Molecular Cell, Volume *70*

Supplemental Information

Capturing the Onset of PRC2-Mediated Repressive Domain Formation

Ozgur Oksuz, Varun Narendra, Chul-Hwan Lee, Nicolas Descostes, Gary LeRoy, Ramya Raviram, Lili Blumenberg, Kelly Karch, Pedro P. Rocha, Benjamin A. Garcia, Jane A. Skok, and Danny Reinberg

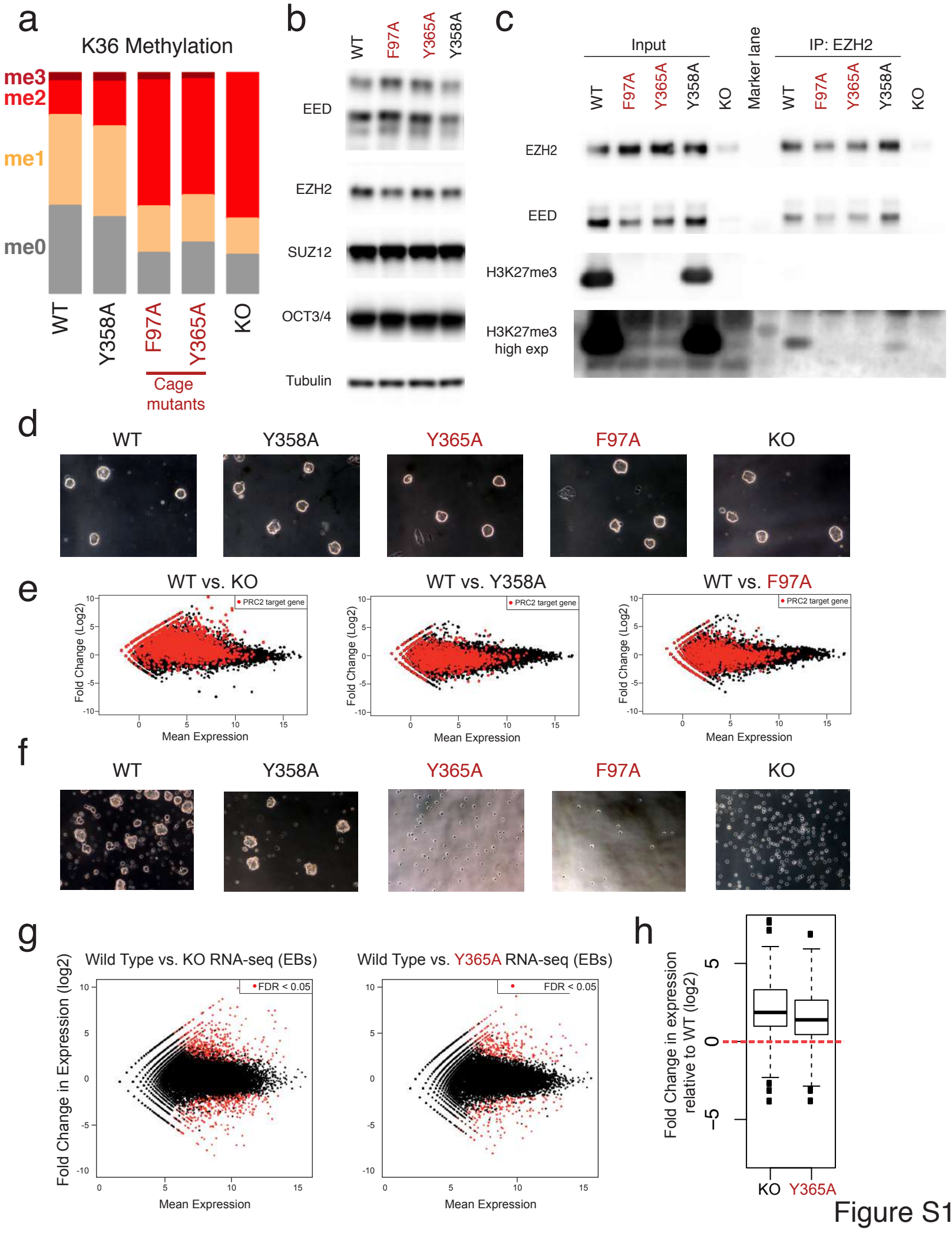

a
K36 Methylation
me3
me2
me1
me0
WT
Y358A
F97A
Y365A
KO
Cage mutants
b
WT
F97A
Y365A
Y358A
EED
EZH2
SUZ12
OCT3/4
Tubulin
c
Input
Marker lane
IP: EZH2
WT
F97A
Y365A
Y358A
KO
EZH2
EED
H3K27me3
H3K27me3 high exp
d
WT
Y358A
Y365A
F97A
KO
e
WT vs. KO
WT vs. Y358A
WT vs. F97A
PRC2 target gene
Fold Change (Log2)
Mean Expression
f
WT
Y358A
Y365A
F97A
KO
g
Wild Type vs. KO RNA-seq (EBs)
Wild Type vs. Y365A RNA-seq (EBs)
FDR < 0.05
Fold Change in Expression (log2)
Mean Expression
h
Fold Change in expression relative to WT (log2)
KO
Y365A

Figure S1

Figure S1: EED cage-mutant mESCs, which have elevated levels of H3K36me2 and intact PRC2 complex, maintained appropriate patterns of gene expression, but failed to differentiate. Related to Figure 1.

a, Methylation status of H3K36 in mESCs of the indicated genotypes as identified by quantitative histone mass-spectrometry. **b,** WB using the indicated antibodies on whole cell extracts from mESCs. **c,** Immunoprecipitation (IP) of EZH2 using nuclear extracts from mESCs. WB was performed on input (left) and IP (right) samples with the antibodies indicated. Cage-mutants are highlighted in red. **d,** Bright-field microscopy images of undifferentiated mESCs comparing the indicated EED genotypes. **e,** MA plots of RNA-seq data comparing mESCs of the indicated EED genotypes. Each dot represents a gene. The x-axis represents its mean abundance across samples, and the y-axis represents the log2 fold enrichment between samples. PRC2 targets are in red. **f,** Bright-field microscopy images of mESCs differentiated into the cardiac lineage. Images were taken 3 days after differentiation. **g,** MA plots of RNA-seq data within 2 days of differentiation into embryoid bodies (EBs) of mESC with the EED genotypes indicated. Differentially expressed genes (FDR<0.05) are in red. **h,** Box plot of genes that were repressed in a WT setting within 2 days of differentiation of mESCs into embryoid bodies (EBs) (*de novo* targets of PRC2). Y-value represents the relative expression of these genes in EED KO and Y365A EBs, relative to the WT background. Error bars represent the SD of the mean.

