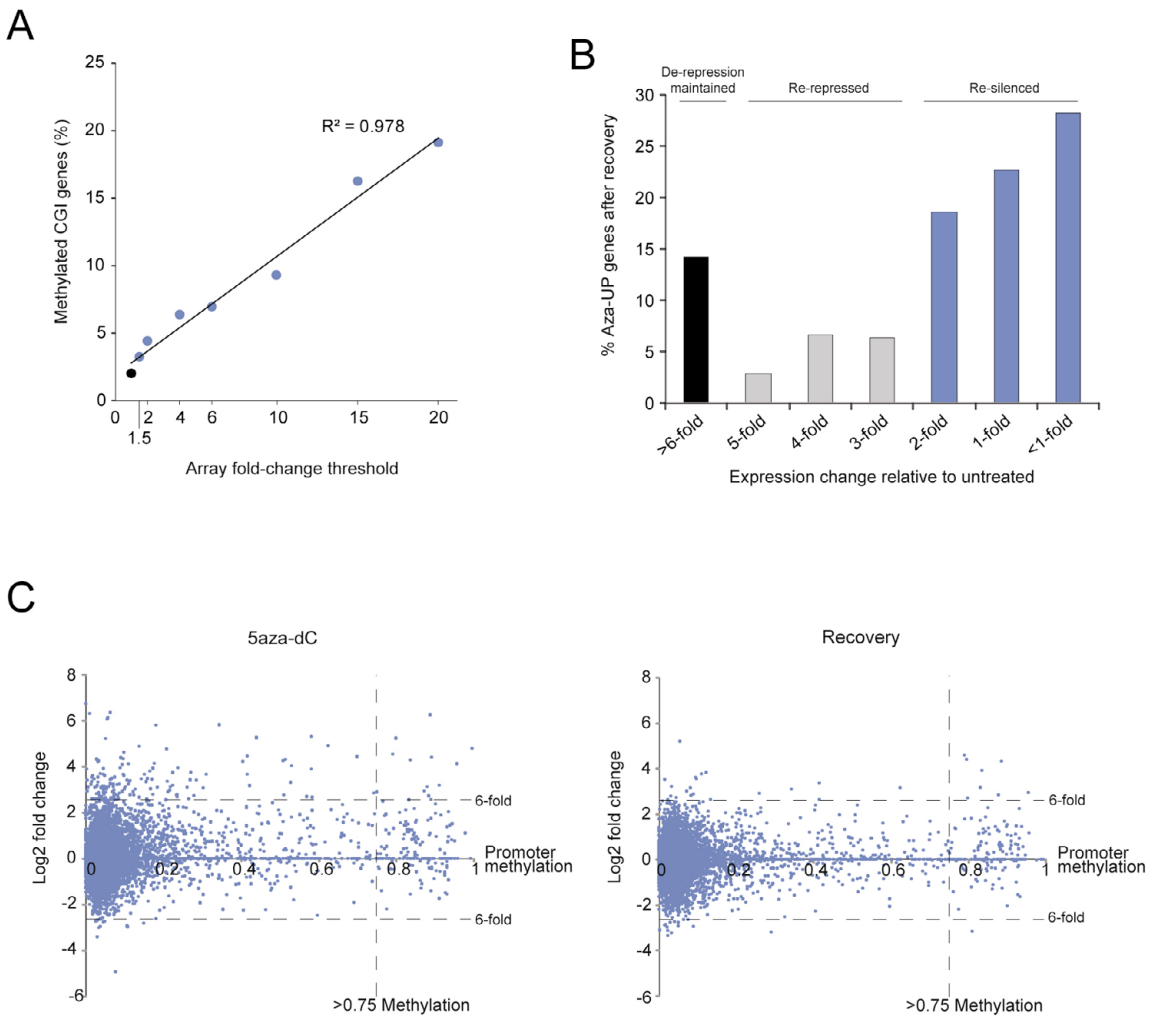

Fig. S1. Increasing the fold-change threshold stringency after 5aza-dC exposure enriches for methylated CGI promoters. (**A**) Scatter plot showing that increasing the fold-change threshold at which genes are considered to be significantly upregulated after 5aza-dC treatment (NIH/3T3) is strongly correlated with an enrichment of methylated CGI promoter genes (Meissner et al., 2008). This indicates that increasing the threshold stringency (i.e. taking the most upregulated genes) selects for somatically methylated CGI genes, which are promising methylation-dependent candidates. The black point indicates CGI promoter methylation genome-wide (i.e. onefold enrichment). (**B**) Expression of Aza-Up genes (activated more than sixfold by 5aza-dC) relative to mock treated after 14 days recovery. There is a bimodal distribution whereby ~85% of genes are re-repressed whereas 15% remain activated (more than sixfold, ~15%) after drug withdrawal. Note that although most genes are re-silenced to less than twofold of original levels, a minority of genes (~16%) are only partially re-repressed (threefold, fourfold, fivefold). However, only genes that maintain more than sixfold induction after recovery are considered further in this study. (**C**) Volcano plot demonstrating the relationship between expression fold-change among CGI genes and promoter methylation after transient exposure to 5aza-dC in NIH/3T3 cells (left) and 14 days cellular recovery (right). Note that upregulated genes exhibit no significant enrichment of highly methylated promoters immediately after 5aza-dC exposure (left), indicating that the drug primarily affects gene expression through indirect mechanisms. By contrast, the recovery set demonstrates that there is preferential recovery of indirect (non-methylated) targets but continual expression of methylation-dependent genes (methylated CGI genes, top right of each graph).

