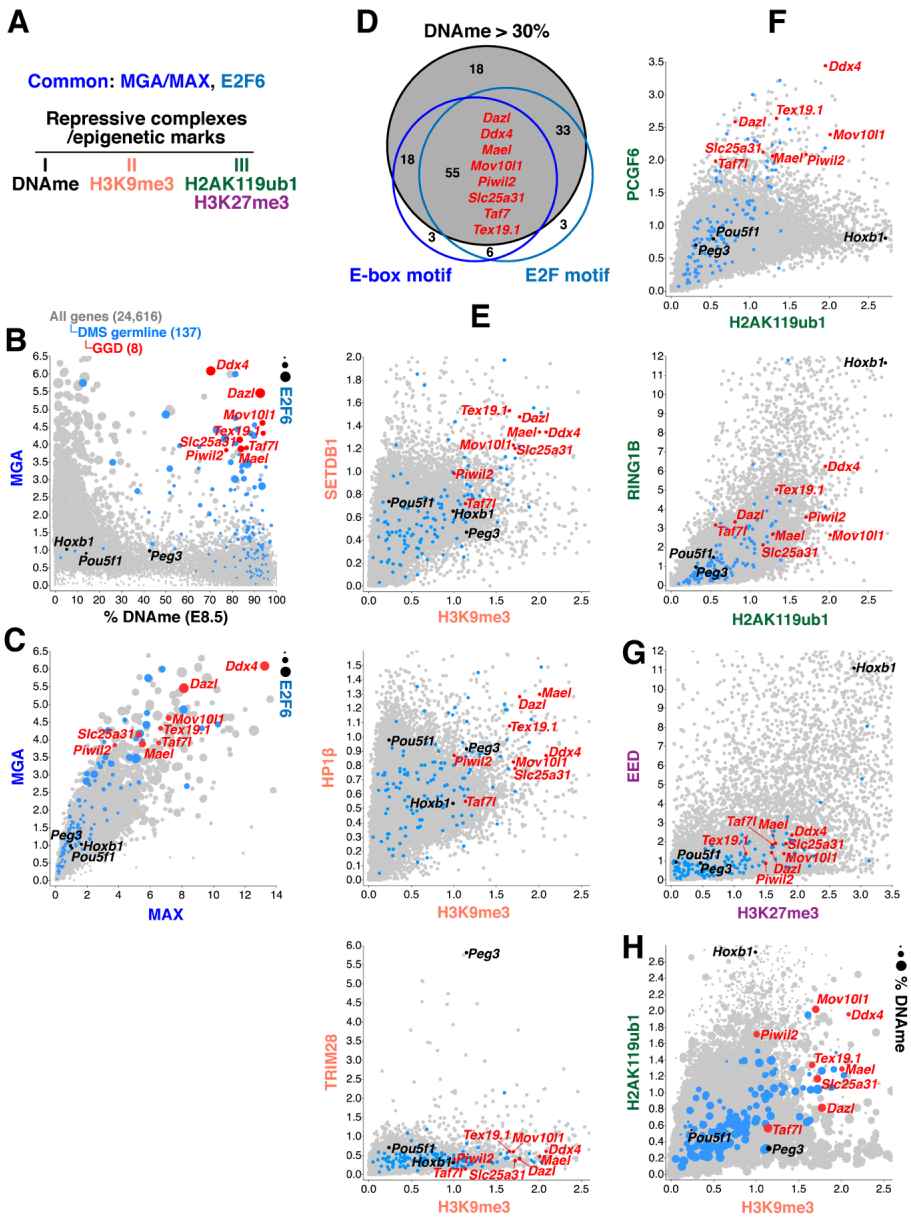
Supplementary Information

Repression of germline genes by PRC1.6 and SETDB1 in the early embryo precedes DNA methylation-mediated silencing

Mochizuki K et al.

