Supplementary Information for

Chromatin Assembly Factor 1 (CAF-1) facilitates the establishment of facultative heterochromatin during pluripotency exit

Liang Cheng[1,2,3], Xu Zhang[3,4,5], Yan Wang[2], Haiyun Gan[3,4,5], Xiaowei Xu[3,4,5], Xiangdong Lyu[3,4,5], Xu Hua[3,4,5], Jianwen Que[6], Tamas Ordog[7,8,9,*], and Zhiguo Zhang[3,4,5,10,*]

[1]Biochemistry and Molecular Biology Track, Mayo Clinic Graduate School of Biomedical Sciences, Mayo Clinic, Rochester, MN, 55902;
[2]Department of Biochemistry and Molecular Biology, Mayo Clinic College of Medicine, Rochester, MN, 55905;
[3]Institute for Cancer Genetics, Columbia University, New York, NY, 10032;
[4]Department of Pediatrics, Columbia University, New York, NY, 10032;
[5]Department of Genetics and Development, Columbia University, New York, NY, 10032;
[6]Department of Medicine, Columbia University, New York, NY, 10032;
[7]Department of Physiology and Biomedical Engineering, Mayo Clinic, Rochester, MN, 55905;
[8]Division of Gastroenterology and Hepatology, Department of Medicine, Mayo Clinic, Rochester, MN, 55905;
[9]Epigenomics Program, Center for Individualized Medicine, Mayo Clinic, Rochester, MN, 55905

[*]Corresponding authors:
Tamas Ordog, Email: ordog.tamas@mayo.edu
Phone: 507-538-3906
Zhiguo Zhang, Email: zz2401@cumc.columbia.edu
Phone: 212-851-4936

[10]Lead contact: Zhiguo Zhang

Running title: p150 and nucleosome assembly during pluripotency gene silencing

Supplementary Figure and legends

Fig. S1

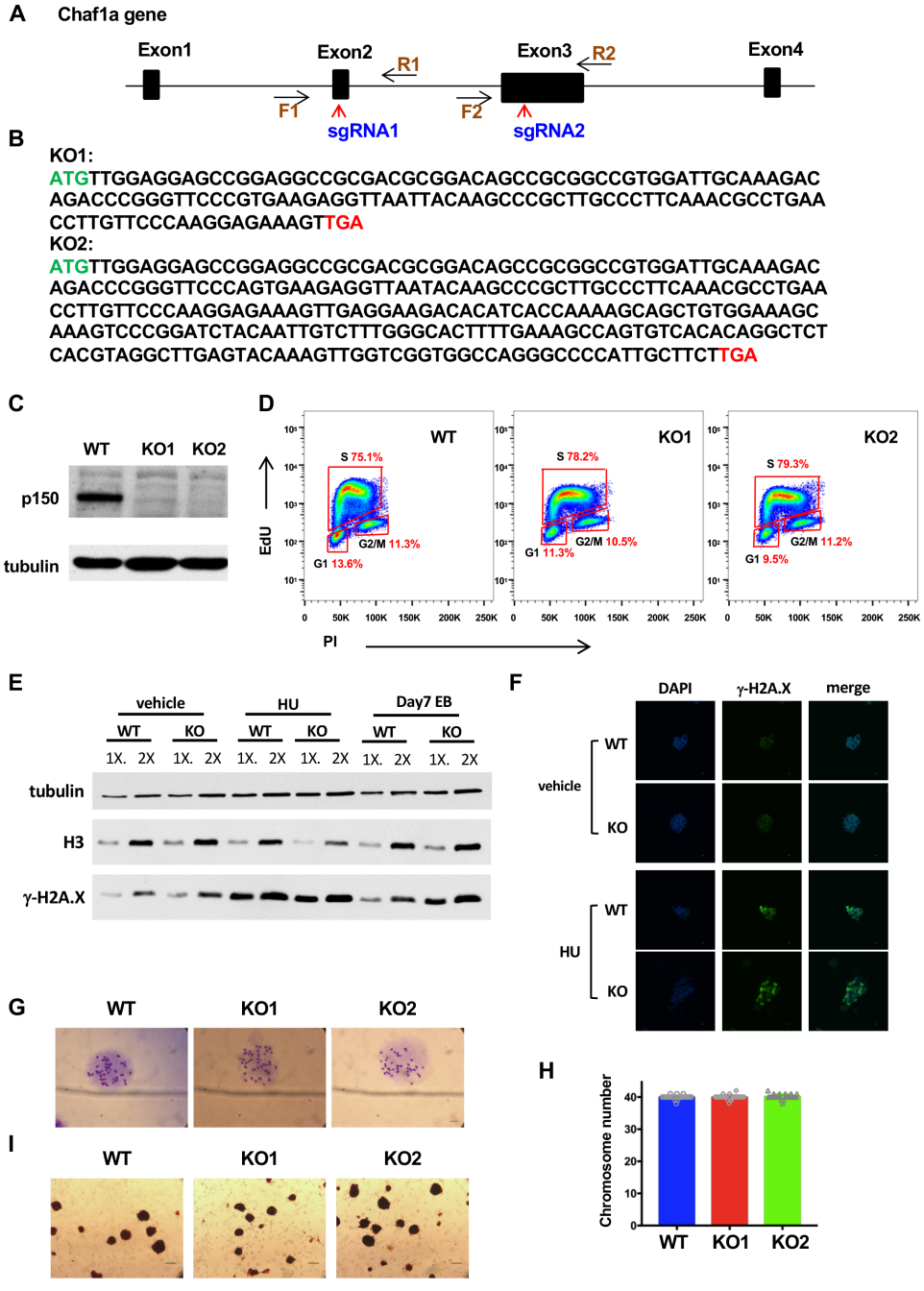


Figure S1. p150 is not required for mouse ESC viability and self-renewal. Related to Figure 1.

