

**Molecular Cell, Volume 74**

**Supplemental Information**

**Functional Landscape of PCGF Proteins**

**Reveals Both RING1A/B-Dependent-**

**and RING1A/B-Independent-Specific Activities**

**Andrea Scelfo, Daniel Fernández-Pérez, Simone Tamburri, Marika Zanotti, Elisa Lavarone, Monica Soldi, Tiziana Bonaldi, Karin Johanna Ferrari, and Diego Pasini**

Figure S1

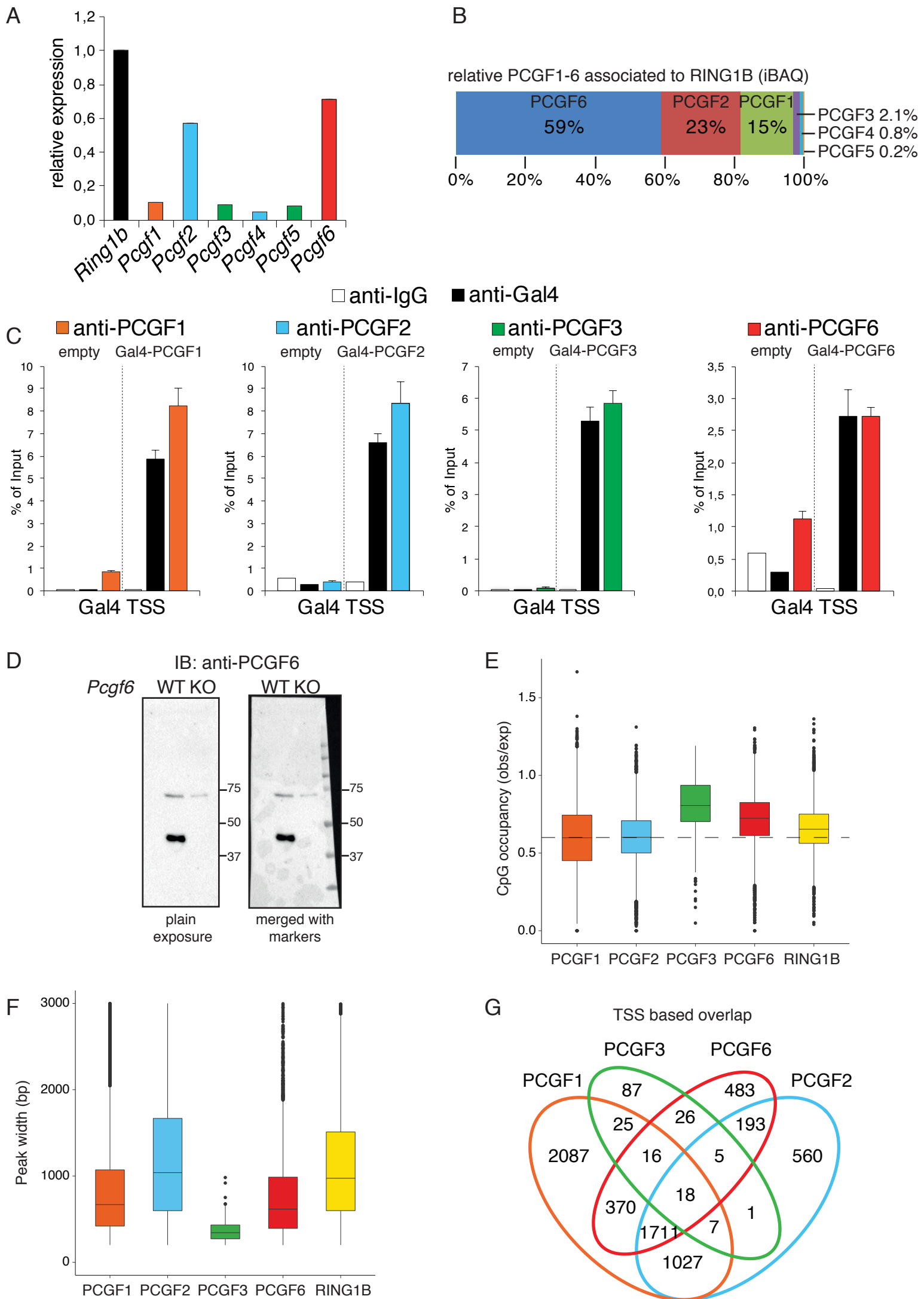
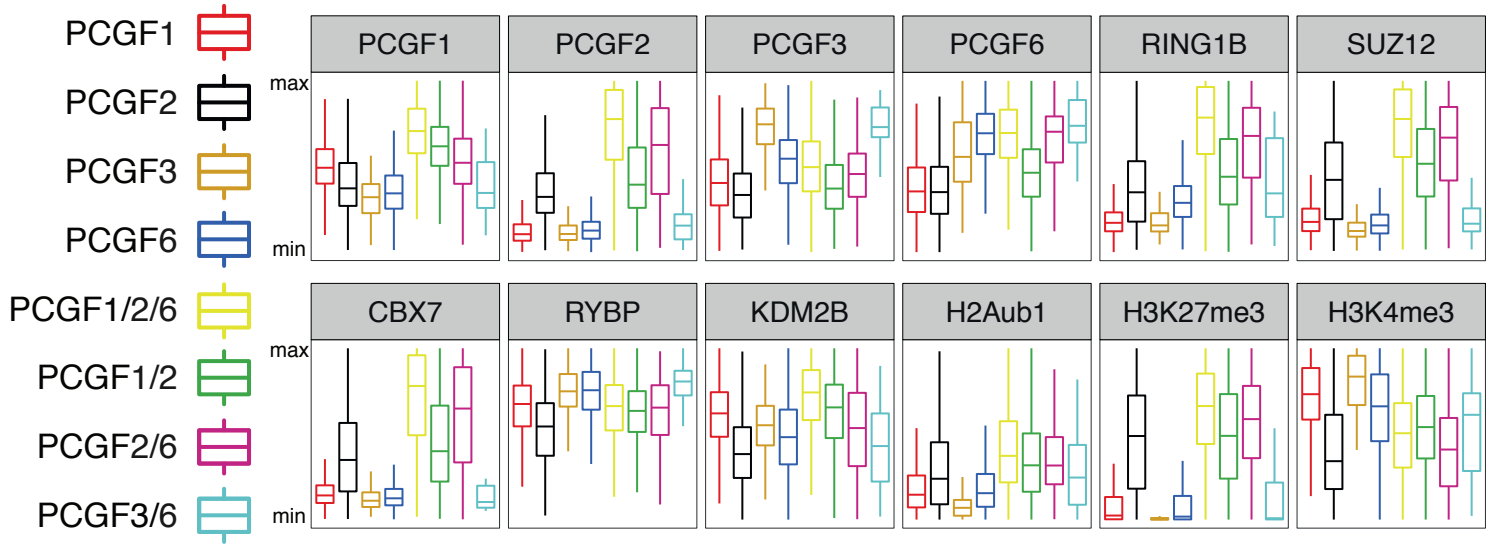