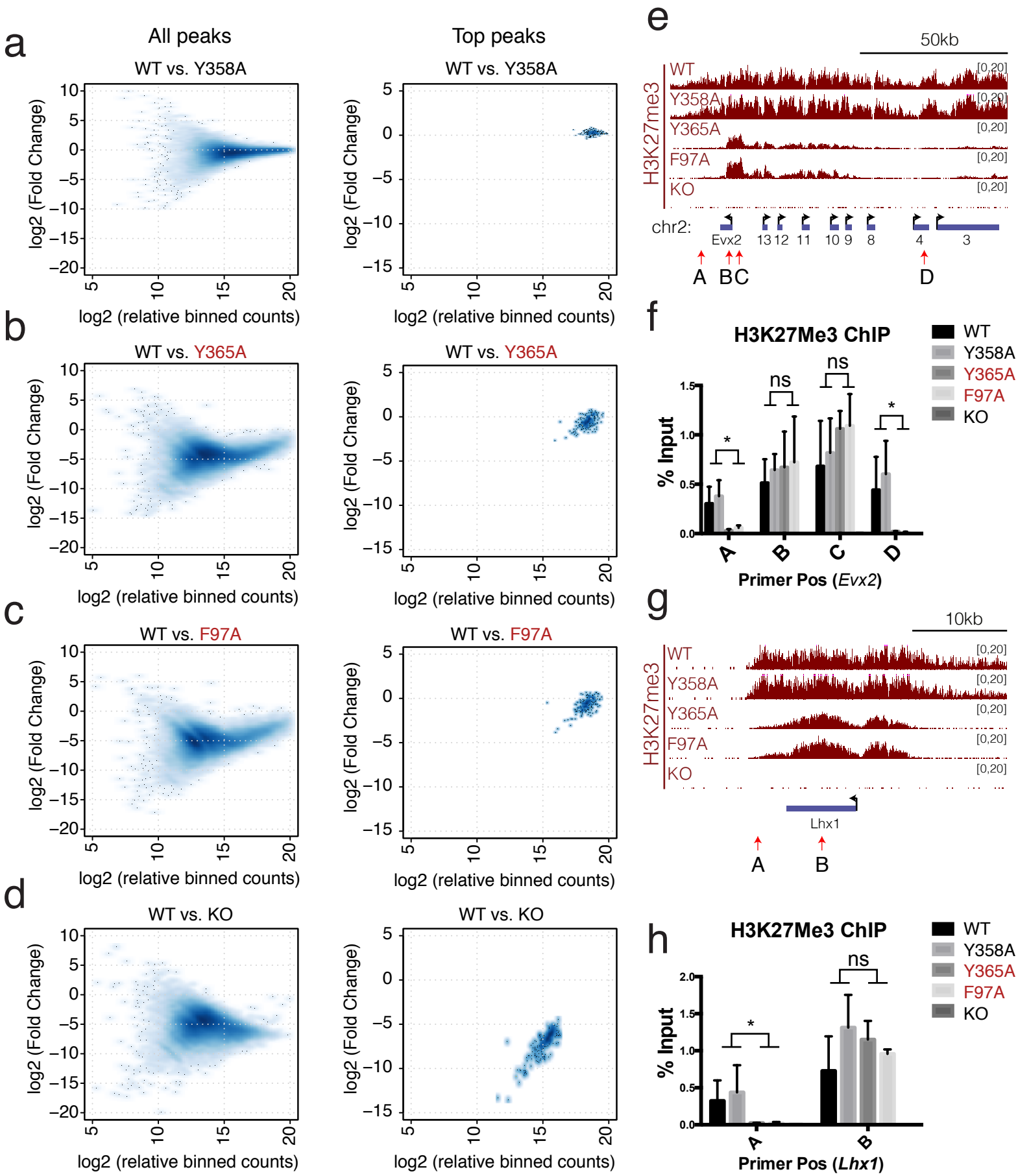

a
All peaks
Top peaks
WT vs. Y358A
WT vs. Y365A
WT vs. F97A
WT vs. KO
log2 (Fold Change)
log2 (relative binned counts)
b
c
d
e
50kb
H3K27me3
WT
Y358A
Y365A
F97A
KO
[0,20]
chr2:
Evx2
13 12 11 10 9 8 4 3
A BC D
f
H3K27Me3 ChIP
% Input
ns
*
Primer Pos (Evx2)
g
10kb
Lhx1
h
Primer Pos (Lhx1)

Figure S2

Figure S2: H3K27me3 levels of the retained peaks in the EED cage-mutants are similar to the WT case, despite widespread loss of this mark in the rest of the genome. Related to Figure 1.

a, b, c, d, MA plots comparing H3K27me3 ChIP-seq intensity over all H3K27me3 peaks found in WT (left) or top peaks in EED cage-mutants (right) from EED WT and Y358A **(a)** Y365A **(b)** F97A **(c)** KO **(d)**. X-axis represents relative binned mean counts transformed onto log2 scale. Y-axis represents log2 fold change of the mean counts of the peaks from the indicated genotypes. **e, g,** ChIP-seq tracks for H3K27me3 along the *Evx2* **(e)** and *Lhx1* **(g)** gene in mESCs with the indicated EED genotypes. Positions of primers used in ChIP-qPCR are indicated by red arrows. **f, h,** H3K27me3 ChIP-qPCR using the primers from **e** and **g**, respectively. Enrichments were calculated as % of input. All ChIP-qPCR experiments were performed in two biological replicates. One sided paired t-test: *p-value<0.05, ns: not significant. Error bars represent the SD of the mean.