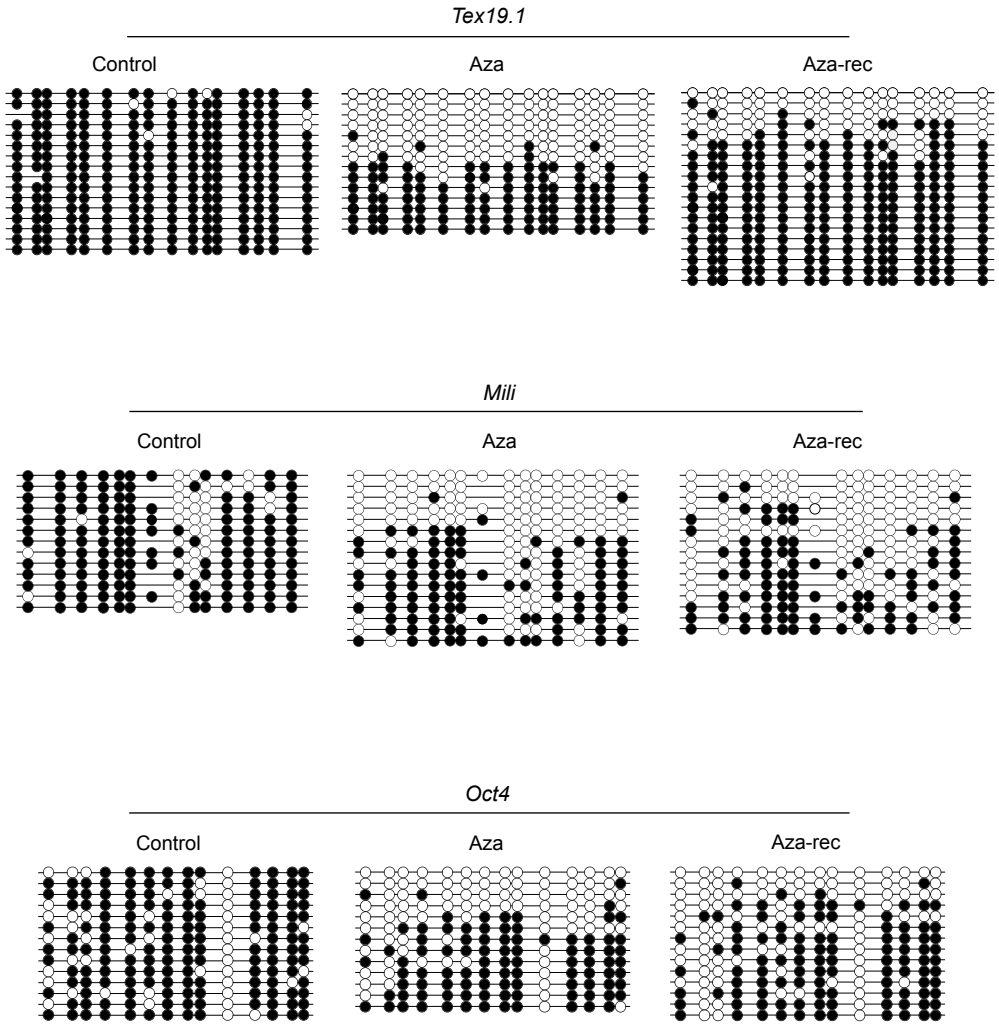

Fig. S2. Aza-recovery assay leads to enduring hypomethylation. Bisulphite methylation analysis of *Tex19.1*, *Mili* and *Oct4* promoters before (control), after 5aza-dC exposure (Aza) and following a 14 day recovery period under normal culture conditions (Aza-rec). There is a significant loss of methylation in a subpopulation of cells following 5aza-dC exposure at each tested locus. Importantly, methylation is not re-acquired during recovery because there is highly limited de novo activity in NIH/3T3 cells. The apparent small gain in methylation during recovery is a consequence of a selective growth advantage in cells that retain global DNA methylation after 5aza-dC exposure (J.A.H., C.E.N., R.R.M., unpublished observations), leading to a gradual titration of demethylated cells over generations. Consistent with this, the apparent gain of methylation is allelic and not mottled (see *Tex19.1*) indicating cells that retain methylation after 5aza-dC exposure are proliferating at a selective advantage and not gaining de novo methylation. *Tex19.1* and *Mili* are activated and remain active after demethylation whereas *Oct4* is not, implying that it is not a methylation-dependent target in somatic cells.