Figure S2

A

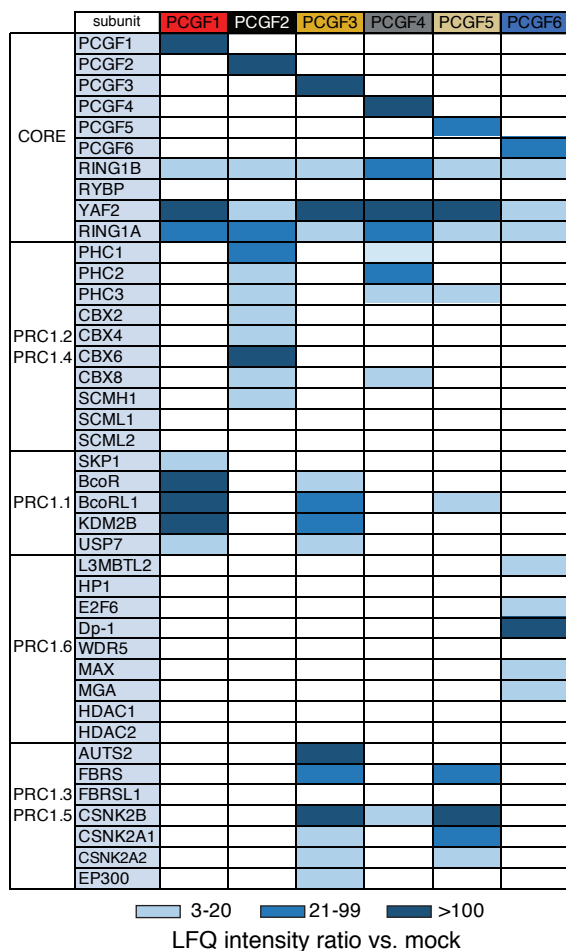
PCGFs bound loci

ChIPseq analyses



B

GAL4-PCGFs fusions - IPs anti-GAL4



C

GO terms

Gene Ontology of PCGFs associated genes

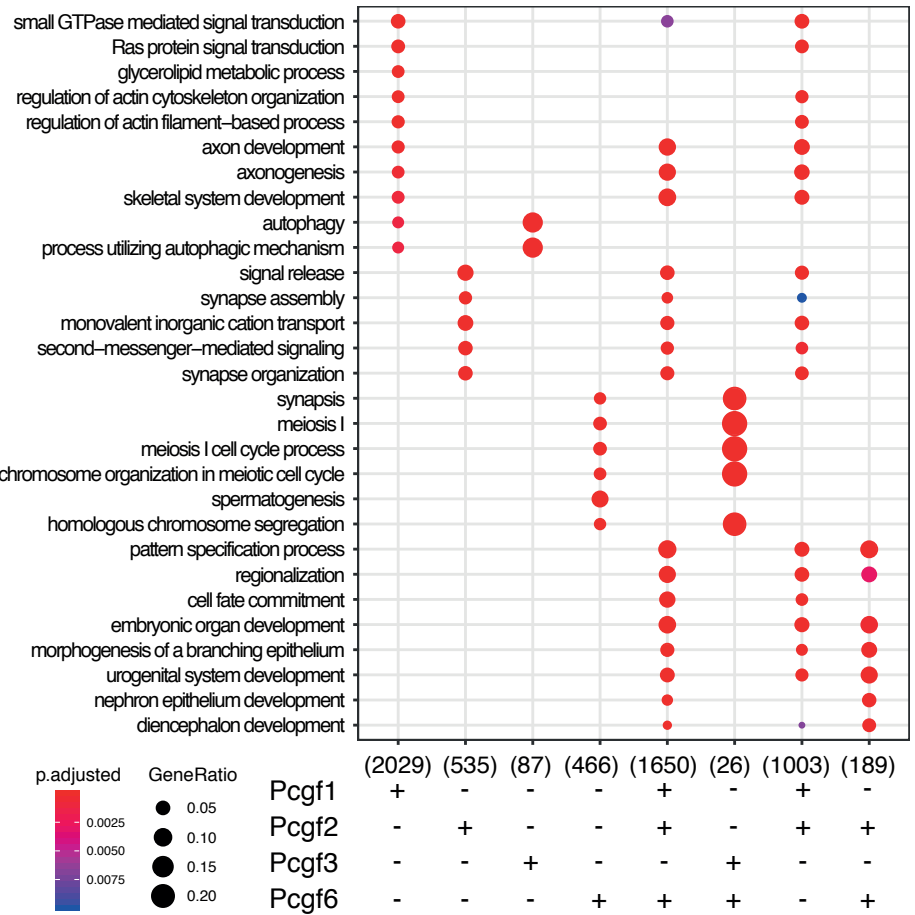


Figure S3