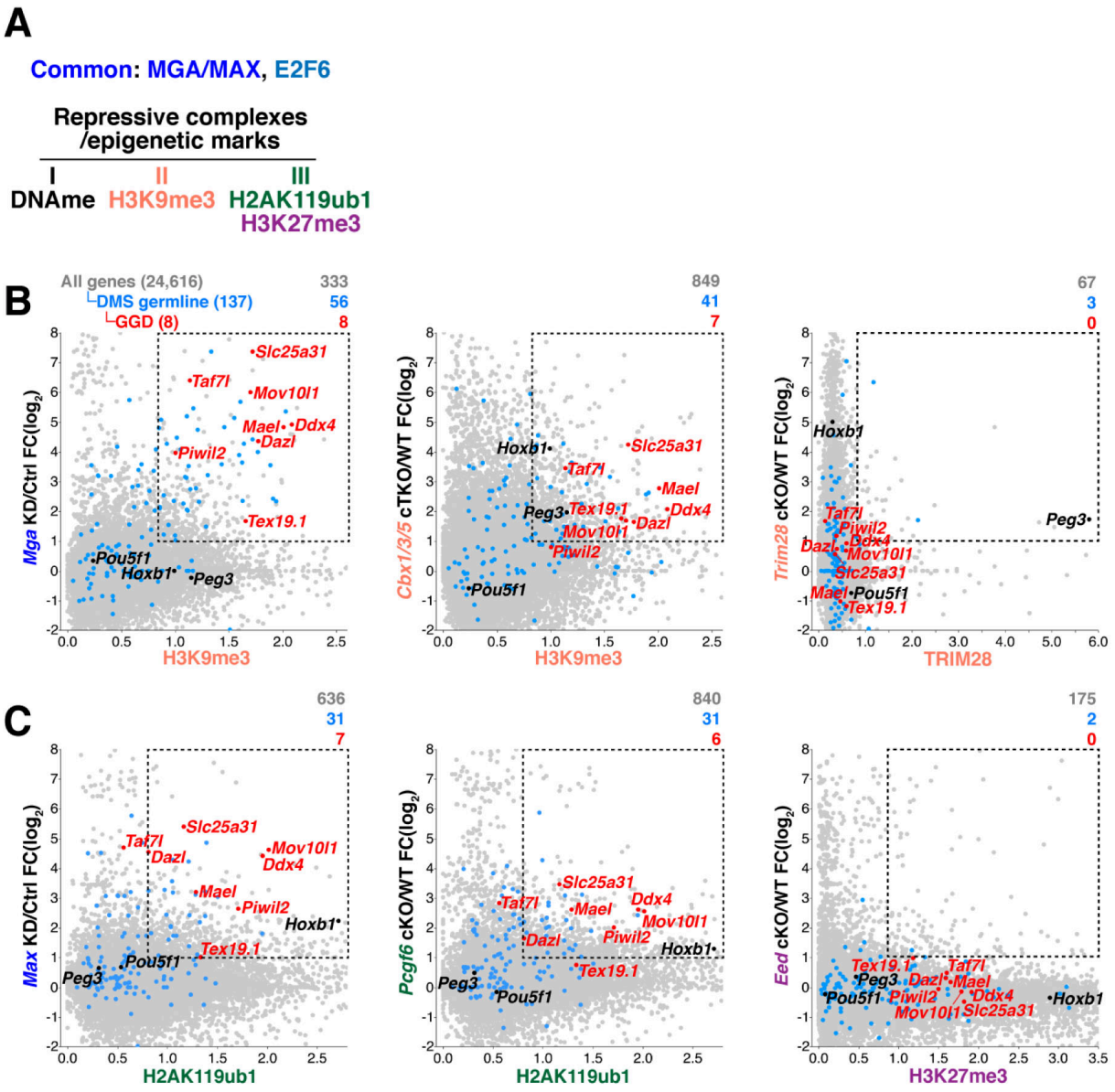
6 Supplementary Figures



Supplementary Fig. 1. Enrichment of repressor complex subunits and associated epigenetic marks at DMS germline genes in pESCs. Related to Fig. 1.