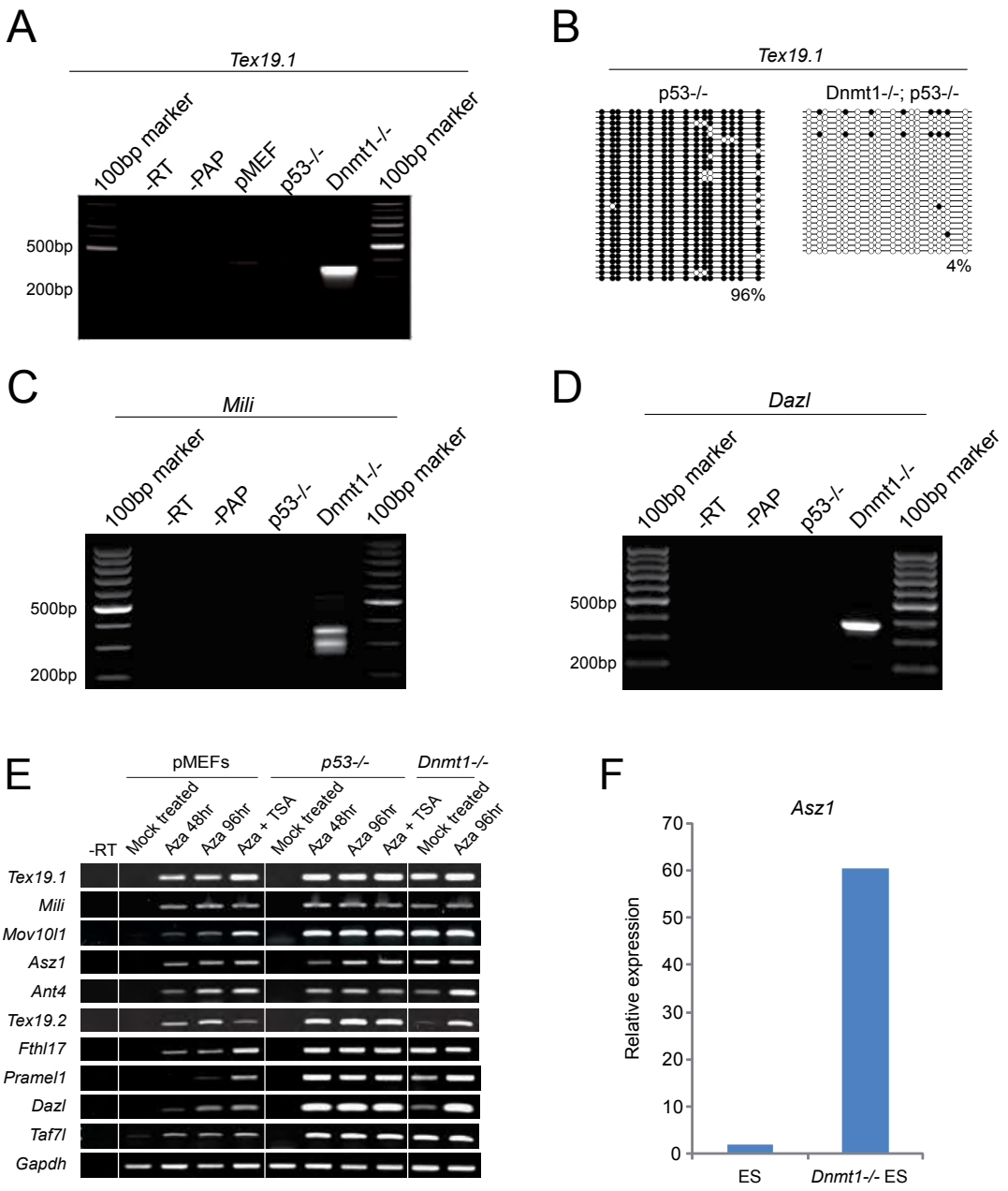

Fig. S3. Transcript mapping and germline gene expression in demethylated cells. (**A**) Mapping of the *Tex19.1* transcriptional start site (TSS) (expected size ≈334 bp). Sequencing of PCR products revealed that *Dnmt1*$^{n/n}$ MEFs express a single *Tex19.1* transcript from the annotated TSS. By contrast, overcycling (60 cycles) the 5′RACE PCR indicates that wild-type and *p53*$^{-/-}$ MEFs transcribe very low background *Tex19.1* from a non-canonical TSS ~51 bp upstream of the annotated TSS. (**B**) Validation that *Tex19.1* is demethylated in *Dnmt1n/n* MEFs. (**C**,**D**) Transcript mapping of *Mili* (expected size ≈336 bp) and *Dazl* (expected size ≈371 bp) in *Dnmt1n/n* cells indicates that expression in hypomethylated cells is driven from the canonical annotated TSS. As expected, *Mili* is additionally transcribed from an annotated promoter downstream of the TSS that also drives expresssion in testes in vivo. These data suggest that alternative promoters are not active in experimentally hypomethylated cells, which is consistent with promoter CpG methylation at the canonical TSS regulating expression of these genes. (**E**) Extended RT-PCR analysis of germline-specific genes that might have a role in genome-defence and are expressed following demethylation by multiple methods in different cell contexts. This figure is the complete panel from Fig. 1E (main text), which is abridged for clarity. (**F**) Expression of *Asz1* is very low in wild-type ES cells but strongly detected in *Dnmt1*-null ES cells. –RT, –reverse transcriptase; –PAP, –polyA polymerase; pMEF, primary mouse embryonic fibroblasts.