(**A**) A schematic representation of generating p150 KO cell line using the CRISPR/Cas9 system. Red arrowheads indicate the two sgRNA targeting sites. The two targeted regions were amplified by PCR using indicated primer sets and sequenced by Sanger sequencing. F, forward primer. R, reverse primer. (**B**) Predicted reading frame from the *Chaf1a* gene locus after nucleotide insertion in p150 KO lines. Start codon is marked by green and premature stop codon in the p150 KO line is marked by red color. (**C**) WB analysis of p150 proteins in WT and two different p150 KO lines using antibodies raised against an N-terminal peptide. Tubulin (bottom) was used as loading control. (**D**) Cell cycle analysis of p150 WT and KO cells by flow cytometry using the 5-ethynyl-2'-deoxyuridine (EdU)/propidium iodide (PI). The percentage of cells in each phase of the cell cycle is indicated. (**E**) Level of H2A.X phosphorylation (γ-H2A.X) increased slightly after p150 KO. ES cells were treated with vehicle or 1mM hydroxyurea (HU) for 24 hours and whole cell lysate was collected. Embryoid body after 7 days differentiation was collected. Tubulin and H3 were used as loading control. (**F**) Representative immunofluorescence image of γ-H2A.X of ESCs treated with vehicle or 1mM HU for 24 hrs. Scale bar: 10 μM. (**G-H**) Karyotypes of p150 WT and KO cells are normal. (**G**) Representative images of chromosome spread. Scale bar: 5μM. (**H**) Quantification of average chromosome numbers of 30 chromosome spreads from (**G**). (**I**) Representative images of alkaline phosphatase staining of p150 WT and KO clones. ESCs were seeded on top of feeder cells and assayed for alkaline phosphatase activity. Scale bar: 50 μM.

Fig. S2

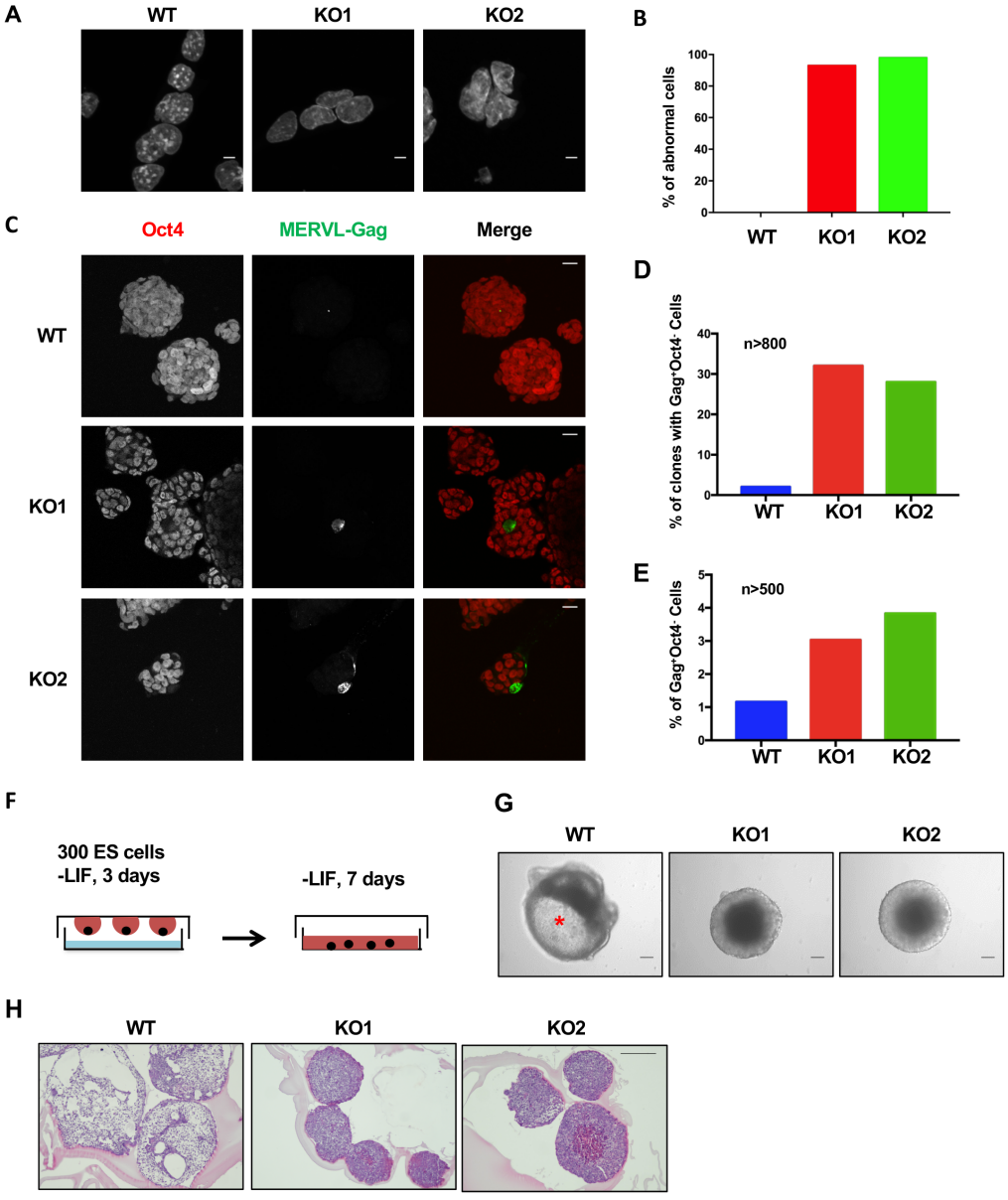


Figure S2. p150 deficiency disrupts chromocenters, induces 2C-like cells and impairs ESC differentiation Related to Figure 1 and 2.

(**A**) Representative images of DAPI staining showing chromocenters in p150 WT and KO cells. Scale bar: 5μM. (**B**) Quantification of cells with perturbed chromocenters. At least 50 cells were counted from randomly selected field. (**C**) Representative confocal immunofluorescence images of Oct4 and MERVL-Gag proteins of p150 WT and KO lines. Scale bar: 20 μM. (**D**) Quantification of clones containing at least one $Gag^{+}Oct4^{-}$ cell using randomly selected fields of confocal images from 3 independent repeats. At least 800 clones were counted in each cell line. (**E**) Percentage of $Gag^{+}Oct4^{-}$ cells in total cell populations. At least 500 cells were counted using randomly selected fields of confocal images. (**F**) Schematic representation of the hanging drop method for *in vitro* embryoid body (EB) formation assay. (**G**) Representative bright field images of day 10 EBs derived from p150 WT and KO lines. WT EBs showed maturation and cavitation marked by red asterisk, while KO EBs remained densely packed. Scale bar: 100μM. (**H**) H&E staining of paraffin sections of day 7 EBs derived from p150 WT and KO lines. Scale bar: 200 μM.