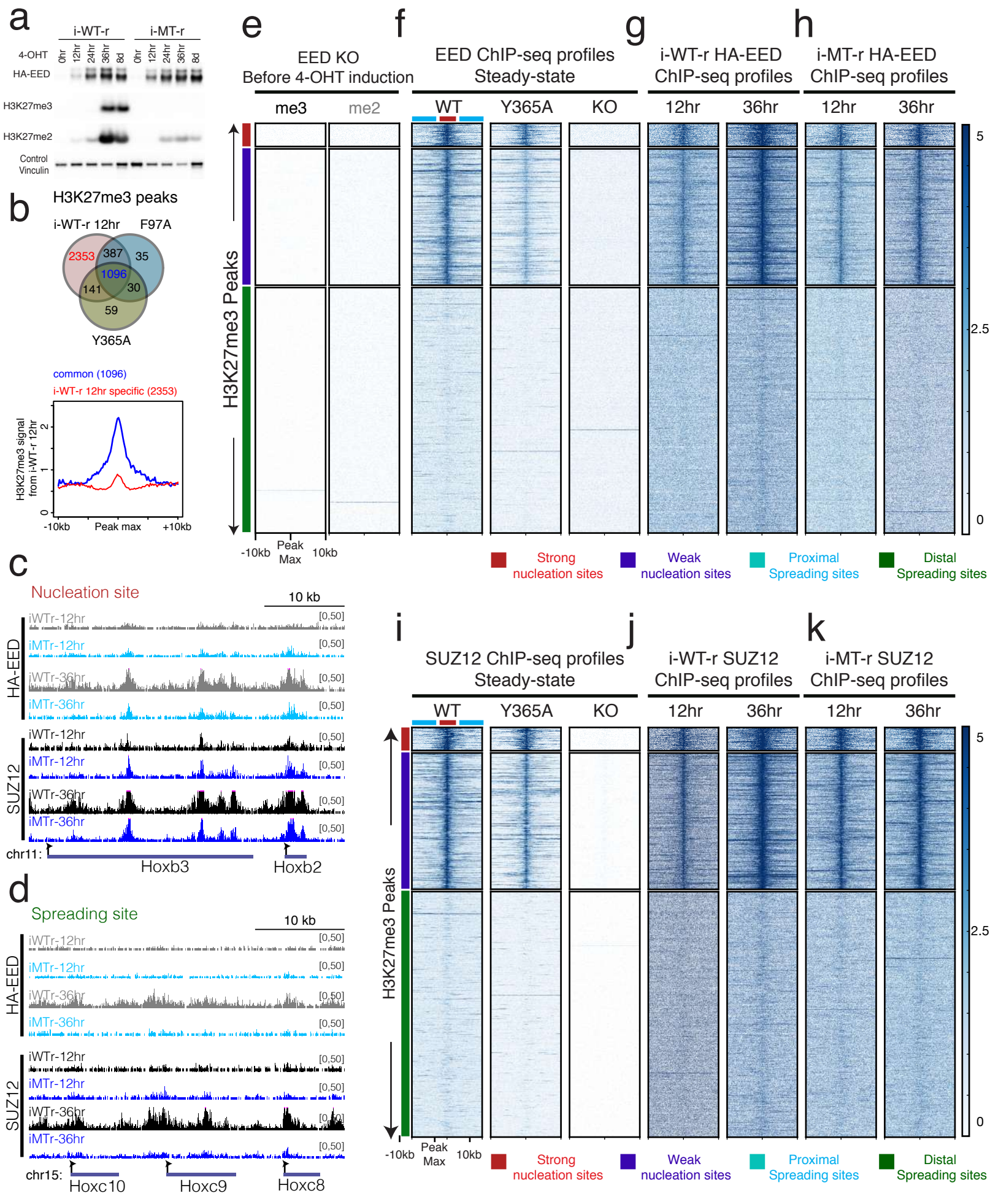

a
i-WT-r
i-MT-r
4-OHT
0hr
12hr
24hr
36hr
8d
HA-EED
H3K27me3
H3K27me2
Control
Vinculin
b
H3K27me3 peaks
i-WT-r 12hr
F97A
Y365A
2353
387
35
1096
141
30
59
common (1096)
i-WT-r 12hr specific (2353)
H3K27me3 signal from i-WT-r 12hr
-10kb
Peak max
+10kb
c
Nucleation site
10 kb
HA-EED
SUZ12
iWTr-12hr
iMTr-12hr
iWTr-36hr
iMTr-36hr
[0,50]
chr11:
Hoxb3
Hoxb2
d
Spreading site
chr15:
Hoxc10
Hoxc9
Hoxc8
e
EED KO
Before 4-OHT induction
me3
me2
f
EED ChIP-seq profiles
Steady-state
WT
Y365A
KO
g
i-WT-r HA-EED
ChIP-seq profiles
h
i-MT-r HA-EED
ChIP-seq profiles
12hr
36hr
H3K27me3 Peaks
-10kb
Peak
Max
10kb
5
2.5
0
Strong
nucleation sites
Weak
nucleation sites
Proximal
Spreading sites
Distal
Spreading sites
i
SUZ12 ChIP-seq profiles
Steady-state
j
i-WT-r SUZ12
ChIP-seq profiles
k
i-MT-r SUZ12
ChIP-seq profiles

Figure S3

Figure S3: Initial deposition of H3K27me3 and PRC2 in i-WT-r and i-MT-r mESCs strongly overlapped with those sites that retained H3K27me3 and PRC2 in the steady-state EED cage-mutants. Related to Figure 1 and 2.