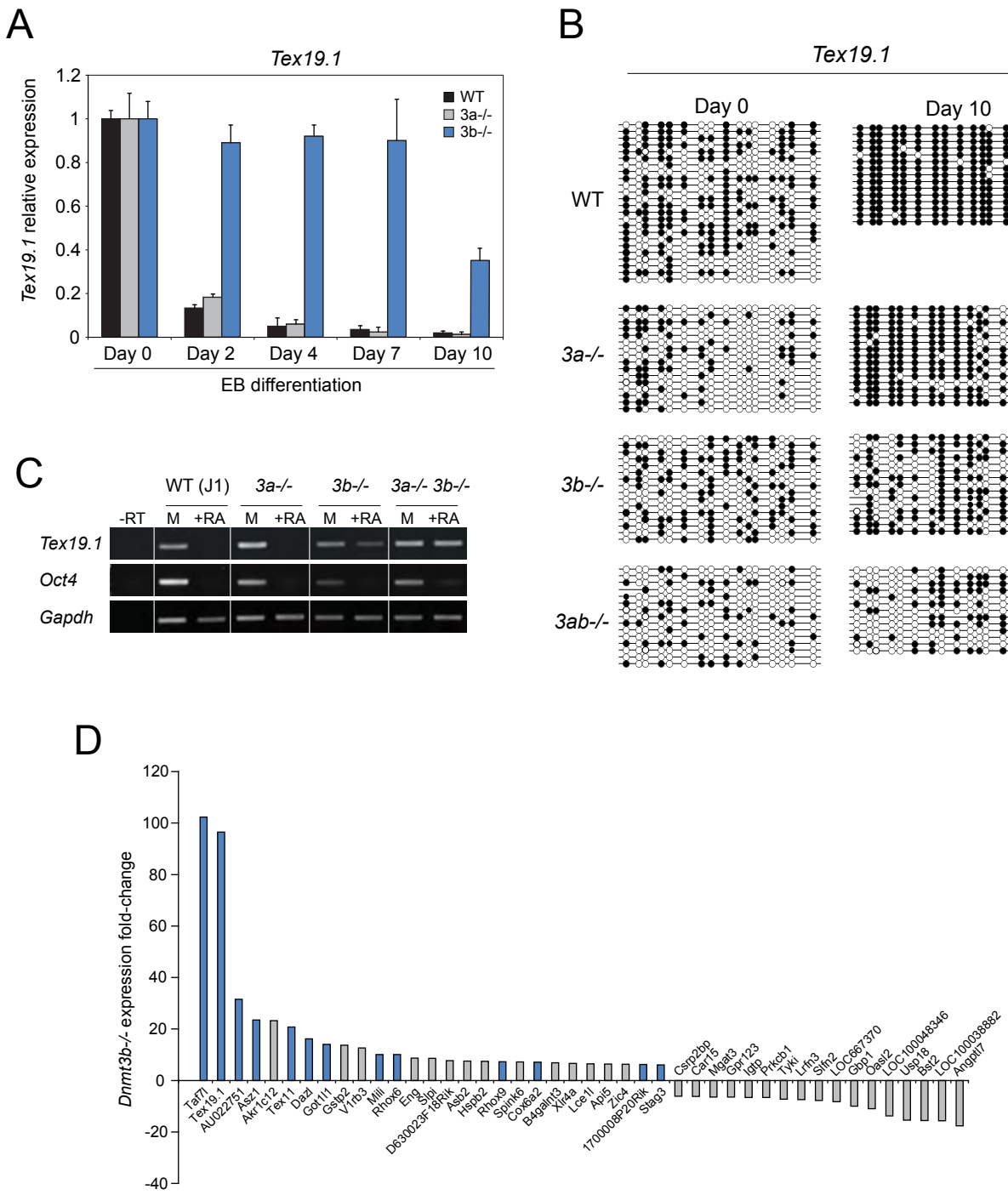

Fig. S4. DNMT3B specifically targets *de novo* methylation and gene silencing to *Tex19.1.* (A) Quantitative RT-PCR of *Tex19.1* expression in wild-type (WT) J1 and mutant ES cells during embryoid body (EB) differentiation. WT and *Dnmt3a*$^{-/-}$ ES cells can impose strong silencing on *Tex19.1* but *Dnmt3b*$^{-/-}$ ES cells are unable to silence expression. Error bars represent s.e.m. (B) Bisulphite methylation analysis of *Tex19.1* before (day 0) and after (day 10) EB differentiation of the indicated genotype ES cells. In the absence of DNMT3B, *Tex19.1* fails to acquire *de novo* methylation and gene silencing. (C) RT-PCR demonstrating that *Tex19.1* silencing also fails in *Dnmt3b*$^{-/-}$ and *Dnmt[3a, 3b]*$^{-/-}$ ES cells differentiated with retinoic acid (RA) relative to mock (M) treated. (D) Genes upregulated and downregulated more than sixfold in a global expression analysis of *Dnmt3b*$^{-/-}$ embryos at E13.5. Genes in blue are germline specific.

