© 2017. Published by The Company of Biologists Ltd | Development (2017) 144, 2606-2617 doi:10.1242/dev.143347

Setd1b, encoding a histone 3 lysine 4 methyltransferase, is a maternal effect gene required for the oogenic gene expression program

David Brici[1], Qinyu Zhang[1], Susanne Reinhardt[2], Andreas Dahl[2], Hella Hartmann[3], Kerstin Schmidt[4], Neha Goveas[1], Jiahao Huang[5], Lenka Gahurova[5,*], Gavin Kelsey[5], Konstantinos Anastassiadis[4], A. Francis Stewart[1,‡] and Andrea Kranz[1,‡]

ABSTRACT

Germ cell development involves major reprogramming of the epigenome to prime the zygote for totipotency. Histone 3 lysine 4 (H3K4) methylations are universal epigenetic marks mediated in mammals by six H3K4 methyltransferases related to fly Trithorax, including two yeast Set1 orthologs: Setd1a and Setd1b. Whereas Setd1a plays no role in oogenesis, we report that Setd1b deficiency causes female sterility in mice. Oocyte-specific *Gdf9-iCre* conditional knockout (*Setd1b*^{Gdf9}cKO) ovaries develop through all stages; however, follicular loss accumulated with age and unfertilized metaphase II (MII) oocytes exhibited irregularities of the zona pellucida and meiotic spindle. Most *Setd1b*^{Gdf9}cKO zygotes remained in the pronuclear stage and displayed polyspermy in the perivitelline space. Expression profiling of *Setd1b*^{Gdf9} cKO MII oocytes revealed (1) that Setd1b promotes the expression of the major oocyte transcription factors including Obox1, 2, 5, 7, Meis2 and Sall4; and (2) twice as many mRNAs were upregulated than downregulated, suggesting that Setd1b also promotes the expression of negative regulators of oocyte development with multiple Zfp-KRAB factors implicated. Together, these findings indicate that *Setd1b* serves as maternal effect gene through regulation of the oocyte gene expression program.

KEY WORDS: Histone methylation, Oogenesis, Pre-implantation development, Transcriptional regulation, Mouse

INTRODUCTION

Resetting the epigenome during germ cell development establishes the *tabula rasa* essential for early development. In the mouse, primordial germ cells originate from the base of the allantois and migrate through the hindgut into the genital ridge where they proliferate and form the gonads (Allegrucci et al., 2005). By the time primordial germ cells have migrated to the genital ridge, they appear to be largely devoid of DNA methylation (Reik, 2007). In the female gonad, primary oocytes are arrested in prophase of meiosis I at birth (Edson et al., 2009). Subsequently they enlarge, accumulate transcripts needed for early embryonic development and undergo notable changes in chromatin configuration accompanied by epigenetic changes including alterations of lysine methylation on the histone 3 tail and re-establishment of DNA methylation (Stewart et al., 2015; Dahl et al., 2016; Liu et al., 2016; Zhang et al., 2016).

The histone 3 (H3) tail serves a central role in epigenetic regulation primarily through alterations of lysine methylation (Zhang et al., 2015). Trimethylation of lysine 4 (H3K4me3) relates to trithorax-Group (trx-G) action (Roguev et al., 2001) and is a hallmark of active promoters (Santos-Rosa et al., 2002; Bannister and Kouzarides, 2011). H3K4me3 is opposed by repression through H3K27me3 mediated by Polycomb-Group (Pc-G) action (Voigt et al., 2012) and H3K9me2/3, which defines heterochromatin (Peters et al., 2001), whereas H3K36me3 is mainly found in transcribed regions due to the association of the major H3K36 methyltransferase with the elongating RNA Polymerase II through phosphorylation of serine 2 of its CTD tail (Li et al., 2003; Schaft et al., 2003; Dambacher et al., 2010). Recent data on growing oocytes indicates that local deposition of H3K36me3 and removal of H3K4me2/me3 sets the stage for *de novo* DNA methylation (Stewart et al., 2015). Development of the preovulatory follicle is finalized by global transcriptional silencing in the oocyte, which requires the H3K4 demethylases Kdm1a (Lsd1) and Kdm1b (Lsd2) (Stewart et al., 2015) and, paradoxically, also the H3K4 methyltransferase Mll2 (Kmt2b, Wbp7) (Andreu-Vieyra et al., 2010).

Shortly after fertilization, the male and female pronuclei have comparable levels of H3K4me1; however, H3K4me2/me3 is initially lacking in the male pronucleus (Lepikhov and Walter, 2004; van der Heijden et al., 2005). The acquisition of H3K4me2/me3 in the male pronucleus, pronuclear fusion and cell division precede zygotic genome activation (Aoshima et al., 2015), although in the mouse minor zygotic transcription activity has been detected at the pronuclear stage (Abe et al., 2015).

To elucidate the roles played by epigenetic mechanisms in oocyte and early embryo development, we are focusing on the six mammalian Set1/Trithorax-type H3K4 methyltransferases (Glaser et al., 2006). These six come in three pairs: Setd1a and Setd1b, which are orthologous to the sole yeast H3K4 methyltransferase Set1 (Roguev et al., 2001); Mll1 (Kmt2a) and Mll2, which are homologous to *Drosophila* Trithorax (Francis and Kingston, 2001), and Mll3 (Kmt2c) and Mll4 (Kmt2d), which are homologous to a fusion of *Drosophila Cara Mitad* (*Lpt*) and *trithorax-related* (Chauhan et al., 2012; Herz et al., 2012). Each of the six appears

---

[1]Genomics, Biotechnology Center, Technische Universität Dresden, BioInnovationsZentrum, Tatzberg 47, Dresden 01307, Germany. [2]Deep Sequencing Group SFB 655, Biotechnology Center, Technische Universität Dresden, 01307 Dresden, Germany. [3]Light Microscopy Core Facility of the Biotechnology Center/Center for Regenerative Therapies Dresden, Technische Universität Dresden, 01307 Dresden, Germany. [4]Stem Cell Engineering, Biotechnology Center, Technische Universität Dresden, BioInnovationsZentrum, Tatzberg 47, Dresden 01307, Germany. [5]Epigenetics Programme, The Babraham Institute, Cambridge, CB22 3AT, UK.
*Present address: Laboratory of Developmental Biology & Genetics, Department of Molecular Biology, University of South Bohemia, 37005 Ceske Budejovice, Czech Republic.

‡Authors for correspondence (andrea.kranz@biotec.tu-dresden.de; stewart@biotec.tu-dresden.de)

L.G., 0000-0002-9412-7971; A.K., 0000-0002-7481-0220

Received 5 August 2016; Accepted 2 June 2017

to reside in its own protein complex. Concordantly, all six are essential for mouse development for different reasons even though all six appear to be expressed ubiquitously (Ernst et al., 2004; Glaser et al., 2006; Lee et al., 2013; Bledau et al., 2014) (A.K., Q.Z., K.A. and A.F.S., unpublished). However, all six protein complexes share the same scaffold as the yeast Set1 complex, Set1C, based on four proteins, Wdr5, Rbbp5, Ash2l and Dpy30 (Wysocka et al., 2005; Dehe et al., 2006; Dou et al., 2006; Lee et al., 2007; Lee and Skalnik, 2012; Gu and Lee, 2013), termed WRAD (Ernst and Vakoc, 2012) or, less precisely, COMPASS (Shilatifard, 2012).

Recently, the Set1 H3K4 methyltransferases in *Drosophila* and *Caenorhabditis elegans* (named dSet1 and SET-2, respectively) have been shown to be essential for germline stem cells (GSCs) (Xiao et al., 2011; Robert et al., 2014; Yan et al., 2014). In mouse germ cell development, conditional mutagenesis identified Mll2 as essential for both oogenesis and spermatogenesis (Glaser et al., 2009; Andreu-Vieyra et al., 2010). Mll2 is the major di- and tri - but not mono-methyltransferase in developing oocytes and continues to be required during the first cell divisions after fertilization. However, by the blastocyst stage, Mll2 is neither the major H3K4 methyltransferase nor required (Andreu-Vieyra et al., 2010). These roles are acquired by Setd1a; however, Setd1a is not required for oogenesis or early mouse development (Bledau et al., 2014). Here, we report that Setd1b is required during oogenesis for the oocyte-to-embryo transition.

