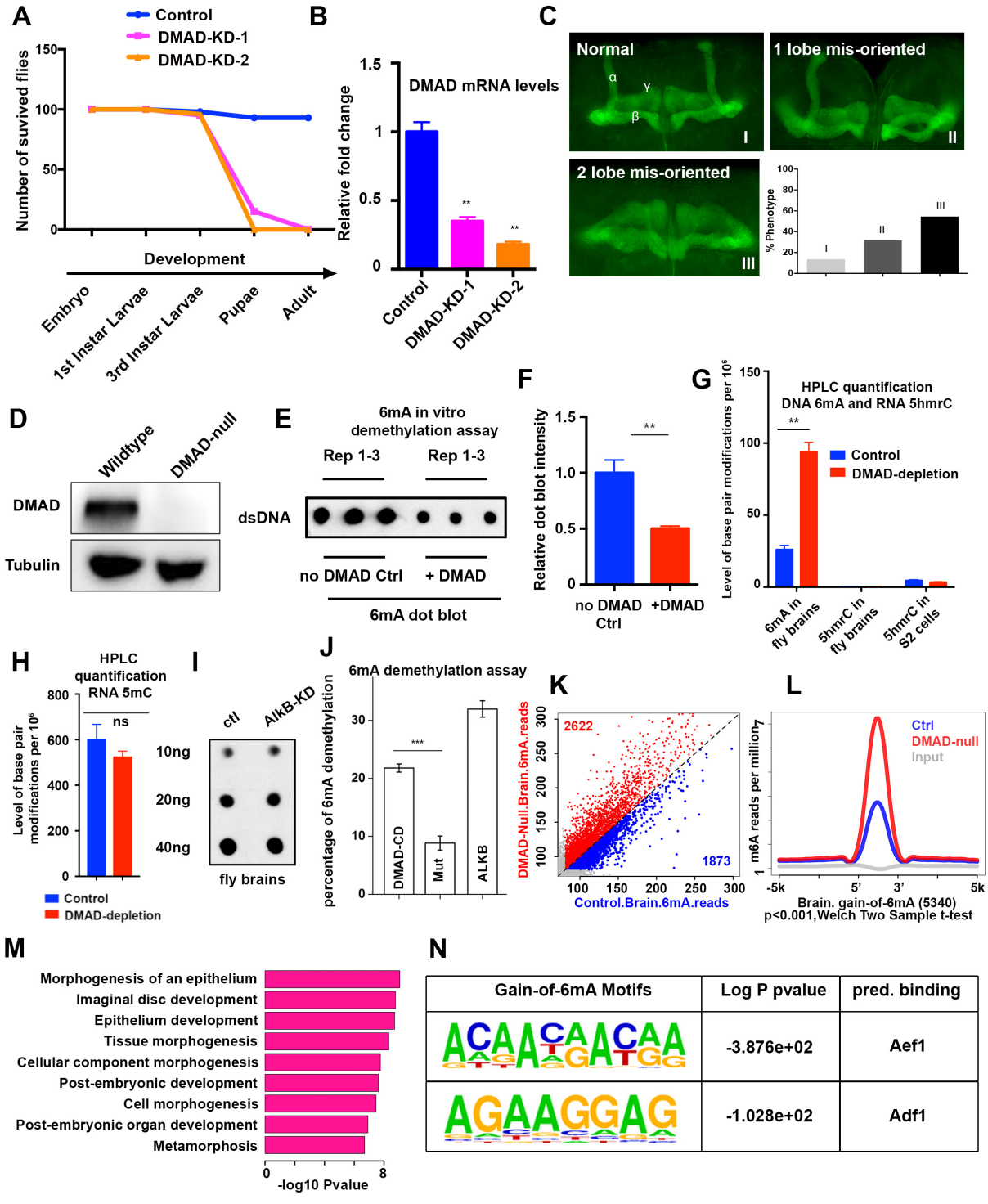
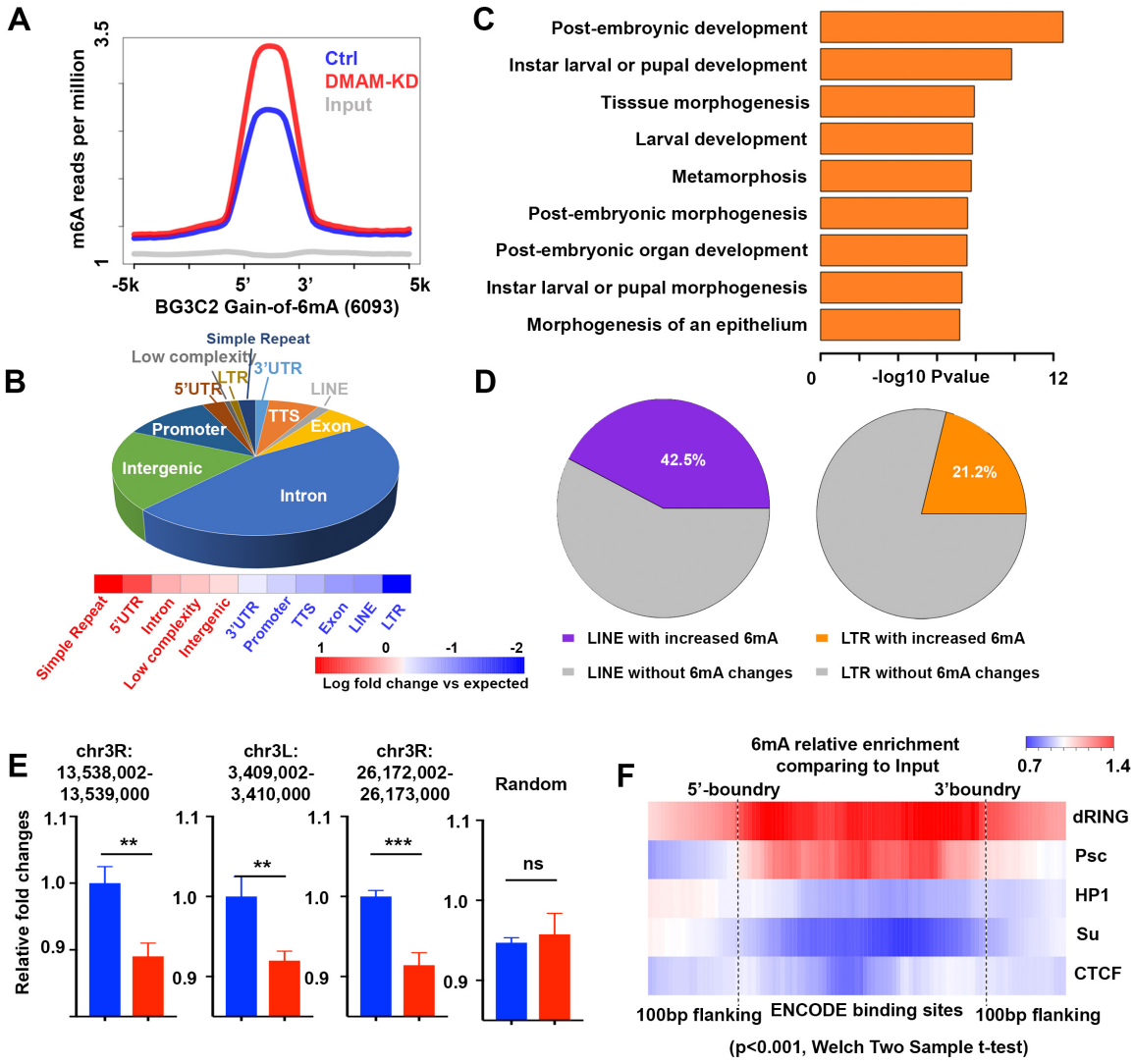