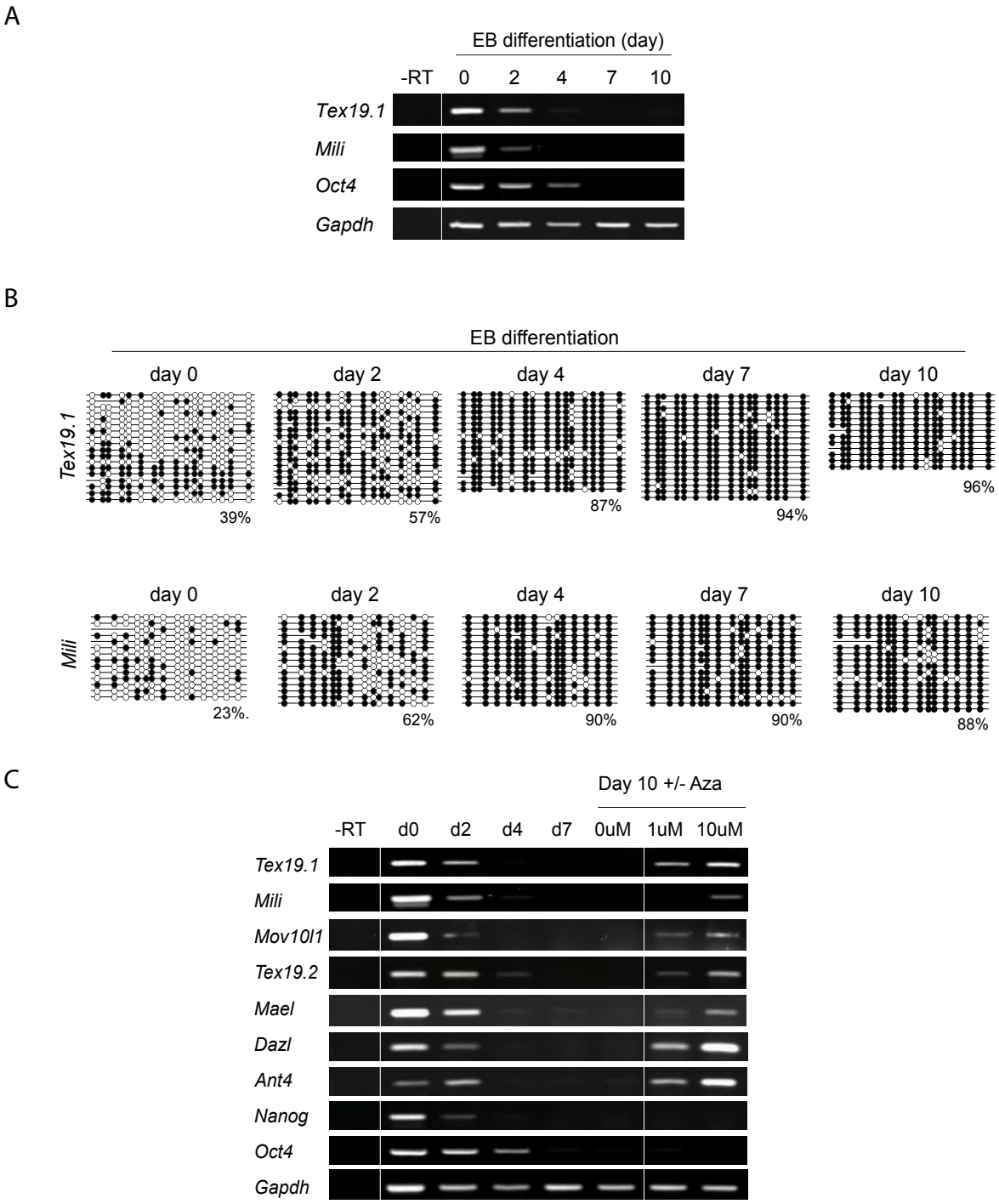

Fig. S5. DNA methylation dynamics during ES cell differentiation. (**A**) Expression of *Tex19.1* and *Mili* during embryoid body (EB) differentiation of ES cells. *Oct4* confirms differentiation and *Gapdh* is a loading control. (**B**) Bisulfite analysis of *Tex19.1* and *Mili* promoter DNA methylation during EB differentiation. Onset of de novo methylation temporally coincides with the initiation of gene silencing. (**C**) RT-PCR showing germline-specific genes with potential roles in genome defence are silenced during EB differentiation, as confirmed by *Nanog* and *Oct4.* Importantly, germline genes are reactivated by exposure to 5aza-dC after silencing, whereas pluripotency genes (*Nanog* and *Oct4)* and not reactivated, suggesting that multiple epigenetic silencing mechanisms operate at pluripotency genes but that only DNA methylation prevents reactivation of silenced germline/genome defence genes.

