GPIF constrains Polycomb Repressive Complex 2 activity in germ cells

Ragazzini R[1], Pérez-Palacios R[1], Baymaz HI[2], Zielinski D[1, 3], Michaud A[1,] Givelet M[4], Borsos M[5,6], Jansen PWTC[2], Servant N[1, 3], Torres-Padilla ME[5,6], Bourc'his D[1], Fouchet P[4], Vermeulen M[2] and Margueron R[1,*]

[1] Institut Curie, PSL Research University, INSERM U934/ CNRS UMR3215, 11-13 Rue Pierre et Marie Curie, 75005 Paris (France).

[2] Institute for Molecular Life Sciences Radboud umc, Route 274 M850.03.87 Geert Grooteplein 28, Nijmegen 6525 GA (Netherlands).

[3] INSERM U900, Mines ParisTech, Paris, France.

[4] Commissariat à l'énergie atomique et aux énergies alternatives, Institut de radiobiologie cellulaire et moléculaire, Laboratoire des Cellules Souches Germinales, INSERM U967, 92260 Fontenay-aux-Roses (France).

[5] Institute of Epigenetics and Stem Cells (IES), Helmholtz Zentrum München, Munich, Germany.

[6] Faculty of Biology, Ludwig Maximilians Universität, Munich, Germany.

[*]Corresponding author:

raphael.margueron@curie.fr

11-13, Rue Pierre et Marie Curie, 75005, Paris, France.

Tel: +33 (0)156246551

ABSTRACT

The Polycomb machinery is required for the proper orchestration of gene expression by virtue of its critical role in maintaining transcriptional silencing. It is composed of several chromatin modifying complexes, including Polycomb Repressive Complex 2 (PRC2), which deposits H3K27me2/3. Here, we report the identification of a new cofactor of PRC2, GPIF (Gonads PRC2 Inhibitory Factor), expressed predominantly in the gonad. GPIF limits the enzymatic activity of PRC2 by lessening the interaction between the core complex and its accessory subunits, but without interfering with its recruitment to chromatin. Deletion of *Gpif* leads to a global increase of H3K27me2/3 deposition both during spermatogenesis and at late stages of oocyte maturation. This alteration of the epigenetic content of mature oocytes does not affect the initial pool of follicles but is associated with a reduction in follicle number in aging mice. In addition, fertility is strongly impaired in *Gpif* -/- females indicative of developmental defects emerging around fertilization or during embryo development. Our study uncovers GPIF as a novel functional player in the comprehensive chromatin remodeling that occurs in the gonad.

INTRODUCTION

Early in development, cells commit to specific lineages and acquire precise identities that require maintenance throughout the lifespan of the organism. Polycomb group proteins play an important role in this process by maintaining transcriptional repression through the regulation of chromatin structure [1]. In mammals, this machinery is composed of two main complexes: Polycomb Repressive Complex 1 and 2 (PRC1 and 2). The core PRC2 complex is composed of four subunits: the catalytic subunit EZH1/2, SUZ12, EED and RbAp46/48 [1]. PRC2 catalyzes the di- and tri- methylation of lysine 27 on histone H3 (H3K27me2/3), an enzymatic activity which is required for its function. Indeed, the mutation of lysine 27 of histone H3 to arginine leads to loss of gene repression and mutant flies display a phenotype similar to deletion of PRC2 components [2]. H3K27me3 is generally enriched around the promoter of transcriptionally silent genes and contributes to the recruitment of PRC1 [1]. H3K27me2 is widely distributed, covering 50-70% of histones, and its role is less defined but may be to prevent aberrant enhancer activation [3].

The question of how PRC2 is targeted to chromatin and how its enzymatic activity is controlled has received ongoing attention [4]. Cumulative evidence suggests that PRC2 may not be actively recruited to chromatin and that instead its activity is promoted by the recognition of its own mark H3K27me3 [5], ubiquitination of lysine 119 of H2A [6], GC-richness [7] or by condensed chromatin [8]. Conversely, some histone modifications negatively influence PRC2 function, particularly those associated with active transcription, such as H3K4me3 and H3K36me3 [9]. PRC2 binding to chromatin may also be inhibited by DNA methylation [10], although other reports suggest that PRC2 is compatible with DNA methylation [11].

A number of accessory subunits have now been shown to influence PRC2 function [12]. Recent comprehensive proteomic analyses suggest that they might form around two main PRC2 subtypes, PRC2.1 and PRC2.2 [13]. The subunit SUZ12 plays a central role by orchestrating the cofactor interactions [14]. PRC2.1 includes one of the three Polycomb-like proteins (PHF1, MTF2 or PHF19) together with the recently identified PRC2 partners EPOP and PALI1 [15 16]. The three Polycomb-like proteins harbor one Tudor domain and two PHD finger domains each [4]. Their Tudor domain is able to recognize H3K36me3 decorated genes, which could be important for PRC2 association with transcribed targets [4]. The function of EPOP remains ambiguous since, in vitro, it stimulates PRC2 catalytic activity while, *in vivo*, it limits PRC2 binding, likely

through interaction with Elongin BC [17]. In contrast, PALI1 is required for H3K27me3 deposition both *in vitro* and *in vivo* [16]. The other complex, PRC2.2, includes JARID2 and AEBP2 subunits in equal stoichiometry [18,19]. Both are able to stimulate PRC2 catalytic activity *in vitro* with JARID2 being also able to bind nucleosomes [20]. JARID2 also appears to be necessary for PRC2 targeting at its loci, possibly through its DNA binding domain or as a result of its methylation by PRC2 [21,22,23,24]. AEBP2 binds to DNA *in vitro*, but appears to negatively modulate PRC2 *in vivo* [25,26]. Of note, AEBP2 was reported to stimulate PRC2 through a mechanism independent of PRC2 allosteric activation [27,28]. While we now have a good picture of the accessory subunits interacting with PRC2, their precise roles are only partially understood. This might be due to compensatory mechanisms, such that interfering both with PRC2.1 and PRC2.2 is required to inhibit PRC2 recruitment [29] as observed upon loss of SUZ12 [30].

The regulation of chromatin structure in germ cells is pivotal as these cells are the bridge between generations and therefore potential vector of epigenetic information. In particular, H3K27me3 has been shown to be involved in parental imprinting [31,32,33]. Yet, in contrast to the extensive characterization of PRC2 in models such as mouse embryonic stem cells (ESC), much less is known about the regulation of its enzymatic activity in germ cells. Deletion of PRC2 core components during spermatogenesis results in the progressive loss of germ cells, indicating that its activity is required for this process [34,35]. At later stages of spermatogenesis, when round spermatids differentiate into mature sperm, histones are progressively replaced by protamines. A variable fraction of the genome retains a nucleosomal structure (1% in mice, 10-15% in human), with histones carrying post-translational modifications, including H3K27me3 [36]. During oogenesis, histones are maintained and H3K27me3 is detected throughout this process [37-39]. However, H3K27me3 displays a peculiar pattern of enrichment in the growing oocyte, showing broad enrichment in intergenic regions and gene deserts (reviewed in [40,41]). Genetic interference with PRC2 function in growing oocytes does not prevent their maturation but has been linked to a postnatal overgrowth phenotype in the progeny [42] possibly through the control of imprinting [33].

Here, we report the identification of a new cofactor of PRC2, GPIF, functioning primarily in mammalian gonads and show that this cofactor limits PRC2-mediated H3K27me3 deposition. Inactivation of this cofactor results in excessive deposition of this mark, altering the epigenetic content of oocytes and impairing mouse female fertility.

RESULTS

Identification of a new cofactor of PRC2 in the gonad

PRC2 recruitment and enzymatic activity is controlled by a set of cofactors interacting in a partially mutually exclusive manner with the core subunit SUZ12 but little is known about its regulation in germ cells. To tackle this question, we first focused on testes (more abundant material than ovaries) and took advantage of knock-in mouse models expressing an N-terminal Flag-tagged version of either EZH1 or EZH2 from their respective endogenous locus (this study [22]). We verified the expression of the tagged-EZH1 by western blot on mouse testis nuclear extract and were able to detect the presence of a slowly migrating polypeptide, which is specifically pulled down by Flag-Immunoprecipitation (Flag-IP) (Fig. S1A). We then isolated nuclei from adult mouse testes (WT-control, EZH2-Flag or EZH1-Flag), performed Flag-IP, and subjected the samples to mass spectrometry. The results of three independent IPs are represented as volcano plots (Fig. 1A and S1B). As expected, both EZH1 and EZH2 proteins interact with the other PRC2 core components and with known accessory subunits: AEBP2, JARID2, PHF1 and MTF2 (Fig. 1A and S1B). Interestingly, our experiments also reveal the existence of a new partner, the uncharacterized protein AU022751 (ENSMUST00000117544; NM_001166433.1), which we retrieved in both EZH1 and EZH2 pull-downs. We called this new cofactor "GPIF" for Gonad PRC2 Inhibitory Factor. Of note, this protein was previously identified in PRC2 interactomes of mouse embryonic stem cells but was not further investigated [19,43,44]. In order to confirm this interaction, we overexpressed Flag-tagged versions of the mouse and human homologs in HeLa-S3 cells (Fig. S1C) and performed IP followed by mass spectrometry. These reverse IPs confirmed the interaction between PRC2 and GPIF (Fig. 1B and S1D). Additional putative partners were identified in both IPs, but with the exception of USP7, they were not common to both homologs. We therefore did not pursue their study further. Importantly, these reverse-IPs also indicate that GPIF interacts with both PRC2 complex subtypes.

GPIF is located on the X chromosome. In most species, it is a mono-exonic gene -that may indicate that it was generated by retroposition- but in the mouse, splicing also creates a shorter isoform. Using PAML (Phylogenetic Analysis by Maximum Likelihood), we observed that *GPIF* homologs are present across *Eutheria* but we did not identify any homologs outside of this clade based on either sequence conservation or on synteny. *GPIF* genes have rapidly evolved both at the nucleotide and amino acid levels, the rodent homologs being particularly

distant from the rest (Fig. 1C and S1E). This contrasts with the other PRC2 components, such as *EZH2*, which are highly conserved across mammals (Fig. S1E). No known protein domain was predicted for GPIF and the only distinguishing feature is a serine-rich region (Fig. 1C in green) including a short amino-acid stretch that is fully conserved in all orthologs identified (Fig. 1C in purple). To characterize *Gpif* expression, we performed RT-qPCR on various tissues (3-month-old females and males). *Gpif* mRNA expression was particularly high in ovaries; it was also expressed in testes and much less in other tissues (Fig. 1D). This pattern of expression is distinct from that of *Ezh2*, which is expressed tissue-wide, with the strongest expression observed in spleen. Analysis of public gene expression datasets from fetal gonads [45] indicates *Gpif* is preferentially expressed in E13.5 primordial germ cells (PGCs) compared to somatic cells, correlating with germ cell markers such as *Piwil2* or *Prdm14* (Fig. S1F). Interestingly, *Gpif* belongs to a set of genes referred to as "germline-reprogramming-responsive" that become active following PGC DNA demethylation [46], as they are associated with strong CpG island promoters. Similarly, in humans *GPIF* is highly transcribed in male and female PGCs from week 5 until week 9 of pregnancy, while almost absent in ESCs and somatic cells (Fig. S1G) [47]. We confirmed this observation at the protein level by performing immunohistochemistry on sections of testes and ovaries of human origin. hGPIF protein was detected in male germ cells inside the seminiferous tubules, especially in spermatogonia and round spermatids (Fig. 1E). In ovaries, GPIF antibody stained primordial follicles and oocytes (red arrows) but not the external follicle cells in contrast to EZH2 antibody, which stained both zones (Fig. 1F). To summarize, GPIF is a novel cofactor of PRC2 in placental mammals. It is a fast-evolving protein with no known protein domain, it is expressed primarily in PGCs during development and remains present in the adult gonad.

GPIF is a negative regulator of PRC2 activity

To study the molecular role of GPIF, we sought a model cell line that would express this factor endogenously. The *GPIF* transcript is undetectable from most cell lines, with the exception of U2OS, an osteosarcoma-derived cell line (Fig. S2A and 2A). We used genome editing to generate U2OS clonal cells that were knockout for *GPIF* or for *EED* as a control for PRC2 inactivation (U2OS *GPIF* -/- and U2OS *EED* -/- respectively). Both cell lines were viable and their proliferation was unaffected by the knockouts (data not shown). Western blot showed that deletion of *EED* destabilized the other PRC2 core components, such as EZH2, while inactivation of *GPIF* had no discernible effect on the accumulation of these proteins (Fig. 2A). We then assessed H3K27 methylation and observed a robust increase in H3K27me2/3 upon

GPIF deletion (Fig. 2B), while H3K27me1 was stable. Interestingly, H3K27me3 level was very low in U2OS compared to extract prepared from HEK-293T cells, which do not express *GPIF* (Fig. S2B). To confirm that *GPIF* deletion was directly responsible for the increased H3K27me3 in U2OS, we stably restored its expression using either full-length (FL) or deletion mutants (Fig. 2C) as verified by western blot (WB) and RT-q-PCR (Fig. S2C). Upon re-expression of FL and mutant GPIF, H3K27me3 returned to basal levels (Fig. 2C/2D) with the notable exception of mutant M5 that lacks the conserved amino-acids stretch (Fig. 2C). Given that such deletion abolishes GPIF interaction with PRC2 in co-IP (mutant 6 *versus* mutant 7; Fig. S2D), it stands to reason that GPIF likely regulates H3K27me3 deposition through direct interference with PRC2 activity.