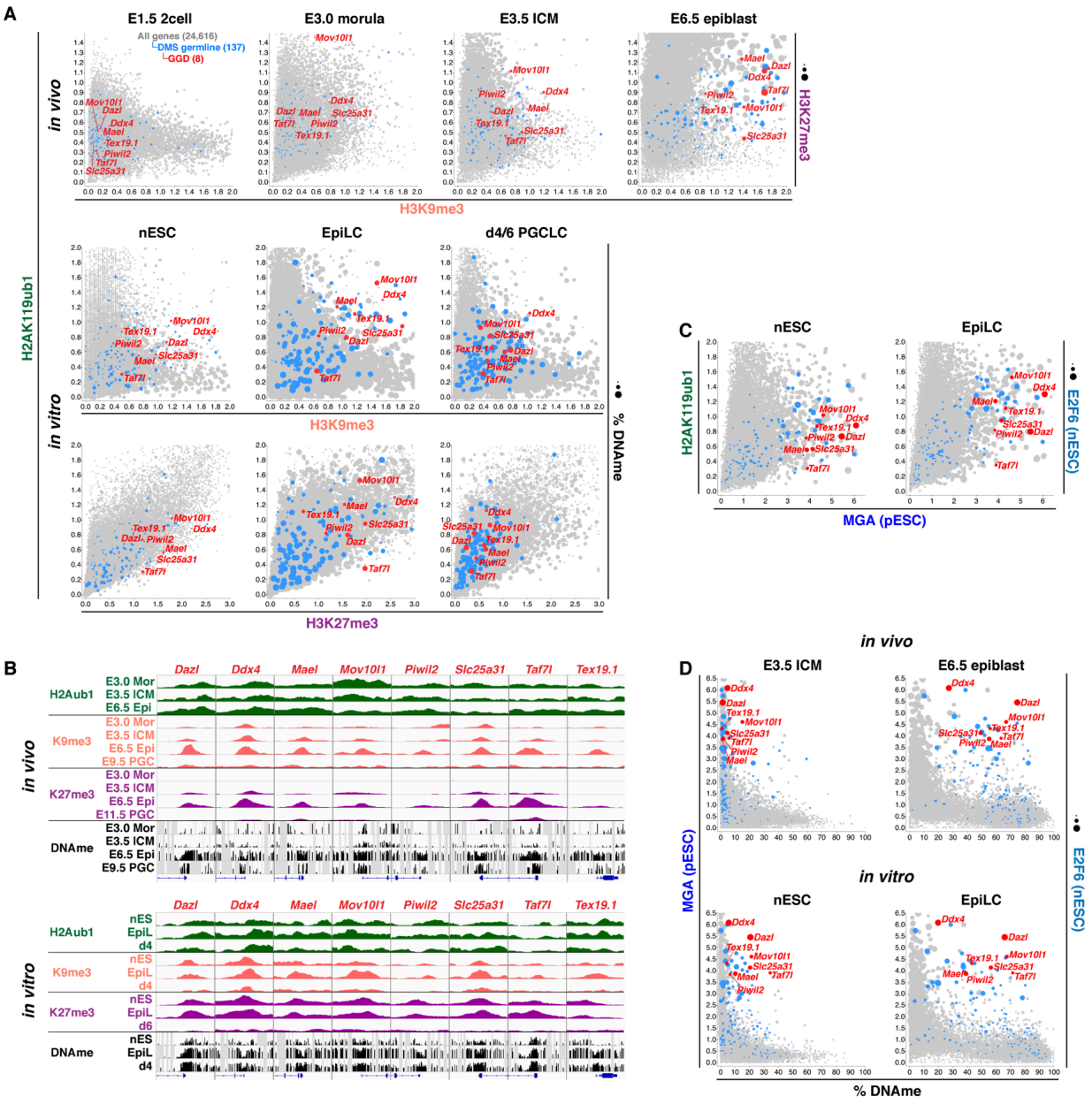
(**A**) Color codes for repressive complex subunits or the marks they deposit and gene categories (color-coded as in Fig. 1) used in panels **B-H** are shown. (**B**) Scatterplot showing the relationship between the enrichment of MGA in pESCs or E2F6 in nESCs and % DNAme (TSS -0.9/+0.4 kb) in E8.5 embryos. (**C**) Scatterplot showing the relationship between the enrichment (RPKM) of MGA, MAX and E2F6 in pESCs in genic TSS regions (+/-2 kb). (**D**) Venn diagram showing the overlap between the 124 (of 137 total) DMS germline genes showing >30% DNAme in the TSS region (-0.9/+0.4 kb) in pESCs with the presence of an E-box and/or E2F consensus motif. All GGD genes

are included among the 55 genes with both motifs. (**E-G**) Scatterplots showing the relationship between the enrichment of (**E**) repressor complex II or (**F-G**) repressor complex III subunits or associated histone PTMs in genic TSS regions (+/-2 kb). (**H**) Scatterplot showing the relationship between H2AK119ub1 and H3K9me3 enrichment, with the % DNAme (TSS -0.9/+0.4 kb) depicted by dot size. All ChIP-seq data are presented as RPKM values.