a, Validation of the i-WT-r and i-MT-r systems. WB using indicated antibodies on whole cell extracts after 4-OHT treatment to induce expression of EED, in i-WT-r or i-MT-r, at the time points indicated. **b, top,** Venn diagram showing the overlap of H3K27me3 peaks between the 12 hr i-WT-r, Y365A and F97A mESCs. **b, bottom,** average H3K27me3 density profiles of 1096 common peaks, and 2353 i-WT-r 12 hr-specific peaks indicated in Venn diagram above, in i-WT-r 12 hr cells. The peaks that are present in i-WT-r at 12 hr, but not in Y365A or F97A are due to minor spreading of H3K27me3 at this time point, as their intensity is much lower than that of the common peaks in i-WT-r at 12 hr. **c**, **d**, ChIP-seq tracks for HA-EED and SUZ12 along the *HoxB* cluster (nucleation sites) **(c)**, or the *HoxC* cluster (spreading sites) **(d)** in i-WT-r and i-MT-r cells after 4-OHT treatment at the time points indicated. **e**, Heatmaps of H3K27me3 (me3) and H3K27me2 (me2) ChIP-seq performed on the EED KO cells before induction of EED expression by 4-OHT treatment. **f,** Heatmaps of EED ChIP-seq performed on the indicated EED genotypes at steady-state. **g, h**, Heatmaps of HA-EED ChIP-seq upon EED rescue, either WT **(g)** or cage-mutant **(h)** by 4-OHT treatment for the specified time points. **i**, **j**, **k**, Heatmaps of SUZ12 ChIP-seq performed on the indicated EED genotypes at steady-state **(i)** or upon rescue with EED, either WT **(j)** or cage-mutant **(k)** with 4-OHT at the specified time points. (Parameters in Fig. 2c are used to generate the Heatmaps). (Scales: 0-5 for EED, HA-EED and SUZ12).