We then investigated H3K27me3 genomic distribution in the absence of *GPIF* by chromatin immunoprecipitation followed by sequencing (ChIP-seq). We used the U2OS *EED* -/- in which H3K27me3 is not detectable as a negative control and compared it to the U2OS wild-type (WT) and *GPIF* -/- conditions. Replicates were well correlated and the U2OS WT and U2OS *EED* -/- clustered together, away from the U2OS *GPIF* -/- (Fig. S2E). This agrees with our earlier observation that H3K27me3 is very low in U2OS, as in the *EED* knockout (Fig. S2E). In contrast, there was a genome-wide increase in H3K27me3 deposition upon deletion of *GPIF* (Fig. 2D), as demonstrated by the large number of peaks detected in this context (Fig. 2E). At the transcriptome level, only a few genes were differentially expressed in U2OS *EED* -/- as compared to WT (FDR<0.05), whereas approximately 500 genes were differentially expressed in *GPIF* -/- *versus* WT, the majority of which were downregulated, as expected upon gain of a repressive mark. Not surprisingly, most genes downregulated upon of *GPIF*-knockout displayed a robust increase in H3K27me3 deposition (Fig. S2F). Altogether, these results reveal that GPIF inhibits the activity of PRC2 without interfering with the integrity of the core complex.

GPIF impairs PRC2 enzymatic activity but not its binding to chromatin

Considering the inhibitory action of GPIF on PRC2, we next sought to explore the underlying mechanisms. First, we hypothesized that GPIF could limit PRC2 binding to chromatin. To test this hypothesis, we performed ChIP against EED to monitor PRC2 recruitment to chromatin by comparing U2OS cells (expressing *GPIF*) and U2OS *GPIF-/-*. This analysis revealed that EED recruitment was comparable in WT and *GPIF* -/- contexts (Fig. 3A) even at genes such as *DRGX* where H3K27me3 is almost absent in the wild type condition (Fig. 2D). Our results

show that *GPIF* deletion does not appear to alter PRC2 binding to chromatin, thus prompting to consider the possibility that it may instead interfere with PRC2 enzymatic activity. We first evaluated whether a titration of purified GPIF (Fig. S3A) inhibits the enzymatic activity of the recombinant PRC2 core complex in a histone methyltransferase assay. However, even at molar excess, GPIF did not impact the enzymatic activity of PRC2 (Fig. S3B). We then reasoned that GPIF might regulate PRC2 activity only in the presence of its cofactors. To test this hypothesis, we purified the core PRC2 and its cofactors from U2OS and U2OS *GPIF* -/- cells that stably over-express a Flag-tagged version of EZH2 (Fig. S3C). EZH2 was immunoprecipitated and further purified through an ion-exchange column before monitoring its activity on native histones. While we observed very low methyltransferase activity towards H3 with PRC2 purified from WT cells, the complex purified from U2OS *GPIF* -/- was much more active (Fig 3B. left *versus* central panel). Furthermore, in contrast to our observation with the recombinant core PRC2 complex, the titration of GPIF on PRC2 purified from U2OS *GPIF* -/- inhibited PRC2 enzymatic activity (Fig 3B. right panel). These results suggest that GPIF might regulate PRC2 by mitigating its interaction with its cofactors. To test this hypothesis, we analyzed PRC2 interactome by mass spectrometry, depending on *GPIF* expression status. Overall, PRC2 displayed the same composition (Fig. S3D); however, the stoichiometry of the cofactors appeared substantially different in the absence of GPIF (label-free quantification based on iBAQ). Namely, several cofactors -AEBP2, JARID2 and PALI1- were present at a higher stoichiometry in the IPs from *GPIF* -/- cells. (Fig. 3C). We confirmed this result by co-IP/WB investigating the interaction of AEBP2 and JARID2 with EZH2 in IPs performed with nuclear extracts prepared from U2OS wild type or *GPIF* -/- cells (Fig. 3D). Our results suggest that GPIF does not prevent PRC2 binding to chromatin but limits the stimulatory action of cofactors, such as AEBP2 and JARID2 on its enzymatic activity.

Gpif **Knockout males are fertile despite a global increase in H3K27me3**

To study the role of GPIF in a more physiological environment, we generated a knockout mouse model in which a CRISPR-Cas9 induced-deletion of 1.5 Kb removes most of the gene body (Fig. S4A). Accordingly, *Gpif* mRNA and protein were absent from testis and from ovaries (Fig. S4B, C and D), two organs where *Gpif* is preferentially expressed. Of note, expression of the genes flanking *Gpif* (*Nudt10* and *Nudt11*) were unaffected by the deletion (Fig. S4B and S4D). *Gpif* mice (-/- or -/Y) did not show any overt developmental defect, with adults appearing undistinguishable from the wild type.

We first investigated the expression of *Gfip* during spermatogenesis in the different sub-populations of germ cells sorted from adult mice based on staining for α6-integrin, the tyrosine kinase receptor c-Kit, and DNA content, as previously described [48,49]. *Gpif* was mostly expressed in spermatogonia (α6-integrin positive, Fig S4F). Its expression was very low in spermatocytes I and II, consistent with the global transcriptional inactivation of the X chromosome at these stages [50], in contrast to *Ezh2* expression, which increases at the final stages of differentiation (4n, 2n and n; Fig. S4E).

We then tested whether deletion of *Gpif* could enhance H3K27me3 deposition during spermatogenesis, as it does in U2OS cells. For this, we probed nuclear extracts from whole testes of adult mice by western blot. As shown in Figure 4A, amounts of H3K27me2 and me3 increased by about two-fold in the absence of GPIF, whereas other histone marks remain unchanged. This observation suggests that H3K27me2/3 might not be invading other chromatin domains but more likely accumulate further where it was already present. Consistent with our previous observations, this effect was not due to a direct effect on the protein accumulation of PRC2 core components (Fig 4A. bottom panel).

To identify the cellular origin of this H3K27me2/me3 upregulation, we performed immunofluorescence on testis sections. Triangle-shaped somatic Sertoli cells -identified by the presence of two satellite chromocenters in their nuclei after DAPI-staining- were strongly positive for H3K27me3 in both WT and KO condition (Fig. 4B, yellow star; [51]). In contrast, germ cells -identified by expression of the germ cell marker TRA98- displayed much stronger H3K27me3 signal in *Gpif* -/Y mice compared to WT littermates (Fig. 4B, yellow arrows). The germ cell specificity of GPIF deletion was confirmed in germ cell-free *Dnmt3l* mutant testes [52], for which H3K27me3 patterns were similar in presence or absence of GPIF (Fig. S4F).

To evaluate the functional consequences of aberrant H3K27me3 deposition, we profiled gene expression of $\alpha6^{+}$c-kit^{-} undifferentiated spermatogonia in WT and *Gpif* -/Y mice. Gene expression was moderately affected by the absence of GPIF: about 125 genes differentially expressed (FDR<0.05, Fig. 4C), the majority of which were downregulated. To determine the impact on spermatogenesis, we analyzed the different germ cell subpopulations from WT and *Gpif* -/Y testes by cell cytometry. The relative sizes of these subpopulations were unaffected (Fig. 4D), in agreement with the normal testis-to-body weight ratio and normal fertility of *Gpif* -/Y males (Fig. S4G). Finally, we evaluated sperm quality through analysis of computer-

assisted spermatozoa images. Spermatozoa motility felt within normal standards, although spermatozoa from *Gpif* -/Y males showed slightly less progressive motility and were a bit more static (Fig. 4E, left and right panel respectively). While most histones are replaced by protamine in mature spermatozoa, a small minority carrying various histone modifications including H3K27me3 is retained [53-55]. To determine whether *Gpif* deletion impacts this residual H3K27me3, we quantified this mark in epidydimal sperm. Western blot of sperm extracts isolated from *Gpif* -/Y mice displayed higher H3K27me3 levels compared to sperm originating from WT animals (Fig. 4F). Whether this upregulation has any functional consequences remains to be investigated, nonetheless, these results confirm the inhibitory activity of GPIF on H3K27me3 deposition in male germ cells. Interestingly, they reveal that an excess of H3K27me3 is not damageable for spermatogenesis and male fertility.

GPIF controls H3K27me3 deposition in growing oocytes

In female, classical assembly of chromatin is conserved throughout oogenesis. While the genome-wide deposition of H3K27me3 in PGCs remains to be investigated, H3K27me3 was reported to be progressively restricted during oogenesis to "non-canonical" locations such as intergenic regions and gene deserts [38]. To assess whether GPIF could play a role in the regulation of H3K27me3 during oogenesis, we first investigated its expression during mouse oocyte development. It is highly expressed at all developmental stages of oocyte maturation (Fig. 5A, data from [56]) but drops sharply post-fertilization (Fig. 5B). Considering this pattern, we evaluated H3K27me3 levels in pre-pubertal female follicles (P17) by immunofluorescence: H3K27me3 levels were slightly higher in *Gpif* -/- primordial follicles compared to WT (Fig. 5C, right panel for quantification). As this difference became more pronounced in secondary follicles (Fig. 5C), we went on investigating H3K27me3 levels at the fully-grown oocyte (FGO) stage in adult females. We first harvested germinal vesicle (GV) oocytes from 3-month-old female siblings and stained for H3K27me3 and DAPI; H3K27me3 levels appeared to be around twice more abundant in *Gpif* -/- oocytes, both in the less condensed chromatin state with no rim surrounding the nucleolus (NSN) and in the fully condensed chromatin state with a DNA-dense rim surrounding the nucleolus (SN), (Fig. 5D and S5A). Of note, this effect was specific to H3K27me3 modifications, as H3K4me3 levels were not lower in NSN oocytes (Fig. S5B). Finally, a strong increase in H3K27me3 deposition was also observed in mature MII oocytes from 4 month-old *Gpif* -/- females (Fig. 5E). However, despite this gain in H3K27me3, RNA-seq analysis revealed a very similar transcriptome compared to WT (Fig. 5F). Importantly, transcription is globally silent in MII oocytes, the expression profiles therefore reflect the

regulation that occurred at the earlier stages of oocyte maturation. Collectively, these findings indicate that GPIF has a crucial role in the regulation of H3K27me3 deposition, in particular at late stages of oocyte maturation.

Impaired fertility of *Gpif* knockout females

The severe alterations in the chromatin environment of *Gpif* -/- oocytes prompted us to investigate whether GPIF might be involved in the control of follicle maturation. We did not observe any significant differences in the number of primordial, primary and secondary/antral follicles of pre-pubertal females (P17) regardless of *Gpif* expression status, indicating that the initial oocyte pool is intact (Fig. 6A top panels). In contrast, sections from older females (16 weeks) showed a global reduction in follicle number in the absence of GPIF (Fig. 6A bottom panels), although the low number of mature follicles (primary and secondary/antral) at this age was insufficient to reach statistical significance. These data nonetheless suggest a progressive, age related, exhaustion of primordial follicle reserve, from which growing follicles develop. Incidentally, ovaries from *Gpif* -/- females appeared smaller, with a weight that was reduced by about 30% compared to wild type and heterozygous counterparts (Fig 6B).

In line with these results, we observed that *Gpif* -/- female mice give rise to fewer progeny. We therefore monitored their fertility by comparing the size and number of litters of WT *versus* mutant females (Fig. 6C). Six-week-old WT and *Gpif* -/- females were mated with a reliably fertile male in the same cage and monitored daily for 20 weeks (Fig. 6C). All litters were genotyped to assign them to the correct mother; the numbers of mice at birth and at 3 weeks of age (time of genotyping) were similar. However, the total number of pups obtained from *Gpif* -/- mothers considerably decreased each month, as the females aged (Fig. 6C). This reflected both a reduction in litter number (WT females gave birth to around 3 litters over a period of 20 weeks while mutant females gave birth to only one litter) and litter size (WT females gave birth to an average of 8 pups/litter while the *Gpif* -/- average litter size was around 3/4 pups/litter) (Fig. 6D). This result is unlikely to result from developmental defects of the reproductive track since uterine horns appeared normal in adult *Gpif* -/- females (data not shown).

We conclude from these experiments that alterations in the epigenetic content of *Gpif* -/- oocytes impairs the maintenance of the oocyte pool and probably the development of the embryo around or after fertilization, eventually leading to a strong reduction in female fertility.

DISCUSSION

Gametogenesis entails significant reprogramming of the epigenome. While histone replacement in spermatogenesis and the progressive loss of DNA methylation during germ cell specification are well documented [57,37], less is known about the regulation of histone post-translational modifications during this process. Here, we focus on the Polycomb complex PRC2 to investigate this question. We identify a new PRC2 interacting protein specific to the gonad and showed that it inhibits PRC2 enzymatic activity. Inactivation of this factor leads to a global increase of H3K27me3 during both spermatogenesis and oogenesis. Alteration of the epigenetic content of oocytes impairs subsequent development resulting in a severely compromised fertility.