A
Control
DMAD-KD-1
DMAD-KD-2
Number of suvived flies
150
100
50
0
Development
Embryo
1st Instar Larvae
3rd Instar Larvae
Pupae
Adult
B
DMAD mRNA levels
Relative fold change
1.5
1
0.5
0
**
**
Control
DMAD-KD-1
DMAD-KD-2
C
Normal
α
γ
β
I
1 lobe mis-oriented
II
2 lobe mis-oriented
III
% Phenotype
100
80
60
40
20
0
I
II
III
D
Wildtype
DMAD-null
DMAD
Tubulin
E
6mA in vitro
demethylation assay
Rep 1-3
Rep 1-3
dsDNA
no DMAD Ctrl
+ DMAD
6mA dot blot
F
Relative dot blot intensity
1.5
1.0
0.5
0
**
no DMAD Ctrl
+DMAD
G
HPLC quantification
DNA 6mA and RNA 5hmrC
Level of base pair modifications per 10^6
150
100
50
0
**
Control
DMAD-depletion
6mA in fly brains
5hmrC in fly brains
5hmrC in S2 cells
H
HPLC quantification
RNA 5mC
Level of base pair modifications per 10^6
800
600
400
200
0
ns
Control
DMAD-depletion
I
ctl
AlkB-KD
10ng
20ng
40ng
fly brains
J
6mA demethylation assay
percentage of 6mA demethylation
30
20
10
0
***
DMAD-CD
Mut
ALKB
K
DMAD-Null.Brain.6mA.reads
300
250
200
150
100
2622
1873
100
150
200
250
300
Control.Brain.6mA.reads
L
Ctrl
DMAD-null
Input
m6A reads per million
7
1
-5k
5'
3'
5k
Brain. gain-of-6mA (5340)
p<0.001,Welch Two Sample t-test
M
Morphogenesis of an epithelium
Imaginal disc development
Epithelium development
Tissue morphogenesis
Cellular component morphogenesis
Post-embryonic development
Cell morphogenesis
Post-embryonic organ development
Metamorphosis
0
8
-log10 Pvalue
N
Gain-of-6mA Motifs
Log P pvalue
pred. binding
-3.876e+02
Aef1
-1.028e+02
Adf1

Figure S1 (Related to Figure 1)