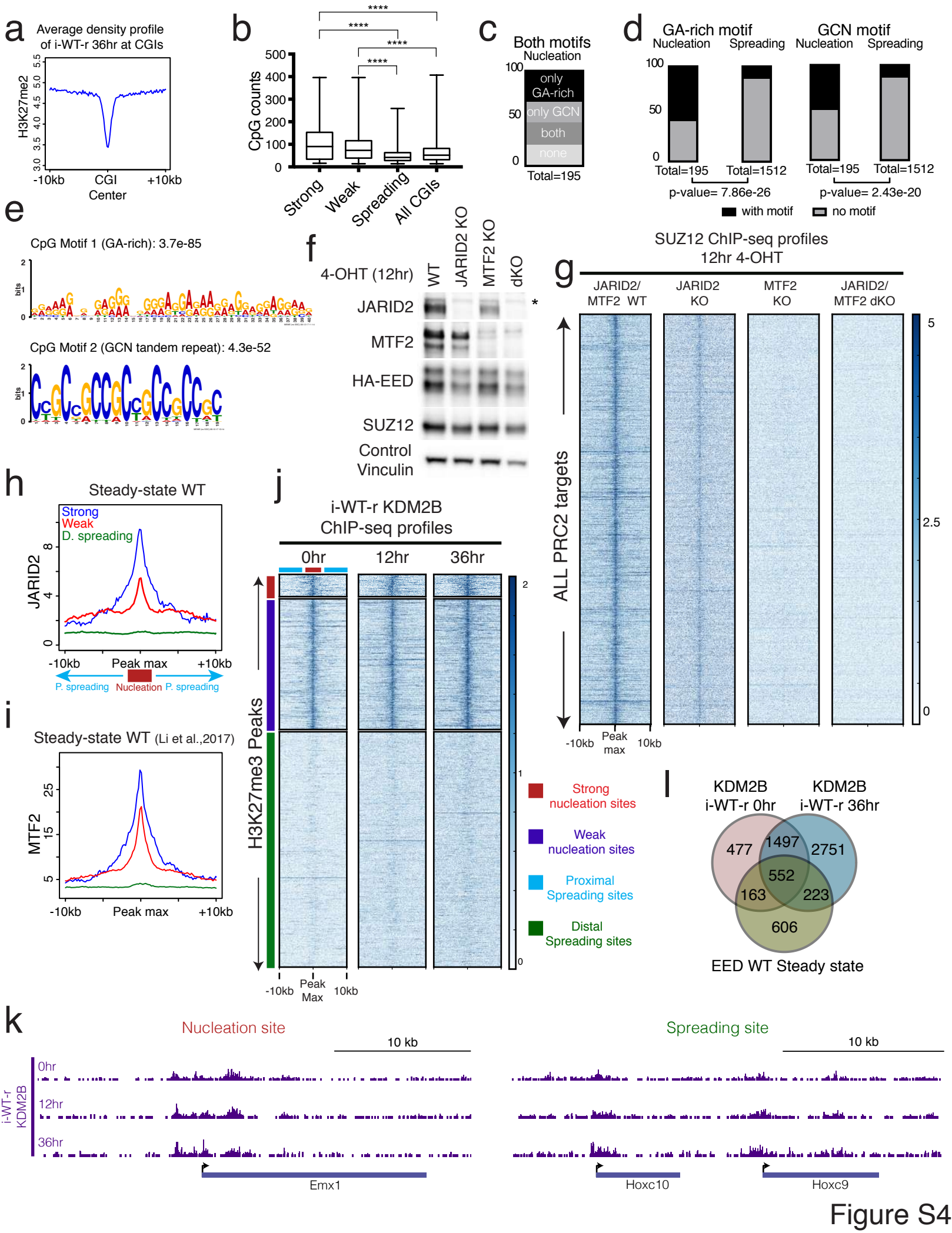

a
Average density profile of i-WT-r 36hr at CGIs
H3K27me2
-10kb
CGI Center
+10kb
b
CpG counts
Strong
Weak
Spreading
All CGIs
****
c
Both motifs
Nucleation
only GA-rich
only GCN
both
none
Total=195
d
GA-rich motif
GCN motif
Nucleation
Spreading
Total=195
Total=1512
p-value= 7.86e-26
p-value= 2.43e-20
with motif
no motif
e
CpG Motif 1 (GA-rich): 3.7e-85
CpG Motif 2 (GCN tandem repeat): 4.3e-52
bits
f
4-OHT (12hr)
WT
JARID2 KO
MTF2 KO
dKO
JARID2
MTF2
HA-EED
SUZ12
Control Vinculin
g
SUZ12 ChIP-seq profiles
12hr 4-OHT
JARID2/ MTF2 WT
JARID2 KO
MTF2 KO
JARID2/ MTF2 dKO
ALL PRC2 targets
-10kb
Peak max
10kb
h
Steady-state WT
Strong
Weak
D. spreading
JARID2
-10kb
Peak max
+10kb
P. spreading
Nucleation
P. spreading
i
Steady-state WT (Li et al.,2017)
MTF2
-10kb
Peak max
+10kb
j
i-WT-r KDM2B ChIP-seq profiles
0hr
12hr
36hr
H3K27me3 Peaks
-10kb
Peak Max
10kb
Strong nucleation sites
Weak nucleation sites
Proximal Spreading sites
Distal Spreading sites
l
KDM2B i-WT-r 0hr
KDM2B i-WT-r 36hr
477
1497
2751
552
163
223
606
EED WT Steady state
k
Nucleation site
Spreading site
10 kb
i-WT-r KDM2B
0hr
12hr
36hr
Emx1
Hoxc10
Hoxc9

Figure S4

Figure S4: Features associated with nucleation versus spreading site CGIs. Related to Figure 1, 2 and 4.

a, Average H3K27me2 density profile of the i-WT-r at 36 hr within CGIs. **b**, Box-plots showing distribution of CpG counts of CGIs at the nucleation sites (strong versus weak) as well as spreading sites. All CGIs serve as a control. X-axis represents CGIs within strong and weak nucleation sites as well as spreading sites and all CGIs. Y-axis represents the CpG counts identified for each sample. The CpG count is the number of CG dinucleotides in the island. Dunn's Multiple Comparison Test: ****p-value < 0.0001; strong vs. spreading: 2.4E-37; strong vs. all: 8.0E-18; weak vs. spreading: 2.3E-74; weak vs. all: 8.2E-39. Error bars represent range of the minimum and maximum values. **c**, % of nucleation sites possessing either or both of the GA-rich and GCN motifs. **d**, % of nucleation and spreading sites containing either the GA-rich (left) or GCN tandem repeat motif (right). P-values represent the enrichment of these motifs within CGIs of nucleation versus spreading sites. **e**, Motif analysis of randomly selected 200 weak nucleation site CGIs (total=1085) using MEME. Iteration of the random selection resulted in identification of similar motifs. **f**, WB using the indicated antibodies on whole cell extracts derived from i-WT-r cells with the indicated genotypes after 12 hr of 4-OHT treatment. Bands exhibiting different sizes in the case of MTF2 and of HA-EED represent different isoforms of the proteins. Non-specific bands are labeled with an asterisk. **g**, Heatmaps of SUZ12 after 12 hr of 4-OHT treatment, within a 20 kb window centered on the maximum value of peak signal in WT mESCs comparing the indicated genotypes. (Scale: 0-5 for SUZ12).

h, i, Average density profiles of JARID2 **(h)** or MTF2 (from Li *et al.,* 2017) **(i)** on nucleation sites (strong and weak) and spreading sites (D=distal, P=proximal), within a 20 kb window centered on maximum value of peak signal in WT mESCs. **j**, Heatmaps of KDM2B ChIP-seq read density in i-WT-r mESCs after 4-OHT treatment at the time points indicated. (Parameters in Fig. 2c are used to generate the Heatmaps). **k**, ChIP-seq tracks for KDM2B along the Emx1 (nucleation site, left) and Hoxc10/9 (spreading site, right) genes in i-WT-r mESCs after 4-OHT treatment at the time points indicated. **l**, Venn diagram showing the overlap of peaks from EED ChIP-seq for the WT steady-state cells, KDM2B ChIP-seq at 12 hr and 36 hr i-WT-r cells.