Fig. S3

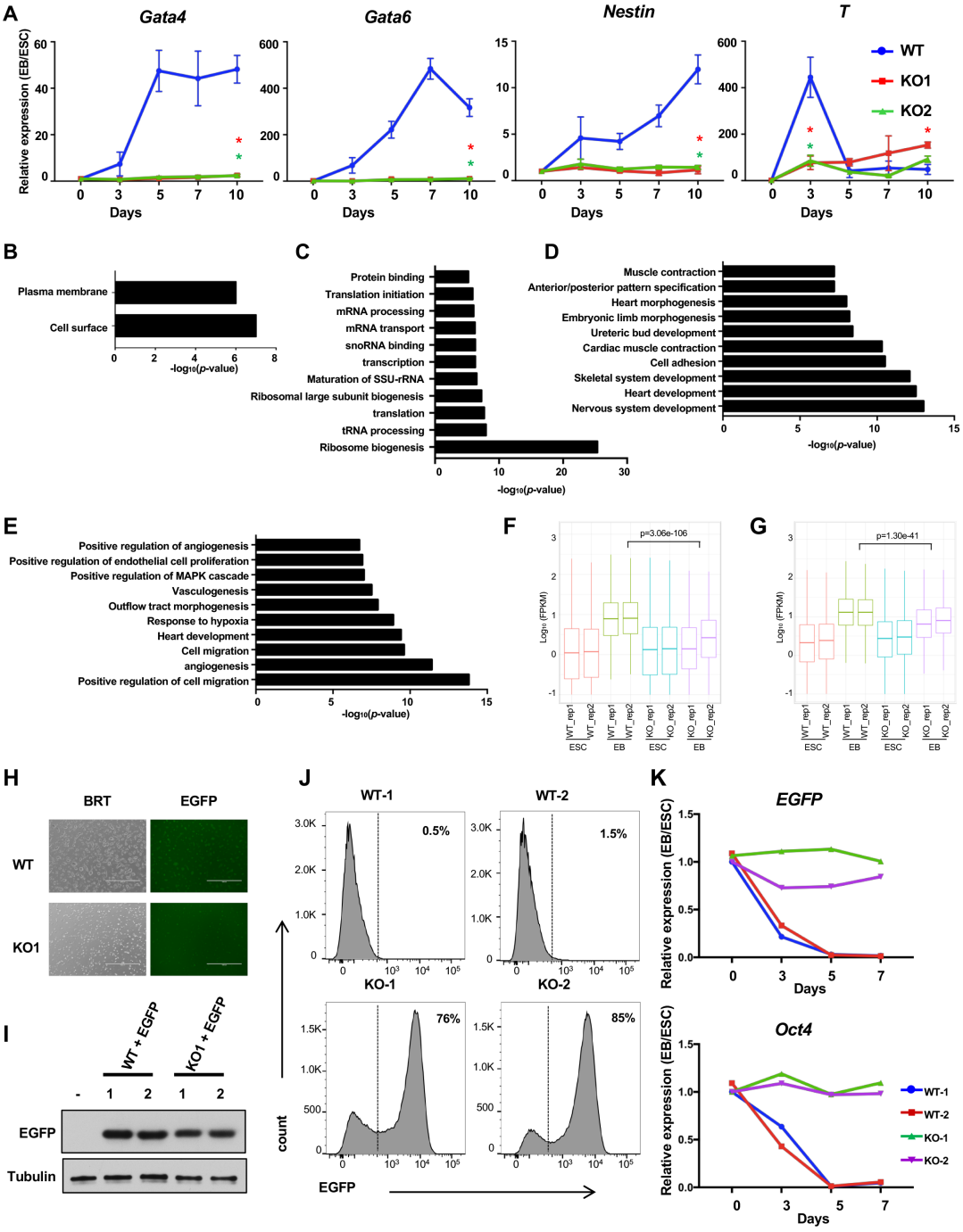

A
Gata4
Gata6
Nestin
T
Relative expression (EB/ESC)
Days
WT
KO1
KO2
B
Plasma membrane
Cell surface
-log10(p-value)
C
Protein binding
Translation initiation
mRNA processing
mRNA transport
snoRNA binding
transcription
Maturation of SSU-rRNA
Ribosomal large subunit biogenesis
translation
tRNA processing
Ribosome biogenesis
-log10(p-value)
D
Muscle contraction
Anterior/posterior pattern specification
Heart morphogenesis
Embryonic limb morphogenesis
Ureteric bud development
Cardiac muscle contraction
Cell adhesion
Skeletal system development
Heart development
Nervous system development
-log10(p-value)
E
Positive regulation of angiogenesis
Positive regulation of endothelial cell proliferation
Positive regulation of MAPK cascade
Vasculogenesis
Outflow tract morphogenesis
Response to hypoxia
Heart development
Cell migration
angiogenesis
Positive regulation of cell migration
-log10(p-value)
F
p=3.06e-106
Log10 (FPKM)
WT_rep1
WT_rep2
KO_rep1
KO_rep2
ESC
EB
G
p=1.30e-41
H
BRT
EGFP
WT
KO1
I
WT + EGFP
KO1 + EGFP
EGFP
Tubulin
J
WT-1
WT-2
KO-1
KO-2
0.5%
1.5%
76%
85%
count
EGFP
K
EGFP
Oct4
Relative expression (EB/ESC)
Days
WT-1
WT-2
KO-1
KO-2

Figure S3. CAF-1 p150 deficiency impairs ESC differentiation Related to Figure 2.