RESULTS

Setd1b is required for oogenesis

As previously reported, *Setd1b*$^{-/-}$ embryos are growth retarded from embryonic day (E) 7.5 and die before E11.5. This result was obtained with both *Setd1b*$^{D/D}$ and *Setd1b*$^{FDC/FDC}$ genotypes (Fig. 1A; Bledau et al., 2014) indicating that both the targeted allele (*D*), which serves as a gene trap, and the Cre recombined allele (*FDC*), which lacks the frameshifting exon 5, are complete knockouts. During pre-implantation development, Setd1b is expressed in all stages from the oocyte to the blastocyst (Bledau et al., 2014). The first indication that Setd1b is required for oogenesis came from conditional knockout (cKO) of *Setd1b* in 2-month-old adults. Using *Rosa26-Cre-ERT2* (*RC*; Seibler et al., 2003), tamoxifen induction of *Setd1b*$^{FD/FD;RC/+}$ females to generate *Setd1b*$^{FDC/FDC;RC/+}$ provoked infertility. After tamoxifen induction, *Setd1b*$^{FDC/+;RC/+}$ and *Setd1b*$^{FDC/FDC;RC/+}$ females were bred to wild-type males. On average, *Setd1b*$^{FDC/+;RC/+}$ females gave birth to 5.3 pups whereas *Setd1b*$^{FDC/FDC;RC/+}$ females never gave rise to offspring. To pursue this indication, we deleted *Setd1b* in oocytes using *Gdf9-iCre* (Fig. 1B; henceforth the Cre-deleted allele is called *Setd1b*^{Gdf9} cKO) resulting in complete loss of Setd1b expression in oocytes as evaluated by qRT-PCR analysis (Fig. 1C), concordant with complete Cre recombination of genomic DNA (Fig. S1).

Breeding to males of verified fertility established that *Setd1b*^{Gdf9} cKO females were sterile (Fig. 1D). The frequency of vaginal plugs and uterine wet weight were comparable to those of control females indicating that sex hormone levels and estrous cycle periodicity were normal in *Setd1b*^{Gdf9} cKO females. Consequently, *Setd1b*^{Gdf9} cKO ovaries were examined at 4, 8, 23 and 40 weeks.

At all analyzed time points, ovaries from *Setd1b*^{Gdf9} cKO females were smaller compared with controls (Fig. 1E,F). At 4 weeks, *Setd1b*^{Gdf9} cKO ovaries appeared otherwise normal and all expected follicular stages including small antral follicles were visible. At 8 weeks of age, the ovaries of control and *Setd1b*^{Gdf9} cKO mice contained corpora lutea, which is evidence of cycling. At 23 weeks, *Setd1b*^{Gdf9} cKO ovaries showed a reduction in follicle count in comparison with that of control mice whereas the number of atretic follicles and oocyte remnants was increased (Fig. 1G). These defects were even more apparent at 40 weeks.

Immunohistochemistry with major oocyte markers was carried out at selected stages; however, no striking differences between control and mutant were observed. Staining with Ki67 (Mki67) antibody to evaluate proliferative activity in the granulosa cells surrounding the oocyte was comparable in the *Setd1b*^{Gdf9} cKO and control ovaries (Fig. 1H). Immunohistochemistry with antibodies against the prominent oocyte markers Zp2, ERα (Esr1), Sox9 and Ddx4 showed that oogenesis in *Setd1b*^{Gdf9} cKO ovaries was apparently normal (Fig. 1H, Fig. S2A).

Defects accumulate during the oocyte-to-embryo transition

Gonadotropin-induced superovulation of pre-pubertal females yielded an average of 37.4 MII oocytes per control female, whereas *Setd1b*^{Gdf9} cKO females yielded an average of 25.2 MII oocytes (Fig. 2A). Notably *Setd1b*^{Gdf9} cKO periovulatory oocytes displayed the distinctive chromatin rearrangement from non-surrounding (NSN) to surrounding (SN) nucleolus configurations, which is the classical marker for oocyte maturation prior to quiescence and occurs concurrently with global transcriptional silencing (Fig. 2B, Fig. S2B).

In the final stages of oocyte maturation, the spindle assembles after nuclear envelope breakdown and segregates half of the chromosomes into the first polar body (Clift and Schuh, 2013). Then a second spindle assembles, relocates to the oocyte surface and the oocyte arrests in metaphase II awaiting follicular stimulation. As visualized by α-tubulin immunostaining, the meiotic spindles of *Setd1b*^{Gdf9} cKO MII oocytes were considerably smaller than control spindles (Fig. 2C,D). Nevertheless, chromosome segregation (Fig. 2C) and metaphase II chromosome spreads (Fig. S2C) were normal indicating that all chromatids had been correctly attached to spindle fibers.

Setd1b^{Gdf9} cKO oocytes and zygotes displayed increased cytoplasmic granularity (Fig. 2B,F) suggesting perturbed cytoplasmic organelles. Evaluation of the Golgi apparatus by immunostaining of GM130 (Golga2), a Golgi matrix protein associated with the *cis*-compartment (Wang et al., 2008), showed that Golgi structures were homogenously distributed and did not colocalize with the darkened cytoplasmic granules of the *Setd1b*^{Gdf9} cKO MII oocytes (Fig. S2D). Furthermore, Tomm20 immunostaining revealed no disorganized mitochondrial distribution (Fig. 2E). Staining zygotes with Nile Red revealed that the granularity was caused by lipid droplets (Fig. 2F).

Single-cell transcriptome profiling identifies disrupted gene expression

The impact of Setd1b removal on oocyte gene expression was evaluated by single-cell RNA profiling of four MII oocytes each from control and *Setd1b*^{Gdf9} cKO superovulated mice (Figs 3 and 4). Using 75-base-pair reads, 16-23 million reads per sample with comparable mappability (80.7-86.4%) and 65.8-73.7% unique reads were obtained (Fig. 3A). All biological replicates clustered into two well-separated groups according to genotype thereby lending confidence regarding data quality (Fig. 3B). Our single-cell RNA-seq dataset is also in good agreement with a previously published dataset made from large numbers of MII oocytes (Park et al., 2013; data available at DNA Data Bank of Japan under accession number DRA001066) with a coverage of almost 80% of the transcripts found in their study (Fig. 3C). As expected, *Setd1b* itself was downregulated as was *Rbm15*, which has been reported to be a