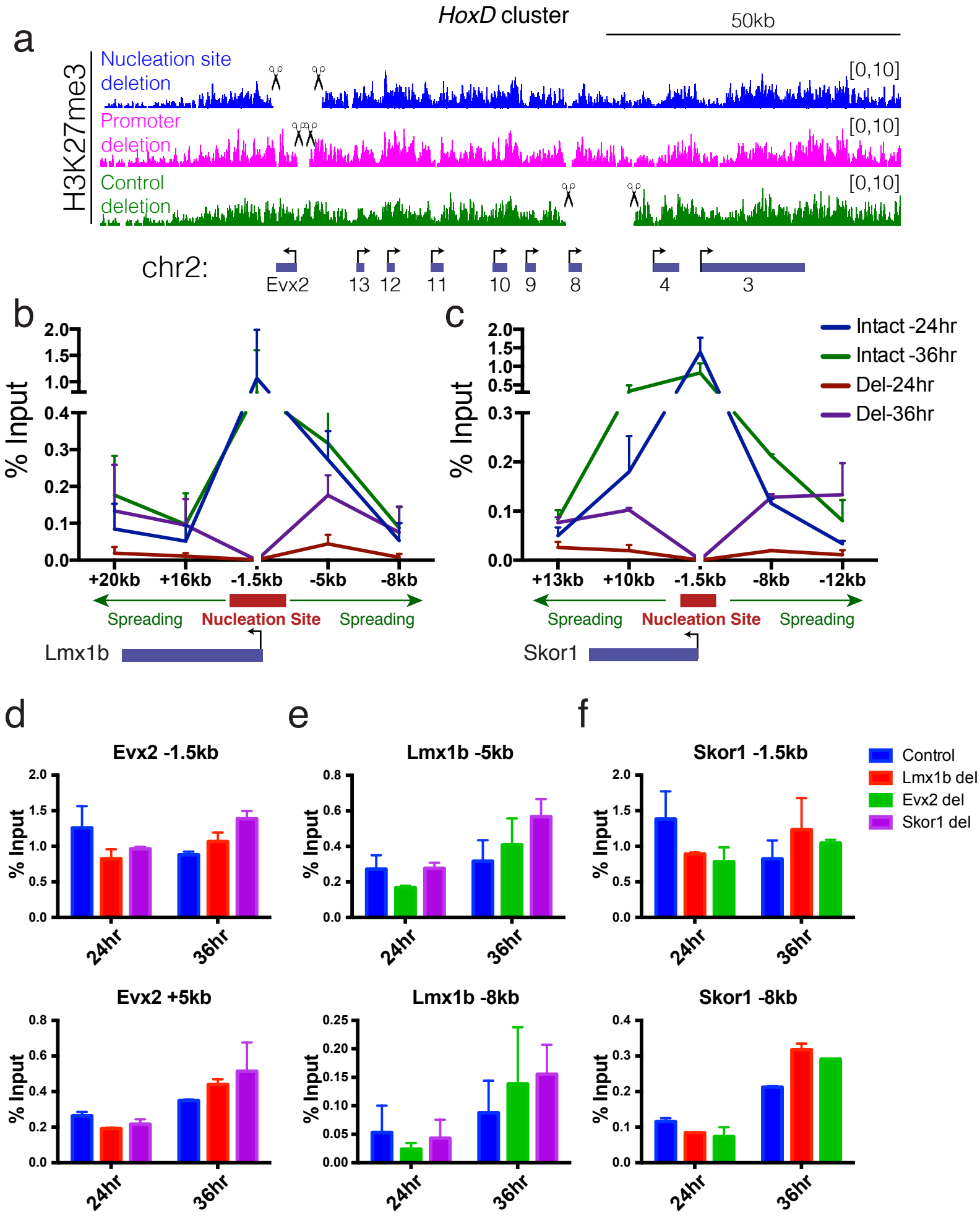
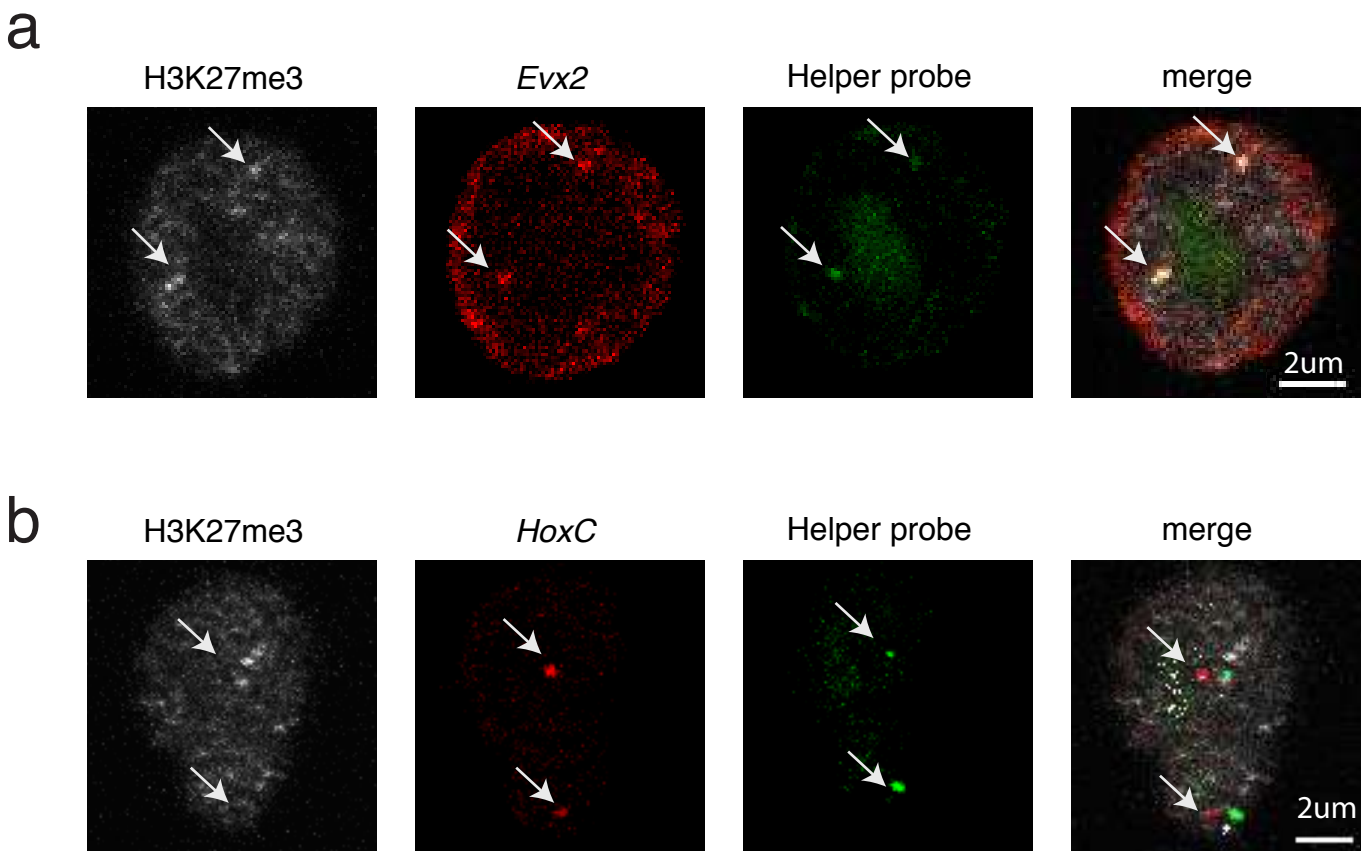
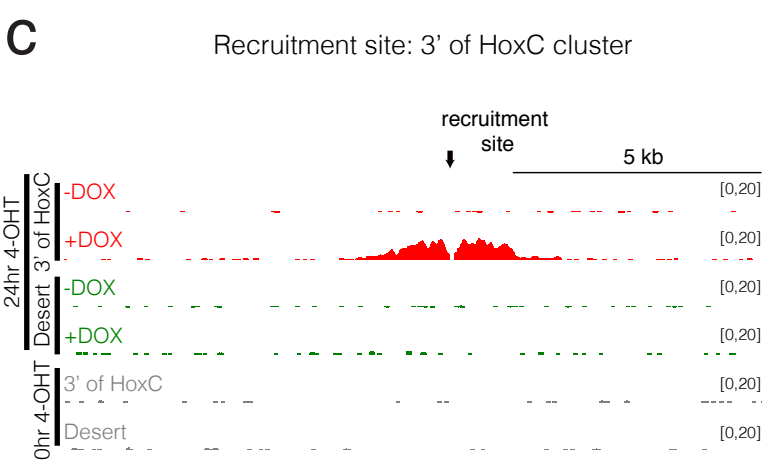