The PRC2 complex is tightly regulated by a set of cofactors. Two subtypes of PRC2 are reported to coexist, defined by which cofactors the core complex interacts with [13]. However, these subcomplexes appear to be at least partially redundant and their respective contributions to PRC2 function remain under investigation. GPIF contrasts with most of the cofactors identified to date: (i) its expression seems mostly restricted to germ cells (ii) homologs have only been found in *Eutherians* and it is a fast-evolving protein (iii) it is a robust inhibitor of PRC2 enzymatic activity and (iv) it interacts with both PRC2 complex subtype. These last two characteristics are likely linked: it is expected that effective inhibition of PRC2 requires the function of both subtypes to be limited. The poor sequence conservation of GPIF sequence and its rather disordered structure prediction (data not shown) are more surprising considering that PRC2 and its cofactors are, in contrast, very well conserved. This suggests that the specificity of action of GPIF on PRC2 could be primarily conferred by the conserved stretch of 13 amino acids. Such a mechanism involving a short linear motif in direct contact with binding partners (including chromatin modifiers) is a common strategy for parasites such as *toxoplasma* to manipulate the host cellular machineries [58]. It will be particularly interesting to perform structural analyses in order to precisely determine how this interaction occurs, how it interferes with the binding of AEBP2, JARID2 or PALI1 to PRC2, and how this impairs the enzymatic activity of PRC2 without impacting its recruitment to chromatin.

Another interrogation raised by this study is the advantage of expressing an inhibitor of PRC2 to limit H3K27me3 deposition in the gonad rather than downregulating the enzyme itself. It is tempting to speculate that an inhibitor enables a tighter control over the timing of the reduction

in PRC2 activity. Consistent with this possibility, *Gpif* was recently identified among a set of genes that is expressed in *PGCs*, in response to the developmental DNA demethylation of the germline genome [46]. We might have expected that when the transcriptional silencing maintained by DNA methylation is attenuated, the role of PRC2 would be more pronounced, instead, we observe that PRC2 activity has to be restrained while DNA methylation is low. Furthermore, disrupting PRC2 activity through the deletion of *Eed* has little immediate consequences on growing mouse oocytes [42]. In contrast, complete deletion of PRC1 at the same stage leads to substantial alterations of the oocyte transcriptome and developmental arrest at the two cell-stage [59]. This suggests that PRC1 might play a unique role in transcriptional control during oocyte growth. To further examine this hypothesis, it would be very informative to determine the distribution of H2AK119ub in oocytes and see whether it remains enriched at the promoters of developmental genes, unlike H3K27me3. It is worth noting that DNA methylation, which is established *de novo* during oocyte growth, is specifically found across transcribed regions [41] while H3K27me3 shows broad enrichment in intergenic regions or at gene deserts [40].

Since transcription is globally silent in the mature oocyte, PRC2 activity might be limited in order to prevent it from invading genomic regions that should be amenable to transcriptional activation during maternal-to-zygotic gene activation. Indeed, DNA methylation (at promoter CpG islands) and active transcription, both of which have been described as antagonistic to PRC2-mediated H3K27me3 deposition, are attenuated in mature oocytes thus freeing the way for PRC2. A recent report has shown that H3K27me3 on the maternal genome is important for the regulation of allele-specific gene expression [33]. Further investigation will be required to analyze the early development of embryos derived from *Gpif* -/- females focusing on genomic imprinting. Our results also suggest that excessive H3K27me3 levels resulting from *Gpif* deletion in testicular germ cells are partially retained in mature spermatozoa. Although it does not seem to impact on the fertilizing properties of the spermatozoa, it would be interesting to determine whether embryos derived from oocytes fertilized with *Gpif* -/- sperm develop normally. If they do, it would be consistent with the report that paternally inherited H3K27me3 is rapidly erased in the zygote and carries limited if any intergenerational potential [38].

Finally, both gain and loss of PRC2 function are a rather recurrent observation in cancers. While we were completing this study, another publication reported the identification of GPIF (CXORF67) as an inhibitor of PRC2 in two cancer cell lines (U2OS and Daoy Cells [60]). This constitutes another means by which cancer cell lines might curtail PRC2 function. Further

studies are required to know whether *GPIF* upregulation might be a recurrent event in cancers and act as a driver of tumor progression. Of note, *GPIF* has also been involved in gene translocations occurring in endometrial stromal sarcoma, a rare malignant tumor of the uterus [61]. Previous reports revealed frequent fusion between the transcriptional repressor *JAZF1* and the PRC2 core component *SUZ12* [62] and it was proposed that this fusion could alter PRC2 function [63]. Interestingly, PRC2 cofactors can also be involved in fusions with transcriptional regulators, as it is the case for *PHF1* with *JAZF1*, *MEAF6*, or *EPC1* [64]. Our study extends this observation by showing that the fusion between GPIF and the nuclear protein malignant brain tumor domain-containing 1 (MBTD1) could result in aberrant PRC2 targeting [64]. It will be important to investigate how these fusions contribute to tumor progression and whether the inhibition of PRC2 could constitute a therapeutic strategy.

Materials and Methods:

Cloning

m*Gpif* cDNA clone was obtained from ORIGENE (Ref. MG214772). *hGPIF* cDNA clone was amplified from HEK-293T genomic DNA. h*GPIF* mutant 1 (a.a. 1-420), *GPIF* mutant 2 (a.a.100-503), *GPIF* mutant 3 (a.a.200-503), *GPIF* mutant 4 (a.a.300-503), *GPIF* mutant 6 (a.a. 1-450) and *GPIF* mutant 7 (a.a. 1-395) were generated by PCR and cloned into pMSCV-Hygromycin retroviral vector and/or pCMV4-HA. *GPIF* mutant 5, depleted of 13AA conserved stretch (a.a.1-503, Δ 394-417) was generated by amplifying the two flanking parts by PCR with overlapping overhangs [65]. *GPIF* and *Gpif* cDNA were amplified by PCR and subcloned into pREV retroviral vector (gift form S. Ait-Si-Ali), downstream a 2X-Flag-2X-HA sequence and upstream IRES followed by CD25 cDNA.

Cell lines

U2OS (ATCC) and HEK-293T (Invitrogen) cell lines were grown according to the manufacturer's instructions. Cell lines were tested for the absence of mycoplasma every month. All transfections were performed using PEI (polyethylenimine) and 150mM NaCl at 6:1 ratio to DNA. U2OS *EED-/-* cell line was generated by co-transfecting (i) gRNA targeting *EED*, (ii) hCas9 and (iii) a targeting cassette bearing Hygromycin resistance flanked by 1kb sequences homologous to *EED* locus. Hygromycin B clone selection was performed at 0,2 mg/ml. U2OS *GPIF-/-* cell line was generated by the same strategy with a targeting construct conferring puromycin resistance (selection was performed at 0,5 mg/ml). Selected U2OS *GPIF-/-* clone has also undergone NHJ reparation with around 20bp deletion at N-terminal part of the sequence. Rescue experiments on U2OS *GPIF-/-* cell line was performed by infection with retroviral vectors expressing *GPIF* FL or mutants stably selected with Hygromycin B 0,2 mg/ml.

HeLa-S3 cells (gift form S. Ait-Si-Ali) were grown in DMEM. pREV retroviruses are produced by transfecting of 293T-Phoenix cell line (gift form S. Ait-Si-Ali) and collecting supernatant after 60 hours. HeLa-S3 cells were infected by incubation with viral supernatants for 3h at 37°C. Infected cells were then selected by FACS sorting using an anti-CD25-FITC conjugated antibody and following manufacturer instructions (BD Biosciences 553866). Expression of the recombinant proteins were assessed by WB analysis of nuclear extracts.

Retroviral production

Production of pMSCV-Hygromycin retroviral vectors was performed in 293T cells. Transduction and selection of targeted cells were performed according to the online Addgene procedure. Hygromycin B was added at 0.2 µg/ml.

Nuclear Extract and Flag-IP

For nuclear extract preparation, cells were incubated with buffer A (10mM Hepes pH 7.9, 2.5mM MgCl2, 0.25M sucrose, 0.1% NP40, 0.5mM DTT, 1mM PSMF) for 10 min on ice, centrifuged at 8000 rpm for 10 min, resuspended in buffer B (25mM Hepes pH 7.9, 1.5mM MgCl2, 700 mM NaCl, 0.5mM DTT, 0.1 mM EDTA, 20% glycerol), sonicated and centrifuged at 14000 rpm during 15 min. For immunoprecipitation 1mg of nuclear extract diluted in BC0 to a final salt concentration of 250mM was incubated with 125 µl of Flag M2 Beads (SIGMA-ALDRICH-A4596), washed three times with BC250 (50mM Tris pH7.9, 250mM KCl, 2mM EDTA, 10% Glycerol, and protease inhibitors), and eluted with 0.2 M glycine pH 2.6.

Protein Gel and Immunoblotting

Nuclear extracts were fractionated by SDS-PAGE and transferred to nitrocellulose by semi-dry transfer (Bio-Rad). Membrane were blocked with 5% nonfat milk in PBS-0.1% Tween 20 for 60 min, the membrane was washed once with the same buffer and incubated with antibodies at 4 °C overnight. Membranes were washed three times for 5 min and incubated with either a 1:5000 (Mouse)/1:10000 (Rabbit) dilution of horseradish peroxidase-conjugated secondary antibodies or with fluorophore-conjugated 1:5000 dilution of Starbright700 (Biorad) for 2 h at room temperature. Blots were washed three times and developed with SuperSignal™ West Pico PLUS Chemiluminescent Substrate (Thermo-Fisher) according to the manufacturer's protocols. Immunoblots incubated with fluorescent secondaries antibody were visualized using the BIORAD ChemiDoc MP.

Mass Spectrometry analysis

Mass spectrometry analysis was performed as described in [66] starting from either testis nuclear extracts or U2OS cell extracts. Nuclear extracts were subjected to a single step Flag-IP in triplicate (ipFLAG). Same IP was performed on extract not expressing the Flag-tagged protein [67]. Nuclear extracts from the Flag-tagged cell line were also incubated with beads as control. Thus, nine pull-downs are performed in total, three specific pull-downs and six control pull-downs.

The precipitated proteins were then subjected to on-bead trypsin digestion after which peptide mixtures are analyzed by nanoLC-MS/MS on an Orbitrap-Velos mass spectrometer. After raw data processing using MaxQuant, the obtained label-free (LFQ) intensities were used to determine statistically enriched proteins in the Flag-IP as described previously [18]. Next, iBAQ intensities for statistically enriched proteins were calculated in each of the nine pull-downs. The iBAQ values obtained in the six control samples indicate background binding. These iBAQ intensities were therefore subtracted from the iBAQ intensity in the Flag pull-downs. The resulting corrected iBAQ intensity for the Flag-tagged protein was set to 1 and the iBAQ values of the interacting proteins with their SD were scaled accordingly. This enables stoichiometry determination of all the interactors relative to the bait protein.

Recombinant proteins purification, baculoviruses

Recombinant EZH2, SUZ12, EED, RBAP48, JARID2 1-530, GPIF full length and mutant proteins were produced in SF-9 insect cells after infection with the corresponding baculoviruses as described previously [5,22,68]. Lysates were resuspended in BC300, sonicated and clarified by centrifugation before incubation with either Flag-beads (M2-beads, SIGMA_ALDRICH 4596) and eluted with Flag peptide, or Streptactin-sepharose suspension (IBA, 2-1201-010) and eluted with 2,5mM Desthiobiotin in BC300. hGPIF and mutant form baculoviruses were produced accordingly to Bac-to-Bac Baculovirus Expression Systems (Invitrogen) after cDNA cloning into pFASTbac vectors. Recombinant GPIF proteins were further purified on size exclusion chromatography (S200).

KMT assay

KMT assay with recombinant PRC2 and GPIF proteins were performed as described previously [5]. Briefly, the reaction was performed with 200ng of PRC2 alone or in presence of GPIF, 1ug of substrates, 4mM DTT in methylation reaction buffer (50mM Tris-HCl pH 8.5, 2.5mM MgCl2), ^{3}H-SAM, and incubated at 30°C for 30 min. For KMT assay with PRC2-Flag purified from U2OS WT and *GPIF* -/-, nuclear extracts were first fractionated on High Trap Q (GE Healthcare) prior to Flag-IP. Nucleosomal substrate for the assay was assembled from 5S 12 repeats DNA [69] and purified HeLa cell histone octamers by salt dialysis through a linear gradient (2.2 M NaCl to 0.4 M NaCl) followed by dialysis against TE solution.

Antibodies

Antibodies against EZH1/2, SUZ12, JARID2 and EED have been previously described [22,68].