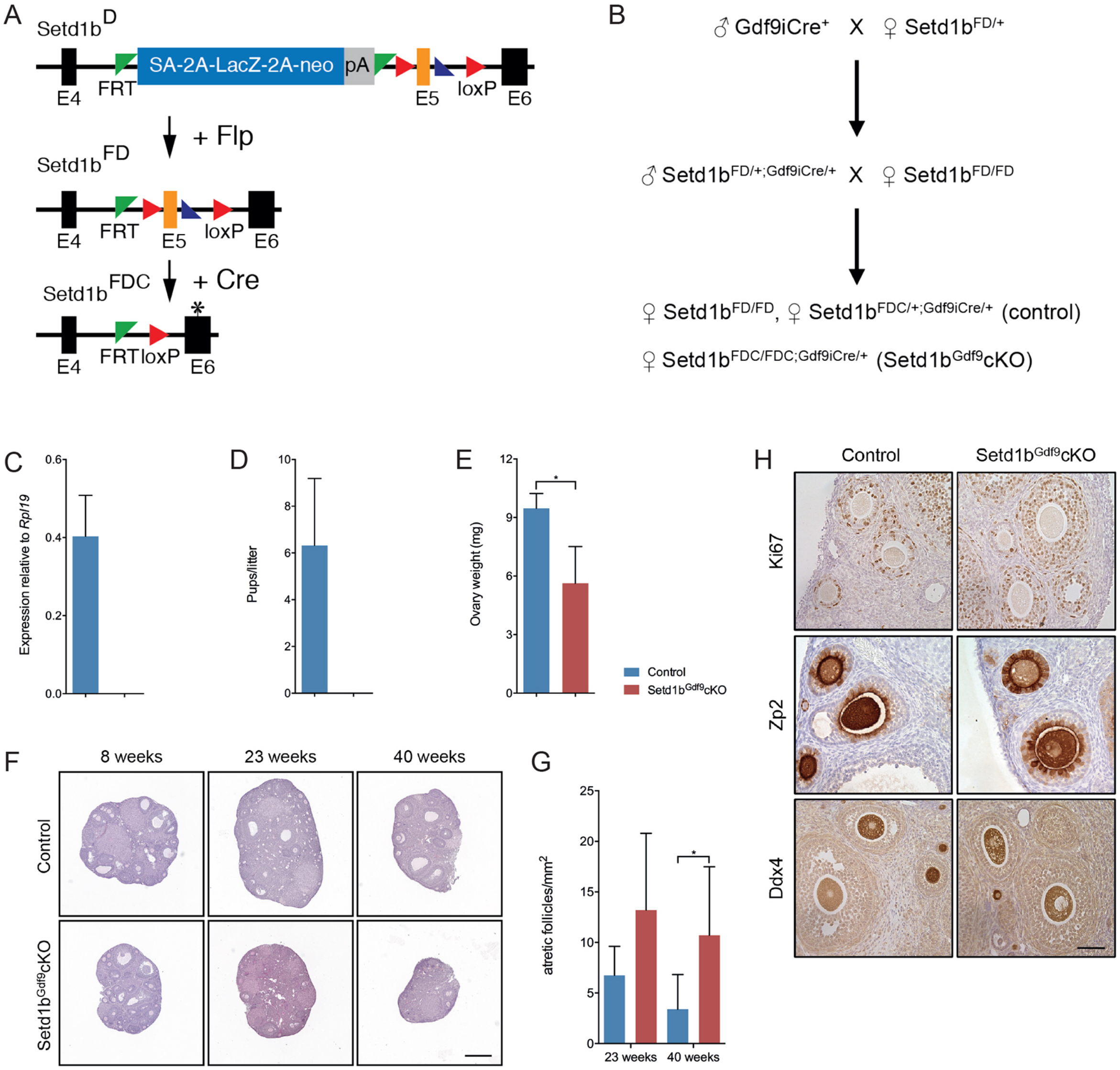

Fig. 1. Setd1b oocyte-specific deletion mediated by the *Gdf9-iCre* allele leads to female infertility. (A) Diagram of *Setd1b* alleles. The targeted allele (*D*) includes a *lacZ*-neo gene trap cassette inserted between exon 4 and 5 (E4, E5) that was used for selection. The *Setd1b* transcript, which is expressed in embryonic stem cells, is captured by a splice acceptor (SA) and terminated by a polyadenylation signal (pA). The cassette is surrounded by FRT sites (green triangles) and exon 5 is surrounded by loxP sites (red triangles). FLP recombination deletes the cassette to generate the *FD* allele, which is wild type. Cre deletion of exon 5 leads to the *FDC* allele and a frameshift in the mRNA when exon 4 is spliced to exon 6. Blue triangle represents the single rox site left after Dre recombination. (B) Breeding schematic to illustrate the control and *Setd1b*^{Gdf9} cKO genotypes. (C) qRT-PCR was used to determine the expression of *Setd1b* in control and *Setd1b*^{Gdf9} cKO MII oocytes (n=90). Relative quantities of mRNA were normalized to *Rpl19*. (D) *Setd1b*^{Gdf9} cKO females are infertile; ♂ C57/BL6×♀ *Setd1b*$^{FD/FD;Gdf9iCre/+}$ (seven breeding pairs never produced offspring), ♂ C57/BL6×♀ *Setd1b*$^{FD/+;Gdf9iCre/+}$ (seven breeding pairs gave offspring in 41 litters). (E) The average weight of ovaries from 40-week-old females. Ovaries from four females of each genotype were analyzed. Mean+s.d. is shown. Statistical significance was tested using an unpaired t-test with Welch's correction: *P<0.05. (F) Ovarian morphology and follicular development. Mice of 8, 23 and 40 weeks of age were sacrificed and ovaries fixed, paraffin embedded and further processed for PAS histochemistry. (G) Counts of atretic follicles from 23- and 40-week-old mice. PAS-stained ovarian sections from 23-week-old *Setd1b*^{Gdf9} cKO (n=4), control (n=4) and 40-week-old *Setd1b*^{Gdf9} cKO (n=5), control (n=7) females were used. Mean +s.d. is shown. Statistical significance was tested using unpaired t-tests with Welch's correction: *P<0.05. (H) Immunohistochemistry for Ki67, Zp2 and Ddx4 using 23-week-old ovaries. Scale bars: 500 µm (F); 50 µm (H).

subunit of the Setd1b complex (Setd1bC; Lee and Skalnik, 2012). Similarly, another potential Set1C subunit, *Bod1l* (van Nuland et al., 2013), was downregulated but all the other known Setd1bC subunits were unaffected.

Unexpectedly for an H3K4 methyltransferase, we observed that more mRNAs were upregulated than downregulated. This observation was sustained in several ways. When we imposed a threshold of ten read counts per gene, about 10% more unique mRNAs (∼1100) were detected in each *Setd1b*^{Gdf9} cKO MII oocyte than in controls (Fig. 3D). To analyze differentially expressed genes, we applied DESeq2 analysis to the RNA-seq dataset. Based on twofold changes in mRNA reads per kilobase million (RPKM), 1231 were upregulated, 650 were downregulated (and 8650 remained unchanged) in *Setd1b*^{Gdf9} cKO oocytes (Fig. 3E). We also compared these differentially expressed mRNAs with a new transcriptome series generated from oocytes at different stages of oogenesis (Gahurova et al., 2017; data available at Gene Expression Omnibus under accession number GSE86297). This revealed that the 650 downregulated genes after deletion of Setd1b tend to be upregulated during normal oocyte development (Fig. 3F,G).