(A) Survival rate of control and two independent DMAD knockdown lines with Tub-Gal4 driver were indicated. Depletion of DMAD resulted in massive lethality in Pupae and Adult stages. **(B)** DMAD qRT-PCR confirmed effective knockdown of DMAD in fly brains. **(C)** Lobe mis-orientation in Mushroom Bodies (MB) of DMAD neuronal KD (DMAD-nKD) flies (n=48) was analyzed by whole-mount Fas II staining of fixed brains. The severity of lobe mis-orientation was blindly scored as normal, one lobe mis-oriented, or two lobes mis-oriented. Representative examples of each category were shown. All control flies (n=36) display normal MB. **(D)** Western blot demonstrated an effective depletion of DMAD in DMAD-null fly brains. **(E)** Double-stranded control and 6mA-modified DNA oligos were either incubated with control reaction buffer without DMAD, or with recombinant DMAD C-terminal catalytic domain (aa 1657-2860) in triplicate to test the direct 6mA demethylation activities of DMAD *in vitro*. 6mA dot blots demonstrated effective and significant DMAD demethylation activity. **(F)** ImageJ quantification of dot blots in (A). Welch Two Sample t-tests were performed. **, $p<0.01$. **(G)** High resolution HPLC quantification of 6mA and 5hmrC in *Drosophila* brains and S2 cells in the presence (control) and absence of DMAD (DMAD-depletion). The absolute level of 6mA or 5hmrC are indicated as percentage per million nucleotide (ppm) (n=2 for 6mA and n=3 for 5hmrC). **(H)** HPLC quantification of RNA 5mC in control and DMAD-null fly brains. No significant changes were found. **(I)** 6mA dot blots showed no changes in the absence of AlkB gene from fly brains. **(J)** *In vitro* conversion assays were performed using purified DMAD catalytic domain (DMAD-CD, aa 1657-2918); DMAD catalytic mutant (DMAD-CD-mut), in which two residues, H1948 and D1950 were mutated to Y and A, respectively. Purified bacterial 6mA demethylase ALKB served as positive control. **(K)** Genome-wide 6mA normalized reads in control and DMAD-null fly brains were counted in a 10kb binned *Drosophila* genome (1000-pooled brains per condition). Bins with more than 80 reads are shown in color. 2622 bins (red) showed more 6mA reads in DMAD-KO than control whereas 1873 bins (blue) contained fewer 6mA reads in DMAD-null than control. **(L)** 5340 dynamic gain-of-6mA regions (active 6mA demethylation sites in control brains) in the absence of DMAD were identified. Each gain-of-6mA region was further separated into 100bp bins, and normalized 6mA mapped reads were calculated in the presence or absence of DMAD to confirm that these regions contained fewer 6mA reads in DMAD-null brains relative to control. Average 6mA reads per million were plotted. **(M)** Gene Ontology analysis was performed on a subset of upregulated genes in fly brains in Figure 1F (pink). Log2 fold change [(KD:WT) > 0.5] was applied as the threshold cut-off. Several biological processes involving development-related pathways were enriched. **(N)** The HOMER suite was applied to predict common motifs in gain-of-6mA regions. Several transcription factors and epigenetic modulators related to Polycomb repressive complex were indicated (Orsi et al., 2014).


A
m6A reads per million
3.5
1
Ctrl
DMAM-KD
Input
-5k
5'
3'
5k
BG3C2 Gain-of-6mA (6093)
B
Simple Repeat
Low complexity
LTR
3'UTR
5'UTR
LINE
TTS
Exon
Promoter
Intergenic
Intron
Simple Repeat
5'UTR
Intron
Low complexity
Intergenic
3'UTR
Promoter
TTS
Exon
LINE
LTR
1
0
-1
-2
Log fold change vs expected
C
Post-embroynic development
Instar larval or pupal development
Tisssue morphogenesis
Larval development
Metamorphosis
Post-embryonic morphogenesis
Post-embryonic organ development
Instar larval or pupal morphogenesis
Morphogenesis of an epithelium
0
-log10 Pvalue
12
D
42.5%
21.2%
LINE with increased 6mA
LINE without 6mA changes
LTR with increased 6mA
LTR without 6mA changes

E
chr3R:
13,538,002-
13,539,000
chr3L:
3,409,002-
3,410,000
chr3R:
26,172,002-
26,173,000
Random
Relative fold changes
1.1
1.0
0.9
**
**
***
ns

F
6mA relative enrichment
comparing to Input
0.7
1.4
5'-boundry
3'boundry
dRING
Psc
HP1
Su
CTCF
100bp flanking
ENCODE binding sites
100bp flanking
(p<0.001, Welch Two Sample t-test)