RBAP48 mouse mAb (GWB-C12FDE) was purchased GenWay Biotech; H3 mAb (39163) has been purchased from Active Motif and polyclonal Rabbit one from cell signalling (9715); H3K9me2 (Ab1220) and H3K27Ac (Ab 4729) have been purchased from Abcam; H3K27me1 mouse mAb C0321 from Active Motif; H3K27me2 Rabbit mAb D18C8 (9728), H3K27me3 Rabbit mAb C36B11 (9733), H3K4me3 Rabbit mAb C42D8 (9751), Rabbit mAb D7C6X (14129) and mouse mAb10E2 HDAC1 (5356S) from Ozyme (Cell Signaling Technologies). hGPIF (HPA006128) and mAb Flag-M2 (F1804) purchased from SIGMA; Anti-Germ cell-specific Rabbit Polyclonal DPP3A/Stella (19878) and TRA98 (Ab82527) Rat monoclonal one from Abcam. Antibody against mGPIF was raised against the two following synthetic peptides: CAESSRAESDQSSPAG (corresponding to a.a. 91-106) and CAQSAGRNLRPRPRSS (corresponding to a.a. 192-206).

By default, primary antibodies were diluted 1:3000 for WB analysis and 1:250 for Immunostaining.

Mouse lines

EZH1- and EZH2-Flag knock-in mice.

EZH2-Flag was previously described [22], EZH1-Flag was generated by homologous recombination at the Institute Clinique de la Souris (ICS).

Gpif knockout mouse

Mice were hosted in pathogen-free Animal Facility. All experimentation was approved by the Institut Curie Animal Care and Use Committee (project APAFIS #14570-2018040917413626-v1) and adhered to European and national regulation for the protection of vertebrate animals used for experimental and other scientific purposes (directives 86/609 and 2010/63). For tissue and collection, mice were euthanized by cervical dislocation. The *Gpif* mouse line was derived by CRISPR/Cas9 engineering of a 1.5-kb deletion spanning the AU022751 locus (Supplementary Figure 4) in embryos at the one-cell stage, according to published protocols [70]. Of the 13 pups generated, 8 carried at least one modified allele. Two founders (N0) carrying the expected 1.5-kb deletion were selected. The absence of *in silico*–predicted off-target mutations was verified by Sanger sequencing, the two founders were bred with C57B6N mice. Two additional backcrosses were performed to segregate out undesired genetic events, following a systematic breeding scheme of crossing *Gpif* heterozygous females with C57B6N males to promote transmission of the deletion. Cohorts of female and male mice were then mated to study complete knockout progeny.

Histological sections and immunostainings

For histological sections, testis and ovary from either human patients from Curie Institute Pathology Platform or mice were dissected, fixed for 6h in 4% paraformaldehyde (Sigma) and washed with 70% ethanol according to pathology platform standard protocols. Organs were paraffin-embedded, sectioned (8µm) and stained with Hematoxylin using standard protocols.
For cryosections, testes and ovaries from adult mice (6 months old males; 2,5 and 5 months old females) were dissected, fixed overnight in 4% paraformaldehyde at 4°C, washed in PBS, followed by two consecutive overnight incubations in 15% and 30% sucrose at 4°C respectively. Testes were embedded in O.C.T. compound (Tissue-Tek), 8-10 µm thick sections were cut and spotted onto Superfrost Plus slides (Thermo Fisher Scientific).

For immunofluorescence detection, testis slides were brought to room temperature, blocked and permeabilized for 1h (10% donkey serum, 3% BSA and 0.2% Triton). Sections were incubated with primary antibodies at 4°C overnight, followed by three PBS-0,1% Tween-20 washes and 2h incubation with Alexa Fluor-conjugated secondary antibodies. Slides are washed three times again in PBS-0,1% Tween-20 and incubated with DAPI 1 ug/ul for 5 minutes. After a quick wash in PBS, slides were mounted with Mounting Media (Life technologies). Images were acquired with Inverted Laser Scanning Confocal LSM700 UV Zeiss microscope with a 40x objective and Z-step in the case of Z- stack scanning. For immunofluorescence on p17 female mice ovary section, antigen retrieval is first performed in 10mM Sodium Citrate, 0.05% Tween-20, pH=6 at 80°C for 20 minutes, before proceeding with following steps as indicated above. Images of the section are acquired with Leica SP8 Confocal microscope, with open pinhole and 10X objective.

Immunohistochemistry on paraffin embedded human samples

Slides were baked 1h at 65°C, before deparaffinization and hydration in xylene and graded ethanol to distilled water. Endogenous peroxidase was blocked for 5 minutes in 1,5% H_2O_2 in methanol, antigen retrieval step was performed by boiling slides for 20 min in "Antigen unmasking solution" (Vector Laboratories) and cooling down 1h at RT. Slides are quickly washed in PBS for 5 min, blocked in PBS-2%BSA-5% FBS 1h at RT and incubated overnight at 4°C in humid chamber with Primary Antibody. After washing 3 times for 5 min in PBS, slides were incubated with Biotinylated secondary antibody (Vector Laboratories) for 30 min. Slides were washed in PBS three times for 5 min, before incubating with ABC substrate for 30 min at RT. After washing again with PBS, DAB was prepared according to manufacturer

instruction (Vector Laboratories) and the staining reaction monitored from 1 to 5 minutes. Slides were stained with H/E following standard methods, dehydration steps from 90% Ethanol solution to Xylene is performed and slides were mounted in VectaMount permanent Mounting Media (Vector Laboratories).

Immunomagnetic, flow cell sorting and flow cytometry analysis of mice testis cell populations

Testicular single-cell suspensions were prepared from 2–3-months-old from WT and *Gpif* -/- mice as described previously [48]. The albuginea was removed and the seminiferous tubules were dissociated using enzymatic digestion by collagenase type I at 100 U/ml for 25 minutes at 32°C in Hanks' balanced salt solution (HBSS) supplemented with 20 mM HEPES pH 7.2, 1.2 mM $MgSO_4$, 1.3 mM $CaCl_2$, 6.6 mM sodium pyruvate, 0.05% lactate. Next, a filtration step was performed with a 40 µm nylon mesh to discard the interstitial cells. After HBSS wash, tubules were further incubated in Cell Dissociation Buffer (Invitrogen) for 25 minutes at 32°C. The resulting whole cell suspension was successively filtered through a 40 µm nylon mesh and through a 20 µm nylon mesh to remove cell clumps. After an HBSS wash, the cell pellet was resuspended in incubation buffer (same as previously plus glutamine and 1% fetal calf serum). Cell concentrations were estimated using Tryptan Blue staining (>95% viable cells).

Hoechst staining (5 µg/ml) of the cell suspensions was performed as described previously [48,71]. Cells were labelled with anti-β2m-FITC (Santa Cruz), anti-α-6 integrin-PE (GoH3) and anti-CD117 (c-KIT)-APC (2B8) antibodies (BD Pharmingen). For purification of undifferentiated spermatogonia, MACS (Miltenyi Biotech), α-6 integrin positive fraction of cells was obtained using anti-α-6 integrin-PE (GoH3) and anti-PE microbeads according to the manufacturer's protocol. This fraction was labelled with anti-β2m-FITC (Santa Cruz) and anti-CD117 (c-KIT)-APC (2B8), and then sorted. Propidium iodide (Sigma) was added before cell sorting to exclude dead cells. Analyses and cell sorting were respectively performed on LSR II and ARIA flow cytometers (Becton Dickinson).

Mouse sperm quality test

Adult male mice were euthanized by dislocation and cauda epididymis was collected post-mortem after carefully removing fat pad. Epididymis was opened and sperm released in IVF media (Vitrolife). 1:100 sperm dilution was loaded on Ivos (Hamilton Thorne machine) and

sperm parameters were evaluated by Remote Capture software.

Mouse fertility evaluation

6 week old WT and *Gpif* -/- females (N=13 each genotype) were crossed and monitored for 20 weeks. One WT female and one *Gpif* -/- female mouse were housed with an adult breeder male tested previously. Cages were monitored daily and pup numbers and litters were constantly registered. Adult females were euthanized at the end of the study and gonad morphology analysed.

Organ Phenotypic Analysis

Adult males and females starting from 3/4 months old have been euthanized, ovaries/testis collected and individually weighted. The whole organ weight has been considered for comparison among the three genotypes for females and testis weight ratio for males.

Follicle counting

Sections were prepared as described above:

For p17 mice, follicles were counted from at least 2 sections each organ/genotype (N=4 each genotype). Primary antibody against DPPA3 (Stella) was used to stain germ cells and DAPI staining to stain nuclei. Different types of follicles were classified by surrounding follicular cells shape.

For 16 week old mice, follicles number were counted from at least 13 sections each organ/genotype (N=3 each genotype). 5uM paraffin sections have been stained with Hematoxylin and Eosin. Follicle classification was based on Pedersen and Peters [72]. Ovaries were serially sectioned, and one every 3 section was counted. Follicles were counted from 13 sections and represented as average number per slide per genotype. Measurements were done using Leica Epifluorescence microscope.

Nuclei isolation and extraction from tissues

Mice tissues are rapidly extracted in PBS and homogenized with dounce (cut into small pieces with scissors, then 6x up-down with loose and 4x with tight pestle) adding sucrose solution 2.2 (Sucrose 2.2M, Hepes 1M pH7.6, KCl 3M, EDTA 0.5M, Spermine 0.1M, Spermidine 1M and protease inhibitors). Homogenized is mixed and added onto sucrose solution 2.05 (Sucrose 2.05 M, Hepes 1M pH7.6, KCl 3M, EDTA 0.5M, Spermine 0.1M, Spermidine 1M and complete set of protease inhibitors added last minute) in ultracentrifuge Beckemann tubes. Spin 45 min, 24k

at 1°C in SW28 rotor. Nuclei pellet were resuspended in an equal volume of Nuclear Lysis Buffer (Hepes 1M, pH7.6, KCl 3M, EDTA 0.5M, Glycerol 87%, Spermine 0.1M, Spermidine 1M, NaF 0.5M, Na_2VO_4 0.5M, $ZnSO_4$ 50mM and complete set of protease inhibitors). Samples were snap-frozen in liquid nitrogen.

GV and MII oocytes isolation from female mice

Germinal vesicle (GV) stage oocytes were obtained from 12 weeks old females. The ovaries were removed, passed in pre-warmed PBS and transferred to M2 medium supplemented with 100ug/mL of dibutyryl cyclic AMP (dbcAMP; Sigma-Aldrich) at 38°C. The ovarian follicles were punctured with a 21-gauge needle and GV oocytes (fully grown oocytes exhibiting a centrally located GV) have been washed five times through M2 droplets in order to ensure that dbcAMP is removed. Zona pellucida has been removed by 3 passages in tyroide acid solution, followed by 3 washes in M2 medium. GV oocytes was further washed in PBS and processed for Immunostaining as described below.

16-week-old mice were super-ovulated by intraperitoneal injection using 5 IU pregnant mare's serum gonadotropin (PMSG) and 5 IU of human chorionic gonadotropin (hCG) 48 hours later. MII oocytes were collected from ovaries of 12 weeks old mice in M2 medium. Cumulus-oocyte complexes were collected from infundibulum at 14 hours after hCG treatment and recovered in M2 medium. Cumulus cells were dispersed by hyaluronidase (300 IU/mL) for 5 minutes in M2 medium and oocytes were washed twice with M2 medium and left in PBS for 5 min. No zona pellucida removal treatment has been performed.

Immunofluorescent staining of GV and MII oocytes

GV or MII Oocytes were fixed for 20 minutes in PBS containing 2,5% paraformaldehyde (PFA) at room temperature and washed with 1% BSA-PBS three times. The cells were permeabilized by incubating in 1% BSA-PBS containing 0.5% Triton X-100 for 30 min at RT. After washing with 1% BSA-PBS three times, oocytes are incubated with primary antibodies in 1% BSA-PBS containing 0.1% Triton X-100 O/N at 4°C. They are subsequently washed once with 1% BSA-PBS-0.1% Triton X-100 and incubated for 1 hour in the dark with Alexa Fluor 488-conjugated IgG secondary antibody (dilution 1:250) in the same buffer. DNA was stained twice for 15 minutes with 4,6-diamidimo-2-phenylindole (DAPI) prior to a 15 min wash in PBS-0,1% Triton X-100. Cells are passed through increasing percentages of glycerol solution for increasing times (2,5% for 5 min- 5% for 5 min-10% for 10 min- 20% for 5 min-50% for 15 min-DTG for 15

min) were then mounted on glass slides with ProLong Gold mounting medium (Life Technologies) for sequential Z-stack imaging. Fluorescence was detected using an Inverted Laser Scanning Confocal LSM700 UV Zeiss microscope with a 63x objective and Z- stack scanning. More than 8 oocytes were examined for each condition unless otherwise specified.

RNA extraction from Mature MII oocytes

Oocytes were incubated in M2 containing tyroide acid's solution for 2–3 min to remove their ZP (zona pellucida). ZP-free oocytes were carefully washed several times with M2 and were pooled prior to lysis in XB buffer from Arcturus PicoPure RNA isolation Kit (Applied Biosystems). We then added the spike-in control External RNA Control Consortium (ERCC) molecules (Invitrogen). Normalization was performed using ERCC spike in at 1:1000000. The purified total RNA concentration was measured using Agilent High Sensitivity RNA ScreenTape on Agilent 2200 TapeStation. First-strand cDNA (from total RNA) was synthesized according to the SMART-Seq™ v4 Ultra™ Low Input RNA Kit protocol (Clontech Laboratories). The PCR-amplified cDNA was purified using SPRI beads (Beckmann Coulter).