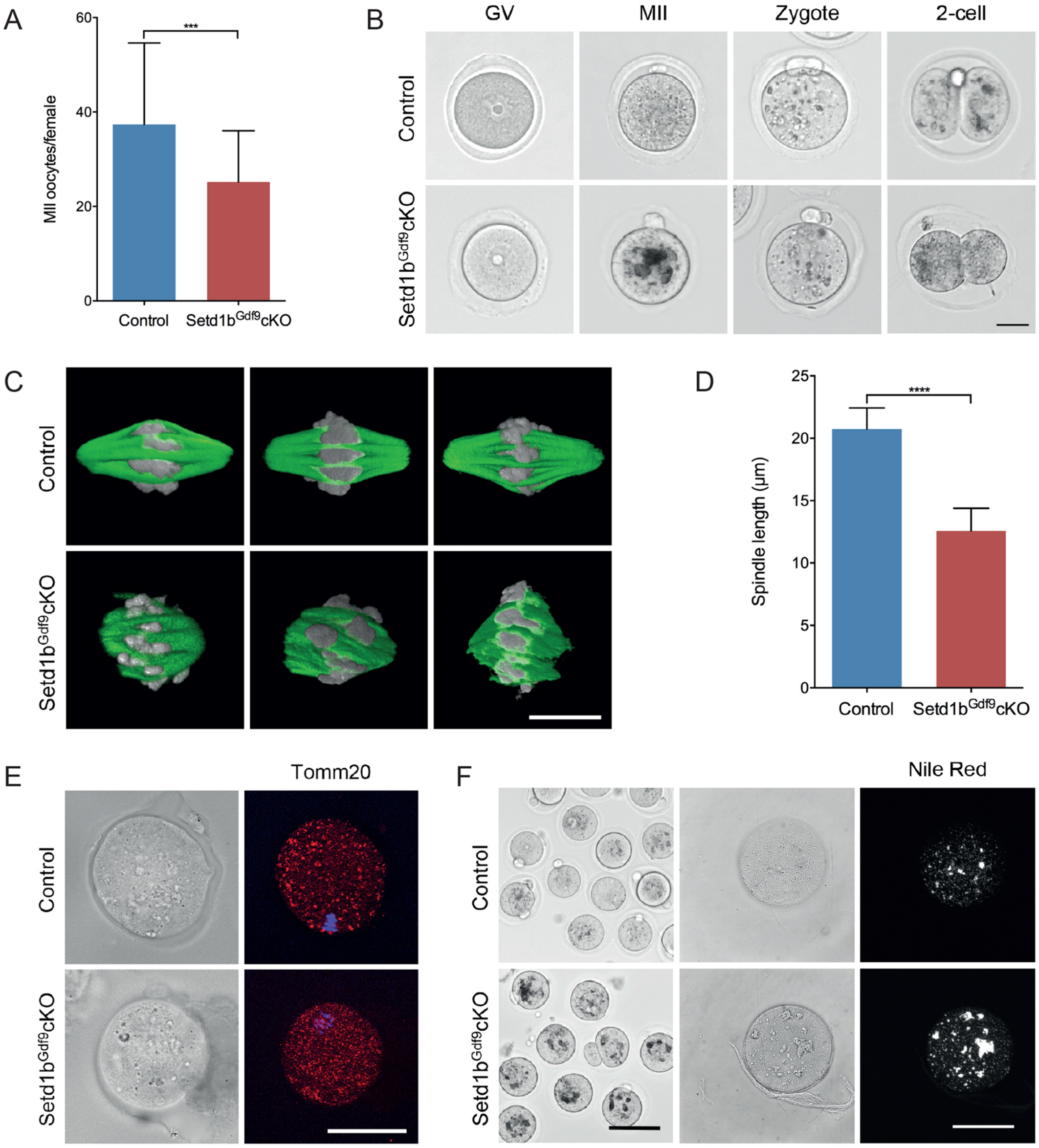

Fig. 2. Morphology of oocytes and early embryos. (A) *Setd1b*^{Gdf9} cKO ovulation rates were significantly reduced in 24- to 28-day-old superovulated *Setd1b*^{Gdf9} cKO females. Mean+s.d. is shown, *n*=49 (control females, blue bar), *n*=36 (*Setd1b*^{Gdf9}cKO females, red bar). Statistical significance was tested using an unpaired *t*-test with Welch's correction: ****P*<0.001. (B) Representative differential interference contrast (DIC) images of periovulatory (GV) oocytes, metaphase II (MII) oocytes, zygotes and two-cell-stage embryos. (C) Spindle defect in *Setd1b*^{Gdf9} cKO oocytes. Representative confocal laser micrographs of the meiotic metaphase II spindle after 3D reconstruction. Green, spindle microtubules; gray, chromosomes. Four independent experiments; total 157 (control), 125 (*Setd1b*^{Gdf9} cKO) oocytes. (D) Meiotic spindle length measurements from 23 control and 14 *Setd1b*^{Gdf9} cKO MII oocytes. Mean+s.d. is shown. Statistical significance was tested using an unpaired *t*-test with Welch's correction: *****P*<0.0001. (E) Tomm20 staining (right) revealed no disorganized distribution of mitochondria in MII oocytes; total 37 (control) and 29 (*Setd1b*^{Gdf9} cKO) MII oocytes. DIC images shown on the left. (F) Lipid droplet accumulation in *Setd1b*^{Gdf9} cKO zygotes. Representative DIC (left and middle) and fluorescence (right) analysis of zygotes at 26 h post-hCG injection (PN4-5) from control and *Setd1b*^{Gdf9} cKO stained with Nile Red; total 56 (control) and 32 (*Setd1b*^{Gdf9} cKO) zygotes were evaluated in two independent experiments. Scale bars: 33 μm (B); 10 μm (C); 50 μm (E; F, middle and right); 100 μm (F, left).

Specifically, 375 of the 650 are normally upregulated, whereas only 43 of this group are downregulated and the remaining 214 show no change during oogenesis. This correlation accords with expectations for action by an H3K4 methyltransferase in transcriptional activity. Conversely, the 1231 upregulated mRNAs after deletion of Setd1b slightly correlated with downregulated mRNAs during normal oogenesis (357 normally downregulated, 243 upregulated and 574 unchanged) (Fig. 3F,G, Fig. S3, Table S3).

Remarkably, the most significant terms associated with downregulated mRNAs after Setd1b removal are 'Regulation of transcription', 'Sexual reproduction' and 'Cell cycle processes' (Fig. 4A). In contrast, we found an over-representation of terms related to 'Redox' and 'Ribosome biogenesis' for the significantly upregulated genes (Fig. 4B).

Notably, transcripts for the oocyte-specific homeobox genes (*Obox1*, *Obox2*, *Obox5* and *Obox7*) were downregulated (Fig. 4C, G, Fig. S4A, Table S3). Concordantly, oocyte-specific genes such as *Omt2a* and *Omt2b* (oocyte maturation 2), *Oosp1* (oocyte secreted protein 1) and *Oog3* (oogenesin 3), were also downregulated (Fig. 4C,G). These genes are likely to be directly regulated by the

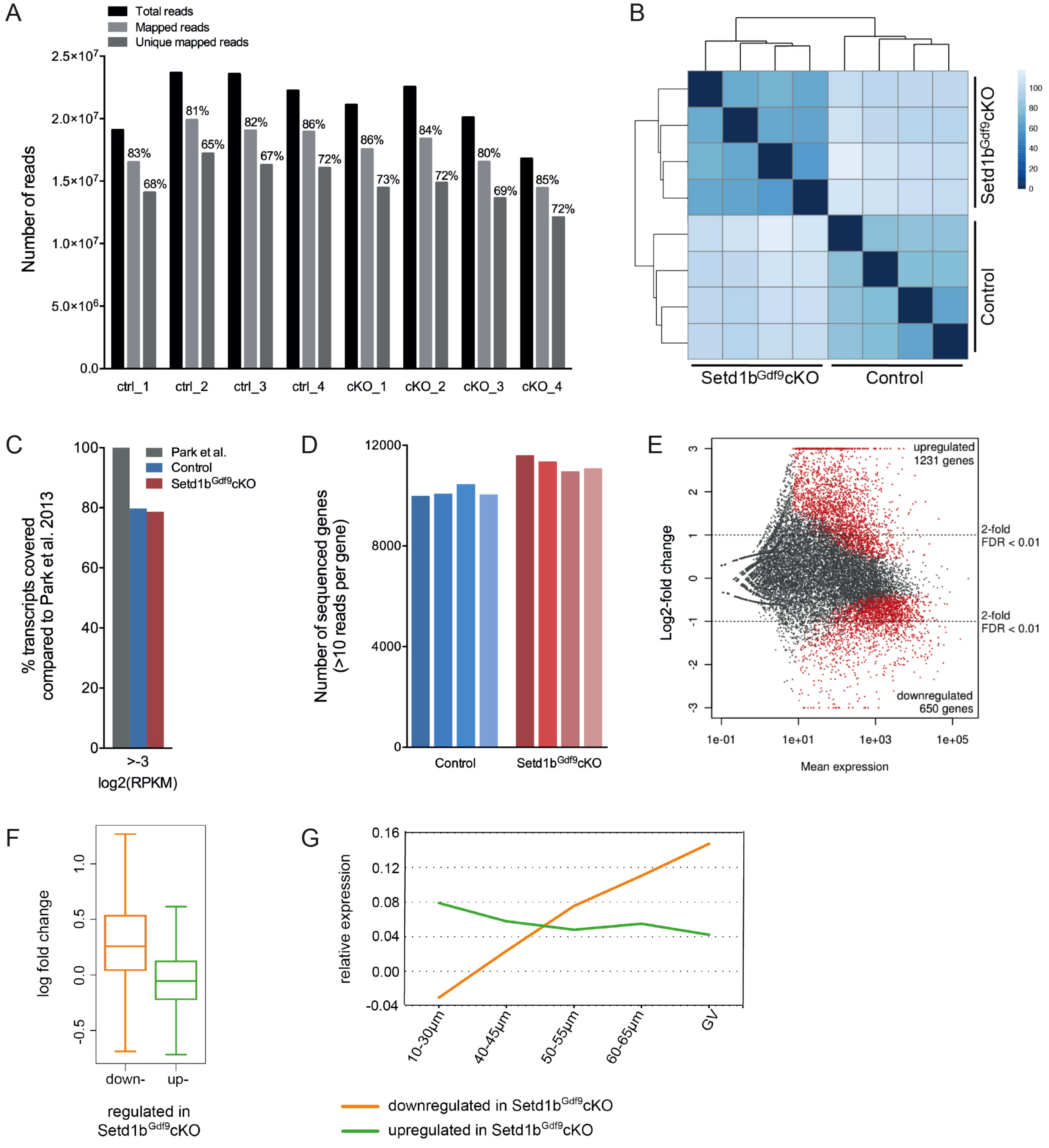

Fig. 3. Single-cell RNA profiling of *Setd1b^{Gd9}*cKO and control oocytes. (A) Mappability of reads per oocyte. (B) Heat map of the sample-to-sample Euclidean distance with dendrogram. (C) Comparison with the whole-transcriptome analysis of Park et al. (2013). With a threshold of >–3 log2(RPKM) almost 80% of the transcripts identified by Park et al. were found in our study. (D) Number of detected genes per oocyte. (E) MA plot visualizing the log2-fold change differences according to expression level. Red dots represent significant DEGs at a 1% false discovery rate (FDR). (F) Change in expression level of genes downregulated and upregulated in *Setd1b^{Gdf9}* cKO during oocyte growth in wild-type oocytes. The *y*-axis shows change between oocytes at the beginning of their growth (10-30 μm) and fully grown oocytes (GV stage). (G) Relative expression levels (mean-centered log transformed FPKM values) of genes down- and upregulated in *Setd1b^{Gdf9}* cKO during oocyte growth in wild-type oocytes.