(**A**) RT-qPCR analysis of lineage-specific gene expression during EB differentiation. Data are means ± SEM (n = 3 of independent experiments, *p < 0.05, **p < 0.01, two-tailed Student *t* test). (**B**) GO analysis of genes with differential expression between p150 WT and KO ESCs. Terms' *p*-values are at the X-axis. (**C**) GO analysis of the Group 2 genes identified in Figure 2C. (**D-E**) GO analysis of Group 3 (**D**) and Group 4 (**E**) genes identified in Figure 2C. Top 10 GO biological processes based on each term's *p*-values are shown. (**F-G**) Effect of p150 KO on the expression of Group 3 and Group 4 genes based on the average of RNA-seq signals at each group of genes identified in Figure 2C. The *p* values were calculated using Wilcoxon test. (**H**) Representative fluorescence images of EGFP signals in p150 WT and KO ES cell lines expressing EGFP driven by the Oct4 distal enhancer. BRT, bright field. Scale bar: 400μM. (**I**) WB analysis of EGFP levels in two independent ES cell clones of p150 WT and KO lines. Tubulin was used as loading control. (**J**) Flow cytometry analysis of $EGFP^+$ cells derived from day 7 EBs. Results of two independent clones from p150 WT and KO ESCs are shown. The dashed line separates the EGFP-negative and EGFP-positive cells. (**K**) RT-qPCR analysis of EGFP as well as endogenous Oct4 during EB differentiation.

Fig. S4

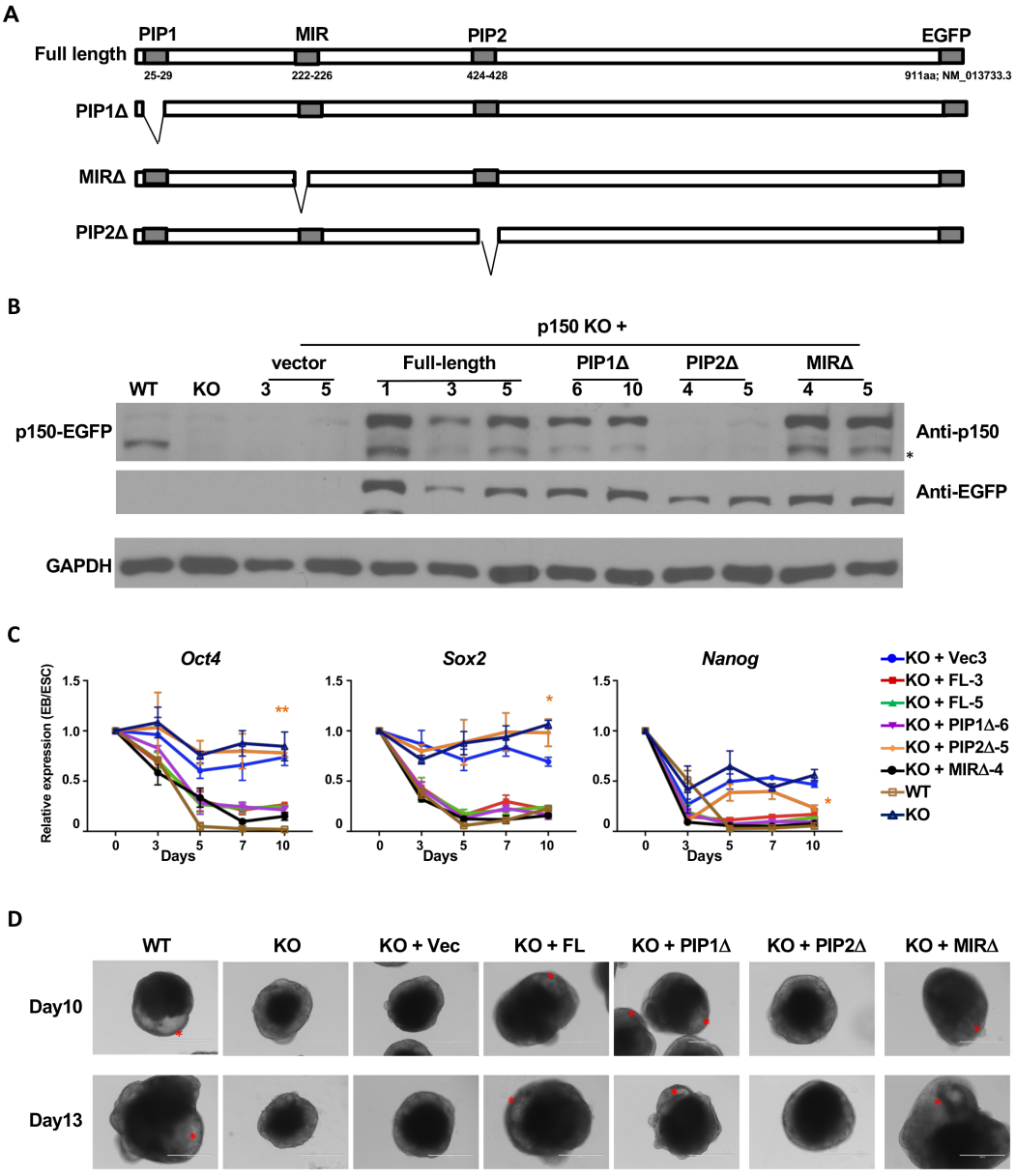


Figure S4. The p150 PIP2 domain is critical for silencing of Oct4, Sox2 and Nanog. Related to Figure 3.

(**A**) Scheme of the EGFP-tagged full-length p150 and p150 mutants with deletion of the indicated regions. Full-length, FL; PIP1Δ, PIP1 domain deletion; PIP2Δ, PIP2 domain

deletion; MIRΔ, HP1-interacting domain MIR deletion. (**B**) WB analysis of the exogenously expressed full-length p150 and p150 mutants in p150 KO cells. Two primary antibodies (anti-p150 and anti-GFP) were used for Western blot analysis of p150 levels in different clones. Failure to detect the p150 proteins in ΔPIP2 mutant cells by antibodies against p150 is likely due to impairment of antibody recognition. Protein levels were compared to endogenous p150 levels. * indicates the degradation of exogenously expressed protein. GAPDH is used as loading control. All clones were used in further experiments. (**C**) RT-qPCR analysis of Oct4, Sox2 and Nanog from p150 KO lines expressing p150 mutants during EB differentiation. Results were from three independent experiments and are presented as means ± SEM. *$p < 0.05$, **$p < 0.01$ (two-tailed Student's *t* test between p150 WT and KO + PIP2Δ lines). (**D**) Representative bright field images of day10 and day 13 EBs derived from p150 WT ESC, p150KO ESC and p150 KO ESC expressing various p150 mutants. WT EBs showed maturation and cavitation marked by red asterisk, while KO EBs remained densely packed. Expression of p150 FL, p150-PIP1Δ or p150-MIRΔ could rescue the defect observed in KO EBs, while p150-PIP2Δ could not rescue the defect. Scale bar: 400μM.