For sequencing, 75 bp paired-end reads were generated using the Illumina MiSeq. Raw reads were trimmed for adapters with cutadapt (1.12) using the Trim Galore! (0.4.4) wrapper (default settings) and subsequently mapped to the complete mouse rRNA sequence with Bowtie2 (2.2.9). Reads that did not map to rRNA were then mapped with STAR (2.5.3a) to the full reference genome (UCSC build GRCm38/mm10) (including the RNA spike-in control sequence ERCC92, ThermoFisher cat. no. 4456740) using the following parameters: --outSAMtype BAM SortedByCoordinate --runMode alignReads --outFilterType BySJout --outFilterMultimapNmax 20 --alignSJoverhangMin 8 --alignSJDBoverhangMin 1 --outFilterMismatchNmax 999 --outFilterMismatchNoverLmax 0.04 --alignIntronMin 20 --alignIntronMax 1000000 --alignMatesGapMax 1000000 --outSAMprimaryFlag OneBestScore --outMultimapperOrder Random --outSAMattributes All. Gene counts were generated using STAR --quant_mode (uniquely mapped, properly paired reads that overlap the exon boundaries of each gene).

For differential expression analysis, genes were filtered to include those with CPM > 0.2 in at least 2 samples and filtered counts were transformed to log2-CPM and normalized with the TMM method. A linear model was fit to the normalized data and empirical Bayes statistics were computed. Differentially expressed genes for the KO versus WT were identified from

the linear fit after adjusting for multiple testing and filtered to include those with FDR < 0.05 and absolute log2fold-change > 1.

RT-qPCR from mouse tissues

Total RNA was isolated using the Rneasy Mini Kit (Qiagen). cDNA was synthetized using High Capacity cDNA RT kit (4368814-Applied Biosystems) and quantitative PCR was performed with technical triplicate using SYBR green reagent (Roche) on a ViiA7 equipment (Applied Biosystems). At least three biological independent experiments were performed for each assay. Primers sequences are provided in table S1.

RNA-seq in Spermatogonial stem cells

Kit- spermatogonial population was isolated by FACS as specified above from adult mice testis WT and *Gpif* -/Y. Sorted cells were resuspended directly in XB lysis buffer from Arcturus PicoPure RNA isolation Kit (Applied Biosystems). The purified total RNA was stored in nuclease-free water and RNA concentration was measured using Agilent High Sensitivity RNA ScreenTape. cDNA synthesis and library preparation were performed using SMARTer Stranded Total RNA-Seq Kit-Pico Input Mammalian. 100bp paired-end reads were generated using the HiSeq 2500 platform. Raw reads were trimmed for adapters with cutadapt (1.12) using the Trim Galore! (0.4.4) wrapper (default settings) and subsequently mapped to the complete mouse rRNA sequence with Bowtie2 (2.2.9). Unmapped reads were then mapped with STAR (2.5.3a) to the full reference genome (GRCm38/mm10). Libraries were confirmed to be stranded according to RSeQC after sampling 200000 reads with MAPQ>30. Gene counts were generated with STAR --quant_mode (uniquely mapped, properly paired reads that overlap the exon boundaries of each gene) using the Ensembl GTF annotation (vM13).

For differential expression analysis, genes were filtered to include those with CPM > 1 in at least 2 samples. Raw count data was normalized with the TMM method and transformed to log2-CPM. A linear model was fit to the normalized data, adjusting for batch effects, and empirical Bayes statistics were computed. Differentially expressed genes for each KO versus WT were identified from the linear fit after adjusting for multiple testing and filtered to include those with FDR < 0.05 and absolute log2fold-change > 1.

RNA sequencing in U2OS cells

Total RNA from U2OS cells was extracted with TRIzol. cDNA were generated according to manufacturer protocols (Illumina). 50bp single-end reads were generated using the HiSeq 2500 platform. Reads were first mapped to the complete human rRNA sequence with Bowtie2 (2.2.9). Unmapped reads were then mapped with STAR (2.5.2b) to the complete human reference genome (GRCh37/hg19). Gene counts were generated using STAR --quant_mode. Libraries were confirmed to be strand-specific according to RSeQC (2.6.4) after sampling 200000 reads with MAPQ>30.

For differential expression analysis, genes with CPM < 1 in 2 or less samples were removed. Raw count data was normalized with the TMM method and converted to log2-CPM. A linear model was fit to the normalized data and empirical Bayes statistics were computed for each comparison. Differentially expressed genes for each comparison were identified from the linear fit after adjusting for multiple testing and filtered to include those with FDR < 0.05.

ChIP q-PCR and ChIP sequencing

ChIP was performed as previously described [68]. EED ChIP-qPCR was performed by double-fixation with 1/10 ESG (Sigma-Aldrich) solution (1.5mM, DMSO stock) in PBS at RT 30 min first, followed by the incubation of final 1% formaldehyde 10 min at RT. Fixation step was blocked by adding 0,25M Glycine at RT 5 min. Cells are washed 2X in PBS and spun 500g 5 min RT. Fixed cells were lysed in 0.5mL NLB (0.5% Triton X-100, 0.1 M sucrose, 5 mM MgCl2, 1 mM EDTA, 10 mM Tris-HCl pH 8.0+Prot Inhibitors) for 10min on ice. Samples were diluted to final volume of 2ml NLB and douce-homogenized 20xA and 20xB. Fixed nuclei were pelleted (2000rpm, 15 min at 4°C) and resuspended in 0.3 mL Lysis Buffer (1 mM EDTA, 0.5 mM EGTA, 10 mM Tris pH 8.0, 0.5% sarcosyl plus protease inhibitors). Chromatin was sheared using using a BioruptorPlus (Diagenode) set to high for 45 (3x15) cycles of 30 s and 30 s off setting High and spun 15 min at 4°C max speed. The supernatant is subsequently diluted with 1.2 ml of (1.25% Triton, 0.125%NaDoc, 6 mM EDTA, 10mM Tris pH8 plus, protease inhibitor). IP is performed O/N at 4°C after pre-washing Prot. A Dynabeads previously coated with antibodies for min 2hrs at +4°C. The rest of the protocol is similar to the one above. ChIP-qPCR primers sequences are provided in the Supplementary Table S1.

100 bp single-end reads were generated using the HiSeq2500 platform. Reads were mapped to the human reference genome (GRCh37/hg19) with Bowtie2 (2.2.9) using end-to-end alignment with the preset --very-sensitive. PCR duplicates were removed with Picard Tools

MarkDuplicates (1.97) and BAM files were filtered to exclude common artifact regions (http://mitra.stanford.edu/kundaje/akundaje/release/blacklists/hg19-human). Reads were counted in bins of length 50 and RPKM normalized and converted to bigWig format using DeepTools bamCoverage (2.4.1). Peaks were called using SICER (1.1) on combined replicates using the EED KO as the control for both dCXorf67 and WT with the following parameters: Window size=200, Gap Size=600, E-value=500, FDR=1e-8. 24112 regions were significantly increased in the KO and 326 in the WT after excluding artifact regions.

Statistical analysis

The significance of the data obtained was determined by performing unpaired, non-parametric test of Kologorov-Smirnov. P values were considered significant when ≤0.05: ***$P \leq 0.001$, **$P \leq 0.01$, *$P \leq 0.05$.

References:

1. Margueron, R. & Reinberg, D. The Polycomb complex PRC2 and its mark in life. *Nature* **469**, 343-9 (2011).
2. Pengelly, A.R., Copur, O., Jackle, H., Herzig, A. & Muller, J. A histone mutant reproduces the phenotype caused by loss of histone-modifying factor Polycomb. *Science* **339**, 698-9 (2013).
3. Ferrari, K.J. et al. Polycomb-dependent H3K27me1 and H3K27me2 regulate active transcription and enhancer fidelity. *Mol Cell* **53**, 49-62 (2014).
4. Holoch, D. & Margueron, R. Mechanisms Regulating PRC2 Recruitment and Enzymatic Activity. *Trends Biochem Sci* **42**, 531-542 (2017).
5. Margueron, R. et al. Role of the polycomb protein EED in the propagation of repressive histone marks. *Nature* **461**, 762-7 (2009).
6. Blackledge, N.P. et al. Variant PRC1 complex-dependent H2A ubiquitylation drives PRC2 recruitment and polycomb domain formation. *Cell* **157**, 1445-59 (2014).
7. Mendenhall, E.M. et al. GC-rich sequence elements recruit PRC2 in mammalian ES cells. *PLoS Genet* **6**, e1001244 (2010).
8. Yuan, W. et al. Dense chromatin activates Polycomb repressive complex 2 to regulate H3 lysine 27 methylation. *Science* **337**, 971-5 (2012).
9. Yuan, W. et al. H3K36 methylation antagonizes PRC2-mediated H3K27 methylation. *J Biol Chem* **286**, 7983-9 (2011).
10. Bartke, T. et al. Nucleosome-interacting proteins regulated by DNA and histone methylation. *Cell* **143**, 470-84 (2010).
11. Cooper, S. et al. Targeting polycomb to pericentric heterochromatin in embryonic stem cells reveals a role for H2AK119u1 in PRC2 recruitment. *Cell Rep* **7**, 1456-1470 (2014).
12. Vizan, P., Beringer, M., Ballare, C. & Di Croce, L. Role of PRC2-associated factors in stem cells and disease. *FEBS J* **282**, 1723-35 (2015).
13. Hauri, S. et al. A High-Density Map for Navigating the Human Polycomb Complexome. *Cell Rep* **17**, 583-595 (2016).
14. Chen, S., Jiao, L., Shubbar, M., Yang, X. & Liu, X. Unique Structural Platforms of Suz12 Dictate Distinct Classes of PRC2 for Chromatin Binding. *Mol Cell* **69**, 840-852 e5 (2018).
15. Alekseyenko, A.A., Gorchakov, A.A., Kharchenko, P.V. & Kuroda, M.I. Reciprocal interactions of human C10orf12 and C17orf96 with PRC2 revealed by BioTAP-XL cross-linking and affinity purification. *Proc Natl Acad Sci U S A* **111**, 2488-93 (2014).
16. Conway, E. et al. A Family of Vertebrate-Specific Polycombs Encoded by the LCOR/LCORL Genes Balance PRC2 Subtype Activities. *Mol Cell* **70**, 408-421 e8 (2018).
17. Zhang, Z. et al. PRC2 complexes with JARID2, MTF2, and esPRC2p48 in ES cells to modulate ES cell pluripotency and somatic cell reprogramming. *Stem Cells* **29**, 229-40 (2011).
18. Smits, A.H., Jansen, P.W., Poser, I., Hyman, A.A. & Vermeulen, M. Stoichiometry of chromatin-associated protein complexes revealed by label-free quantitative mass spectrometry-based proteomics. *Nucleic Acids Res* **41**, e28 (2013).
19. Kloet, S.L. et al. The dynamic interactome and genomic targets of Polycomb complexes during stem-cell differentiation. *Nat Struct Mol Biol* **23**, 682-690 (2016).
20. Son, J., Shen, S.S., Margueron, R. & Reinberg, D. Nucleosome-binding activities within JARID2 and EZH1 regulate the function of PRC2 on chromatin. *Genes Dev* **27**, 2663-77 (2013).
21. Peng, J. et al. Jarid2/Jumonji Coordinates Control of PRC2 Enzymatic Activity and Target Gene Occupancy in Pluripotent Cells. *Cell* **139**, 1290-1302 (2009).