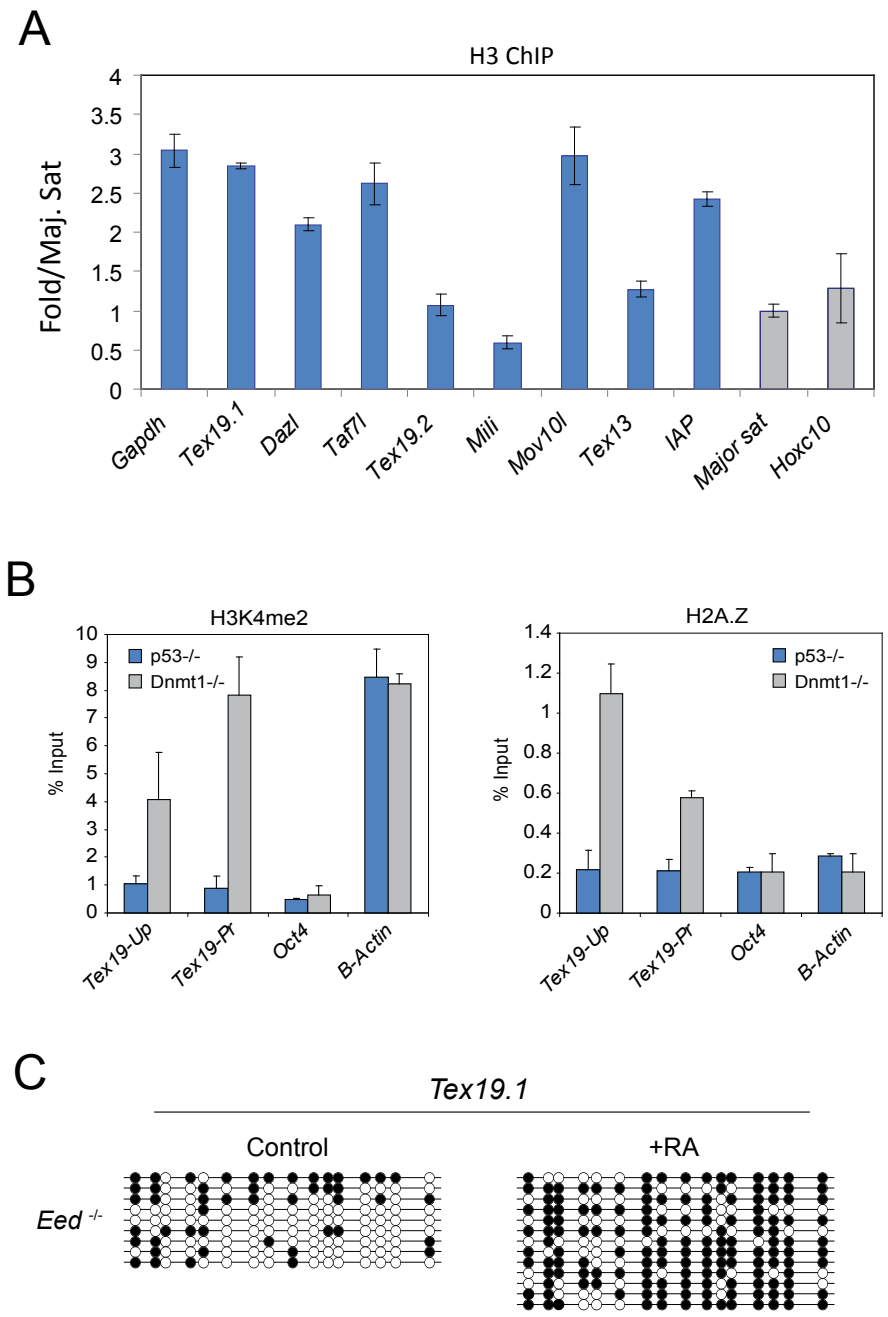

Fig. S6. Germline genes have nucelosome-dense promoters that become modified after demethylation. (**A**) ChIP for histone H3 at methylation-dependent germline genes. Assayed germline genes are enriched in H3 comparably or greater than positive control loci (*HoxC10*, Maj Sat, IAP), indicating that they are not nucleosome depleted regions. (**B**) Native-ChiP for the indicated active histone marks/variants at the promoter proximal (Tex19-Pr) and upstream promoter region (Tex19-Up) of the *Tex19.1* promoter in non-expressing *p53*$^{-/-}$ cells and expressing *Dn mt1*$^{n/n}$ *(p53*$^{-/-}$*)* cells. Inactive *Oct4* and active β-*actin* are controls. The enrichment of H3K4me2 at *Tex19.1* in *Dnmt1n/n* cells is likely to be a downstream consequence of gene activation by DNA demethylation, which may directly occlude H3K4me2 thereby maintaining gene silencing. (**C**) The *Tex19.1* promoter can still acquire *de novo* methylation (and silencing) in the absence of EED, an essential component of the PRC2 complex, indicating that polycomb is functionally dispensable for *Tex19.1* regulation. Error bars represent s.e.m.

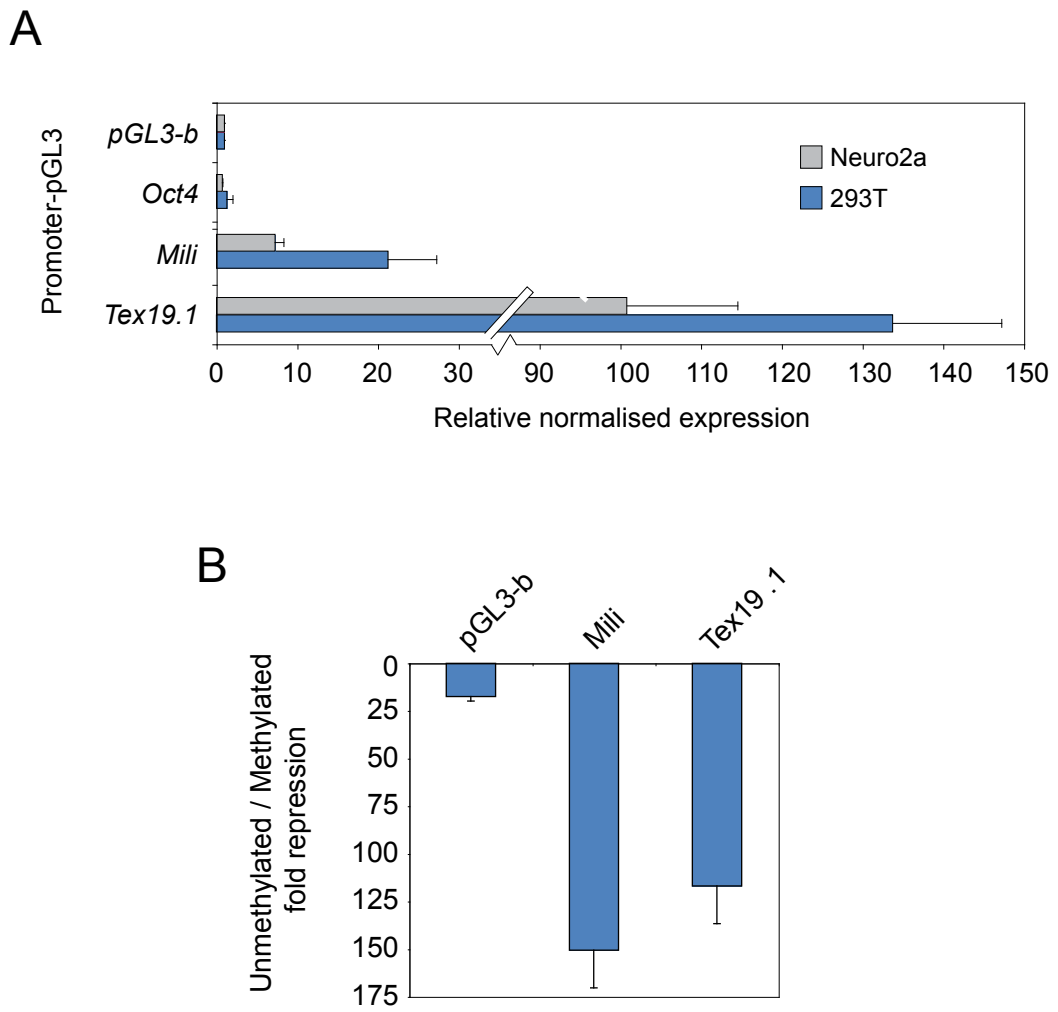

Fig. S7. *Tex19.1* and *Mili* reporters drive strong expression in somatic cells. (A) *Tex19.1* (–1034 to +224 bp) and *Mili* (–878 to +126 bp) promoters, but not the *Oct4* promoter, can drive strong expression of a reporter in neural (Neuro2a) or mesodermal (293T) derived cells. This indicates that these promoters are not regulated by germline-specific trans-acting activators or somatic repressors. Thus, an epigenetic mechanism, such as DNA methylation, must prevent activation of the endogenous *Tex19.1* and *Mili* genes in somatic cells. (B) *In vitro* methylation of the *Tex19.1* and *Mili* reporters strongly silences (up to 150-fold) their capacity to drive expression in somatic cells. pGL3-basic is empty reporter vector. Error bars represent s.e.m.