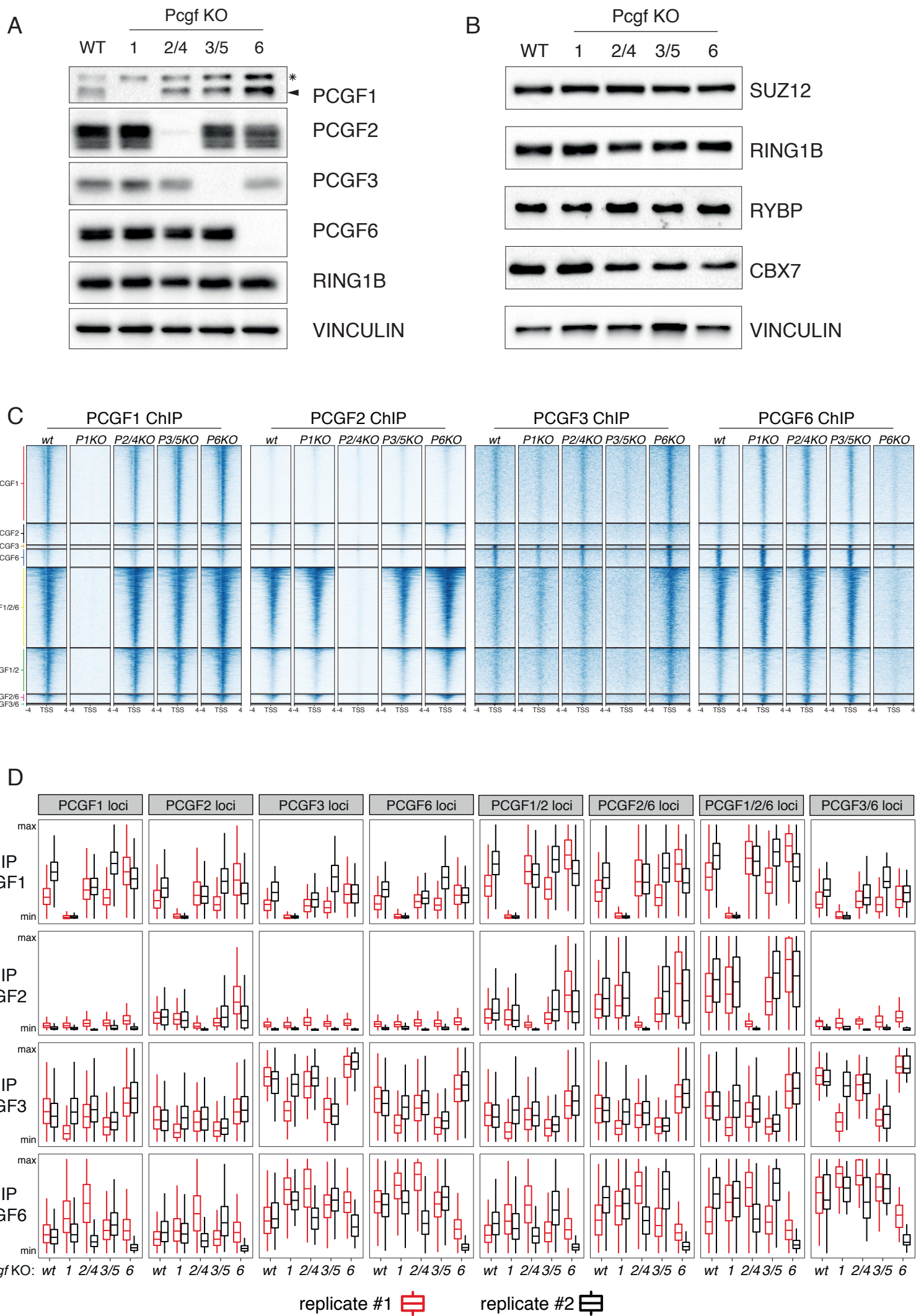


Figure S4

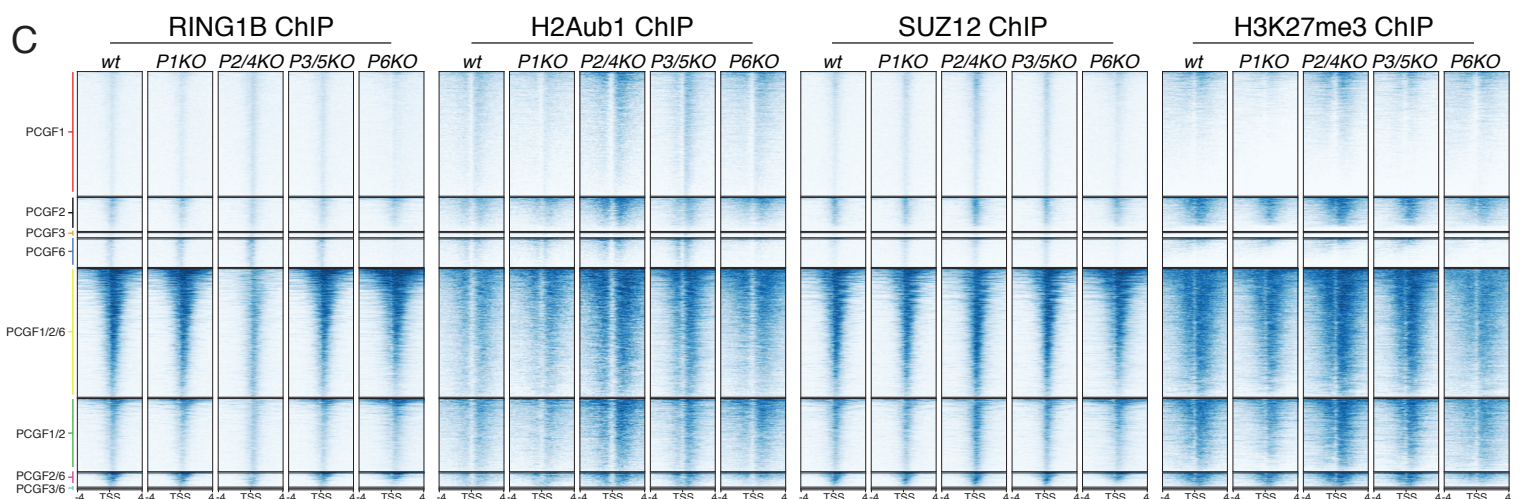
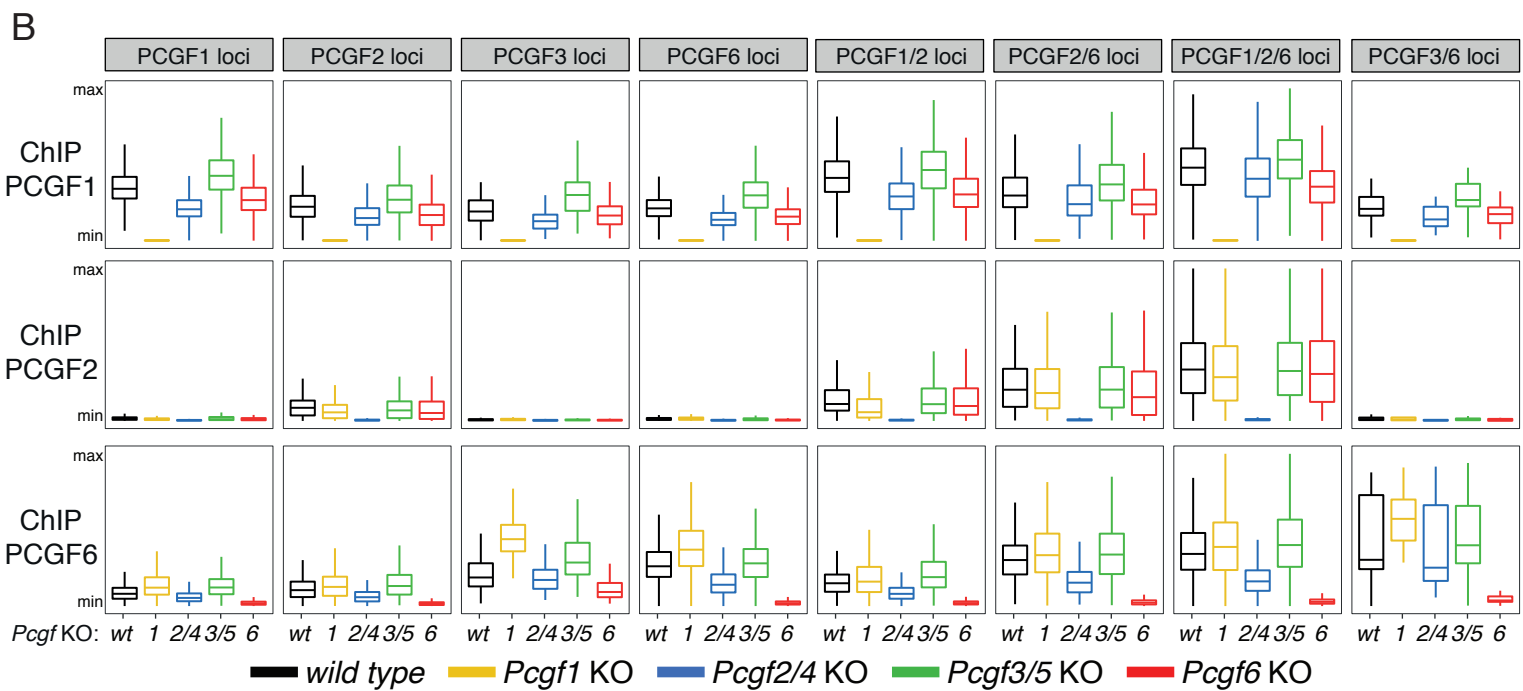
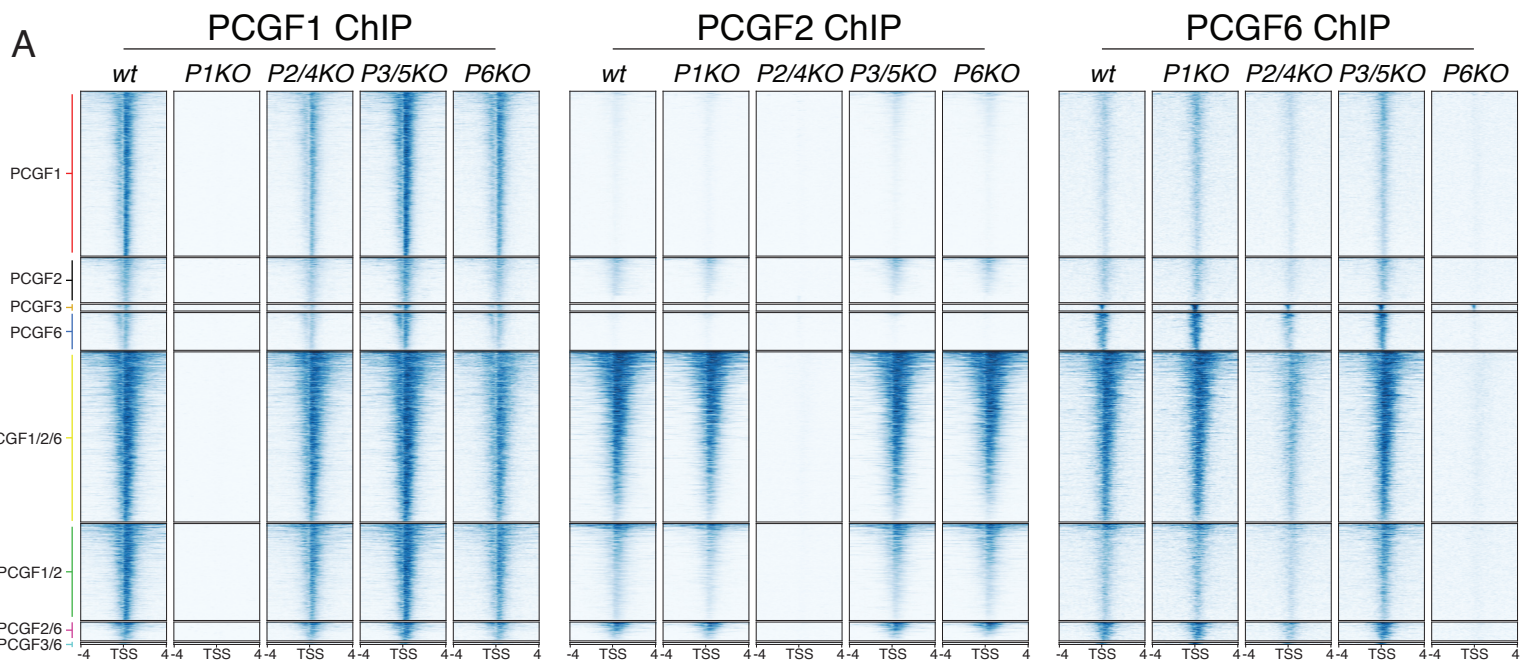


Figure S5

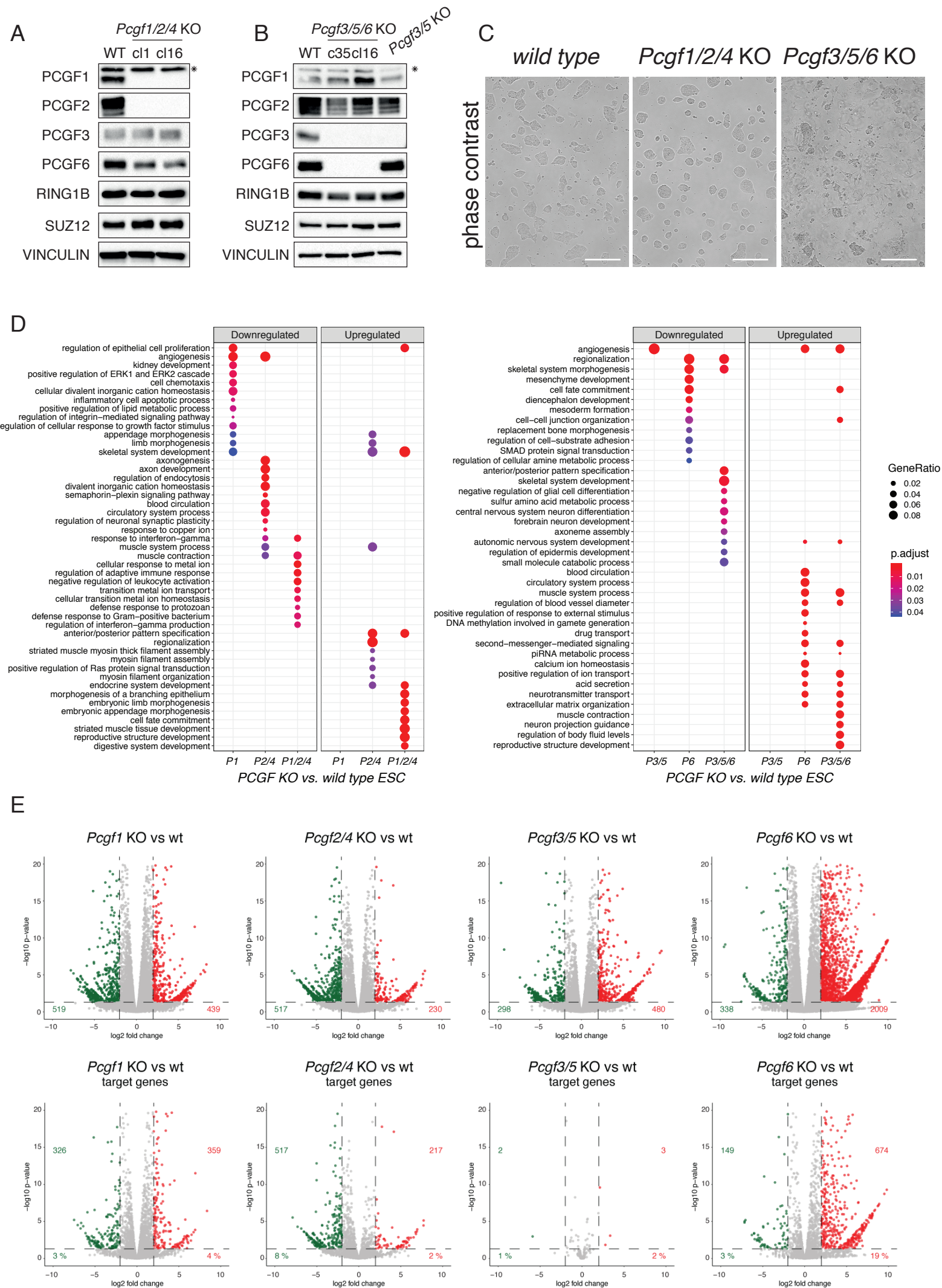
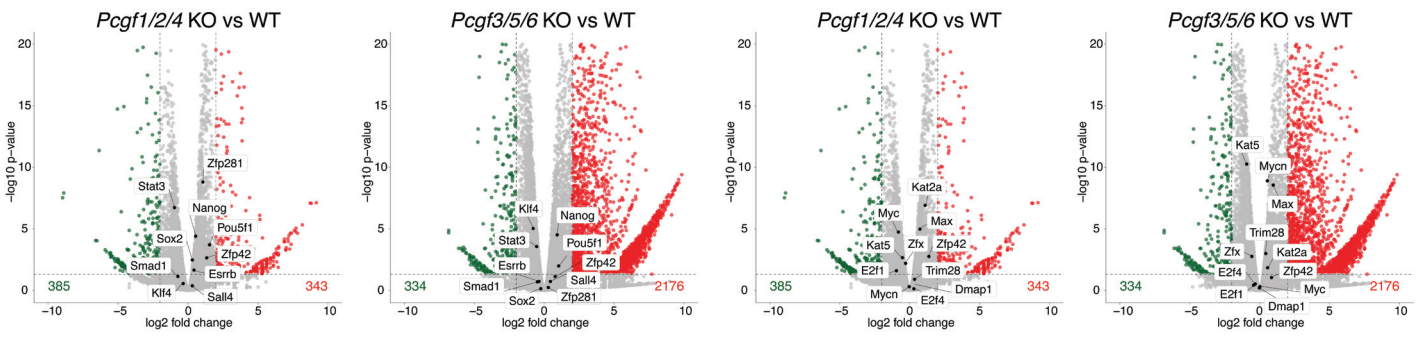