22. Li, G. et al. Jarid2 and PRC2, partners in regulating gene expression. *Genes Dev* **24**, 368-80 (2010).
23. Pasini, D. et al. JARID2 regulates binding of the Polycomb repressive complex 2 to target genes in ES cells. *Nature* **464**, 306-10 (2010).
24. Sanulli, S. et al. Jarid2 Methylation via the PRC2 Complex Regulates H3K27me3 Deposition during Cell Differentiation. *Mol Cell* **57**, 769-83 (2015).
25. Kim, H., Kang, K. & Kim, J. AEBP2 as a potential targeting protein for Polycomb Repression Complex PRC2. *Nucleic Acids Res* **37**, 2940-50 (2009).
26. Grijzenhout, A. et al. Functional analysis of AEBP2, a PRC2 Polycomb protein, reveals a Trithorax phenotype in embryonic development and in ESCs. *Development* **143**, 2716-23 (2016).
27. Lee, C.H. et al. Allosteric Activation Dictates PRC2 Activity Independent of Its Recruitment to Chromatin. *Mol Cell* **70**, 422-434 e6 (2018).
28. Kasinath, V. et al. Structures of human PRC2 with its cofactors AEBP2 and JARID2. *Science* **359**, 940-944 (2018).
29. Oksuz, O. et al. Capturing the Onset of PRC2-Mediated Repressive Domain Formation. *Mol Cell* **70**, 1149-1162 e5 (2018).
30. Hojfeldt, J.W. et al. Accurate H3K27 methylation can be established de novo by SUZ12-directed PRC2. *Nat Struct Mol Biol* **25**, 225-232 (2018).
31. Umlauf, D. et al. Imprinting along the Kcnq1 domain on mouse chromosome 7 involves repressive histone methylation and recruitment of Polycomb group complexes. *Nat Genet* **36**, 1296-300 (2004).
32. Lewis, A. & Murrell, A. Genomic imprinting: CTCF protects the boundaries. *Curr Biol* **14**, R284-6 (2004).
33. Inoue, A., Jiang, L., Lu, F., Suzuki, T. & Zhang, Y. Maternal H3K27me3 controls DNA methylation-independent imprinting. *Nature* **547**, 419-424 (2017).
34. Mu, W., Starmer, J., Fedoriw, A.M., Yee, D. & Magnuson, T. Repression of the soma-specific transcriptome by Polycomb-repressive complex 2 promotes male germ cell development. *Genes Dev* **28**, 2056-69 (2014).
35. Mu, W., Starmer, J., Shibata, Y., Yee, D. & Magnuson, T. EZH1 in germ cells safeguards the function of PRC2 during spermatogenesis. *Dev Biol* **424**, 198-207 (2017).
36. Rathke, C., Baarends, W.M., Awe, S. & Renkawitz-Pohl, R. Chromatin dynamics during spermiogenesis. *Biochim Biophys Acta* **1839**, 155-68 (2014).
37. Hajkova, P. et al. Chromatin dynamics during epigenetic reprogramming in the mouse germ line. *Nature* **452**, 877-81 (2008).
38. Zheng, H. et al. Resetting Epigenetic Memory by Reprogramming of Histone Modifications in Mammals. *Mol Cell* **63**, 1066-79 (2016).
39. Liu, X. et al. Distinct features of H3K4me3 and H3K27me3 chromatin domains in pre-implantation embryos. *Nature* **537**, 558-562 (2016).
40. Xu, Q. & Xie, W. Epigenome in Early Mammalian Development: Inheritance, Reprogramming and Establishment. *Trends Cell Biol* **28**, 237-253 (2018).
41. Hanna, C.W., Demond, H. & Kelsey, G. Epigenetic regulation in development: is the mouse a good model for the human? *Hum Reprod Update* **24**, 556-576 (2018).
42. Prokopuk, L. et al. Loss of maternal EED results in postnatal overgrowth. *Clin Epigenetics* **10**, 95 (2018).
43. Kalb, R. et al. Histone H2A monoubiquitination promotes histone H3 methylation in Polycomb repression. *Nat Struct Mol Biol* **21**, 569-71 (2014).
44. Maier, V.K. et al. Functional Proteomic Analysis of Repressive Histone Methyltransferase Complexes Reveals ZNF518B as a G9A Regulator. *Mol Cell Proteomics* **14**, 1435-46 (2015).

45. Percharde, M., Wong, P. & Ramalho-Santos, M. Global Hypertranscription in the Mouse Embryonic Germline. *Cell Rep* **19**, 1987-1996 (2017).
46. Hill, P.W.S. et al. Epigenetic reprogramming enables the transition from primordial germ cell to gonocyte. *Nature* **555**, 392-396 (2018).
47. Tang, W.W. et al. A Unique Gene Regulatory Network Resets the Human Germline Epigenome for Development. *Cell* **161**, 1453-67 (2015).
48. Corbineau, S. et al. Spermatogonial stem cells and progenitors are refractory to reprogramming to pluripotency by the transcription factors Oct3/4, c-Myc, Sox2 and Klf4. *Oncotarget* **8**, 10050-10063 (2017).
49. Barroca, V. et al. Mouse differentiating spermatogonia can generate germinal stem cells in vivo. *Nat Cell Biol* **11**, 190-6 (2009).
50. Lifschytz, E. & Lindsley, D.L. The role of X-chromosome inactivation during spermatogenesis (Drosophila-allocycly-chromosome evolution-male sterility-dosage compensation). *Proc Natl Acad Sci U S A* **69**, 182-6 (1972).
51. Iwamori, N., Iwamori, T. & Matzuk, M.M. H3K27 demethylase, JMJD3, regulates fragmentation of spermatogonial cysts. *PLoS One* **8**, e72689 (2013).
52. Bourc'his, D. & Bestor, T.H. Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. *Nature* **431**, 96-9 (2004).
53. Hammoud, S.S. et al. Distinctive chromatin in human sperm packages genes for embryo development. *Nature* **460**, 473-8 (2009).
54. Hisano, M. et al. Genome-wide chromatin analysis in mature mouse and human spermatozoa. *Nat Protoc* **8**, 2449-70 (2013).
55. Carone, B.R. et al. High-resolution mapping of chromatin packaging in mouse embryonic stem cells and sperm. *Dev Cell* **30**, 11-22 (2014).
56. Veselovska, L. et al. Deep sequencing and de novo assembly of the mouse oocyte transcriptome define the contribution of transcription to the DNA methylation landscape. *Genome Biol* **16**, 209 (2015).
57. Kimmins, S. & Sassone-Corsi, P. Chromatin remodelling and epigenetic features of germ cells. *Nature* **434**, 583-9 (2005).
58. Hakimi, M.A. & Bougdour, A. Toxoplasma's ways of manipulating the host transcriptome via secreted effectors. *Curr Opin Microbiol* **26**, 24-31 (2015).
59. Posfai, E. et al. Polycomb function during oogenesis is required for mouse embryonic development. *Genes Dev* **26**, 920-32 (2012).
60. Pajtler, K.W. et al. Molecular heterogeneity and CXorf67 alterations in posterior fossa group A (PFA) ependymomas. *Acta Neuropathol* **136**, 211-226 (2018).
61. Puliyath, G. & Nair, M.K. Endometrial stromal sarcoma: A review of the literature. *Indian J Med Paediatr Oncol* **33**, 1-6 (2012).
62. Hrzenjak, A. JAZF1/SUZ12 gene fusion in endometrial stromal sarcomas. *Orphanet J Rare Dis* **11**, 15 (2016).
63. Ma, X. et al. The JAZF1-SUZ12 fusion protein disrupts PRC2 complexes and impairs chromatin repression during human endometrial stromal tumorogenesis. *Oncotarget* **8**, 4062-4078 (2017).
64. Dewaele, B. et al. Identification of a novel, recurrent MBTD1-CXorf67 fusion in low-grade endometrial stromal sarcoma. *Int J Cancer* **134**, 1112-22 (2014).
65. Gibson, D.G. et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* **6**, 343-5 (2009).
66. Baymaz, H.I., Spruijt, C.G. & Vermeulen, M. Identifying nuclear protein-protein interactions using GFP affinity purification and SILAC-based quantitative mass spectrometry. *Methods Mol Biol* **1188**, 207-26 (2014).

67. Boufassa, F. et al. CD4 dynamics over a 15 year-period among HIV controllers enrolled in the ANRS French observatory. *PLoS One* **6**, e18726 (2011).
68. Margueron, R. et al. Ezh1 and Ezh2 maintain repressive chromatin through different mechanisms. *Mol Cell* **32**, 503-18 (2008).
69. Dorigo, B. et al. Nucleosome arrays reveal the two-start organization of the chromatin fiber. *Science* **306**, 1571-3 (2004).
70. Yang, H. et al. One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell* **154**, 1370-9 (2013).
71. Barroca, H., Castedo, S., Vieira, J., Teixeira, M. & Muller-Hocker, J. Altered expression of key cell cycle regulators in renal cell carcinoma associated with Xp11.2 translocation. *Pathol Res Pract* **205**, 466-72 (2009).
72. Pedersen, T. & Peters, H. Proposal for a classification of oocytes and follicles in the mouse ovary. *J Reprod Fertil* **17**, 555-7 (1968).

Acknowledgments:

Work in the laboratory of R.M. is supported by the labelisation ARC (Fondation pour la Recherche sur le Cancer), the Labex DEEP and the Institut Curie. R.R. was recipient of fellowship from the FRM (Fondation pour la Recherche Médicale). High-throughput sequencing was performed by the NGS platform of the Institut Curie, supported by grants ANR-10-EQPX-03 and ANR10-INBS-09-08 from the Agence Nationale de le Recherche (investissements d'avenir) and by the Canceropôle Ile-de-France. We thank very much Fatima El Marjou and the Mice Transgenic Facility in Curie Institute for the creation of Gpif mouse model. Mice sperm quality analysis was realized with the help of Céline Daviaud (IVF platform at Curie Institute) and Franck Bourgade (Pasteur Institute). Tissue sections were prepared by the Histo-pathology platform (Cochin Institute). We acknowledge the Institute Curie imagery facility for the support in microscopy analysis. The Curie Pathology department donated human gonadal samples. We thank Joan Barau, Katia Ancelin, Rafael Galupa and Edith Heard for stimulating discussions and members of the Margueron lab for valuable comments on the manuscript.

Author Contributions:

R.R. performed most of the experiments, R. M. conceived the study. R.P.P and D.B. helped with oogenesis characterization. M.G. and P.F. helped with spermatogenesis characterization. M.B. and M.E.T.P performed in vivo analysis on P17 females. I.B., P.J. and M.V. performed IP/Mass Spectrometry analysis of Ezh1/Ezh2-Flag testis, GPIF-Flag HeLa and U2OS cell lines. D.Z. performed bioinformatic analysis. A.M. provided technical support. R.R. and R.M. prepared the manuscript. All authors contributed to experimental designed and edited the manuscript.

Declaration of Interests:
The authors declare no competing interests.

Figures and legends:

Figure 1. GPIF interacts with PRC2 in gonads.

(A) Volcano plot representation of EZH2 interactome from EZH2-Flag mice testis IP compared to WT. In red are all the core complex subunits, in green the cofactors and in blue GPIF. (B) Volcano plot representation of GPIF interactome after Flag-IP from HeLa-S3 overexpressing GPIF compare to WT control. Same color codes as in (A). (C) Schematic representation of GPIF protein sequence from *Mus Musculus* (upper part) and *Homo Sapiens* (middle part). Serine-rich region is colored in beige and conserved amino acid stretch in green. The conserved sequence stretch is displayed as well as protein residues conservation between the 2 sequences in green (Sequence Homology determined using Genious software) (D) *Gpif* and *Ezh2* mRNA relative abundance normalized to *Tbp* in various mice tissues (mean ± s.d., n≥2). (E) GPIF and EZH2 IHC representative staining on human adult seminiferous tubules sections. (F) GPIF and EZH2 IHC representative staining on human adult ovaries sections, red arrows indicate the follicles.

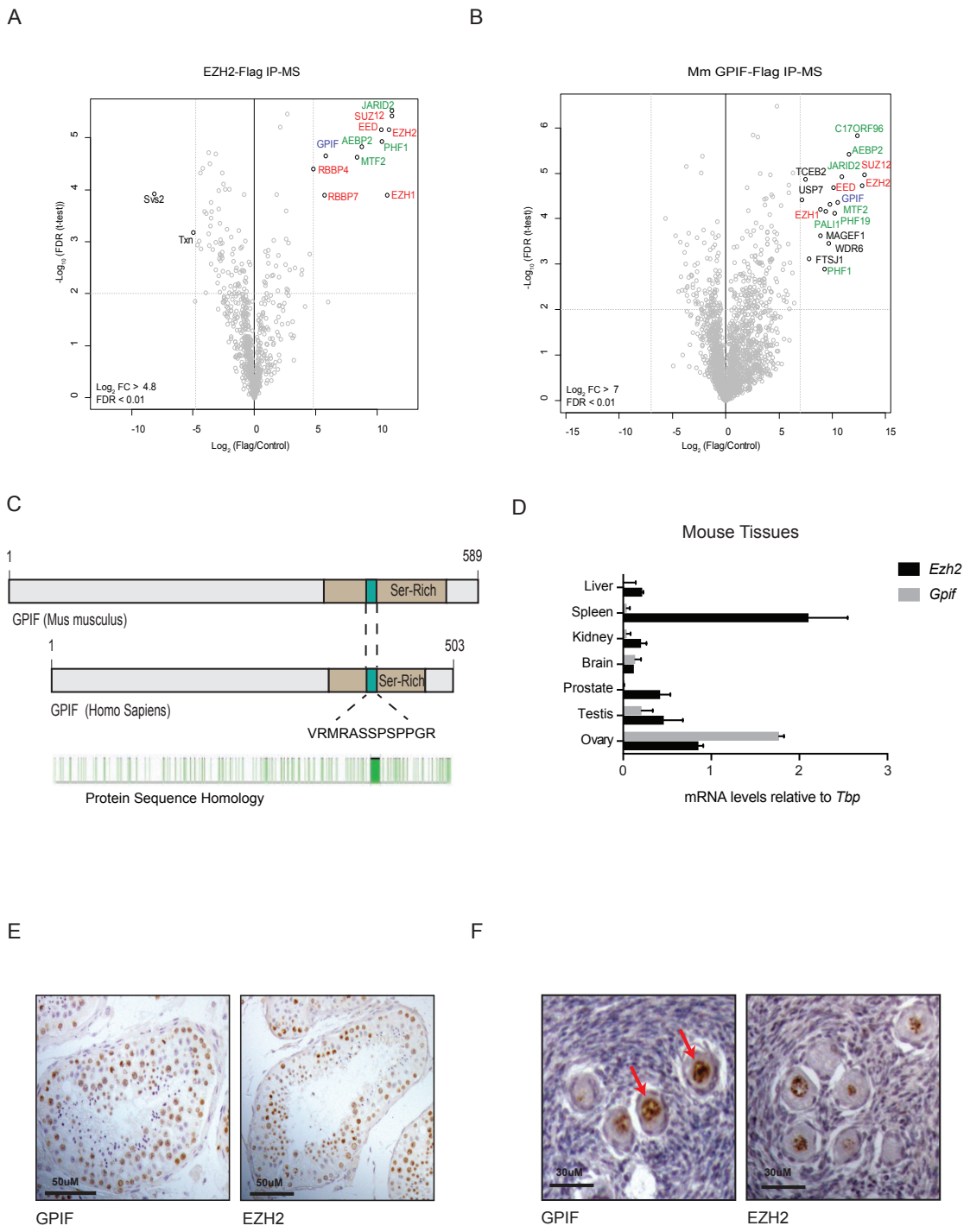


Fig. 1 Ragazzini et al

Figure 2. GPIF inhibits H3K27me3 deposition.