a
HoxD cluster
50kb
H3K27me3
Nucleation site deletion
Promoter deletion
Control deletion
[0,10]
chr2:
Evx2
13 12 11 10 9 8 4 3
b
c
% Input
Intact -24hr
Intact -36hr
Del-24hr
Del-36hr
+20kb +16kb -1.5kb -5kb -8kb
+13kb +10kb -1.5kb -8kb -12kb
Spreading
Nucleation Site
Spreading
Lmx1b
Skor1
d
e
f
Evx2 -1.5kb
Lmx1b -5kb
Skor1 -1.5kb
Evx2 +5kb
Lmx1b -8kb
Skor1 -8kb
Control
Lmx1b del
Evx2 del
Skor1 del
24hr
36hr

Figure S5

Figure S5: Deletion of a nucleation site delays H3K27me3 deposition to spreading sites. Related to Figure 5.

a, ChIP-seq tracks for H3K27me3 along the *HoxD* cluster in mESCs harboring the depicted genomic deletions in steady-state mESCs. **b**, **c**, H3K27me3 ChIP-qPCR comparing intact versus deleted nucleation sites at 24 hr and 36 hr after 4-OHT addition to i-WT-r cells, at the indicated regions relative to the TSS of *Lmx1b* (**b)**, and *Skor1* **(c)**. Enrichments were calculated as % of input. One-tailed paired t-test comparing the following sites, excluding the deleted regions: *Lmx1b* Del vs. Intact site: p values: 24 hr; 0.0002*** and 36 hr; 0.0613 (ns). *Skor1* Del vs. Intact site: p values: 24 hr; 0.0126* and 36 hr; 0.2187 (ns). Error bars represent the SD of the mean. **d**, **e, f,** H3K27me3 ChIP-qPCR at 24 hr and 36 hr of WT rescue in mESCs with intact nucleation site (control) and the indicated nucleation site deletions using primers specific to *Evx2* (**d**) *Lmx1b* (**e**) and *Skor1* (**f**) regions. Enrichments were calculated as % of input. All ChIP-qPCR experiments were performed in two biological replicates. Deletion of the indicated nucleation sites did not result in significant reductions in H3K27me3 occupancy at other locations in the genome (One-tailed paired t-test was performed comparing control to each of the deletion sites, p value cut off is 0.05). Error bars represent the SD of the mean.


a
H3K27me3
Evx2
Helper probe
merge
2um
b
H3K27me3
HoxC
Helper probe
merge
2um


c
Recruitment site: 3' of HoxC cluster
recruitment site
5 kb
24hr 4-OHT
3' of HoxC
-DOX
+DOX
Desert
-DOX
+DOX
0hr 4-OHT
3' of HoxC
Desert
[0,20]

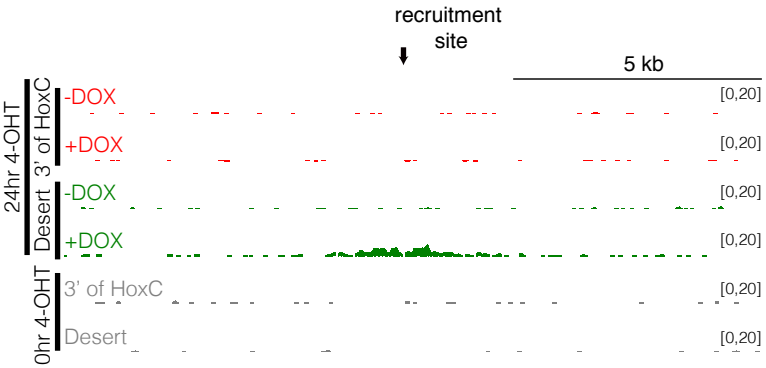

d
Recruitment site: Gene Desert
recruitment site
5 kb
24hr 4-OHT
3' of HoxC
-DOX
+DOX
Desert
-DOX
+DOX
0hr 4-OHT
3' of HoxC
Desert
[0,20]

Figure S6

Figure S6: Nucleation site overlaps with H3K27me3 foci and artificial recruitment of TetR-EZH2 to the 3' end of the *HoxC* cluster or a gene desert results in formation of ~5 kb of H3K27me3 domains. Related to Figure 6.

a, b, ImmunoFISH using H3K27me3 antibody, a nucleation site probe (*Evx2*) **(a)** or a spreading site probe (*HoxC*) **(b)** at 12 hr after WT EED expression in i-WT-r cells. A helper probe, which is specific to a region in the vicinity of *Evx2* or *HoxC,* was used to determine the real FISH signal. Arrows indicate the localization of the nucleation site and spreading site alleles. Images are examples of 2 biological replicates that are quantified in Figure 5b. **c**, **d**, ChIP-seq tracks for H3K27me3 showing deposition of H3K27me3 after induction of TetR-EZH2 recruitment to 3' of *HoxC* **(c)** and a gene desert **(d)** upon DOX treatment in i-WT-r cells after WT EED expression (24 hr, 4-OHT induction). As controls, H3K27me3 levels in the absence of WT EED expression (0 hr, without 4-OHT) at each recruitment site are indicated.