Supplementary Fig. 2. Impact of depletion of key subunits of previously implicated repressive complexes on expression of DMS germline genes in pESCs. Related to Fig. 2.

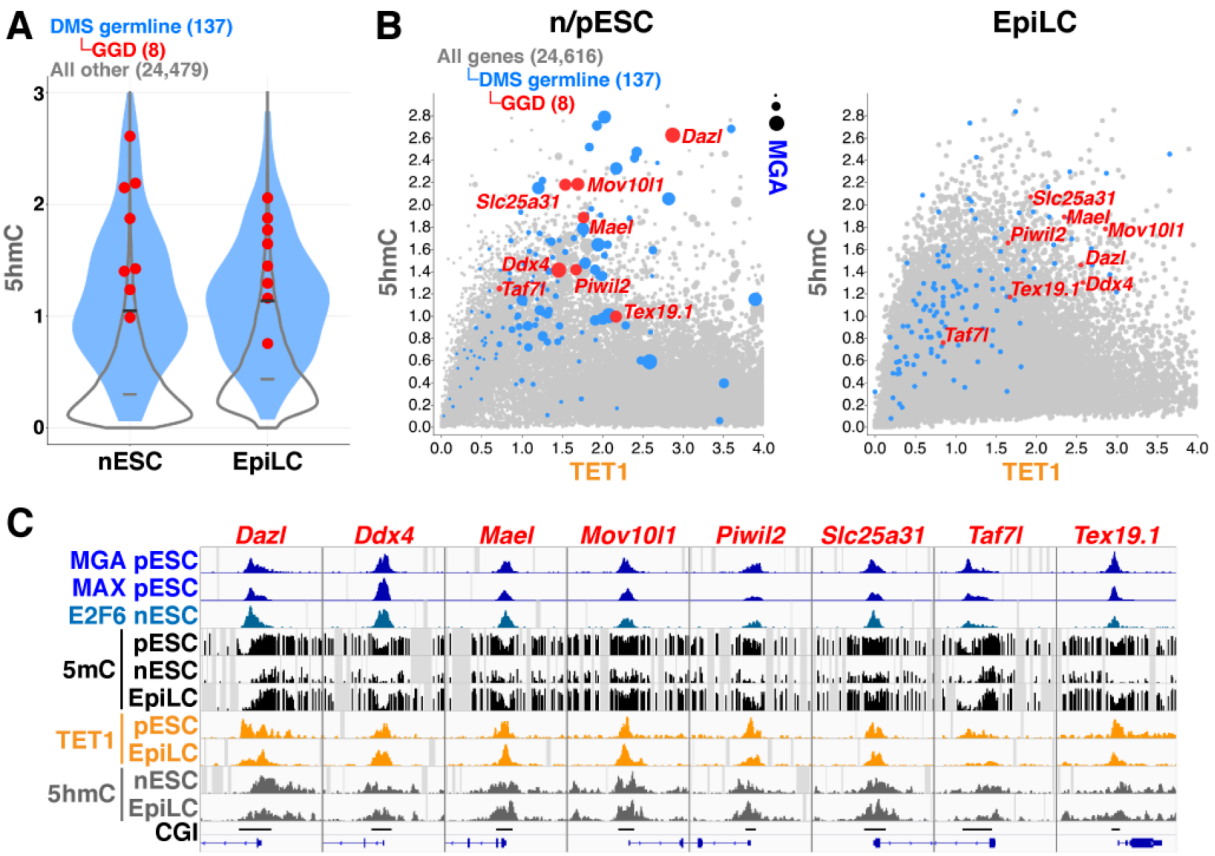
(**A**) Color codes for repressive complex subunits or the marks they deposit shown in panels **B-C**. (**B-C**) Scatterplots showing the fold-change (FC) of gene expression in pESCs following KO or KD of common factors or repressor complex subunits (y-axis) versus enrichment (RPKM) of relevant histone PTMs/subunits (TSS +/-2 kb) (x-axis). For each category of genes (color-coded as shown), the total number of genes is shown in parentheses. The number showing a >2-fold increase in expression and ChIPseq enrichment (RPKM >0.8), is also shown at the top right of each plot.