Obox and other oocyte-specific transcription factors. Other notable transcription factors amongst the 90 that were downregulated more than twofold (Table S4) include *Meis2* and *Sall4*. Additionally prominent were *Klf17* and fifteen more Zfps (zinc finger proteins) with KRAB (Krüppel-associated boxes), which likely connect Setd1b regulation with the repression of transcription (Fig. 4D,H).

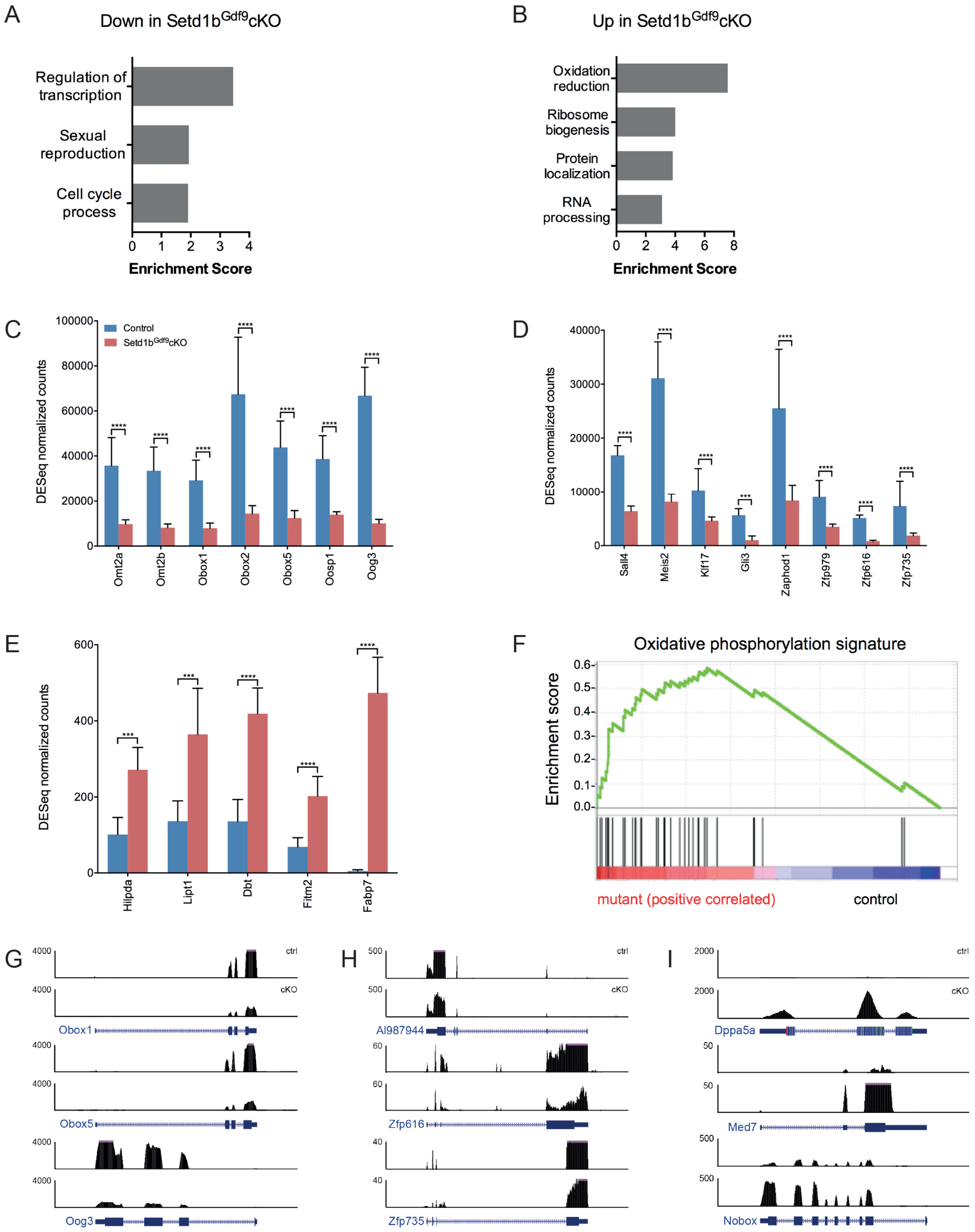

Fig. 4. Differential gene expression in *Setd1b^{Gd9}* cKO and control oocytes. (A,B) Biological processes enriched in genes downregulated (A) or upregulated (B) in *Setd1b^{Gdf9}*cKO oocytes compared with control oocytes using the DAVID GO/BP database. (C-E) DESeq-normalized counts for selected genes expressed in control and *Setd1b^{Gdf9}* cKO oocytes. Mean+s.d. is shown; *n*=4; ****P*<0.001; *****P*<0.0001. (F) Gene set enrichment analysis (GSEA) of control versus *Setd1b^{Gdf9}* cKO for the oxidative phosphorylation gene signature. (G-I) RNA-seq signals (reads per million) of different gene loci for control (top line) and *Setd1b^{Gdf9}* cKO (bottom line) with gene diagrams below (exons, thick lines). For AI987944 we suggest the name *Zaphod1*.

In contrast, expression of oogenesis regulators such as *Dppa5a* and *Nobox* was sustained (Fig. 4I, Fig. S4B). Concordant with increased Nile Red staining (Fig. 2F), transcripts related to storage and binding of lipids were upregulated (Fig. 4E). Mature oocytes are characterized not only by transcriptional and translational silencing but also by respiratory quiescence (Sieber et al., 2016). Strikingly, transcripts of genes belonging to all five of the respiratory chain complexes were significantly upregulated (Fig. 4F, Fig. S4C). To validate the RNA-seq results, we performed quantitative RT-PCR on four selected genes (*Sdhc*, *Sdhd*, *Cox7b*, *Ndufs2*) and found that all of them were in good agreement (Fig. S4D). Transcripts encoding genes involved in ribosome biogenesis (Fig. S4E), purine and pyrimidine metabolism (Fig. S4F; data not shown) were also more prevalent in *Setd1b*^{Gdf9} cKO oocytes. These transcriptional changes suggest a hyperactive metabolic state in *Setd1b*^{Gdf9} cKO oocytes.

Defective zona pellicuda and polyspermy

To examine fertility defects of *Setd1b*^{Gdf9} cKO further, females were superovulated and mated with wild-type males. After verification of vaginal plugs, oocytes and zygotes were harvested for overnight culture. Most (69%) of the oocytes were fertilized and this proportion was comparable between control and *Setd1b*^{Gdf9} cKO except for a few apoptotic mutant eggs (Fig. 5A).

Irregularities in the zona pellucida were consistently observed with *Setd1b*^{Gdf9} cKO periovulatory oocytes and the irregularities persisted during the oocyte-to-embryo transition (Fig. 2B). Immunostaining of two-cell stage embryos with a Zp2 antibody revealed a disordered distribution in the zona matrix, which likely contributes to the irregularities and ruffling (Fig. 5B). Furthermore, we often observed multiple sperms (in more than half of the cases) within the *Setd1b*^{Gdf9} cKO perivitelline space. Concordant with the polyspermy, polyploid embryos constituted approximately 10% of all fertilization events in *Setd1b*^{Gdf9} cKO oocytes (Fig. 5C,D). Paradoxically, transcript levels for the major zona pellucida proteins were strongly elevated (Fig. 5E).

Most *Setd1b*^{Gdf9}cKO embryos fail before the two-cell stage

Almost all control zygotes advanced into the two-cell stage embryo compared with 26% for *Setd1b*^{Gdf9} cKO zygotes (Fig. 6A). The two-cell *Setd1b*^{Gdf9} cKO embryos displayed cellular and genetic asymmetries resulting in blastomeres and nuclei of different sizes (Fig. 2B, Fig. 5B, Fig. 6B) and none progressed to the four-cell stage.

Of the zygotes that did not engage the first cleavage, most (87%) were arrested in the pronuclear stages (PN2-5) (Fig. 6C). A further 13% arrested shortly after nuclear envelope breakdown and pronuclear fusion. Therefore, the majority arrested between the beginning of DNA replication and nuclear envelope breakdown. Together, these data indicate Setd1b is essential for progression through the first cell cycle and the predominant one-cell-stage arrest indicates that the requirement for Setd1b operates before zygotic genome activation.