Figure S6

A

Pluripotency network

MYC Transcriptional network



B

Embryonic signature    Mesoderm signature    Endoderm signature    Ectoderm signature    Germ cell signature

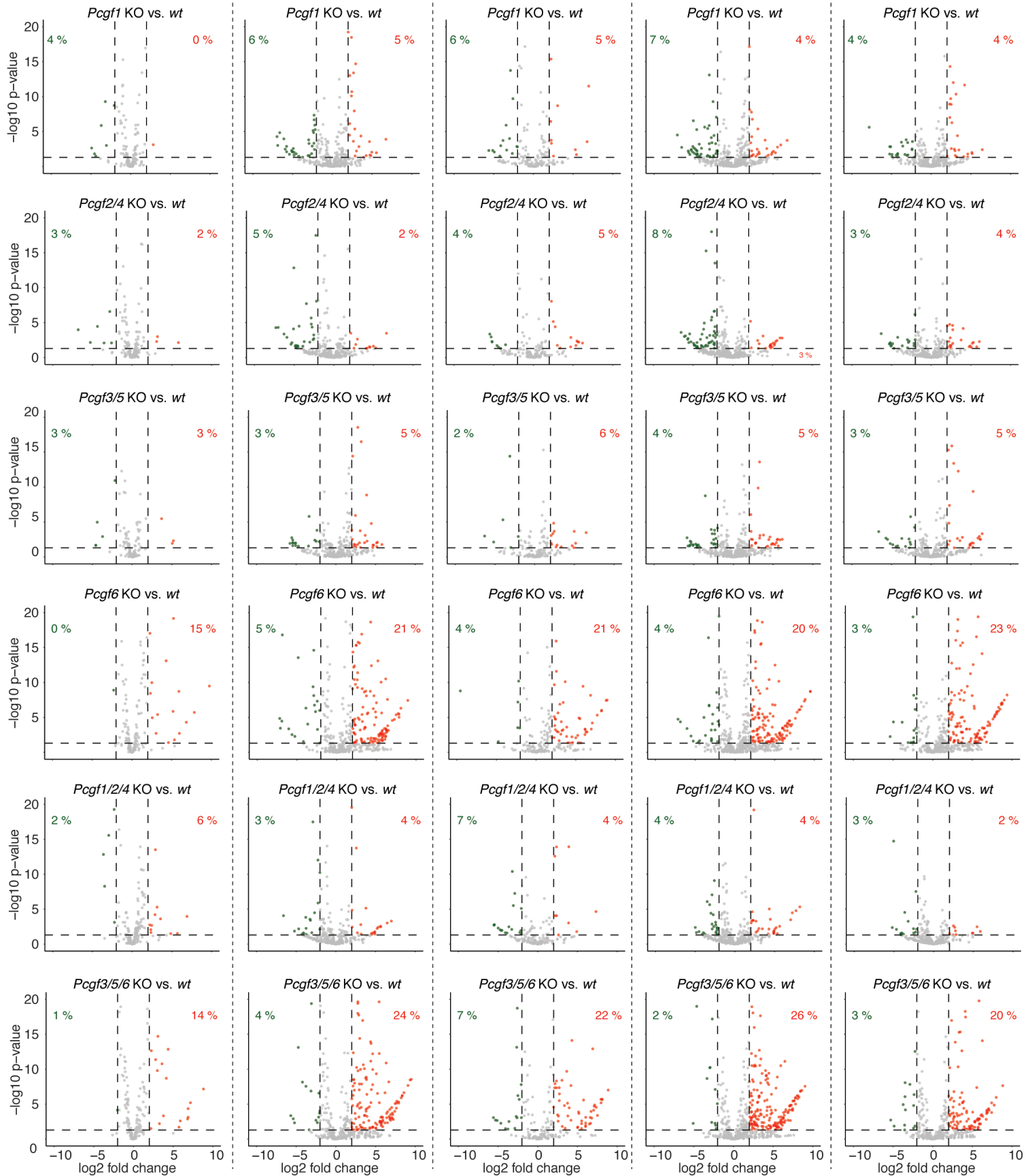
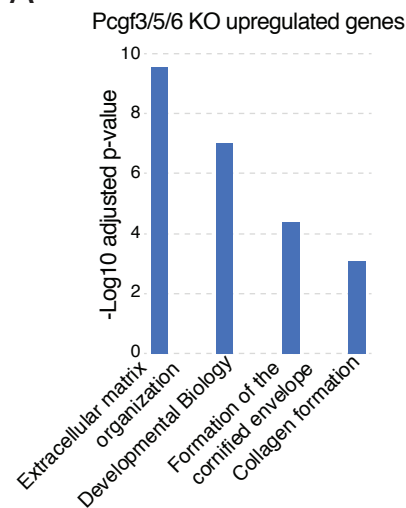
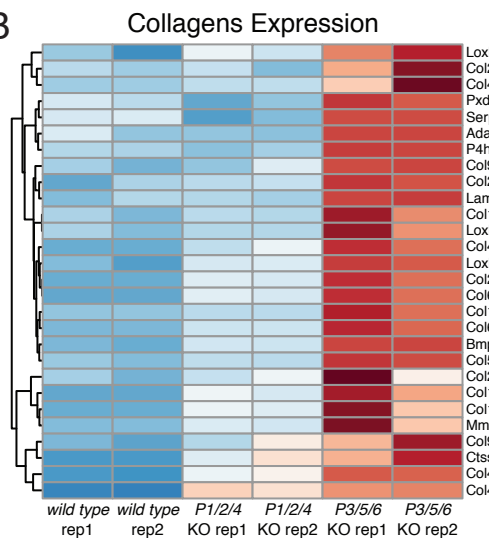