Supplementary Fig. 3. GGD genes are hypomethylated and enriched for H2AK119ub1, H3K9me3 and H3K27me3 in the pre-implantation embryo and nESCs.

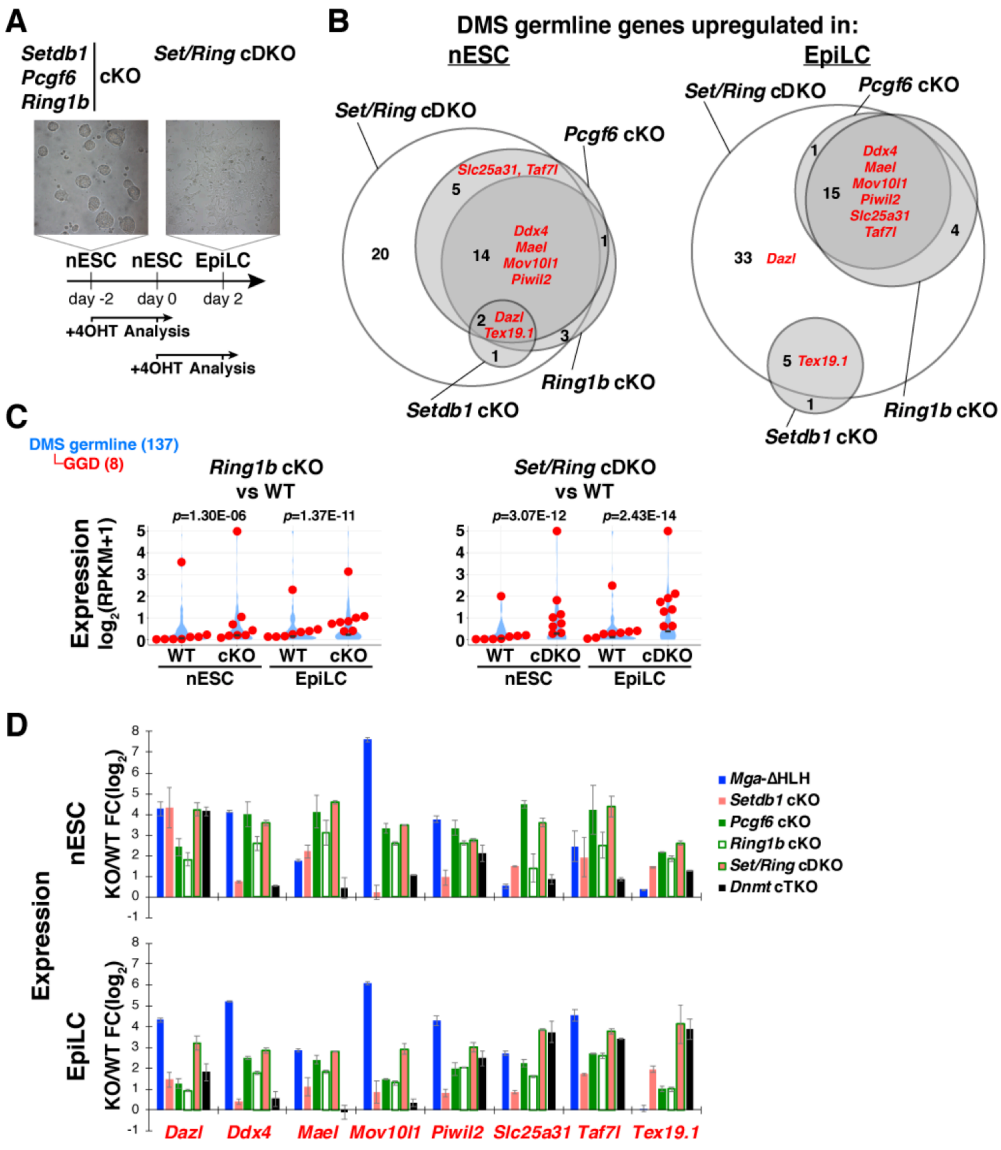
(**A**) (Top) Scatterplots showing the relationship between the enrichment (RPKM) of H2AK119ub1 and H3K9me3 during early embryonic development in regions flanking all genic promoters (TSS +/-2 kb). The range of enrichment levels (RPKM) of H3K27me3 over the same developmental stages is depicted by dot size. (Bottom) Scatterplots showing the relationship between the enrichment (RPKM) of H2AK119ub1 and H3K9me3 or H3K27me3 in nESCs, EpiLCs and d4/6 PGCLCs at all genic promoters (TSS +/-2 kb). The range of % DNAme in the promoter region (TSS -0.9 kb/+0.4 kb) over the same developmental stages is depicted by dot size. All genes, DMS germline genes and