Normal H3K4 methylation in *Setd1b*^{Gdf9} cKO oocytes and zygotes

No differences in global levels of H3K4me1, me2 or me3 were observed between control and *Setd1b*^{Gdf9} cKO oocytes and zygotes (Fig. S5A-D). In *Setd1b*^{Gdf9} cKO zygotes, both male and female pronuclei were established with apparently normal decondensed chromatin and extrusion of the second polar body (Fig. 2B), followed by successful migration from the cell periphery towards each other (Fig. S5B-E). Immunofluorescence showed comparable levels of H3K4me1 in female and male pronuclei of PN4 control and *Setd1b*^{Gdf9} cKO zygotes. Also in both control and *Setd1b*^{Gdf9} cKO zygotes, H3K4me2 and H3K4me3 immunofluorescence displayed the expected asymmetry between female and male

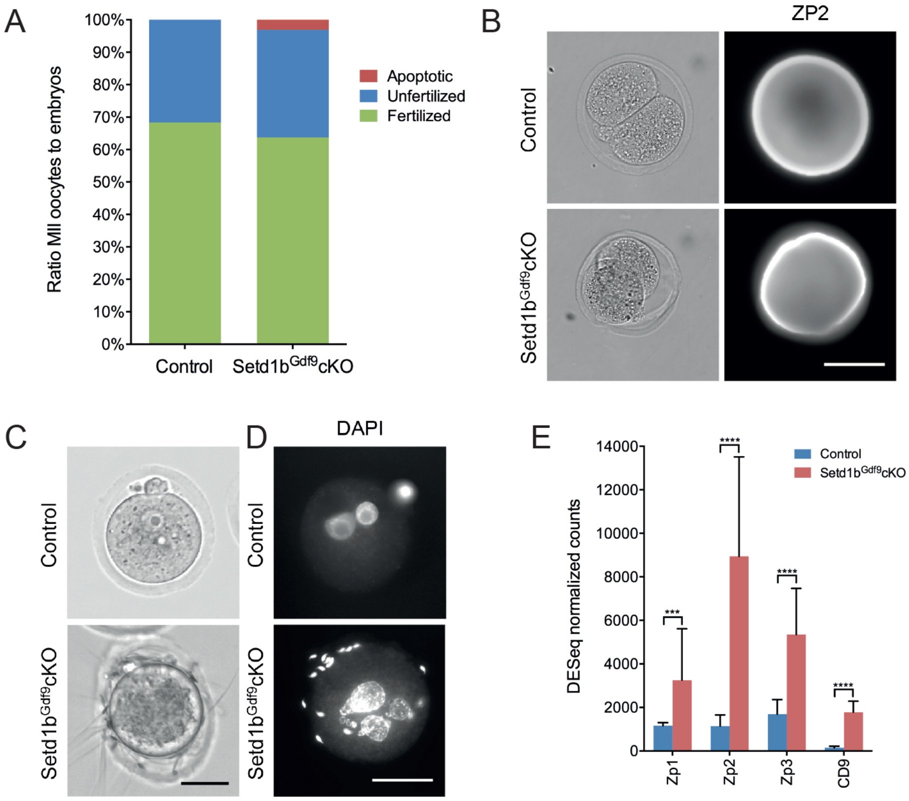

Fig. 5. Defects in *Setd1b*^{Gdf9} cKO zona pellucida. (A) Percentages of MII oocytes compared with embryos collected from plugged female mice, determined at 44 h post-hCG injection. Note that 3% of the *Setd1b*^{Gdf9} cKO oocytes are apoptotic; *n*=139 (control), *n*=193 (*Setd1b*^{Gdf9} cKO). (B) DIC (left) and fluorescence microscopy (right) analysis of two-cell-stage embryos immunostained with Zp2 illustrating the defective zona pellucida in *Setd1b*^{Gdf9} cKO embryos; total 34 (control) and 35 (*Setd1b*^{Gdf9} cKO) oocytes/embryos were evaluated in two independent experiments. (C) DIC microscopy of superfluous sperm in the perivitelline space of *Setd1b*^{Gdf9} cKO zygotes. (D) Fluorescence microscopy of a *Setd1b*^{Gdf9} cKO zygote displaying three pronuclei in addition to superfluous sperm in the perivitelline space of *Setd1b*^{Gdf9} cKO zygotes. (E) DESeq-normalized counts for genes expressed in the zona pellucida in control and *Setd1b*^{Gdf9} cKO oocytes. Mean+s.d. is shown; *n*=4; ****P*<0.001; *****P*<0.0001. Scale bars: 50 µm (B,D); 33 µm (C).

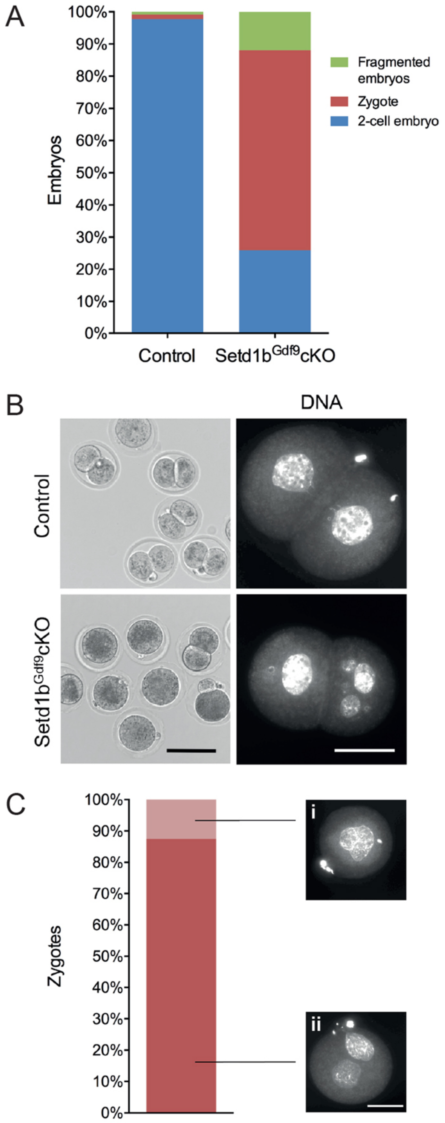

Fig. 6. Maternal Setd1b is essential for the first cleavage. (A) Percentages of two-cell embryos, zygotes (PN2-5) and fragmented embryos determined at 44 h post-hCG injection; *n*=227 (control), *n*=251 (*Setd1b*^{Gdf9} cKO). The fertilized oocytes from Fig. 5A were included. (B) *Setd1b*^{Gdf9} cKO two-cell embryos display asymmetric divisions. Representative DIC and fluorescent images of embryos at 44 h post-hCG injection. The small percentage of *Setd1b*^{Gdf9} cKO zygotes that advance to the two-cell embryo stage displays asymmetric and multinuclear blastomeres. (C) *Setd1b*^{Gdf9} cKO zygotes arrest at different zygotic stages. Respective percentage of zygotes at different stages analyzed after DAPI staining at 44 h post-hCG injection (*n*=64). Inserts show zygotes at the two stages of cell arrest: (i) nuclear envelope breakdown and (ii) pronuclear stage 5. Scale bars: 50 µm (B, right; C); 100 µm (B, left).

pronuclei at PN4 (Fig. S5C,D). At PN5, the male pronucleus acquired H3K4me2 and me3 (data not shown). Hence, no perturbations of H3K4 methylation were observable *in situ* in *Setd1b*^{Gdf9} cKO oocytes and zygotes. In both genotypes, only the female pronucleus was decorated with the silencing mark H3K9me2 (Fig. S5E). Also, DNA replication proceeded in both paternal and maternal nuclei (Fig. S5F).

Setd1b is not required for primordial germ cell migration

This study began because near-ubiquitous conditional mutagenesis of *Setd1b* in 2-month-old females provoked infertility, which indicated that Setd1b is required for an essential aspect of fertility and/or embryonic development after normal MII oocytes had been established. However, our work with *Gdf9-iCre* did not deliver an explanation for this adult Setd1b requirement for fertility and further investigations are required. We also investigated whether Setd1b was required for oocyte development before deletion by *Gdf9-iCre*. Tamoxifen induced *Rosa26-Cre-ERT2* deletion *in utero* at E6.5 and dissection at E10.5 revealed comparable numbers of alkaline phosphatase-positive cells migrating into the anlage of the female genital ridges of control and *Setd1b*$^{FDC/FDC;RC/+}$ embryos (Fig. S6) indicating that Setd1b is not required for specification or migration of primordial germ cells.