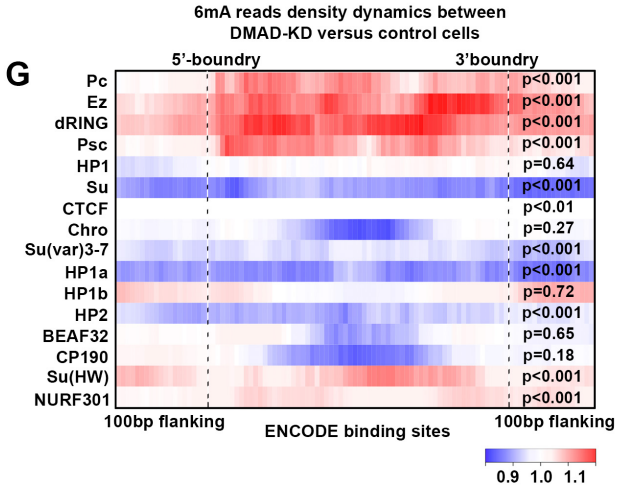
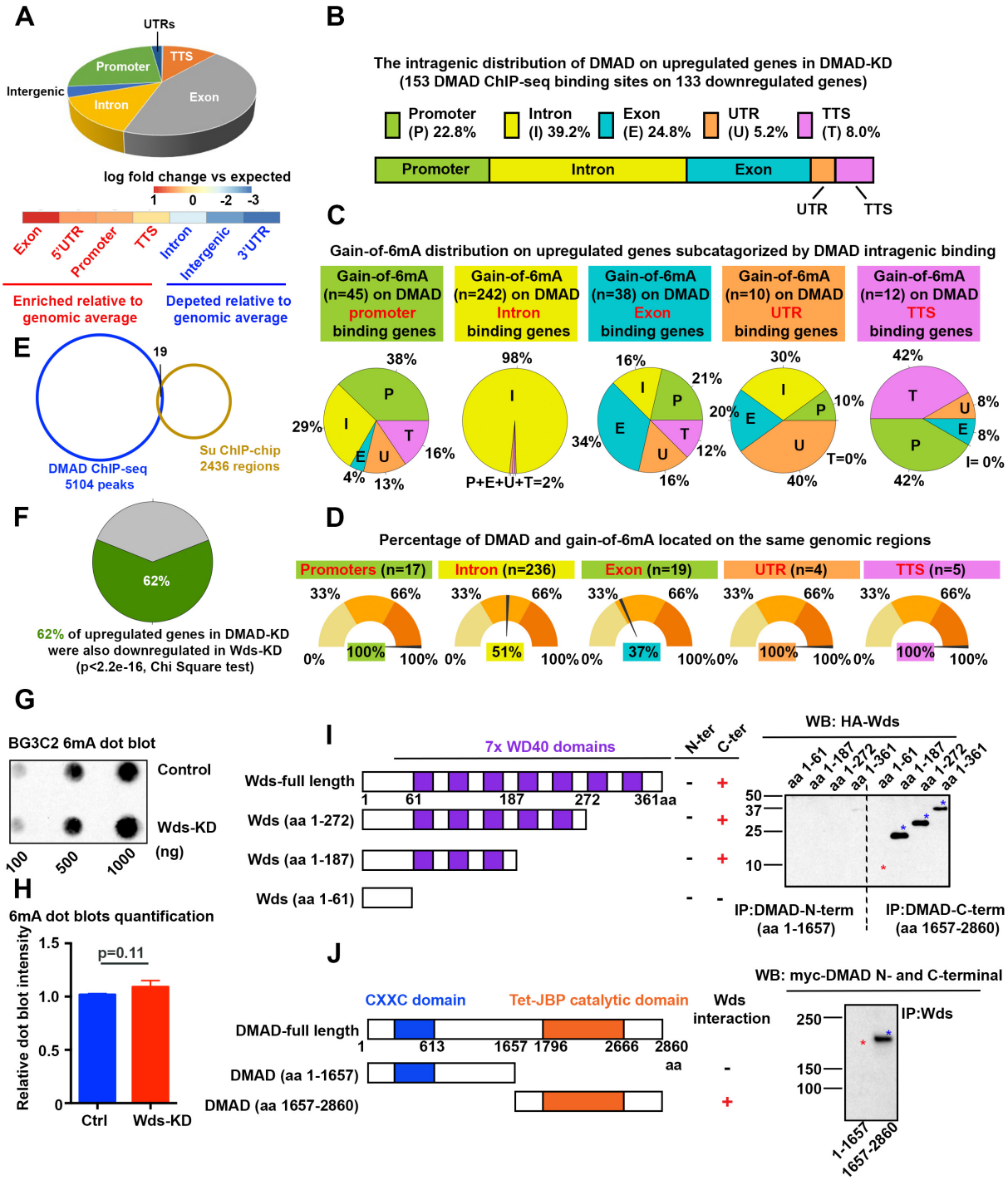

6mA reads density dynamics between
DMAD-KD versus control cells
G
5'-boundry
3'boundry
Pc
Ez
dRING
Psc
HP1
Su
CTCF
Chro
Su(var)3-7
HP1a
HP1b
HP2
BEAF32
CP190
Su(HW)
NURF301
p<0.001
p<0.001
p<0.001
p<0.001
p=0.64
p<0.001
p<0.01
p=0.27
p<0.001
p<0.001
p=0.72
p<0.001
p=0.65
p=0.18
p<0.001
p<0.001
100bp flanking
ENCODE binding sites
100bp flanking
0.9
1.0
1.1

Figure S2 (Related to Figure 2)

(A) 6093 dynamic gain-of-6mA regions (active 6mA demethylation sites in control brains) in BG3C2 cells were identified. Each gain-of-6mA region was further separated into 100bp bins, and normalized 6mA mapped reads were calculated in the presence or absence of DMAD to confirm the increase of 6mA on these regions when DMAD was depleted. Average 6mA reads per million were plotted. P<0.01, Welch Two sample t-tests**. (B)** Genomic annotation of gain-of-6mA regions in BG3C2 cells revealed intragenic characterization, with many of these regions aligning to introns, exons, promoters, untranslated regions (UTRs) and transcription termination sites (TTS). 6mA dynamics on introns, UTRs and some repetitive elements were higher than expected values. The enrichment of each genomic features versus expected values are indicated by heatmap. **(C)** Gene Ontology analysis was performed on up-regulated genes carrying gain-of-6mA regions in the absence of DMAD. Log2 fold change [(KD:WT) > 0.5] was applied as the threshold cut-off. GO terms were enriched for several development-related pathways. **(D)** Transposon expression was altered with DMAD depletion. 42.5% and 21.2% of neuronal-expressing LINEs and LTRs correlated with increase of 6mA when DMAD was knocked down. **(E)** 6mA-sensitive restriction enzyme DpnI that preferentially cleaves methylated adenosine at GATC/CATC/GATG sites was used to validate regions identified by 6mA-IP in BG3C2 cells. The percentages of 6mA in either control or DMAD-KD were assessed by qPCR amplification of digested DNA using undigested DNA as a control (digested/undigested). Inverse correlation between 6mA reads distribution and qPCR fold changes were observed to confirm the 6mA dynamic changes in these regions. p values were indicated (unpaired *t*-test). **(F)** Average fold change in 6mA mapped reads versus non-enriched input DNA was calculated for various binned ChIP-chip regions of epigenetic regulators available from the modENCODE database and plotted in Heatmap view. Red (fold change >1) indicates enrichment over input whereas blue (fold change <1) indicates depletion over input. 6mA was specifically enriched at Polycomb protein binding sites including dRING and Psc. Enrichment and depletion were significant with p-value < 0.001, Welch Two sample t-tests. **(G)** Average fold change in 6mA mapped reads of DMAD-KD over control were calculated for various binned ChIP-chip regions of epigenetic regulators available from the modENCODE database and plotted in Heatmap view. Red (fold change >1) indicates enrichment over control whereas blue (fold change <1) indicates depletion over control. Knockdown of DMAD led to further 6mA accumulation primarily on Polycomb binding sites. p-values were indicated, Welch Two sample t-tests.