GGD genes are color-coded as shown. (**B**) Genome browser track showing RPM values of H2AK119ub1, H3K9me3 and H3K27me3 data as well as % DNAme in the promoter regions (TSS+/-3kb) of GGD genes at the developmental time points shown. For each WGBS track, regions highlighted in grey reflect the absence of DNAme data. (**C**) Scatterplots showing the relationship between the enrichment (RPKM) of MGA (in pESCs) and H2AK119ub1 in nESCs and EpiLCs. E2F6 enrichment (in nESCs) is depicted by dot size (RPKM). (**D**) Scatterplots showing the relationship between the enrichment (RPKM) of MGA in pESCs and % DNAme (TSS -0.9/+0.4 kb) at the developmental time points shown. E2F6 enrichment (in nESCs) is depicted by dot size (RPKM).



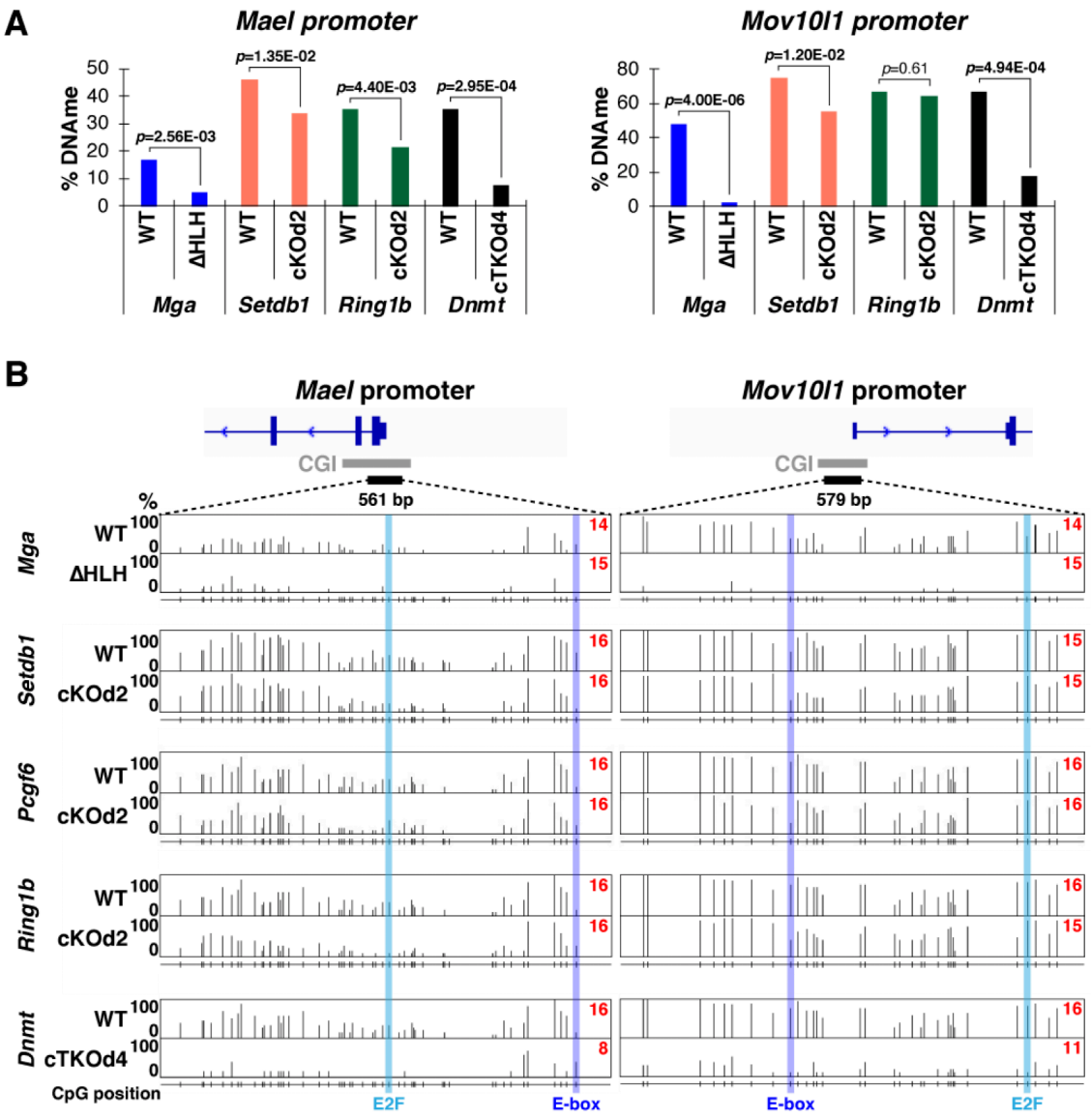
Supplementary Fig. 4. DMS germline genes are enriched for TET1 and 5hmC in nESCs and EpiLCs. Related to Fig. 4.