(A) Western blot analysis of PRC2 core complex subunits (SUZ12, RBAP48, EED and EZH2), GPIF and HDAC1 (loading control) on U2OS nuclear extracts WT, *EED -/-* and *GPIF -/-*. (B) Western blot analysis of H3K27 methylation: H3K27me1, H3K27me2, H3K27me3 and H3 (loading control). (C) Scheme representing GPIF mutants (left panel) stably reintroduced in U2OS *GPIF -/-* line. The red lambda indicates the epitope recognized by antibody detecting GPIF protein. Western blot analysis (right panel) of cell lines expressing GPIF FL or mutants probed with antibodies against GPIF (top panel), H3K27me3 (middle panel) or LAMIN B1 (loading control, bottom panel). (D) Genome-browser image of H3K27me3 enrichment in U2OS WT, *EED -/-* and *GPIF -/-*. (E) Venn diagram showing overlap between peaks enriched for H3K27me3 in U2OS WT and *GPIF -/-* after ChIP-seq analysis. (F) Scatterplot showing log2 fold-change (logFC) expression between U2OS WT and *EED-/-* versus average log2 counts per million (logCPM), right corner: number of genes significantly differently expressed (FDR<0,05) corresponding to the dots colored in red. Bottom, same as above comparing expression between U2OS WT and *GPIF-/-*.

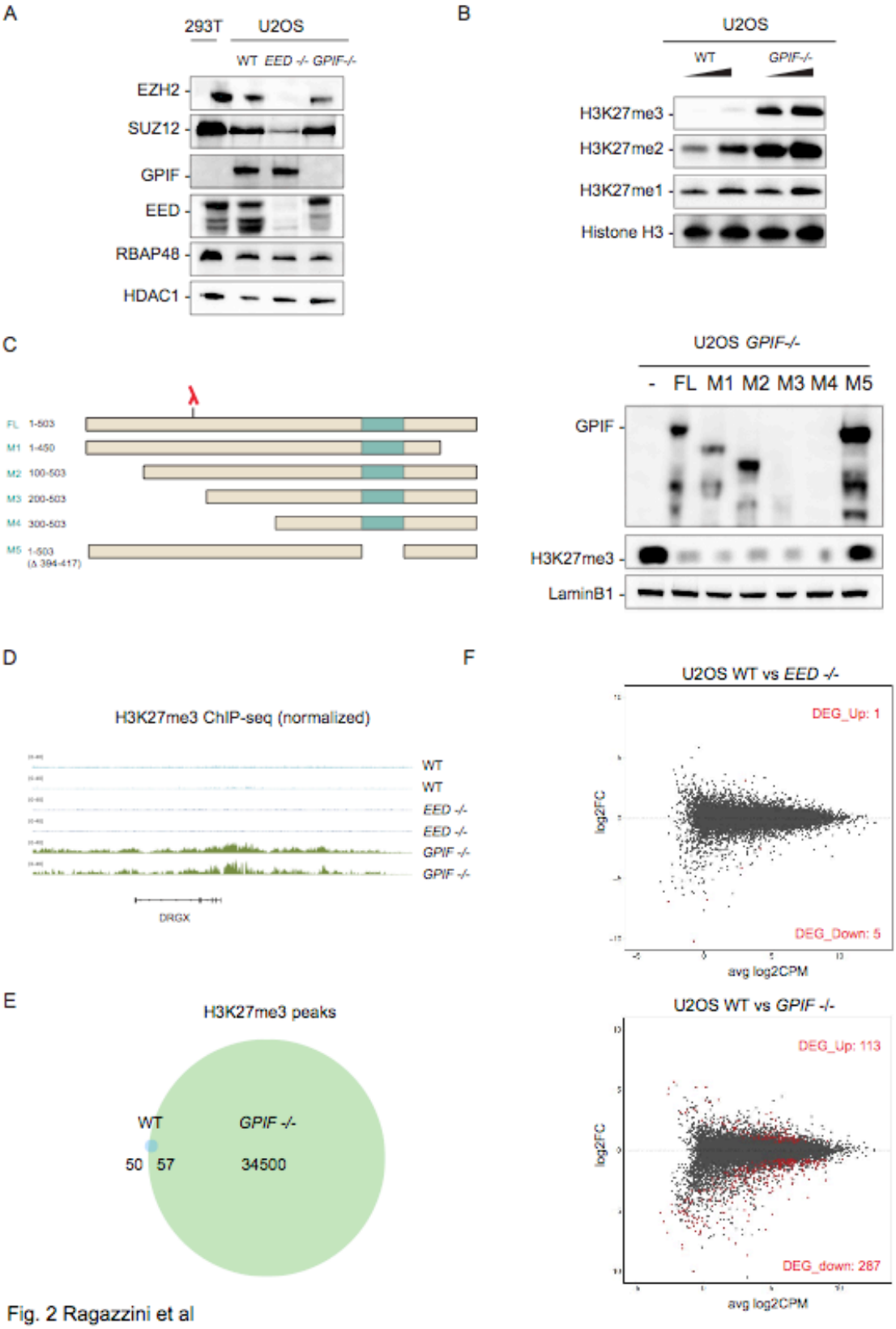
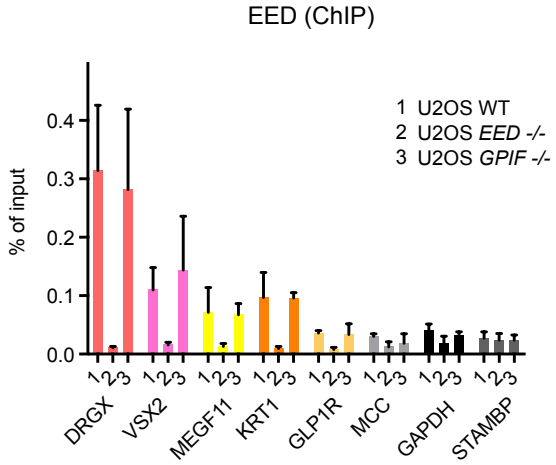


Fig. 2 Ragazzini et al

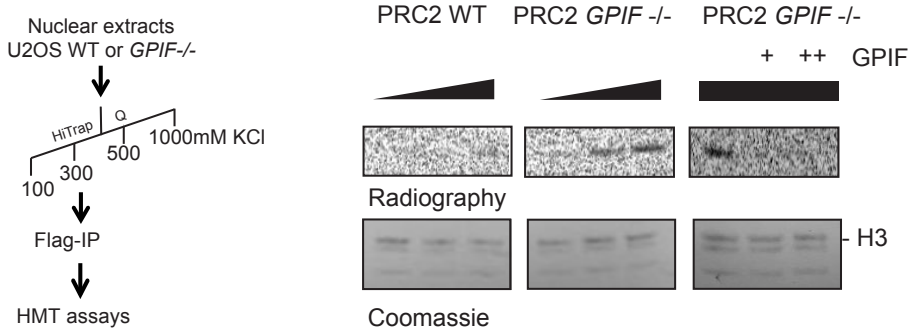
Figure 3. GPIF mitigates PRC2-cofactors interactions.

(A) ChIP-qPCR detecting EED at the genes indicated below the X axis in U2OS WT, *EED* -/- and *GPIF* -/- (mean ± s.d., n=3) (B) Left, Purification scheme for PRC2. Right, histone methyltransferase (HMT) assay to monitor the enzymatic activity of PRC2 purified from WT (left panel) or *GPIF*-/- (middle panel) U2OS cells (Titration: 1, 2, 5X) on native nucleosomes. Right panel, same assay as previously but this time titrating recombinant hGPIF on PRC2-purified from *GPIF*-/- U2OS cells (PRC2 quantity 5X). The upper panels are autoradiography and the lower panels are the corresponding SDS-PAGE coomassie staining. (C) Quantification of EZH2-Flag IP through mass spectrometry (iBAQ values). Heatmap representing the Log2-transformed median centered values. Horizontal axis: U2OS WT and *GPIF*-/-, n=3. Vertical axes: PRC2 components. Values are normalized on iBAQ values from untagged U2OS wt and *GPIF*-/-. See also Figure S3D. (D) EZH2-Flag Co-IP from nuclear extracts either WT or *GPIF* -/- and probed with antibodies against EZH2, EED, AEBP2 or JARID2.

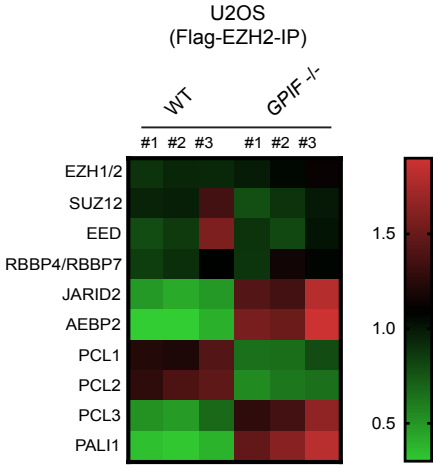
A



B



C



D

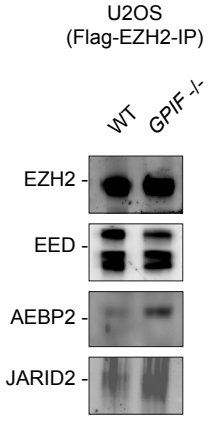


Fig. 3 Ragazzini et al

Figure 4. Global increase of H3K27me2/3 in male germ cells *Gpif* -/-.

(A) Western blot analysis of H3K27ac, H3K4me3, H3K27me3, H3K27me3, H3K9me2 and H3 on whole testis extracts WT and *Gpif* -/Y (Titration 1, 2.5X). Bottom, same extracts probed for EZH2 and SUZ12. (B) Immunofluorescence detection of H3K27me3 (green) and TRA98 (purple) in testis sections (6 months old mice). Nucleus is stained with DAPI. (C) Scatterplot showing log2 fold-change (logFC) expression between adult male WT and *Gpif* -/Y sorted spermatogonial population versus average log2 counts per million (logCPM). Differentially expressed genes are highlighted in red (upregulated: 37; downregulated :86, FDR<0,05). (D) Quantification of spermatocyte I (4N), spermatocyte II (2N), spermatids (N), and differentiating (a-6+ kit+) and undifferentiated spermatogonial (a-6+ kit-) by FACS in percent of the total germinal cell population from WT and *Gpif* -/Y mice (mean ± sem, n=4). (E) Sperm quality was measured using computer-assisted IVO technologies comparing WT and *Gpif* -/Y knockout males. Lower left panel: percent sperm progressivity. Lower right panel: percent sperm rapidity (mean ± s.d., n≥ 2). (F) Western blot analysis of H3K27me3 and H3 on mice sperm extracted from WT and *Gpif* -/Y animals.