Figure S7

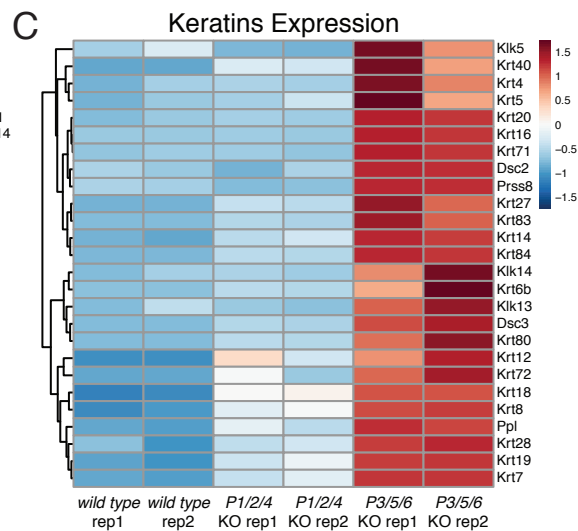
A



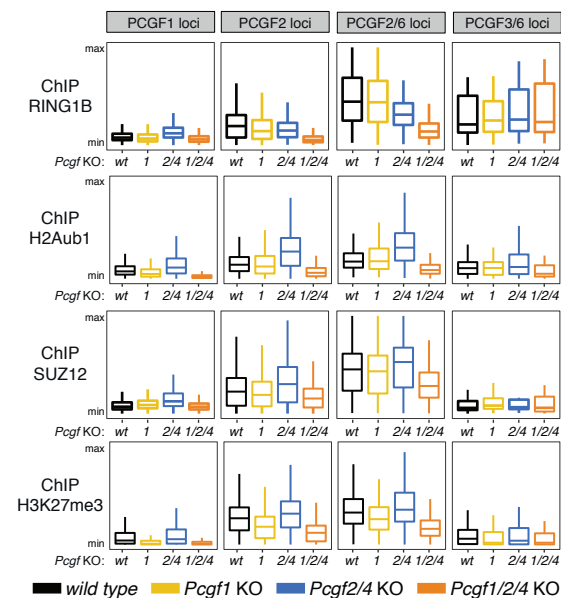
B



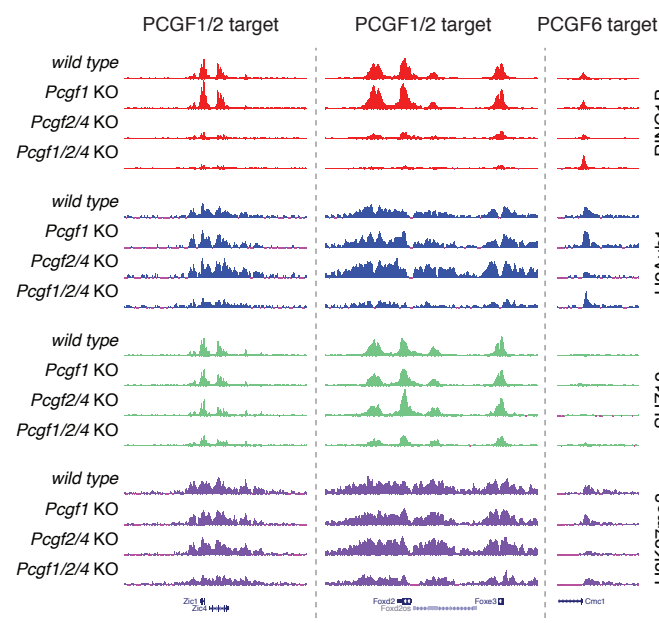
C



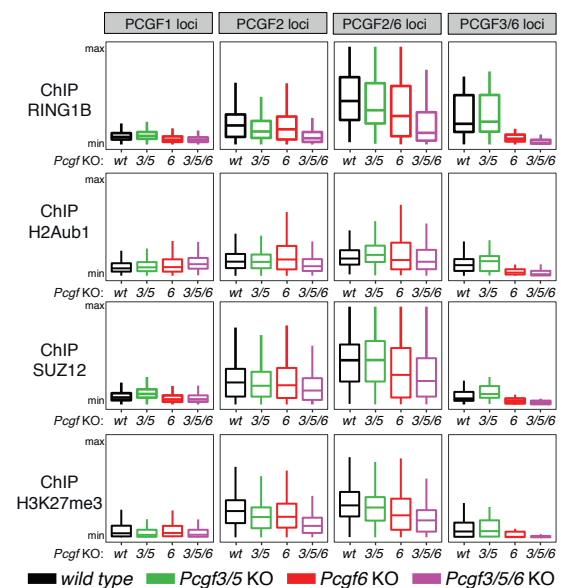
D



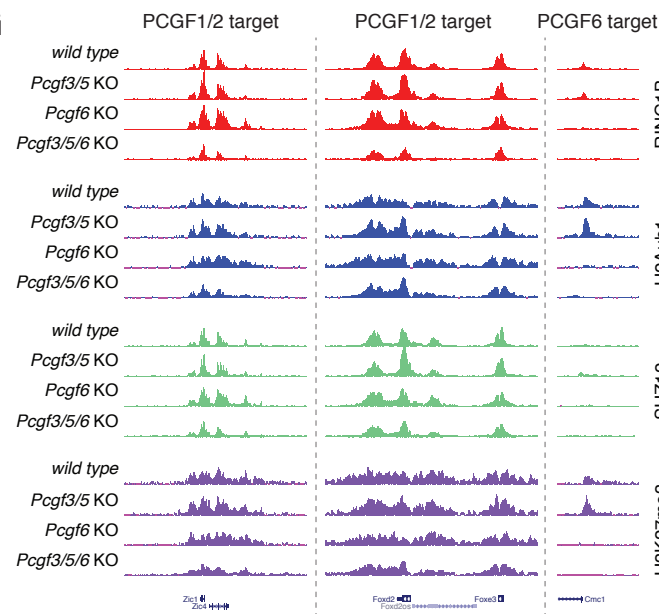
E



F



G



H

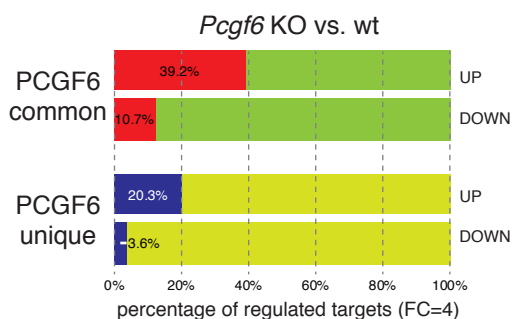
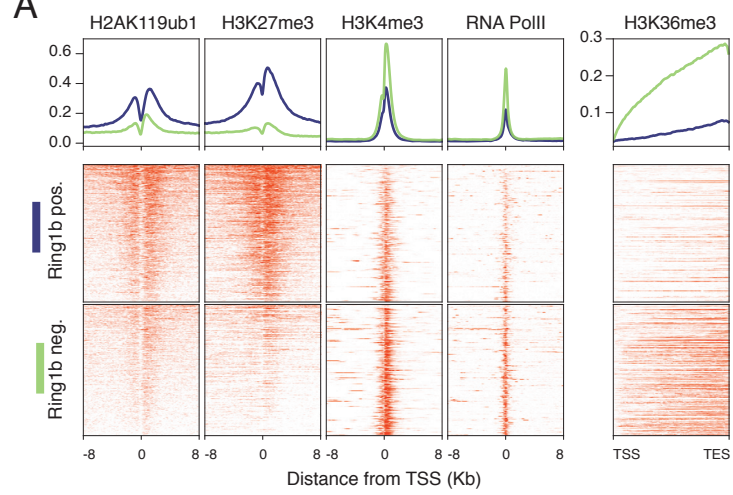


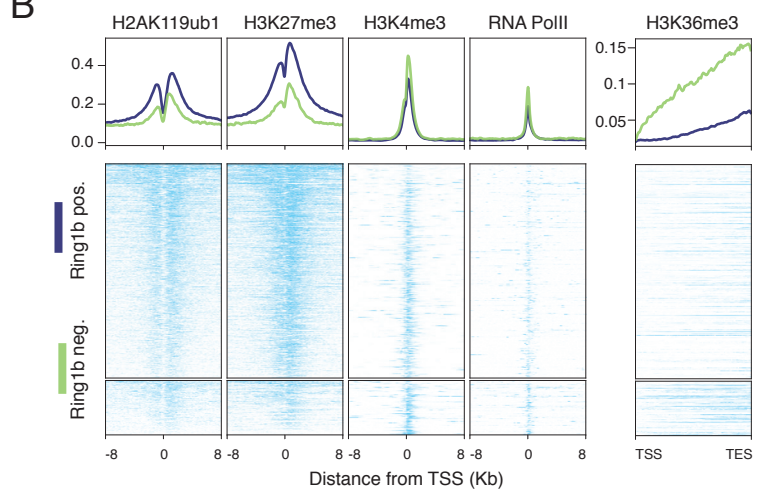


Figure S8

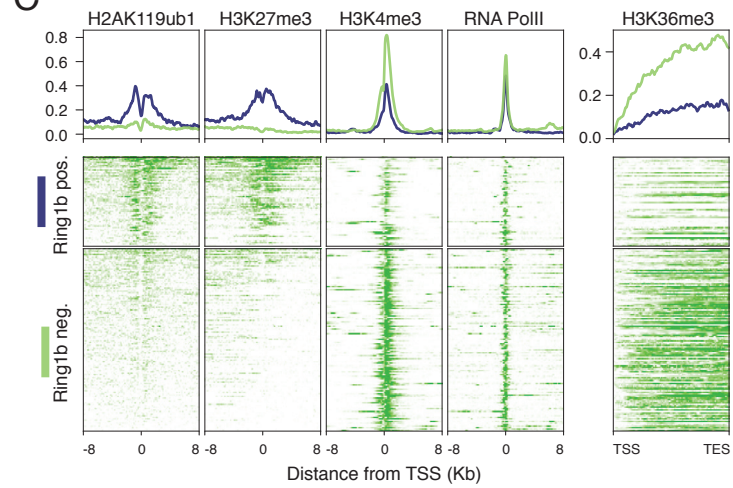
PCGF1 target loci



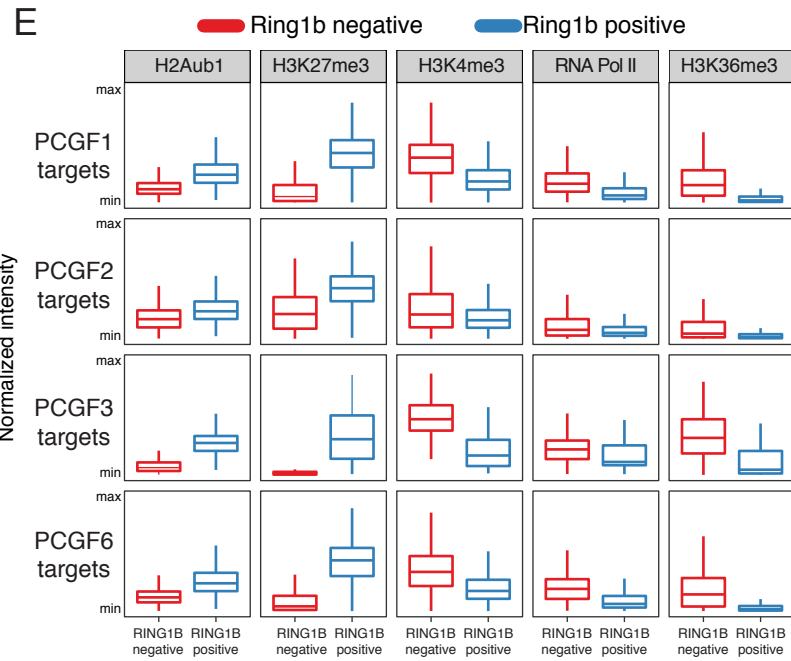
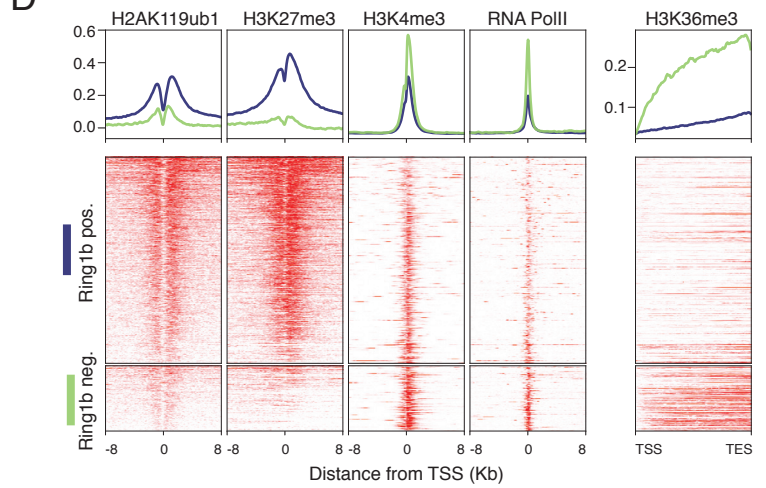
PCGF2 target loci