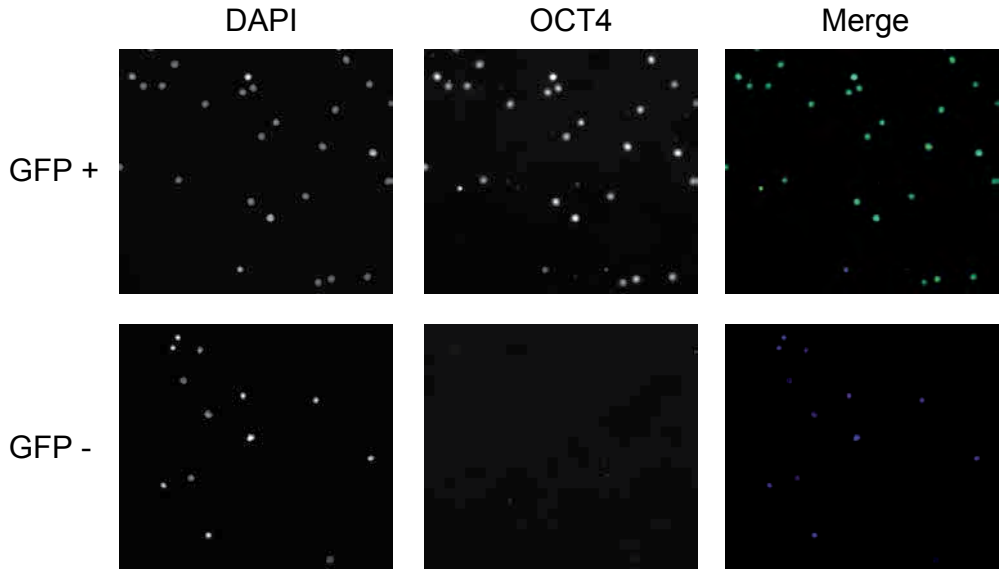

Fig. S8. PGC purity after FACS sorting. Staining for endogenous OCT4 expression in FACS sorted E13.5 germ cells expressing an *Oct4-GFP* transgene. This analysis demonstrated that sorted PGCs were >98% pure.