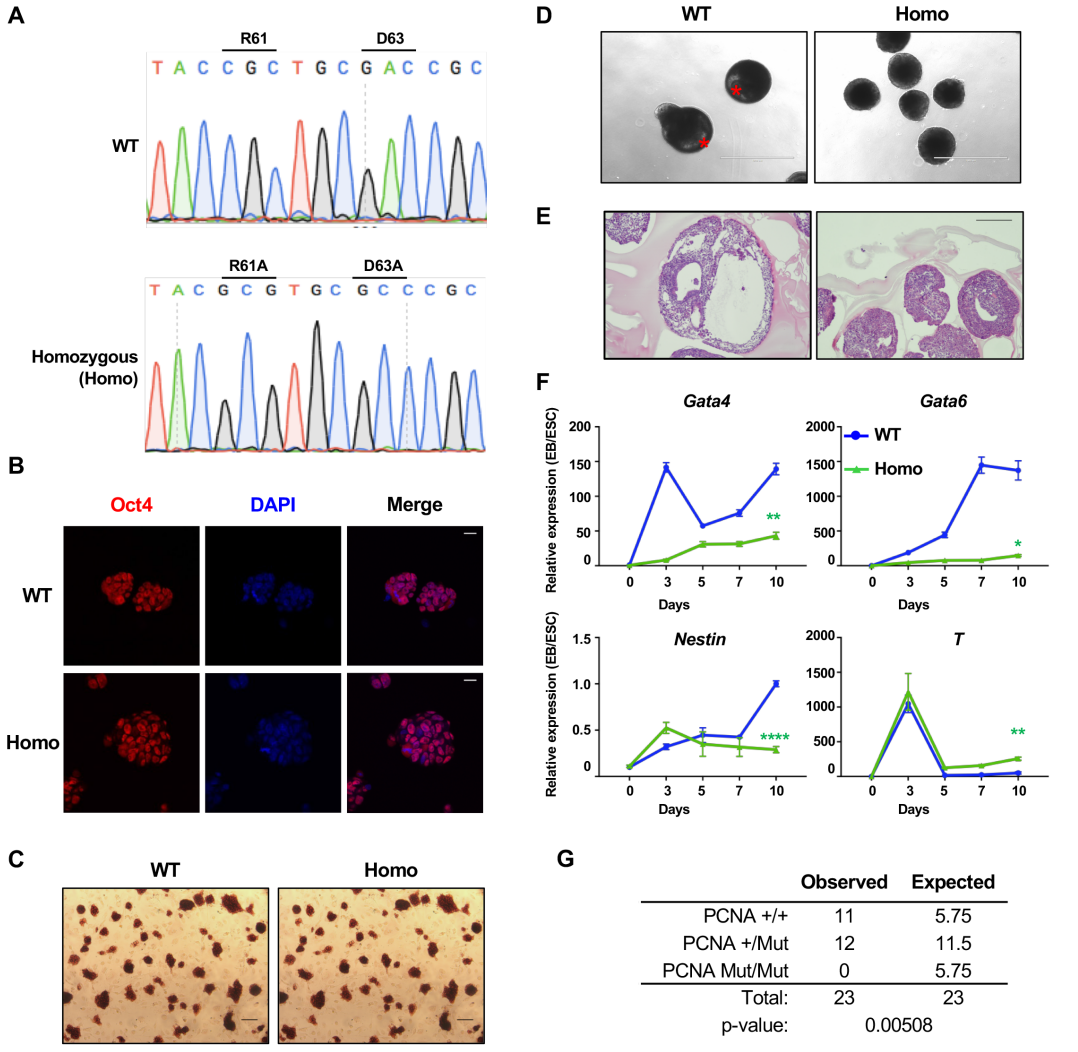


Figure S5. The PCNA R61A D63A mutation impairs ESC differentiation. Related to Figure 4.

(**A**) Sanger sequencing analysis of the PCNA mutant clones generated via CRISPR/Cas9-mediated gene editing. DNA sequence of a homozygous mutant line shows the expected R61A, D63A mutations. (**B**) Representative immunostaining of Oct4 in PCNA WT and homozygous (Homo) mutant lines. Scale bar: 20 μM. (**C**) Representative alkaline phosphatase staining of PCNA WT and PCNA R61A D63A homo mutant clones. Homo,

homozygous PCNA mutation. Scale bar: 50 μM. (**D**) Representative bright field images of day 10 EBs of PCNA WT and Homo mutant ESC lines. WT EBs show maturation and cavitation marked by red asterisk, while PCNA Homo mutant EBs were densely packed. Scale bar: 100 μM. (**E**) H&E staining of paraffin sections of day 7 EBs derived from PCNA WT and Homo mutant lines. Scale bar: 200 μM. (**F**) RT-qPCR analysis of lineage-specific genes expression during EB differentiation. Error bar represents means ± SEM (n=3, *p < 0.05, **p< 0.01, two-tailed Student's *t* test). (**G**) Expected and observed frequency of PCNA mutation status in offspring from crossing of heterozygous PCNA +/Mut mice. Genotyping was performed after weaning. The *p*-value was calculated by chi-square test.