PCGF3 target loci



PCGF6 target loci



**F** *Ring1a* KO - *Ring1b* FL

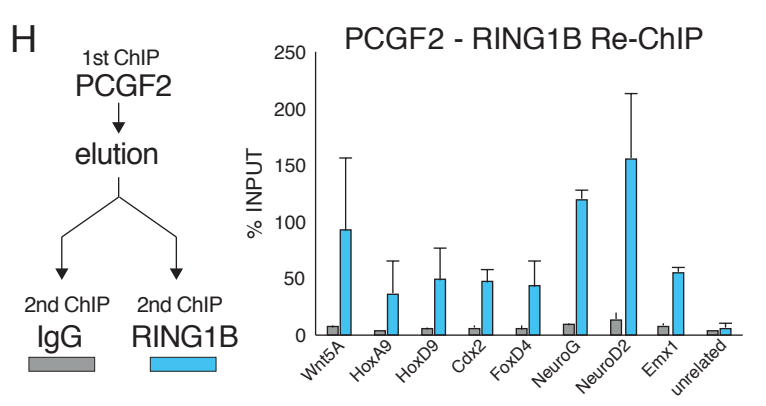
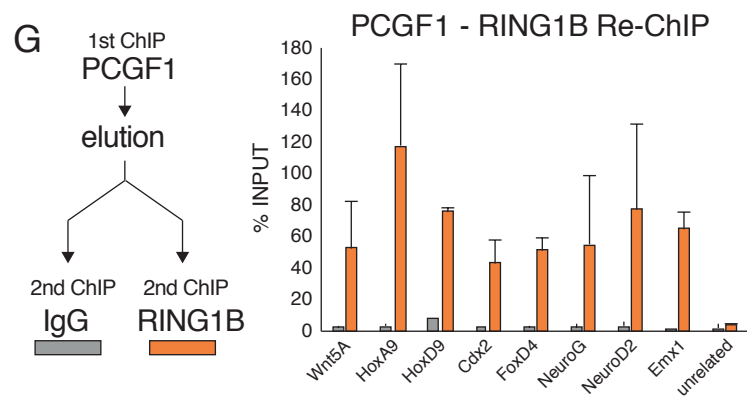
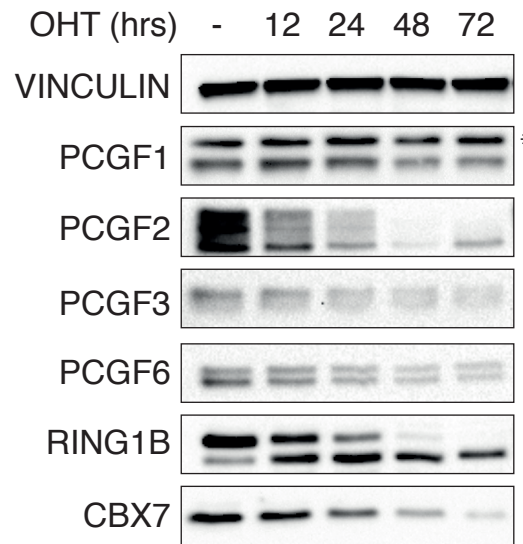


Figure S9

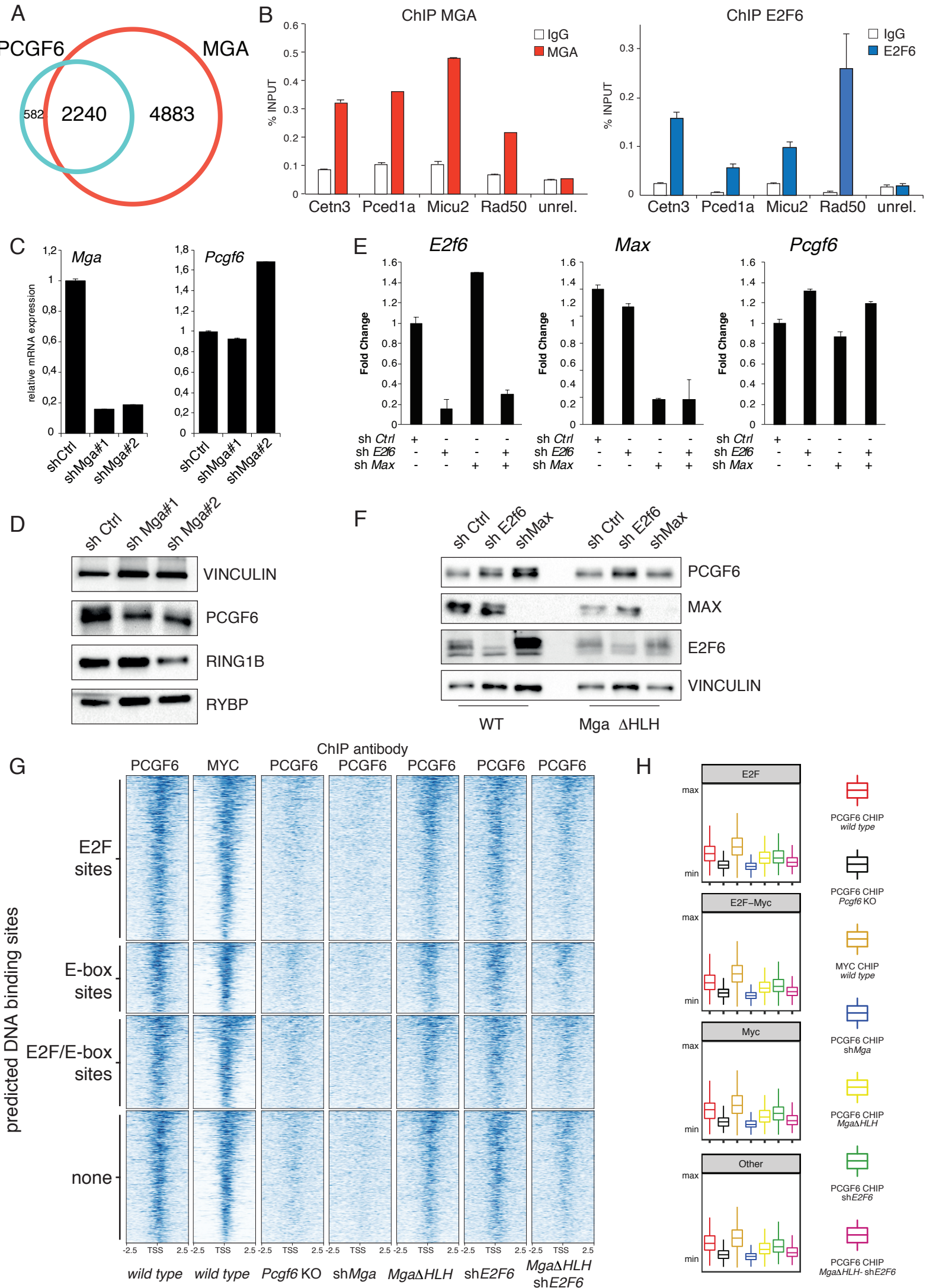
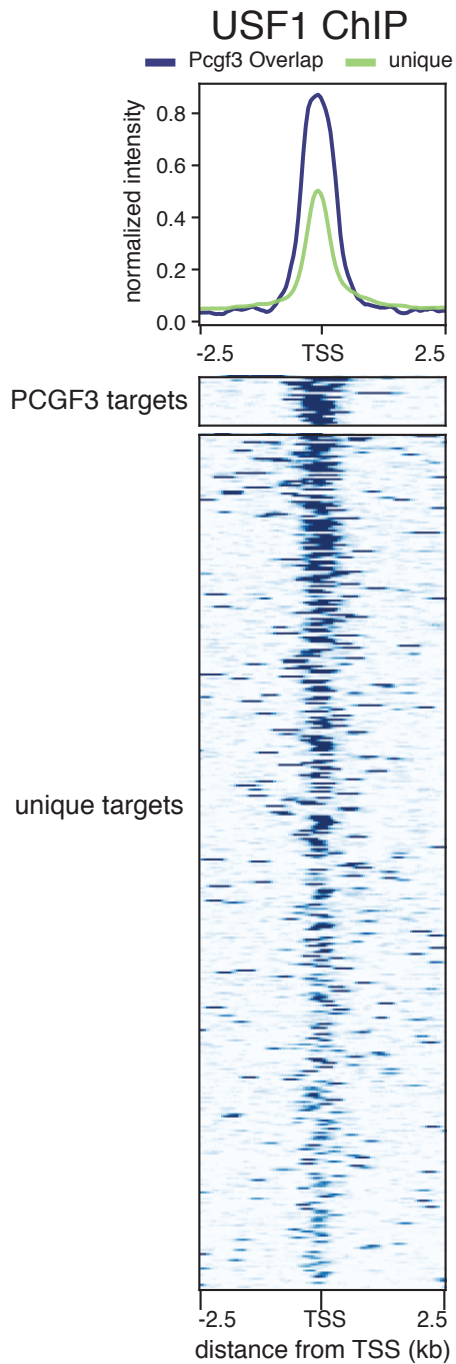
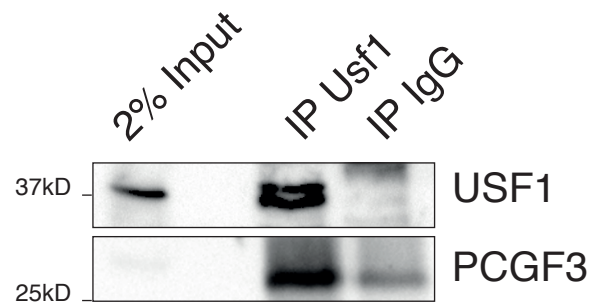