A
UTRs
TTS
Promoter
Intergenic
Intron
Exon
log fold change vs expected
1 0 -2 -3
Exon 5'UTR Promoter TTS Intron Intergenic 3'UTR
Enriched relative to genomic average
Depeted relative to genomic average
B
The intragenic distribution of DMAD on upregulated genes in DMAD-KD
(153 DMAD ChIP-seq binding sites on 133 downregulated genes)
Promoter (P) 22.8%
Intron (I) 39.2%
Exon (E) 24.8%
UTR (U) 5.2%
TTS (T) 8.0%
Promoter Intron Exon UTR TTS
C
Gain-of-6mA distribution on upregulated genes subcatagorized by DMAD intragenic binding
Gain-of-6mA (n=45) on DMAD promoter binding genes
Gain-of-6mA (n=242) on DMAD Intron binding genes
Gain-of-6mA (n=38) on DMAD Exon binding genes
Gain-of-6mA (n=10) on DMAD UTR binding genes
Gain-of-6mA (n=12) on DMAD TTS binding genes
38% 29% 16% 4% 13%
98% P+E+U+T=2%
16% 21% 34% 12% 16%
30% 20% 10% T=0% 40%
42% 8% 8% I= 0% 42%
E
19
DMAD ChIP-seq 5104 peaks
Su ChIP-chip 2436 regions
F
62%
62% of upregulated genes in DMAD-KD were also downregulated in Wds-KD (p<2.2e-16, Chi Square test)
D
Percentage of DMAD and gain-of-6mA located on the same genomic regions
Promoters (n=17)
Intron (n=236)
Exon (n=19)
UTR (n=4)
TTS (n=5)
33% 66% 0% 100% 100%
33% 66% 0% 51% 100%
33% 66% 0% 37% 100%
33% 66% 0% 100% 100%
33% 66% 0% 100% 100%
G
BG3C2 6mA dot blot
Control
Wds-KD
100 500 1000 (ng)
H
6mA dot blots quantification
Relative dot blot intensity
1.5 1.0 0.5 0
p=0.11
Ctrl Wds-KD
I
7x WD40 domains
N-ter C-ter
Wds-full length 1 61 187 272 361aa - +
Wds (aa 1-272) - +
Wds (aa 1-187) - +
Wds (aa 1-61) - -
WB: HA-Wds
aa 1-61 aa 1-187 aa 1-272 aa 1-361 aa 1-61 aa 1-187 aa 1-272 aa 1-361
50 37 25 10
IP:DMAD-N-term (aa 1-1657)
IP:DMAD-C-term (aa 1657-2860)
J
CXXC domain
Tet-JBP catalytic domain
Wds interaction
DMAD-full length 1 613 1657 1796 2666 2860 aa -
DMAD (aa 1-1657) -
DMAD (aa 1657-2860) +
WB: myc-DMAD N- and C-terminal
250 150 100
IP:Wds
1-1657 1657-2860