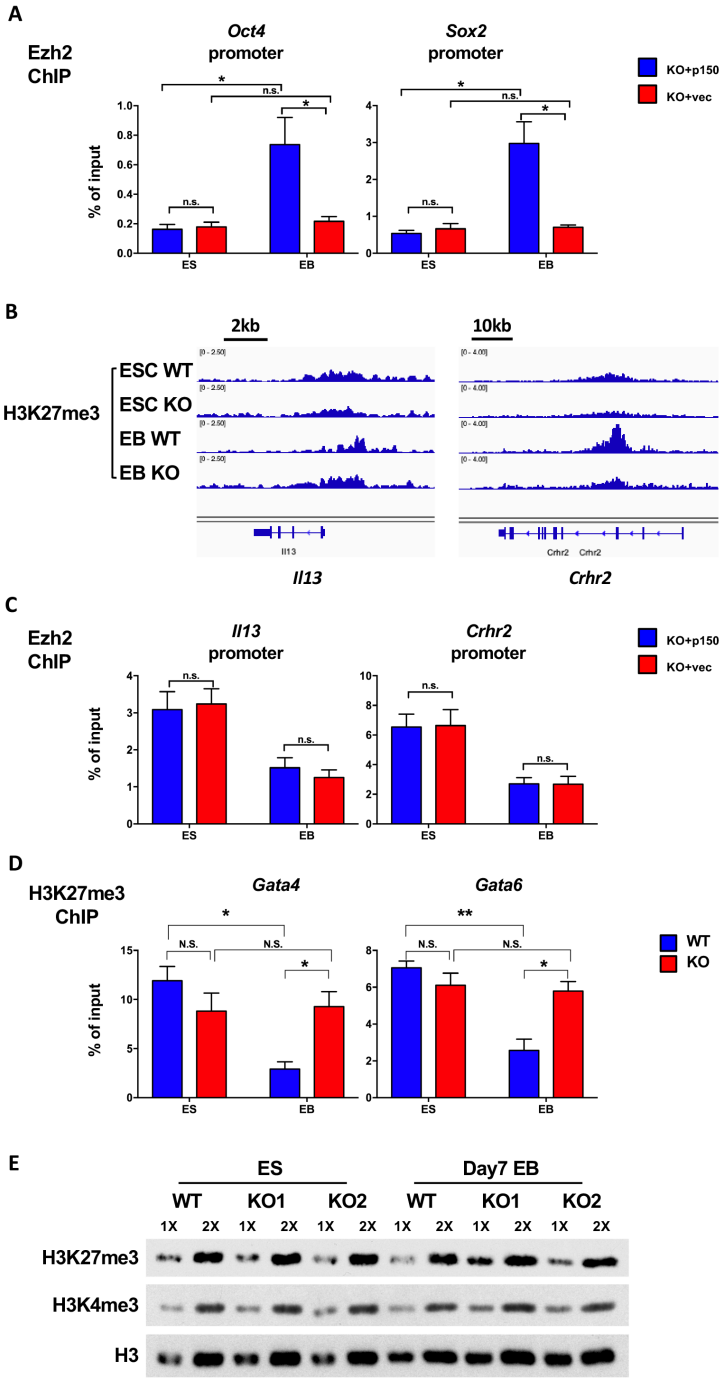


Figure S6. Effect of p150 deletion on Ezh2 and histone modifications at gene promoters. Related to Figure 5.

(**A**) Ezh2 enrichment at two pluripotent genes (*Oct4* and *Sox2*) was compromised in day 7 EB of p150 KO line. Ezh2 ChIP was performed using chromatin from ESC and day 7 EB of p150 KO cells with vector or p150-GFP. ChIP DNA was analyzed by quantitative PCR using primers annealing to the promoter regions of the indicated genes. Results are from three independent experiments (means ± SEM, n=3, *p < 0.05, n.s., not significant, two-tailed Student's *t* test). (**B**) IL13 and Crhr2 are two genes remained silenced during differentiation. Representative browser tracks showing H3K27me3 at the *Il13* and *Crhr2* loci before (ESC) and after (EB, day 7 EB) differentiation. (**C**) p150 knockout does not affect the Ezh2 enrichment at silent genes (*Il13* and *Crhr2*). Ezh2 ChIP was performed as described above and ChIP DNA was analyzed using primers targeting the promoters of *Il13* and *Crhr2*. Results are from three independent experiments (means ± SEM, n=3, n.s., not significant, two-tailed Student's *t* test). (**D**) Reduction of H3K27me3 at bivalent genes (*Gata4* and *Gata6*) was compromised in p150 KO EB. H3K27me3 ChIP was performed using chromatin from ESC and day 7 EB of p150 WT and KO cell lines. Quantitative PCR was performed targeting the promoter region of the indicated genes. Results are from three independent experiments (means ± SEM, n=3, *p < 0.05, **p < 0.01, n.s., not significant, two-tailed Student's *t* test). (**E**) Analysis of global histone modification in p150 WT, p150 KO ESC and day 7 EBs by Western Blot. Histones were isolated using acid extraction. Histone H3 was used as loading control