Figure S10

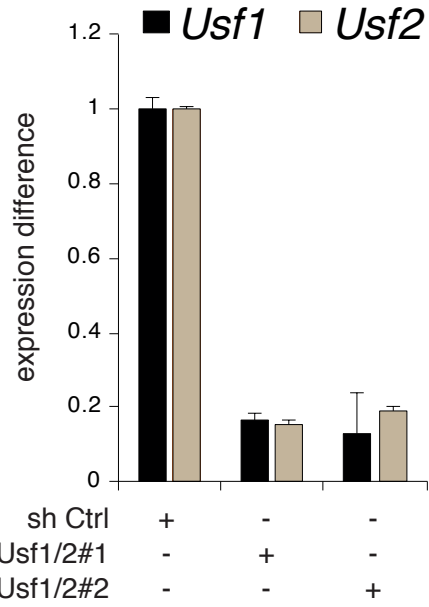
A



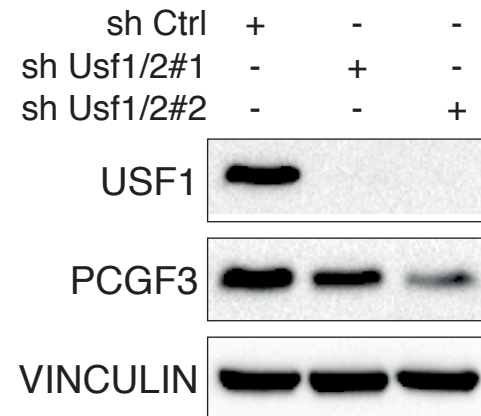
B



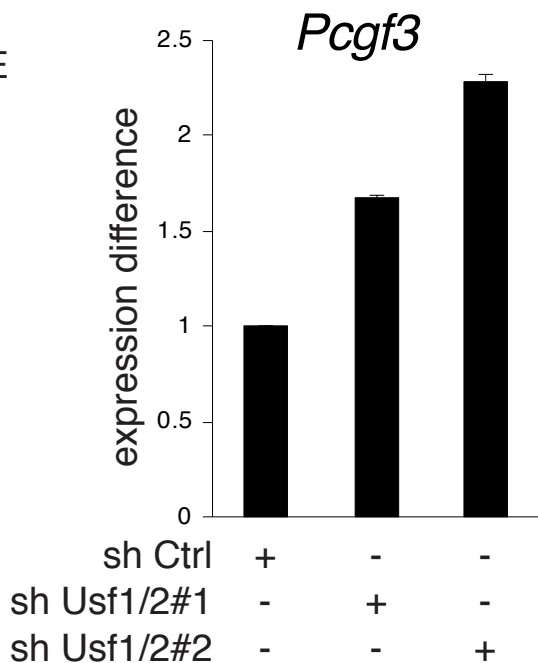
C



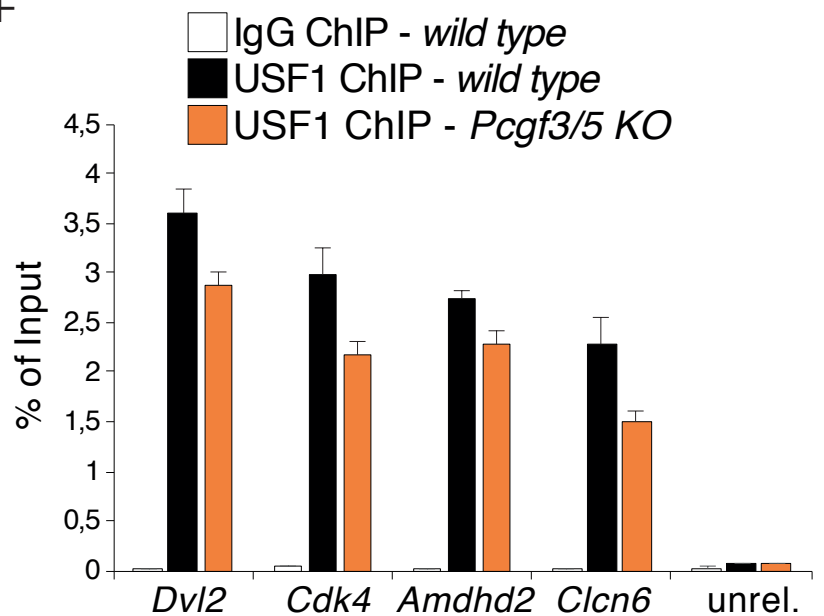
D



E



F



## Supplemental Figure Legends

### Figure S1. Genome-wide PCGF Binding Specificity. Related to Figure 1

- (A) RT-qPCR analysis showing the mRNA expression levels of the indicated *Pcgf* genes in WT mESC relative to *Ring1b* expression. *Gapdh* served as normalizing expression control. Data represent mean  $\pm$  SEM.
- (B) Relative abundance based on intensity Based Absolute Quantification (iBAQ) values of PCGF peptides identified in RING1B-Flag-bio tandem purifications coupled to mass-spectrometry experiment.
- (C) ChIP-qPCR analysis at the *Gal4*-TK-Luciferase promoter of ChIP experiments in the indicated 293Trex clones expressing inducible Gal4 (empty) or Gal4-PCGF fusion proteins, comparing independently the efficiency of the GAL4 to the PCGF1, PCGF2, PCGF3 or PCGF6 antibodies. Rabbit IgG served as a negative control. ChIP enrichments are presented as percentage of input. Data represent mean  $\pm$  SEM.
- (D) Uncropped western blot images of immunoblotting with PCGF6 antibody on WT and *Pcgf6* KO mESC extracts.
- (E) Observed/expected ratio of CpG dinucleotides contained within the peaks of the indicated ChIP-seq analyses. Ratio of 0.6 (black-dashed line) or higher is associated with the presence of CpG islands.
- (F) Average width of the peaks identified in the indicated ChIP-seq profiles.
- (G) Venn diagram showing the overlap between the target genes for the different indicated PCGF proteins.

### Figure S2. Genomic and Proteomic Landscape of PRC1 Family Complexes. Related to Figure 2

- (A) Boxplots of the input subtracted normalized intensity profiles of ChIP-seq analyses for the indicated proteins and histone modifications over  $\pm 4$  kb of the indicated loci stratified for PCGF co-occupancy in wild-type WT mESCs. CBX7 and RYBP datasets from mESCs were obtained from (Morey et al., 2013) and KDM2B dataset from (Farcas et al., 2012).
- (B) Heatmap of PRC1-associated complexes. The GAL4-PCGF (1-6) PRC1 subunits indicated at the top of the table were immunoprecipitated and the recovered proteins were subjected to Liquid Chromatography-Mass Spectrometry (LC-MS). The LFQ (Label Free quantification) ratio is color-coded and displayed as a heatmap with color intensity representing the degree of enrichment obtained in every specific immunoprecipitation. Proteins are grouped in accordance to which GAL4-PCGF they associate with.

(C) Gene ontology analysis for the indicated PCGF unique and co-occupied target genes. The most represented categories are highlighted. Dot size is proportional to the number of genes corresponding to that gene ontology category, color scale indicates statistical significance (adjusted p-value < 0.01 and q-value < 0.01).

**Figure S3. Functional Cross-talk Between Different PCGF Proteins. Related to Figure 3**

(A, B) Western blot analyses with the indicated antibodies of protein lysates from WT and indicated *Pcgf* KO mESC clones. Vinculin served as a loading control. \*, unspecific signal.

(C) Heatmaps of the input subtracted normalized intensity profiles of ChIP-seq analyses for PCGF1, PCGF2, PCGF3, or PCGF6 performed in WT mESCs, *Pcgf1*, *Pcgf2/4*, *Pcgf3/5*, or *Pcgf6* KO ESC clones. The regions plotted correspond to  $\pm 4$  kb around TSS of unique and co-occupied target genes, as indicated.

(D) Boxplots of the input subtracted normalized intensity profiles from two distinct biological replicates of ChIP-seq analyses for PCGF1, PCGF2, PCGF3, or PCGF6 performed in WT, *Pcgf1*, *Pcgf2/4*, *Pcgf3/5*, or *Pcgf6* KO mESC clones. Signal enrichment was calculated over  $\pm 500$  bp around the TSS of the indicated loci stratified for PCGF co-occupancy in WT mESCs.

**Figure S4: Chip-Seq Normalization by Spike-in RX Human. Related to Figure 3**

(A) Heatmaps of the spike-in input subtracted normalized intensity profiles of ChIP-seq analyses for PCGF1, PCGF2, or PCGF6 performed in WT mESCs, *Pcgf1*, *Pcgf2/4*, *Pcgf3/5*, or *Pcgf6* KO ESC clones. The regions plotted correspond to  $\pm 4$  kb around TSS of unique and co-occupied target genes, as indicated.

(B) Boxplots of the spike-in input subtracted normalized intensity profiles of ChIP-seq analyses for PCGF1, PCGF2, or PCGF6 performed in WT mESCs, *Pcgf1*, *Pcgf2/4*, *Pcgf3/5*, or *Pcgf6* KO ESC clones. The regions plotted correspond to  $\pm 500$  bp around TSS of unique and co-occupied target genes, as indicated.

(C) Heatmaps of the input subtracted normalized intensity profiles of ChIP-seq analyses for RING1B, H2Aub1, SUZ12, or H3K27me3 performed in WT mESCs, *Pcgf1*, *Pcgf2/4*, *Pcgf3/5*, or *Pcgf6* KO ESC clones. The regions plotted correspond to  $\pm 4$  kb around TSS of unique and co-occupied target genes, as indicated.

**Figure S5. Phenotypic and Transcriptomic Analyses of *Pcgf* KO mESC. Related to Figure 4.**

(A, B) Western blot analyses with the indicated antibodies of protein lysates from WT and indicated *Pcgf* KO mESC clones. Vinculin served as a loading control. \*, unspecific signal.