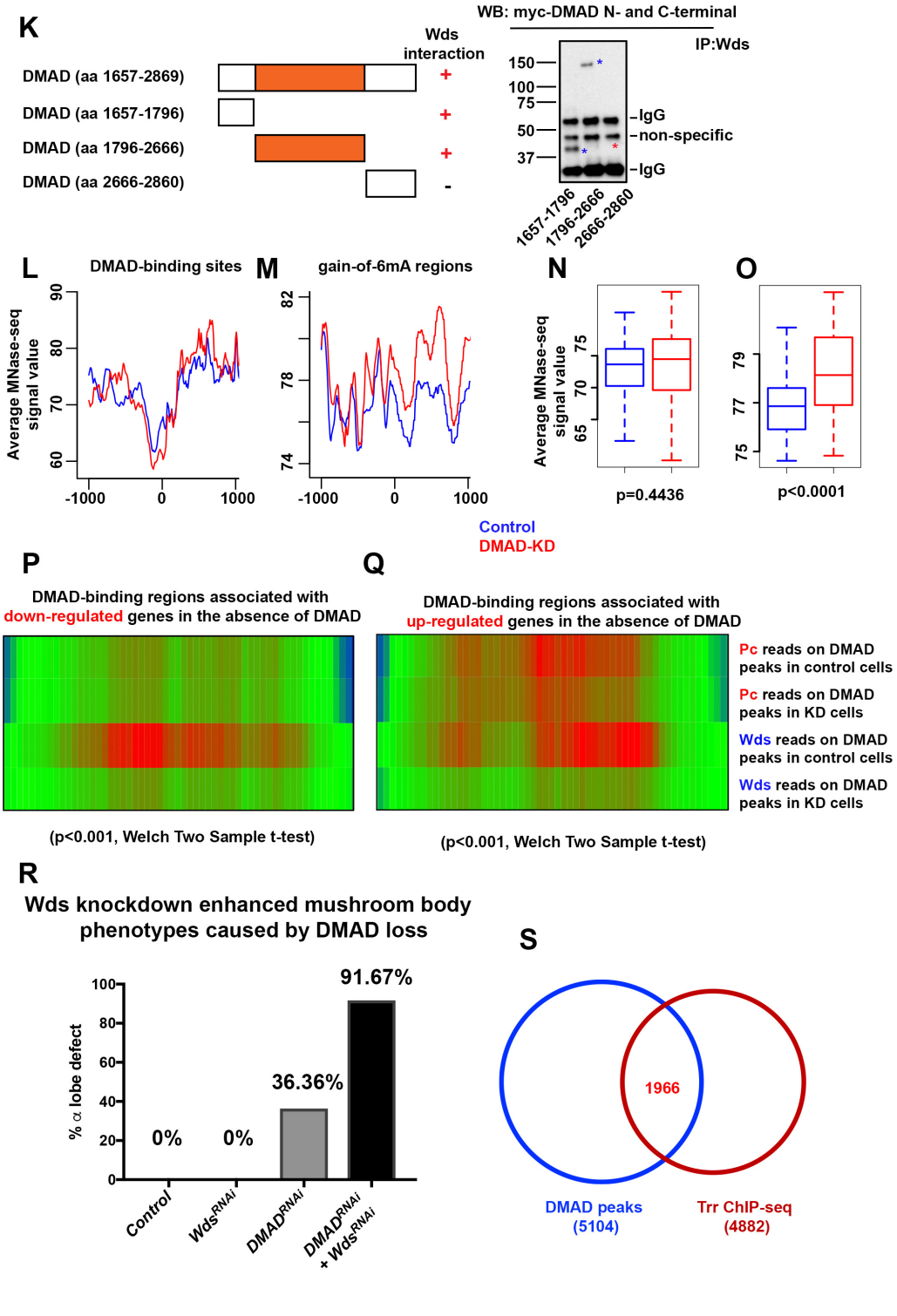

K
WB: myc-DMAD N- and C-terminal
IP:Wds
Wds interaction
DMAD (aa 1657-2869)
DMAD (aa 1657-1796)
DMAD (aa 1796-2666)
DMAD (aa 2666-2860)
+
+
+
-
150
100
75
50
37
IgG
non-specific
IgG
1657-1796
1796-2666
2666-2860
L
DMAD-binding sites
Average MNase-seq signal value
M
gain-of-6mA regions
N
O
p=0.4436
p<0.0001
Control
DMAD-KD
P
DMAD-binding regions associated with down-regulated genes in the absence of DMAD
(p<0.001, Welch Two Sample t-test)
Q
DMAD-binding regions associated with up-regulated genes in the absence of DMAD
(p<0.001, Welch Two Sample t-test)
Pc reads on DMAD peaks in control cells
Pc reads on DMAD peaks in KD cells
Wds reads on DMAD peaks in control cells
Wds reads on DMAD peaks in KD cells
R
Wds knockdown enhanced mushroom body phenotypes caused by DMAD loss
% α lobe defect
0%
0%
36.36%
91.67%
Control
Wds^RNAi
DMAD^RNAi
DMAD^RNAi + Wds^RNAi
S
1966
DMAD peaks (5104)
Trr ChIP-seq (4882)

Figure S3 (Related to Figure 3)