DISCUSSION

H3K4 methyltransferases are known to promote and/or sustain gene expression either through trimethylation of H3K4 on promoter nucleosomes, which presents binding epitopes for proteins involved in transcriptional activity (Ruthenburg et al., 2007; Vermeulen et al., 2007), action at enhancers (Guo et al., 2013; Hu et al., 2013) or by opposing Polycomb group repression, as exemplified by action of the paradigm, fly Trithorax (Klymenko and Müller, 2004). Consequently, loss of an H3K4 methyltransferase is expected to be accompanied by loss of target gene expression. This expectation was partly fulfilled here by the observation that mRNA levels for the oocyte-specific Hox genes *Obox 1*, *2*, *5* and *7* were all decreased in the absence of Setd1b along with oocyte-specific target genes that are presumably regulated by the Obox factors. Indeed, along with the Oboxs, the prominence of *Meis2* and *Sall4* in the top ten most highly expressed transcription factors that were downregulated in the absence of Setd1b leads us to propose that they are also part of the oocyte-specific gene expression program (Fig. 7). Meis factors are known binding partners of other homeodomain transcription factors (Chang et al., 1997; Shen et al., 1999; Moens and Selleri, 2006) suggesting that Meis2 may be an Obox partner. Sall4 is central to early development and pluripotency (Elling et al., 2006), mutated in premature ovarian failure (Wang et al., 2009) and dependent on the early oocyte-specific transcription factor Nobox for expression in oocytes (Choi et al., 2007).

However, against expectations for H3K4 methyltransferases, we found that more mRNAs were upregulated than downregulated in *Setd1b*^{Gdf9} cKO MII oocytes (Fig. 3). Potentially, the widespread mRNA upregulation was due to reactive compensation caused by loss of Setd1b-dependent mRNA expression and consequent functional deficits. This indirect explanation could be part of the reason but it is unsatisfactory because it does not explain why upregulation exceeded downregulation. Furthermore, many of the upregulated mRNAs were either related to metabolic processes, such as ribosome biogenesis, that should have been downregulated in MII oocytes after their peak levels earlier in oogenesis or were mRNAs that should not be expressed in oogenesis. Consequently, we suggest that Setd1b is not only required for appropriate expression of transcriptional activators, such as the Obox factors, but also transcriptional repressors required to maintain the later stages of the oocyte-specific expression program by reducing earlier and unwanted gene expression. Support for this proposition was found in the observation that the most significant term associated with the downregulated mRNAs is ‘Regulation of transcription’ (Fig. 4A). Further investigation revealed that amongst the list of transcription factors downregulated more than twofold in *Setd1b*^{Gdf9} cKO MII oocytes (Table S4), a quarter (16/60) are Zfps with KRAB domains. The KRAB domain mediates gene repression through Trim28/KAP1, the NuRD (nucleosome

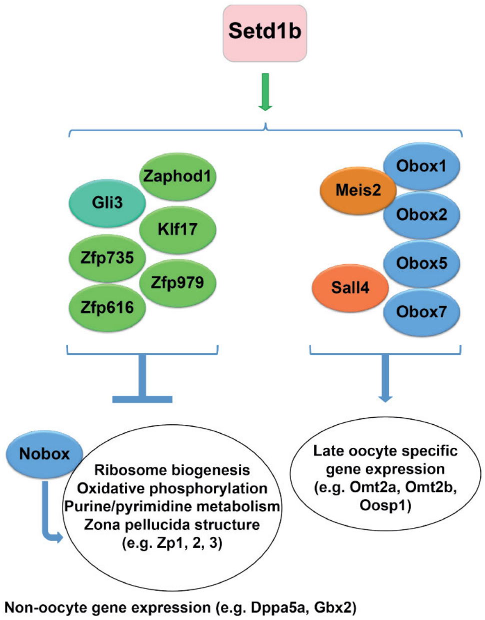
Setd1b
Zaphod1
Gli3
Klf17
Zfp735
Zfp979
Zfp616
Meis2
Obox1
Obox2
Obox5
Sall4
Obox7
Nobox
Ribosome biogenesis
Oxidative phosphorylation
Purine/pyrimidine metabolism
Zona pellucida structure
(e.g. Zp1, 2, 3)
Late oocyte specific gene expression (e.g. Omt2a, Omt2b, Oosp1)
Non-oocyte gene expression (e.g. Dppa5a, Gbx2)

Fig. 7. Model for Setd1b regulation of the oocyte-specific gene expression program. Based on the phenotypic and transcriptome alterations observed after *Gdf9i-Cre* ablation of Setd1b expression, we propose that Setd1b is required to promote expression of the Obox oocyte-specific transcription factors and their potential binding partner Meis2, as well as Sall4 amongst others. These transcription factors in turn promote oocyte-specific gene expression. Setd1b is also required to promote expression of Zpf-KRAB transcriptional repressors that reduce expression of both the early oocyte-specific transcription factor Nobox and mRNAs involved in anabolic metabolism including Zp1, 2, 3, ribosome and mitochondrial biogenesis, and purine metabolism, after their peaks during follicular growth. Sustained expression of Nobox might also sustain expression of the metabolic mRNAs. Setd1b repression through the Zfp-KRAB factors might also be required to keep inappropriate gene expression down as potentially exemplified by elevated Dppa5a and Gbx2 expression.

remodeling deacetylase) complex, Setdb1 (the histone 3 lysine 9 methyltransferase) and HP1 (heterochromatin protein 1; Ayyanathan et al., 2003; Lupo et al., 2013). Potentially, these 16 Znf-KRAB factors mediate a substantial part of the Setd1b repressive capacity. Highest on this list of potential oocyte-specific repressors is the unnamed gene *AI987944* (and we suggest the name Zaphod1 for zinc finger-associated protein highly implicated in oocyte development), followed by *Klf17*, *Zfp979*, *Zfp735*, *Gli3*, *Zfp616* and *Zfp820*. To add further context, we examined the list of 1231 mRNAs that were upregulated more than twofold (excluding poorly expressed mRNAs not upregulated to above 100 RPKM) for transcription factors and nuclear proteins (Table S5). Notably, the key early regulator of oocyte development Nobox (Rajkovic et al., 2004) rated high on the list. Sustained expression of Nobox in MII oocytes implies a failure of appropriate timely downregulation, suggesting that downregulation of Nobox is mediated by Setd1b-associated repressors as a late part of the oocyte-specific gene expression program. In contrast to the 16 Zfp-KRAB factors downregulated more than twofold, only three were upregulated more than twofold thereby lending some support to our proposition regarding the mechanism of Setd1b repression. Of other potential regulators high on the upregulated list we note the presence of *Gbx2* (gastrulation brain homeobox 2), *Glis1*, the CBP/p300 co-factor *Cited2* and five subunits of the Mediator complex, *Med26*, *7*, *17*, *6* and *14*. We also note *Dppa5*, which was upregulated more than 30-fold, has an atypical KH RNA-binding domain like Ooep (oocyte expressed protein) (Pierre et al., 2007). Its substantial overexpression might interfere with Ooep action.

The paradoxical overexpression of zona pellucida components was unexpected because the zona pellucida was obviously compromised. It is possible that an oversupply of components provoked faulty assembly or was provoked by faulty assembly and a zona pellucida quality control feedback mechanism. Alternatively, overexpression could indicate delayed downregulation and persistence of the peak Zp mRNA levels achieved in growing follicles. Notably, ovastacin (*Astl*) expression was unchanged. In normal conditions, zona-penetrated sperm fuse with the oocyte plasma membrane and initiate oocyte activation including the release of cortical granule contents leading to modification of the zona pellucida that prevents subsequent sperm penetration (Clift and Schuh, 2013). Thus, superfluous spermatozoa do not enter the perivitelline space of fertilized oocytes. In *Setd1b*^{Gdf9} cKO zygotes, superfluous sperm accumulated in the perivitelline space thereby testing the second polyspermy block at the oolemma. In 10% of the *Setd1b*^{Gdf9} cKO zygotes this polyspermy block was not sufficient to prevent triploid zygotes. Alternatively, the triploid zygotes might not have been due to breaches of the oolemma but to non-extrusion of the second polar body, which formed a pronucleus. Non-extrusion of the second polar body mechanism is the main cause for tripronuclear zygotes after intracytoplasmic sperm injection (Staessen and Van Steirteghem, 1997).