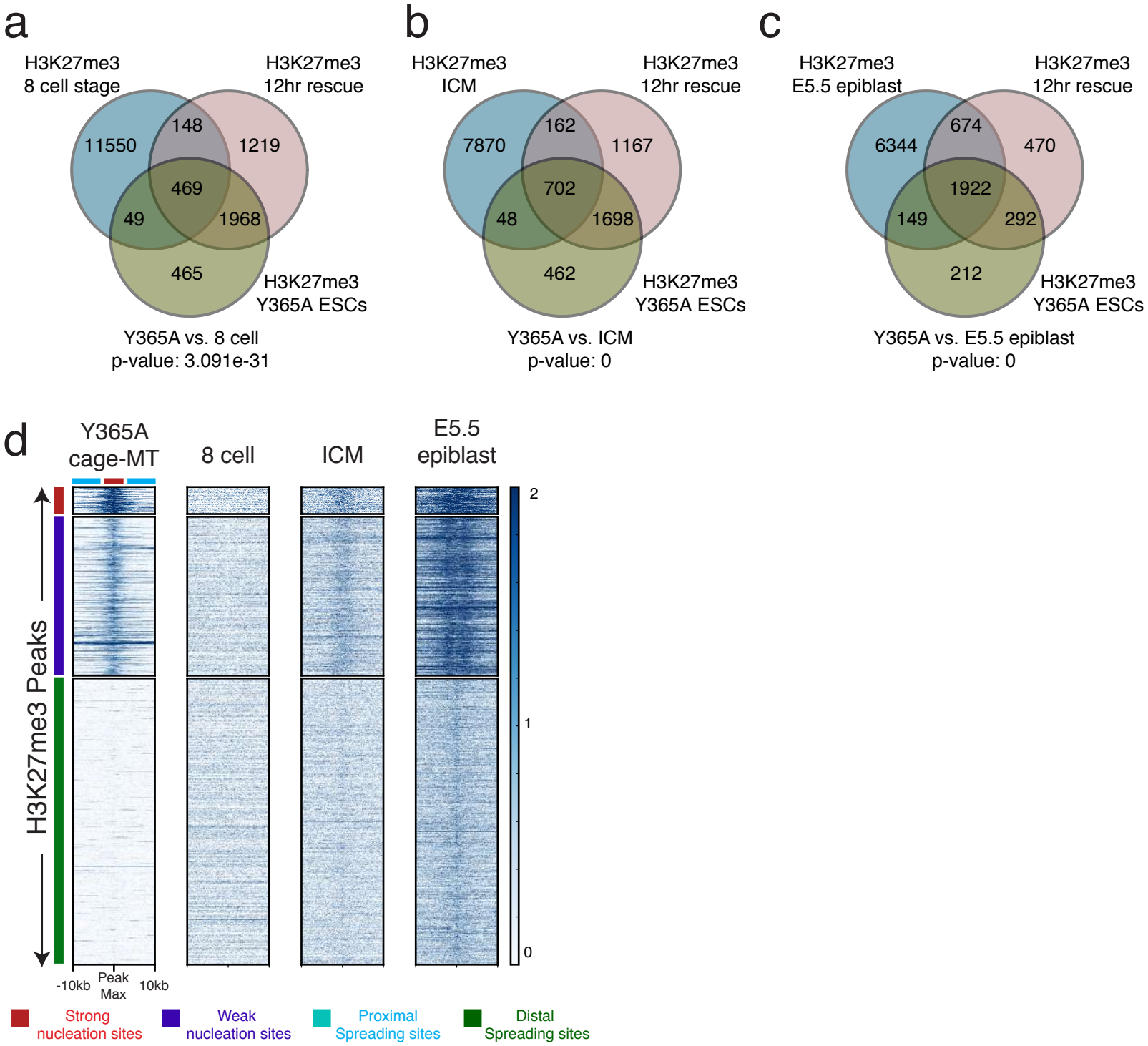

a
H3K27me3
8 cell stage
H3K27me3
12hr rescue
11550
148
1219
469
49
1968
465
H3K27me3
Y365A ESCs
Y365A vs. 8 cell
p-value: 3.091e-31
b
H3K27me3
ICM
H3K27me3
12hr rescue
7870
162
1167
702
48
1698
462
H3K27me3
Y365A ESCs
Y365A vs. ICM
p-value: 0
c
H3K27me3
E5.5 epiblast
H3K27me3
12hr rescue
6344
674
470
1922
149
292
212
H3K27me3
Y365A ESCs
Y365A vs. E5.5 epiblast
p-value: 0
d
Y365A
cage-MT
8 cell
ICM
E5.5
epiblast
H3K27me3 Peaks
2
1
0
-10kb
Peak
Max
10kb
Strong
nucleation sites
Weak
nucleation sites
Proximal
Spreading sites
Distal
Spreading sites

Figure S7

Figure S7: *De novo* deposition of H3K27me3 to nucleation sites occurred as early as the ICM stage during mouse development. Related to Figure 1 and 2.

a, b, c, Venn diagram showing the overlap in H3K27me3 peaks between the 12 hr i-WT-r, Y365A from this study and the 8-cell stage **(a)** ICM **(b)** and E5.5 epiblast **(c)** from Zheng *et al*., 2016. Significance of overlaps were determined by Fisher's exact test (two-tail) and p-values are indicated on each figure panel. **d,** Heatmaps of H3K27me3 ChIP-seq in the EED Y365A cage-mutant and the indicated developmental stages. (Parameters in Fig. 2**c** are used to generate the Heatmaps). (Scale: 0-2 for H3K27me3).