(A) Genomic annotation of overall DMAD ChIP-seq peaks in BG3C2 cells revealed a strong preference for intragenic binding. **(B)** Intragenic distributions of DMAD on DMAD-bound upregulated genes bearing accumulation of 6mA upon DMAD knockdown was demonstrated proportionally. DMAD showed strong intronic enrichment on these genes. (**C**) The genes in (B) were further subcategorized based on the DMAD intragenic association. Gain-of-6mA distributions on each subset of genes were calculated, and the percentages were indicated by pie chart. The DMAD and gain-of-6mA regions largely coordinated on the same genomic features. 95% of gain-of-6mA were found in introns of genes with DMAD also binding to introns. (**D**) The percentage of genes with both DMAD and gain-of-6mA binding to the same genomic feature were calculated. When DMAD bound to introns and UTRs, 76% and 82% of gain-of-6mA were found on the exact same introns and UTRs, respectively. Only 43% of exons showed simultaneous occupancy by both DMAD and gain-of-6mA. **(E)** Venn diagram shows little overlap between DMAD and Su(var)3-7. (**F**) Upregulated genes in the absence of DMAD showed significant overlapping with upregulated genes in the absence of Wds. Chi-square tests were performed. RNA-seq were performed in triplicates. RPKM fold changes >0.1 were included. (**G**) Dot blots using an antibody specific for 6mA showed no 6mA changes in the absence of Wds from BG3C2 cells. (**H**) ImageJ quantification of dot blots in (G). Welch Two Sample t-tests were performed. **(I)** A series of Myc-tagged Wds deletion constructs were generated to determine the importance of WD40 domains to bind to DMAD. Co-IP experiments using HA tagged N- or C-terminal DMAD truncations suggested WD40 domains were critical to facilitate Wds interacting with DMAD, as complete removal of WD40 domains (Wds aa 1-61) abolished its binding to DMAD (red asterisk). Full length and Wds truncation constructs bearing WD40 domains (blue asterisk) interacted with C-terminal truncation of DMAD (aa 1657-2860). **(J)** Co-IP experiments with Myc tagged full length Wds confirmed the C-terminal DMAD containing 6mA catalytic domain (blue asterisk) interacted with Wds. **(K)** Fine mapping of DMAD C-terminal domains indicated the DMAD aa1657-1796 and catalytic domain aa 1796-2666 (blue asterisk) were responsible to interact with Wds. DMAD (aa 2666-2860) did not interact with Wds (red asterisk). **(L-O)** Nucleosome positions on intragenic DMAD binding sites and gain-of-6mA regions in the presence and absence of DMAD were analyzed by MNase-seq and plotted by Dynamic Analysis of Nucleosome Position and Occupancy by Sequencing (DANPOS) program. While no significant changes of nucleosome occupancy were observed on DMAD-binding sites **(L and N)**, the nucleosome occupancy was significantly enriched on the gain-of-6mA regions in the absence of DMAD **(M and O)** ($p<0.001$, Welch Two Sample t-tests). **(P)** Average fold change in Pc and Wds reads of control and DMAD-KD over non-enriched input DNA were calculated at DMAD-binding sites of down-regulated genes shown in Fig. 4A and 4B. Wds but not Pc was enriched at these regions, suggesting a dominant role for Wds in maintaining expression of these genes. Enrichment was significant over input ($p<0.001$, Welch Two Sample t-tests). **(Q)** Average fold change in Pc and Wds reads of control and DMAD-KD over non-enriched input DNA were calculated at DMAD-binding sites on up-regulated genes in Fig. S4A and S4B. Both Pc and Wds occupied these genes in control cells and were reduced in the absence of DMAD, suggesting that Pc and Wds may coordinate to modulate these genes. Enrichment was significant over input ($p<0.001$, Welch Two Sample t-tests). **(R)** Penetrance of α lobe defects in the control (*elav-Gal4*), pan-neuron Wds RNAi (*elav-Gal4,UAS-Wds^{RNAi}*), pan-neuron DMAD RNAi (*elav-Gal4,UAS-DMAD^{RNAi}*), or pan-neuron Wds RNAi + DMAD RNAi (*elav-Gal4,UAS-Wds^{RNAi},UAS-DMAD^{RNAi}*) flies. N=14, 10, 22 and 12,

respectively. **(S)** Trithorax-related (Trr) ChIP-seq in S2 cells were obtained from (Herz et al., 2012). 1966 out of 4882 Trr binding sites (40%) are also overlapping with DMAD binding, suggesting a potential role of DMAD in enhancer activities.