(**A**) Violin plots showing 5hmC (RPKM) profiles in regions flanking the promoters (TSSs +/-1 kb) of the 137 DMS germline (blue filled), 8 GGD (red data points) and all other genes (open) in nESCs and EpiLCs. (**B**) Scatterplots showing the relationship between the enrichment (RPKM) of TET1 in pESCs or EpiLCs versus 5hmC in nESCs and EpiLCs in the promoter region (TSS +/-1 kb). The range of the enrichment (RPKM) of MGA (in pESCs) is depicted by dot size. All, DMS germline and GGD genes are color-coded as shown. (**C**) Genome browser tracks of the promoter regions (TSS +/-3 kb) of GGD genes showing ChIP-seq, WGBS and 5hmCIP-seq in the indicated cells. RPM values are shown for MGA, MAX, E2F6, TET1 and 5hmC and % 5mC for DNAme. For each WGBS track, regions highlighted in grey reflect the absence of DNAme data.



Supplementary Fig. 5. Redundant roles of SETDB1 and RING1B in PRC1.6-mediated silencing of DMS germline genes in nESCs and EpiLCs. Related to Fig. 5.

(**A**) A schematic representation of the timing of 4OHT-induction of *Setdb1*, *Pcgf6*, *Ring1b* or *Setdb1* and *Ring1b* (*Set/Ring)* deletion and harvest for RNA-seq of nESCs and EpiLCs. (**B**) Venn diagram showing the overlap among DMS germline genes upregulated in nESCs or EpiLCs following cKO of the indicated factors. (**C**) Violin plots showing gene expression of the 137 DMS germline (blue filled) and 8 GGD genes (red data points) in control (WT), *Ring1b* cKO or *Set/Ring* cDKO nESCs and EpiLCs (n=2). Two-sided paired-samples *t*-tests were performed each WT/KO pair of all germline gene values. (**D**) Bar graph showing the FC (log_2) in expression of GGD genes in nESCs and EpiLCs for each of the KO lines indicated. Error bars show SE of biological replicates.



Supplementary Fig. 6. Role of MGA, SETDB1 and/or PRC1.6 in DNAme of GGD genes in EpiLCs. Related to Fig. 5.

(**A**) Bar graphs showing the mean levels of DNAme, as determined by Sanger bisulphite sequencing, in the promoter regions of GGD genes *Mael* and *Mov10l1* in *Mga*-ΔHLH, *Setdb1* cKO, *Ring1b* cKO or *Dnmt* cTKO and control (WT) EpiLCs. Mann-Whitney U-tests were performed between each mutant and WT. For *Setdb1*, *Pcgf6* and *Ring1b* lines, cultures were harvested at day 2 post 4OHT (cKOd2). (**B**) DNAme profiles of the CGI promoter regions of *Mael* and *Mov10l1* in WT and *Mga*-ΔHLH, *Setdb1* cKO, *Pcgf6* cKO, *Ring1b* cKO or *Dnmt* cTKO EpiLCs. The number of molecules sequenced for each data set is shown in red and the mean DNAme level at each CpG is shown. E-box and E2F motifs in each locus are also shown.