Fig.S7

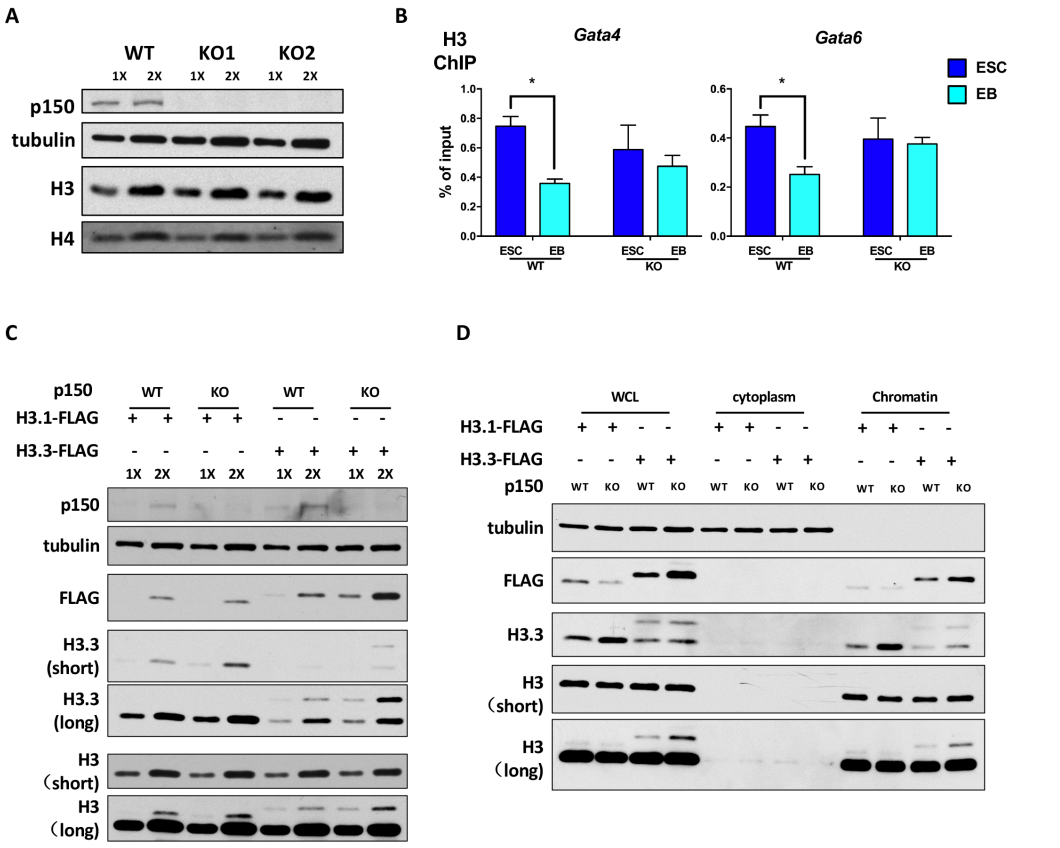
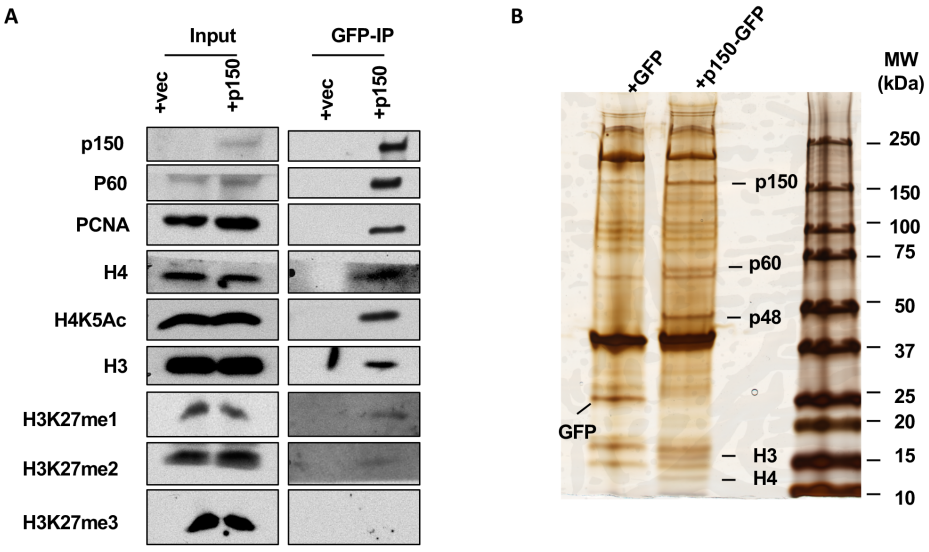


Figure S7. H3.3 on chromatin increased slightly after p150 deletion. Related to Figure 7.

(**A**) Level of histone H3 and H4 was not changed in p150 KO lines. Whole cell extract was used for WB analysis. Tubulin was used as loading control. (**B**) Nucleosome occupancy at two lineage-specific genes (*Gata4* and *Gata6*) in ES cells and day 7 EB. H3 ChIP was performed using chromatin from ESC and day 7 EB of p150 WT and KO cell line. Quantitative PCR was performed targeting the -1 nucleosome upstream of TSS of the indicated genes. Results are from three independent experiments (means ± SEM, n=3, *$p < 0.05$, two-tailed Student's *t* test). (**C**) Level of histone H3.3 increased slightly after p150 KO. Endogenous histone H3.1 and H3.3 was tagged with FLAG epitope via

CRISPR/Cas9 system and p150 KO lines was derived from the tagged cell line. Whole cell extract was used for WB analysis. Tubulin was used as loading control. Short/long, short/long exposure time during film development. (**D**) H3.3 on chromatin increased slightly after p150 KO. Chromatin fraction and cytoplasm fraction was separated from cells with H3.1-FLAG or H3.3-FLAG. Histone level was compared between p150 WT and KO cells. WCL, whole cell lysates. Short/long, short/long exposure time during film development.

Fig.S8



C

Modification	Exp.1	Exp.2	total
H3K27ac	4	20	24
H3K27me1	0	2	2
H3K27me2	3	16	19
H3K27me3	1	10	11
H3K36	3	10	13
H3K36ac	0	4	4
H3K36me1	1	10	11
H3K36me2	4	22	26
H3K36me3	0	2	2
H4K5ac	8	9	17
H4K8ac	10	11	21
H4K12ac	9	9	18
H4K16ac	9	9	18

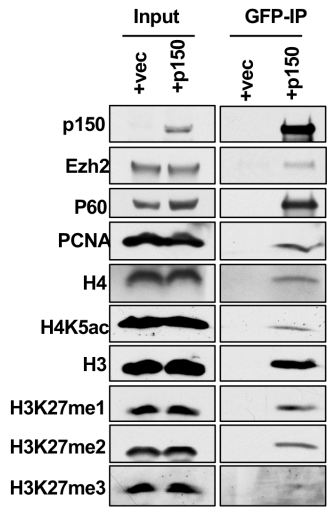


Figure S8. CAF-1 interacts with histones with different modifications in ESC. Related to Figure 7.