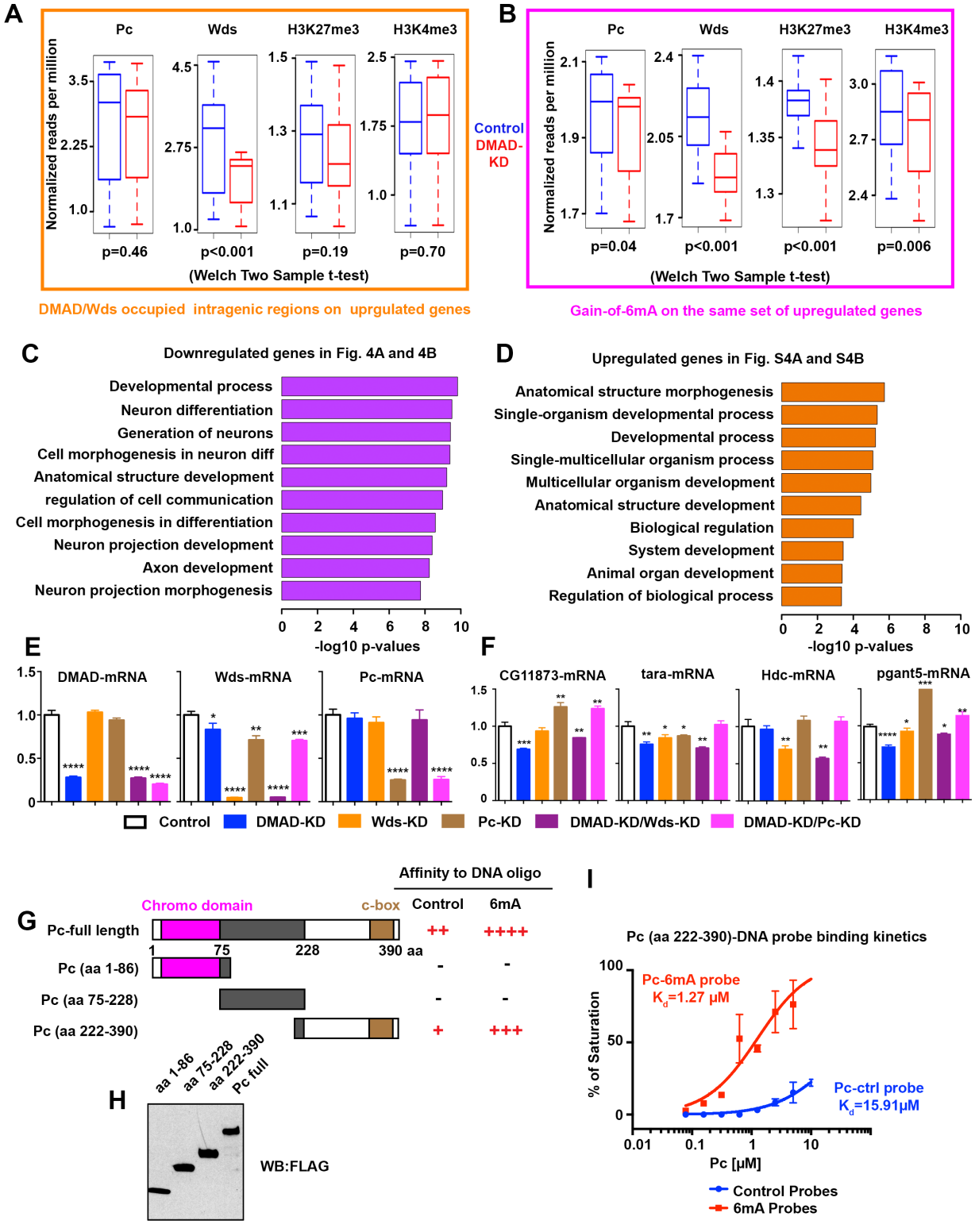

A
Pc
Wds
H3K27me3
H3K4me3
Normalized reads per million
p=0.46
p<0.001
p=0.19
p=0.70
(Welch Two Sample t-test)
DMAD/Wds occupied intragenic regions on uprgulated genes
Control
DMAD-KD
B
p=0.04
p<0.001
p<0.001
p=0.006
(Welch Two Sample t-test)
Gain-of-6mA on the same set of upregulated genes
C
Downregulated genes in Fig. 4A and 4B
Developmental process
Neuron differentiation
Generation of neurons
Cell morphogenesis in neuron diff
Anatomical structure development
regulation of cell communication
Cell morphogenesis in differentiation
Neuron projection development
Axon development
Neuron projection morphogenesis
-log10 p-values
D
Upregulated genes in Fig. S4A and S4B
Anatomical structure morphogenesis
Single-organism developmental process
Developmental process
Single-multicellular organism process
Multicellular organism development
Anatomical structure development
Biological regulation
System development
Animal organ development
Regulation of biological process
-log10 p-values
E
DMAD-mRNA
Wds-mRNA
Pc-mRNA
F
CG11873-mRNA
tara-mRNA
Hdc-mRNA
pgant5-mRNA
Control
DMAD-KD
Wds-KD
Pc-KD
DMAD-KD/Wds-KD
DMAD-KD/Pc-KD
G
Affinity to DNA oligo
Chromo domain
c-box
Control
6mA
Pc-full length
++
++++
1
75
228
390 aa
Pc (aa 1-86)
-
-
Pc (aa 75-228)
-
-
Pc (aa 222-390)
+
+++
H
aa 1-86
aa 75-228
aa 222-390
Pc full
WB:FLAG
I
Pc (aa 222-390)-DNA probe binding kinetics
% of Saturation
Pc-6mA probe
K_d=1.27 μM
Pc-ctrl probe
K_d=15.91μM
Pc [μM]
Control Probes
6mA Probes

Figure S4 (Related to Figure 4)

(A) Average normalized reads (per million) dynamics in Pc, Wds, H3K27me3 and H3K4me3 at DMAD/Wds binding sites in genes upregulated in the absence of DMAD relative to control are shown. Wds signals were significantly reduced when DMAD was depleted, and positively correlated with DMAD changes ($p<0.001$, Welch Two Sample t-tests). Pc, H3K4me3 and H3K27me3 distributions are not significantly altered in DMAD-KD. **(B)** Average normalized reads (per million) dynamics in Pc, Wds, H3K27me3 and H3K4me3 at gain-of-6mA regions of upregulated genes in fig. S13A. Pc significantly reduced on these regions ($p=0.04$, Welch Two Sample t-test), possibly due to the active transcription environment that antagonized its binding to 6mA. **(C)** Gene Ontology analyses were performed on the downregulated genes from Fig. 4A-4B. Several biological processes involved in neurodevelopment and neuronal functions were demonstrated. -log10 p-value of enrichment were plotted. **(D)** Gene Ontology analyses were performed on the upregulated genes from fig. S4A-4B. -log10 p-value of enrichment were plotted. **(E)** Double-stranded RNA (dsRNA) was used to individually knockdown DMAD, Wds and Pc in BG3C2 cells. Combined depletion of DMAD/Wds and DMAD/Pc were also performed. qPCR of DMAD, Wds and Pc mRNA confirmed the effective knockdown of these factors. Pc knockdown induced significant downregulation of Wds. **(F)** Additional loci from Figure 4A and 4B were further tested by qPCR for expression changes in the absence of DMAD, Wds, Pc or combined depletion of DMAD and Pc as well as DMAD and Wds. The expression of these genes shows similar trends when DMAD or Wds are independently removed. Co-knockdown of DMAD and Wds showed a consistent downregulation of these loci. Co-depletion of DMAD and Pc leads to de-repression of these loci (n=3), Welch Two Sample t-tests were performed. *, $p<0.05$; **, $p<0.01$; ***, $p<0.001$; ****, $p<0.0001$. **(G)** A series of Pc deletion constructs with FLAG tag were generated to determine their binding domain(s) and kinetics to control and 6mA-modified DNA. The C-terminal Pc (aa 222-390) was responsible for direct association with both control and 6mA DNA. **(H)** The expression of Pc full length and truncations by Baculovirus-infected SF9 cells demonstrated by Western blot **(I)** Pc (aa 222-390) binding kinetics to control and 6mA-modified probes showed its preferential binding to 6mA-modified DNA probes, as measured by fluorescence polarization assays. Pc (aa 222-390) 6mA probe K_d=1.27μM, and Pc-control probe K_d=15.91 μM. Purified FLAG-Pc (aa 222-390) concentrations range from 0.01 to 10 μM.