(C) Representative phase-contrast field of WT, *Pcgf1/2/4* and *Pcgf3/5/6* triple KO mESCs. Scale bars correspond to 200  $\mu$ m.

(D) Gene ontology analysis for the differentially expressed genes in the indicated *Pcgf* KO. The most represented categories are highlighted. Dot size is proportional to the number of genes corresponding to that Gene ontology category, color scale indicates statistical significance (adjusted p-value < 0.01 and q-value < 0.01).

(E) Volcano plots of  $-\log_{10}$  (P-value) against  $\log_2$  fold change representing the differences in gene expression between *Pcgf1*, *Pcgf2/4*, *Pcgf3/5*, or *Pcgf6* KO mESC clones and WT for all protein coding genes (upper panels) and for PCGF1, PCGF2, PCGF3 and PCGF6 targets, respectively (bottom panels).

**Figure S6. Transcriptomic Analysis of Pluripotency and Differentiation Signatures in *Pcgf* KO mESC. Related to Figure 4.**

(A) Volcano plots of  $-\log_{10}$  (P-value) against  $\log_2$  fold change representing pluripotency related genes, MYC transcriptional network genes from RNA-seq analyses performed in WT and *Pcgf1/2/4* or *Pcgf3/5/6* triple KO mESCs.

(B) Volcano plots of  $-\log_{10}$  (P-value) against  $\log_2$  fold change of genes from multiple embryonic layers signatures obtained from (Hutchins et al., 2017) from RNA-seq analyses performed in WT and *Pcgf1*, *Pcgf2/4*, *Pcgf3/5*, *Pcgf6*, *Pcgf1/2/4* or *Pcgf3/5/6* KO mESCs.

**Figure S7. Analysis of Polycomb Signature at PCGF Targets in Triple KO mESCs. Related to Figure 4.**

(A) Gene ontology enrichments based on adjusted p-values ( $-\log_{10}$  p) of genes upregulated in *Pcgf3/5/6* triple KO mESCs.

(B) Z-score expression heatmaps for collagen related genes from RNAseq analyses performed in WT, *Pcgf1/2/4* or *Pcgf3/5/6* triple KO mESCs

(C) As in (B) for several indicated keratins related genes.

(D) Boxplots of the input subtracted normalized ChIP-seq intensity profiles of RING1B, H2AK119ub1 (H2Aub1), SUZ12, H3K27me3 performed in WT and *Pcgf1*, *Pcgf2/4* and *Pcgf1/2/4* KO mESC clones. Signal enrichment is calculated over  $\pm 500$  bp ( $\pm 4$  kb for H2Aub1 and H3K27me3) around the TSS of the indicated group of PCGF target genes.

(E) Genomic snapshots of the ChIP-seq profiles quantified in (A) at selected PCGF1/2 common and PCGF6 unique target gene loci, performed in WT and in the indicated KO mESCs.

(F) Boxplots of the input subtracted normalized ChIP-seq intensity profiles of RING1B, H2AK119ub1 (H2Aub1), SUZ12, H3K27me3 performed in WT, *Pcgf6*, *Pcgf3/5*, and *Pcgf3/5/6* KO mESC clones. Signal enrichment is calculated over  $\pm 500$  bp ( $\pm 4$  kb for H2Aub1 and H3K27me3) around the TSS of the indicated group of PCGF target genes.

(G) Genomic snapshots of the ChIPseq profiles quantified in C at selected PCGF1/2 common and PCGF6 unique target gene loci performed in WT and the indicated KO mESCs.

(H) Percentage of deregulated PCGF6 targets genes co-occupied by RING1B (first two bars) or PCGF6 unique respect to RING1B (last two bars).

**Figure S8. PCGF3 and PCGF6 Recruit Modest RING1B Activity and Present Features of Active Transcription. Related to Figure 5.**

(A–D) Heatmap analysis representing the input subtracted normalized ChIP-seq intensities over  $\pm 8$  kb centered at TSS of PCGF1 (A), PCGF2 (B), PCGF3 (C), PCGF6 (D) target loci stratified for RING1B co-occupancy in WT mESCs. H3K36me3 intensities are shown along the entire gene length (from TSS to TES).

(E) Boxplots of the input subtracted normalized intensity profiles of the indicated ChIPseq analyses performed in A–D. Signal enrichment is calculated over  $\pm 8$  kb around the TSS of the indicated group of PCGF target genes.

(F) Western blot analysis with the indicated antibodies of protein lysates prepared from *Ring1A*<sup>-/-</sup>; *Ring1B*<sup>fl/fl</sup>; *Rosa26::CreERT2* (R1A KO-R1B FL) mESC treated with 4-OHT at the indicated time points. Vinculin served as loading control. \*, unspecific signal.

(G, H) qPCR analysis at selected target regions of sequential RING1B ChIP (re-ChIP) experiments upon PCGF1 (G) or PCGF2 (H) ChIP in WT mESC. IgG served as control for ChIP assay. ChIP enrichments are

normalized to input. Data represent mean  $\pm$  SEM.

**Figure S9. Loss of E2F6, MGA, or MAX Does Not Affect PCGF6 Expression. Related to Figure 6.**

(A) Venn diagram of PCGF6 (blue) and MGA (red) target genes.

(B) ChIP-qPCR validation of MGA and E2F6 binding at PCGF6 targets. IgG served as control for ChIP assay. ChIP enrichments are normalized to input. Data represent mean  $\pm$  SEM.

(C) RT-qPCR analysis of *Mga* and *Pcgf6* mRNA levels in WT mESC expressing scramble (shCtrl) or specific shRNAs targeting *Mga*. *Gapdh* served as normalizing expression control. Fold change over shCtrl is shown. Data represent mean  $\pm$  SEM.

(D) Western blot analysis with the indicated antibodies of protein lysates from cells described in C. Vinculin served as loading control.

(E) RT-qPCR analysis of *E2f6*, *Max* and *Pcgf6* mRNA levels in WT mESC expressing scramble (sh Ctrl) or shRNAs specifically targeting *E2f6* and *Max* in the indicated combinations. *Gapdh* served as normalizing expression control. Fold change over sh Ctrl is shown. Data represent mean  $\pm$  SEM.

(F) Western blot analysis with the indicated antibodies of protein lysates from WT and *Mga* $\Delta$ HLLH mESC mutant expressing scramble (sh Ctrl) or shRNAs targeting specifically *E2f6* and *Max*. Vinculin served as loading control.

(G) Heatmaps of input subtracted normalized intensity profiles of MYC in WT mESC and PCGF6 binding in WT or *Pcgf6* KO, a sh*Mga*, *Mga* $\Delta$ HLLH mutant, sh*E2f6*, and a sh*E2f6* in *Mga* $\Delta$ HLLH mutant mESCs around  $\pm$ 2.5 kb of the TSS of PCGF6 targets stratified by the presence of E2F, E-box, E2F and E-box and unrelated binding motifs predicted by FIMO.

(H) Boxplots of the input subtracted normalized intensity profiles of ChIP-seq analyses performed in (G) over  $\pm$ 500 bp respect to the TSS.

**Figure S10. USF1 Interacts with PCGF3, which is not Affected by USF1 Loss. Related to Figure 7.**

(A) Input subtracted normalized intensity profiles and heatmap of USF1 binding in WT mESCs around  $\pm$ 2.5 kb of the TSS of USF1-Pcgf3 common and USF1 unique target loci.



(B) Western blot analysis with the indicated antibodies of USF1 immuno-precipitations using WT mESC nuclear extracts. Mouse IgG served as unrelated antibody. Input is shown as loading control.

(C) RT-qPCR analysis of *Usf1* and *Usf2* mRNA levels in WT mESC expressing scramble (sh Ctrl) or shRNAs targeting specifically *Usf1* and *Usf2* in the indicated combinations. *Gapdh* served as normalizing expression control. Fold-change over sh Ctrl is shown. Data represent mean  $\pm$  SEM.

(D) Western blot analysis with the indicated antibodies of protein lysates from cells described in (C). Vinculin served as loading control.

(E) RT-qPCR analysis of *Pcgf3* mRNA levels in WT mESC expressing scramble (sh Ctrl) or shRNAs specifically targeting *Usf1* and *Usf2* in the indicated combinations. *Gapdh* served as normalizing expression control. Fold change over sh Ctrl is shown. Data represent mean  $\pm$  SEM.

(F) USF1 ChIP-qPCR analysis at selected PCGF3 target regions performed in WT and *Pcgf3/5* KO mESCs. IgG served as a negative control for ChIP. ChIP enrichments are normalized to input. Data represent mean  $\pm$  SEM.

## **Supplemental Table Legends**

### **Table S1. shRNA and sgDNA sequences. Related to Star Methods**

List of oligo sequences used for RNAi and CRISPR/cas9 manipulations of mESCs.

### **Table S2. PCGF Enriched Genomic Regions. Related to Figure 1**

List of peaks called for the different PCGF ChIP in WT mESCs.

### **Table S3. PCGF Target genes. Related to Figure 1**

List of target promoters bound by each PCGF protein in WT mESCs.

### **Table S4. Gal4-PCGF fusion proteins MS results. Related to Figure 2**

Values of the LFQ ratios of the different PRC1.1-6 components obtained by MS/MS analyses in the Gal4-PCGF1-6 fusions immuno-purifications (anti-Gal4) from 293TRex cells.

### **Table S5. Gene Expression in PCGF KO ESC. Related to Figure 4**

Expression values of all RefSeq genes in the different PCGF KO mESC lines.

### **Table S6. Gene Ontology of Differentially Expressed Genes. Related to Figure 4**

List of significant GO terms enriched in differentially expressed genes in the different PCGF KO mESC lines.