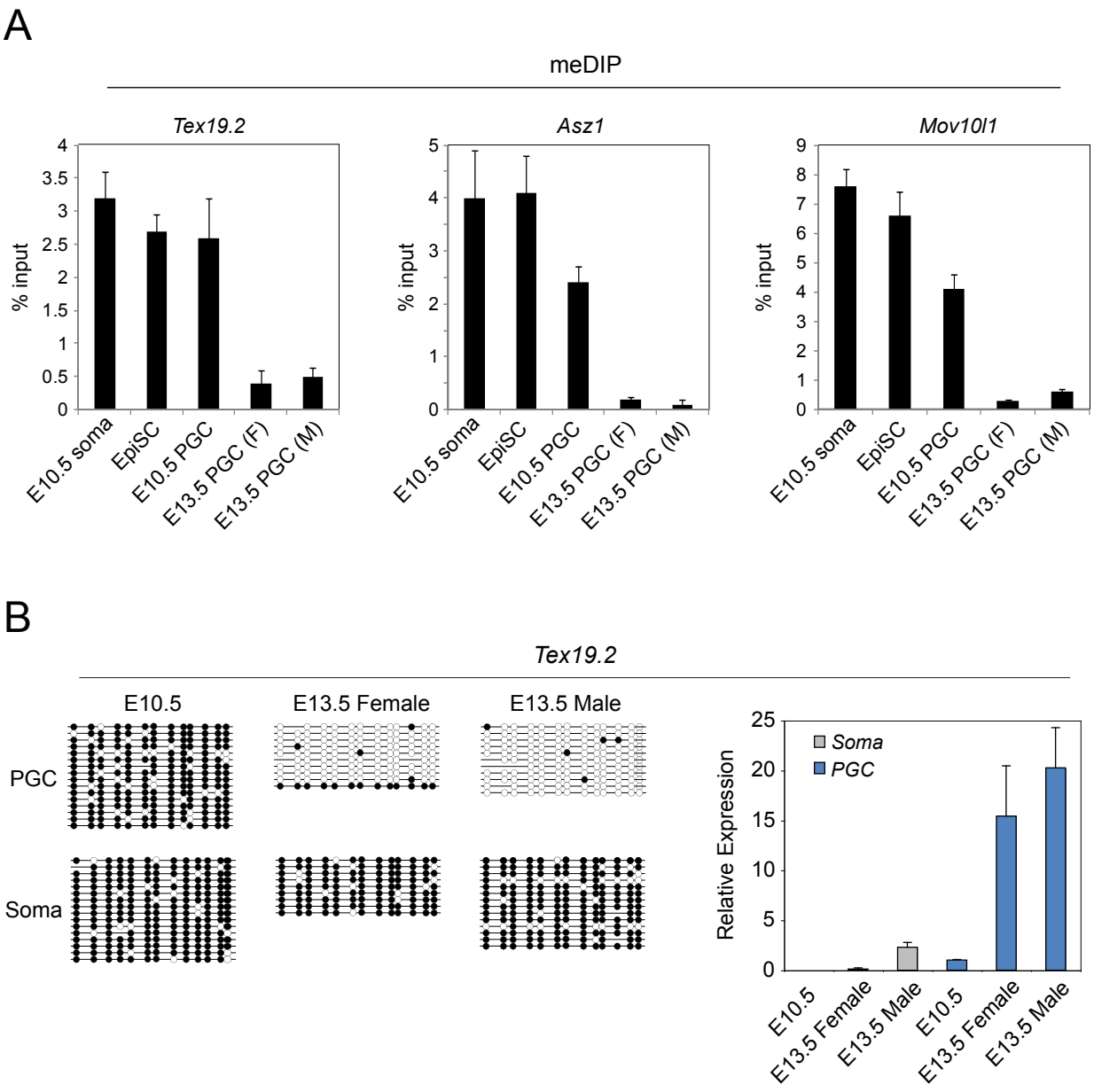

Fig. S9. Promoter demethylation dynamics and gene expression in PGCs. (**A**) meDIP showing relative promoter DNA methylation in E10.5 and E13.5 PGCs, and in epiblast stem cells (EpiSC) and soma as positive controls. The genome defence genes *Asz1*, *Tex19.2* and *Mov10l1* remain at least partially methylated in PGCs at E10.5 and become demethylated by E13.5. By contrast, *Tex19.1* and *Mili* are already demethylated in PGCs by E10.5 (Fig. 4A). (**B**) Left panel: Bisulfite methylation analysis of the *Tex19.2* promoter confirms promoter demethylation in PGCs between E10.5-13.5. Right panel: qRT-PCR analysis of *Tex19.2* expression for the indicated samples. Note that *Tex19.2* shows similar dynamics to *Dazl* whereby the promoter is demethylated and transcriptionally activated between E10.5 and E13.5 in PGCs. Error bars represent s.e.m.

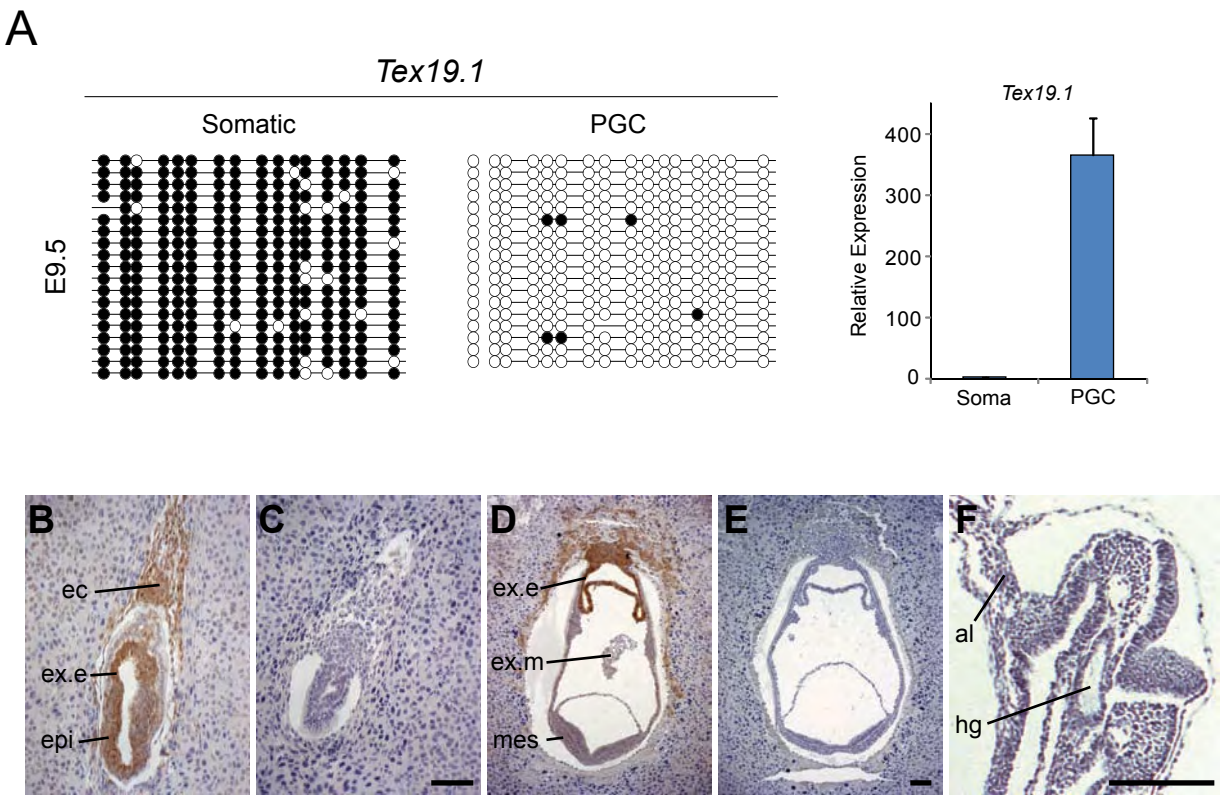

Fig. S10. TEX19.1 is expressed in the post-implantation embryo but silenced after E6.5. (**A**) *Tex19.1* is hypomethylated (left panels) and expressed (right panel) at E9.5 in PGCs. This confirms the hypomethylated state of *Tex19.1* at E10.5. (**B**) Anti-TEX19.1 staining (brown precipitate) in an E6.5 embryo can be seen in the ectoplacental cone (ec), extra-embryonic cell lineages (ex.e) and the epiblast (epi). (**C**) IgG control. (**D**) In E7.5 embryos, TEX19.1 becomes downregulated in epiblast-derived tissues in the embryo with some faint residual staining present in embryonic mesoderm (mes) and extra-embryonic mesoderm (ex.m). TEX19.1 staining is present in extra-embryonic tissues (ex.e) at this stage. (**E**) IgG control. (**F**) TEX19.1 staining is not detectable in embryonic tissues at E8.5. TEX19.1 staining is absent in the hindgut endoderm where primordial germ cells are located at this stage, consistent with expression of *Tex19.1* not occurring until after reprogramming at ~E8.5 in PGCs. Panel shows posterior region of E8.5 embryo. hg, hindgut; al, allantois. Scale bars: 200 μm.