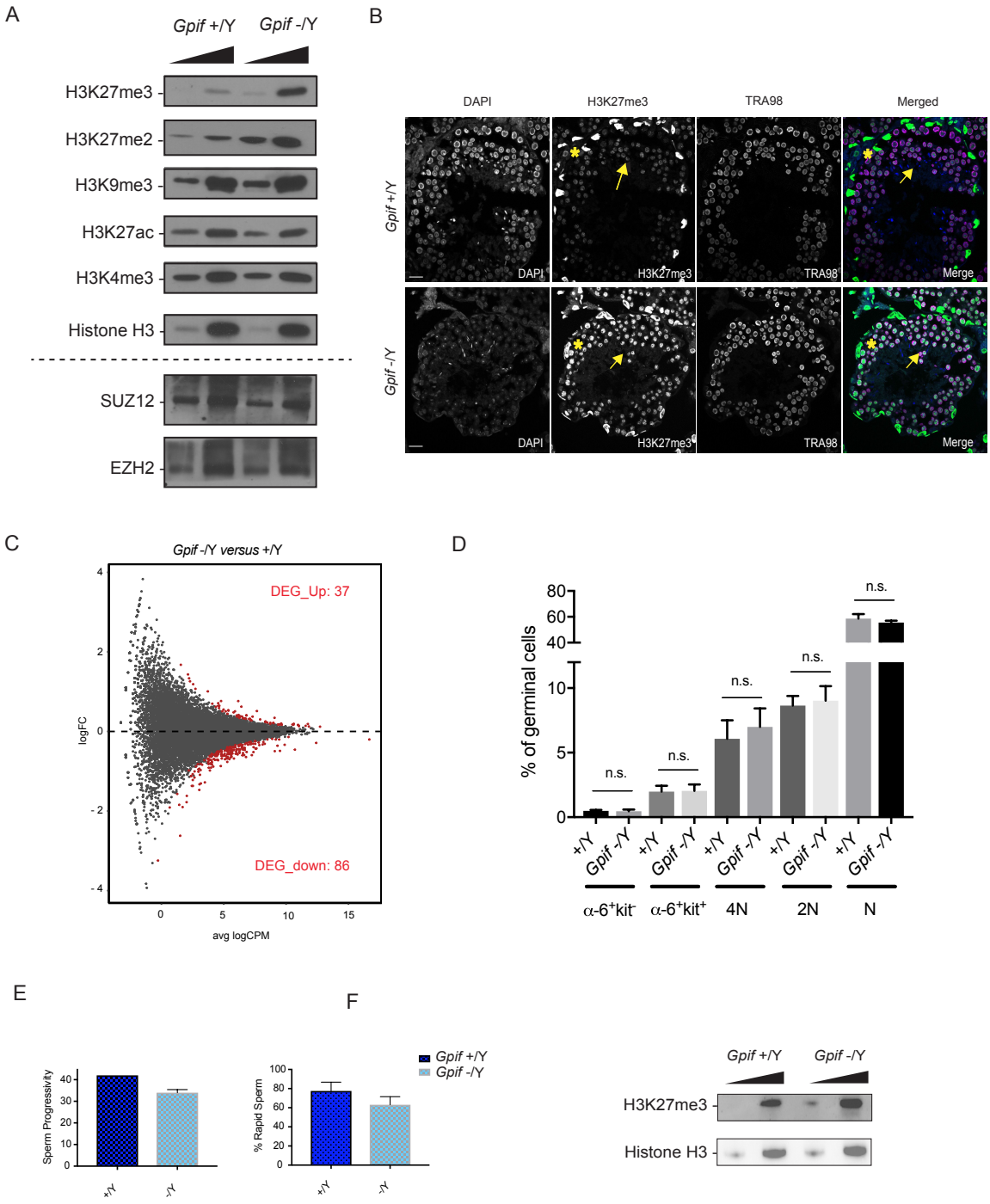


Fig. 4 Ragazzini et al

Figure 5. Mature oocyte *Gpif* -/- displays an altered epigenetic landscape.

(A) *Gpif* and *Ezh2* expression (FPKM) in oocytes isolated at different stages of follicular growth (Data extracted from GEO: GSE70116). (B) Single-cell RNA-seq *Ezh2* and *Gpif* expression data on early embryo developmental phases (oocyte, pronucleus, 2 cells, 4 cell, 8 cells and morula, data from GEO80810). (C) Quantification of H3K27me3 levels in P17 old females primordial and secondary follicles detected by immunofluorescence. Right, H3K27me3 intensities are normalized to DAPI (mean ± s.d., each dot represents a follicle, *n* indicated on the graph). Left, representative image of secondary follicles stained with DAPI, H3K27me3 and merge, *Gpif* +/- versus *Gpif* -/-. (D) Quantification of H3K27me3 levels by immunofluorescence in mouse Surrounded Nuclei (SN) GV oocytes. Top representative picture, bottom quantification (mean ± s.d., *n* indicated on the graph). (E) Quantification of H3K27me3 levels by immunofluorescence in mature mouse MII oocytes. Top representative picture, bottom quantification (mean ± s.d., *n* indicated on the graph). (F) Scatterplot showing log2 fold-change (logFC) expression *versus* average log2 counts per million (logCPM) for MII oocytes obtained from 4 weeks *Gpif* +/- and *Gpif* -/- superovulated females. No additional differentially expressed genes apart from *Gpif* itself.

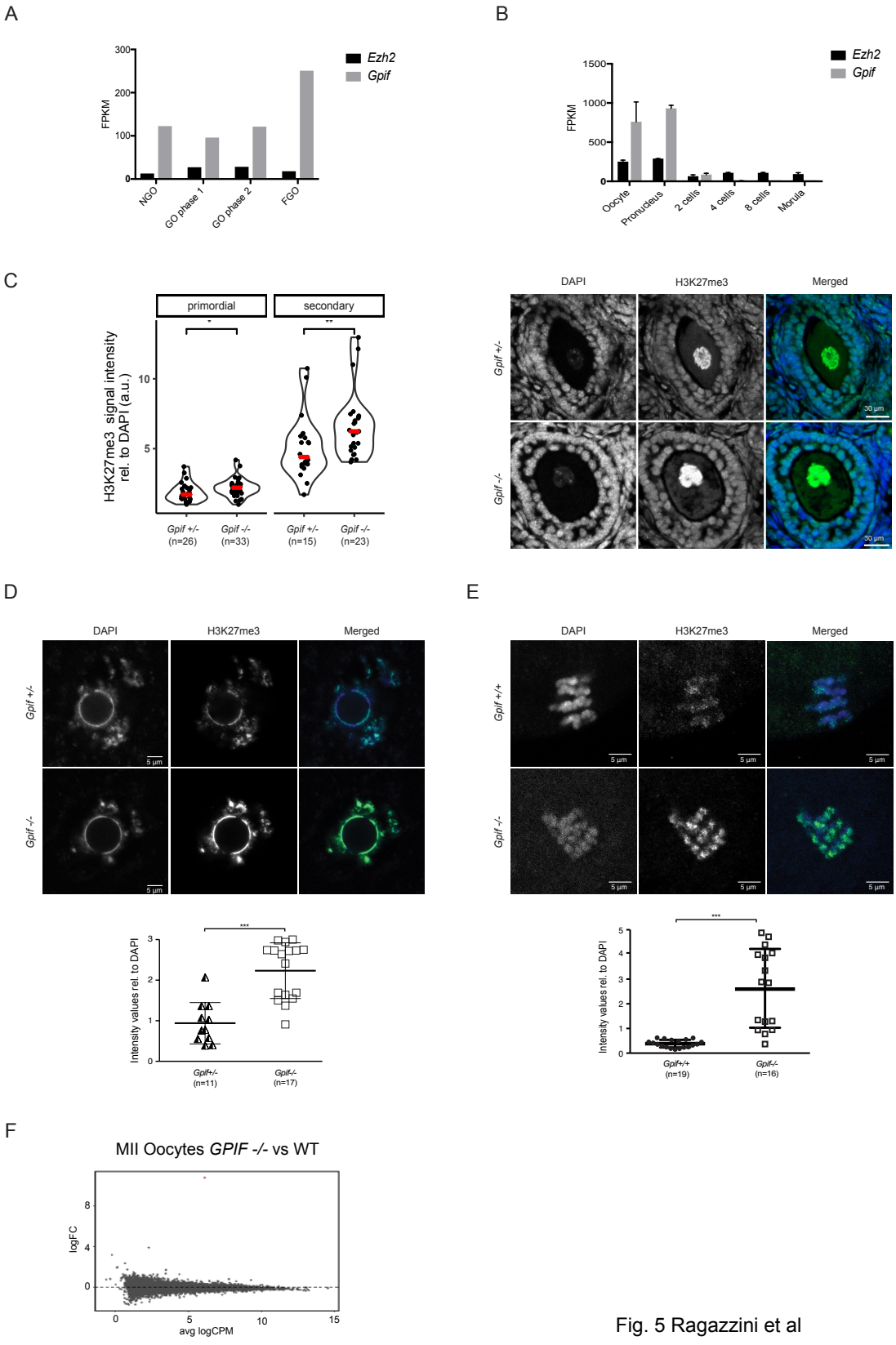


Fig. 5 Ragazzini et al

Figure 6. *Gpif* knockout affects adult female fertility.

(A) Follicle counting on WT versus *Gpif* -/- 2 weeks-old female slides and 16 weeks-old, upper and lower panel respectively (8 slides counted for each genotype, mean ± s.d., n≥ 3, p-value>0,05 if n.s., three stars p-value<0,001). Left panel corresponds to the total number of follicles, then each panel corresponds to a different folliculogenesis step: primordial follicles, primary follicles, secondary/antral follicles. Y axis represent the average follicles number per slide, genotype is indicated in the legend. (B) Average ovaries weight (mg) in adult females (4-5 months old, mean ± s.d., each dot represents an independent female). (C), (D) and (E) Fertility of WT and *Gpif* -/- females monitored during 5 months. (C) Cumulative number of pups *per* genotype depending on time (n=13 cages, breeding scheme represented on top, three stars p value <0,001) (D) Number of litters per female during the 20 weeks monitoring (E) Litter size.

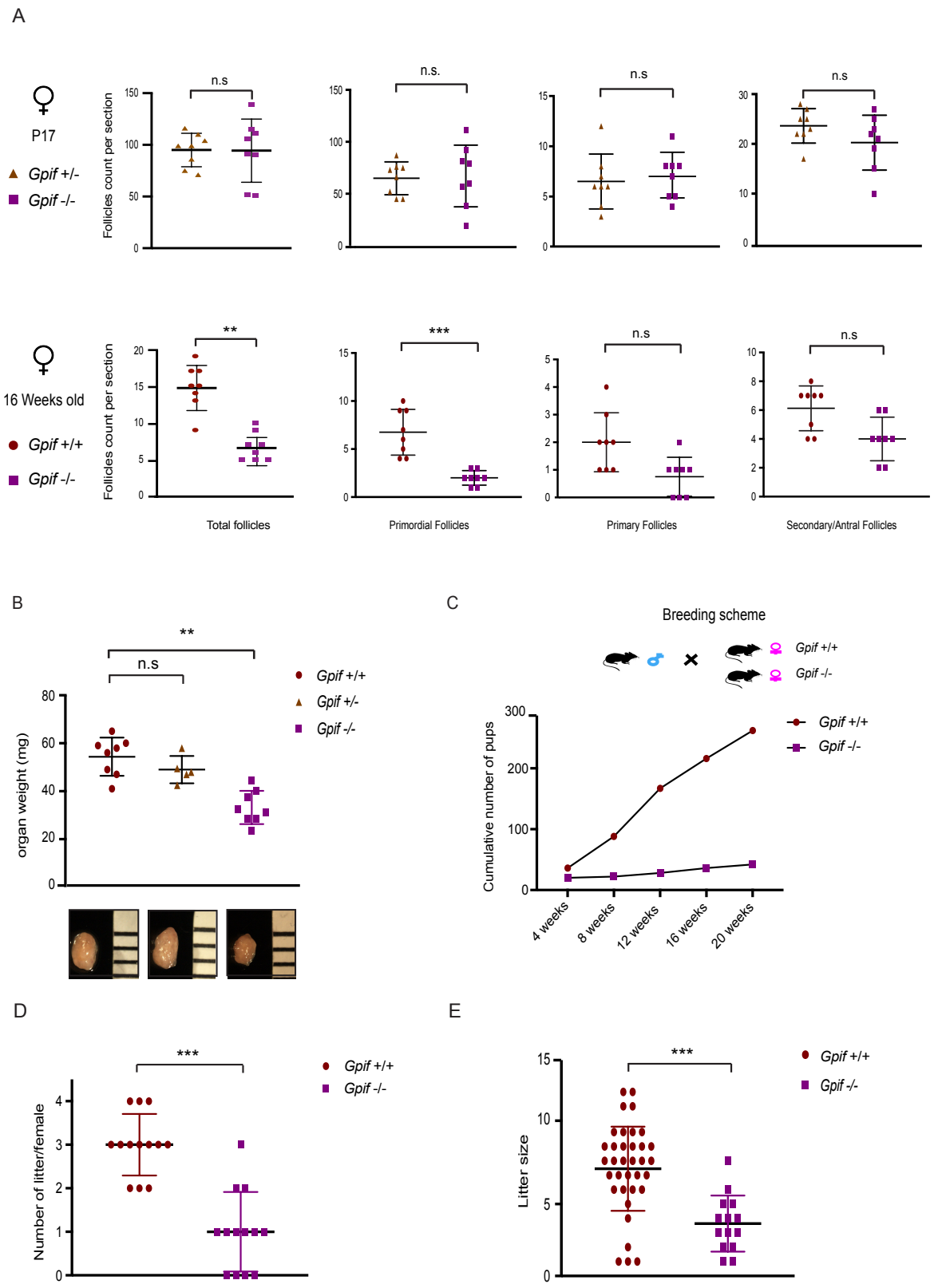


Fig. 6 Ragazzini et al