(**A**) Histone H4 co-purified with CAF-1 is acetylated at lysine 5 (K5). EGFP-tagged full-length p150 expressed in p150 KO cells was immunoprecipitated using antibodies against GFP. Co-purified proteins were analyzed by WB using the indicated antibodies. (**B**) H3-H4 was co-purified with EGFP-p150. EGFP-p150 was purified from ES cells expressing

EGFP-p150 or EGFP. Co-purified proteins were analyzed by SDS-PAGE and detected by silver staining. Subunits of CAF-1 complex (p150-EGFP, p60 and p48) and histone H3/H4 are indicated.). (**C**) Quantification of histone H3 and H4 peptides with different modifications on H3-H4 co-purified with CAF-1. Peptides counts are from two independent mass spectrometry experiments. The H3 and H4 bands on the gel shown in (B) was excised and subjected to mass spectrometry analysis. (**D**) p150 interact with Ezh2 in PCNA mutant cells. EGFP-tagged full-length p150 expressed in PCNA homozygous mutant cells was immunoprecipitated using antibodies against GFP. Co-purified proteins were analyzed by WB using the indicated antibodies.

Table S1. A list of oligonucleotides used in the study

	Forward	**Reverse**
p150 sg1	TGCCCTCAAGAAGCTATCAA	TTGATAGCTTCTTGAGGGCA
p150sg2	CCAGTGAAGAGGTTAATACA	TGTATTAACCTCTTCACTGG
PCNA sg1	CTCCCGCCACCCGCTCACC	GGTGAGCGGGTGGCGGGAG
p150 mut detect1	ACCCCCATATACAAACAAACTAAG	TGCTTAGTGGCCGATACTTG
p150 mut detect2	TTGCTGTCACTAAAGATGCT	TGACAAGTGATGTCTGAAGG
PCNA Mut detect	CCAGACCTCGTTCCTCTTAG	CCCCTCAAATCCATGATGCG
Hist1h3g sg1	TCCGCGGGGAGAGGGCTTAA	TTAAGCCCTCTCCCCGCGGA
H3f3b sg1	GAGAGAGCTTAAGTTGAAG	CTTCAACTTAAGCTCTCTC
Hist1h3g Mut detect	TTTTCCCTACGGTTACTTGCC	CGGTGTGATGAAAGGAGACCT
H3f3b Mut detect	GCGGCTTGATTGCTCCAGTA	TGCAAAAGGACCCTGTACACATT
Nanog pro.	GGGTAGGGTAGGAGGCTTGA	CGGCTCAAGGCGATAGATT
Oct4 pro.	GGGGTGAGAGGACCTTGAA	GGACAGGACAACCCTTAGGAC
Sox2 pro	CCTAGGAAAAGGCTGGGAAC	GTGGTGTGCCATTGTTTCTG
Gata4 pro	AAGCGCTCTTTTCTCCTTCC	GTGAGGGCTACAGGGAGTGA
Gata6 pro	CCTTCCCATACACCACAACC	CCCCTCCTTCCAAATTAAGC
Il13 pro	GTCGATAAGAGGACGCCCG	GACACTGATCCAGCGGTCC
Crhr2 pro	CAGACCTGCACTAGGAGAACG	GAAGAGCTGCTCCTGGACGG
Oct4 -1 nucl.	CATCCGAGCAACTGGTTTGTG	ACAACCCTTAGGACGGGACC
Oct4 -2 nucl.	TGGGCTGAAATACTGGGTTCA	TCCATTGAATGTTCGTGTGCC
Oct4 -3 nucl.	ATCTTGAGGAAAGAGGCCCCG	TCCATCTTGTGTTGTCCAGGTTG
Nanog -1 nucl.	TGTCTTTAGATCAGAGGATGCCC	ACCATGGACATTGTAATGCAAAAGA
Nanog -2 nucl.	ACCAACTTACTAAGGTAGCCCG	ATCTTCCCTCCAAAAGTGCGG
Nanog -3 nucl.	TCTTCCATTGCTTAGACGGC	CCATGCCCAATTTAAGGAGTGT
Gata4 -1 nucl.	CTCCGTTCTGCCCTCAGTG	CTTGAACAGGGCGCTGGA
Gata6 -1 nucl.	CTGGATGATTATGGGTCGCTG	TCCTCAGAGTAGGAGCGAGTC
GAPDH -1 nucl.	GGAGCCAGGGACTCTCCTTT	TCCTGTGTTCTCCCCTCACT
Nanog RT	TCTTCCTGGTCCCCACAGTTT	GCAAGAATAGTTCTCGGGATGAA
Oct4 RT	CACCATCTGTCGCTTCGAGG	AGGGTCTCCGATTTGCATATCT
Sox2 RT	GCGGAGTGGAAACTTTTGTCC	CGGGAAGCGTGTACTTATCCTT
Nestin RT	CCCTGAAGTCGAGGAGCTG	CTGCTGCACCTCTAAGCGA
Fgf5 RT	GCAAGTGCCAAATTTACGGATGAC	TTGCTCGGACTGCTTGAACCTG
Gata4 RT	CCCTACCCAGCCTACATGG	ACATATCGAGATTGGGGTGTCT
Gata6 RT	TGCCTCGACCACTTGCTATGAAAA	CACTGATGCCCCTACCCCTGAG
T RT	TCCCGGTGCTGAAGGTAAATGTGT	TTGGGCGAGTCTGGGTGGATGTAG
Gapdh RT	CTGACGTGCCGCCTGGAGAAAC	CCCGGCATCGAAGGTGGAAGAGT
PCNA targeting ssODN	TTACTCTGCGCTCCGAAGGCTTCGACACATACGCGTGCGCCCGCAACCTAGC CATGGGCGTGAACCTCACGAGGTGAGCGGGTGGCGGGAGCGGGGCCCCACT CTTCCCGCTTCCGCTCTTGGCGGGGCTGTGACTCTGCACGCTCATTG	
Hist1h3g targeting ssODN	TCATGCCCAAGGACATCCAGCTGGCCCGTCGCATCCGCGGGGAGAGGGCTG ACTACAAAGACGATGACGACAAGTAAGGGTTTCTGTTAATCCACACAACCAC TTTAAAGGCTCTTCTTAGAGC	
H3f3b targeting ssODN	CAAAGACATCCAGTTGGCTCGCCGGATACGGGGGGAGAGAGCTGACTACAA AGACGATGACGACAAGTAAGTTGAAGCGGTTTTTATGGCATTTTGTAGTAAA TTCTGTAAAATACTTTGGTTT	