The *Setd1b*^{Gdf9} cKO MII oocyte gene expression profiles also displayed perturbations in metabolic processes particularly related to mitochondria, which concords with the low developmental capacity and abnormal morphology. Mature oocytes enter a quiescent state in which transcription has ceased, translation is suppressed and mitochondrial activity is very low (Sieber et al., 2016). Despite displaying the transition from the NSN to the SN configuration, *Setd1b*^{Gdf9} cKO MII oocytes had more pervasive transcription than normal and appeared to maintain a more active metabolic state.

Based on these observations, we propose that Setd1b plays a crucial role in the regulation of the late oocyte-specific gene expression program. A part of this relationship involves upregulation of oocyte-specific transcripts and another part involves downregulation of metabolic processes particularly after their peak during follicular growth (Fig. 7).

Notably, the sole ortholog of Set1 in *C. elegans*, SET-2, is required for normal germline development and fertility (Xiao et al., 2011) and soma-specific genes were ectopically upregulated in the germline of worms lacking SET-2 with loss of H3K4 methylation and germline transdifferentiation (Robert et al., 2014) again indicating a role for Set1 in repression of inappropriate germline gene expression. Similarly, removal of the *Drosophila* Set1, dSet1, leads to gradual GSC loss with disruption of H3K4 trimethylation in the ovary (Xuan et al., 2013). In an additional study (Yan et al., 2014), using transgenic RNAi in the *Drosophila* female germline, several candidate genes including *dSet1* were identified with a specific requirement in GSCs. Potentially, these similarities indicate conserved aspects of H3K4 methylation action in germ lineages.

In mouse oocytes, mRNAs for all six Set1/Trithorax-type members, *Setd1a*, *Setd1b* and *Mll1-4*, are expressed (Kageyama et al., 2007; Andreu-Vieyra et al., 2010; Pfeiffer et al., 2011; Bledau et al., 2014; Aoshima et al., 2015). To date, we have conditionally mutated *Mll1*, *Mll2*, *Setd1a* and *Setd1b* using *Gdf9-iCre* to establish that only oocyte-specific deletion of Setd1b and Mll2 perturb fertility (Andreu-Vieyra et al., 2010; Bledau et al., 2014; this paper; and A.K., Q.Z., K.A. and A.F.S., unpublished). Mll2 is required during oogenesis and shortly after fertilization when it is the major H3K4 trimethyltransferase. By the blastocyst stage, the embryo and embryonic stem cells are independent of Mll2 but dependent on Setd1a (Bledau et al., 2014). Mll2 is again required after gastrulation; however, from the blastocyst stage onwards it is no longer the major H3K4 trimethyltransferase (Glaser et al., 2006). Setd1b is required for progression through the first cell cycle before zygotic genome activation but like Mll2 is not required in embryonic stem cells. Notably, Setd1b, like Mll2, is again required after gastrulation with the knockout of either gene presenting a very similar retarded and disorganized phenotype (Bledau et al., 2014). These functional shifts amongst the H3K4 methyltransferases suggest aspects of cell-type specificity that are difficult to explain. However, the role of Setd1b in the mouse oocyte potentially provides a new paradigm for action by an H3K4 methyltransferase as an intermediate coordinator in a cell type-specific gene expression program.

MATERIALS AND METHODS

Generation of mouse lines

The *Setd1b* conditional strain (Fig. 1A) was generated previously (Bledau et al., 2014). *Setd1b*$^{D/+}$ mice were crossed to the *CAGGs-Flpo* (Kranz et al., 2010) line to generate *Setd1b*$^{FD/+}$ mice. Subsequently, those mice were crossed to the *Rosa26-Cre-ERT2* (*RC*) line (Seibler et al., 2003) to generate conditional, tamoxifen-inducible *Setd1b*$^{FD/+;RC/+}$ mice. In order to achieve Setd1b deletion *in utero*, tamoxifen was administered to pregnant *Setd1b*$^{FD/+}$ females after crossing to *Setd1b*$^{FD/FD;RC/RC}$ males. Pregnant females at E6.5 received one dose of 3 mg tamoxifen by oral gavage and embryos were dissected 4 days later. To obtain Setd1b-deficient oocytes by Cre deletion from postnatal day 3, *Setd1b*$^{FD/FD}$ females were crossed with *Gdf9-iCre* males (Lan et al., 2004). For brevity, *Setd1b*$^{FDC/FDC;Gdf9iCre/+}$ mice are termed *Setd1b*^{Gdf9} cKO and *Setd1b*$^{FD/FD}$ or *Setd1b*$^{FDC/+;Gdf9iCre/+}$ are termed control (Fig. 1B). Primers for genotyping are provided in Table S1. All animal experiments were performed according to German law.

Oocyte and embryo collection

Oocytes from 24- to 28-day-old females were obtained by injection with 5 IU of pregnant mare serum gonadotropin (PMSG, Intravet Deutschland, 707165) followed by an injection of 5 IU human chorionic gonadotropin (hCG, Sigma-Aldrich, CG-10-1VL) 46 h later. Cumulus-oocyte complexes were isolated from the ampulla 16 h after hCG injection and collected in KSOM containing 1 mg/ml hyaluronidase (Sigma-Aldrich) to dissociate cumulus cells. In order to obtain periovulatory oocytes, ovaries were digested and fully grown oocytes (>60 µm) with a germinal vesicle were collected. To evaluate *in vitro* development, females were mated, embryos collected, cultured in KSOM and scored at the indicated time points.

Histology, immunohistochemistry and immunofluorescence

Ovaries were fixed with 4% paraformaldehyde (PFA). The dehydrated ovaries were embedded in paraffin, and 5 µm-thick sections were prepared and stained with Periodic acid-Schiff (PAS) stain. Immunohistochemistry was performed as described (Bledau et al., 2014). Primary antibodies for immunohistochemistry were used as described in Table S2. Embryos were isolated, fixed in 4% PFA and stained with 25 mM Tris-maleate pH 9.0, 8 mM $MgCl_2$, 0.4% 1-naphthyl phosphate (Sigma-Aldrich) and 1% Fast Red (Sigma-Aldrich).

Oocytes were fixed in 4% PFA in PBS/0.2% bovine serum albumin (BSA) for 20 min at 37°C. After several washes in 0.2% BSA in PBS-Tween 20 (0.05% Tween 20 in PBS), the oocytes were permeabilized in 0.5% Triton X-100 in PBS for 15 min at room temperature (RT). Oocytes were incubated with the desired primary antibody according to Table S2. Secondary antibodies were used as followed: bovine anti-rabbit IgG-Rhodamine (1:50; Santa Cruz, sc-2367), goat anti-rabbit IgG-CFL 488 (1:200; Santa Cruz, sc-362262), fluorescein rabbit anti-rat IgG (1:200; Vector Laboratories, FI-4001), goat anti-mouse IgG-CFL 488 (1:100; Santa Cruz, sc-362257). For detection of lipid droplets, oocytes were incubated with Nile Red (Sigma-Aldrich, 5 mg/ml) in KSOM for 5 min. Oocyte chromosome spreads were performed as described (Hodges and Hunt, 2002).

BrdU labeling for DNA replication

Fertilized oocytes were labeled *in vitro* with 5-bromo-2’-deoxyuridine (BrdU) in order to evaluate DNA replication according to Wossidlo et al. (2010). Zygotes were incubated in KSOM medium supplemented with 1 mM BrdU (Sigma-Aldrich) until 44 h post-hCG injection. After incubation, embryos were fixed for 15 min in 4% PFA in PBS/0.2% BSA at RT, permeabilized with 0.2% Triton X-100 in PBS for 20 min RT, treated with 4 M HCl for 10 min and neutralized for 10 min with 100 mM Tris-HCl (pH 8.5). After a second fixation step, embryos were blocked with 3% BSA/0.1% Triton X-100 in PBS for 2 h at 4°C and then incubated for 2 h at RT with anti-BrdU FITC-conjugated rat monoclonal antibody (1:50; Abcam, ab74545). Embryos were washed and mounted on slides with a small drop of Vectashield mounting medium (Vector Laboratories) containing 5 µg/ml propidium iodide.

Microscopy

Images of the mounted oocytes/embryos were taken using a Zeiss Axiocam MRm camera on an Axiovert 200M microscope with ApoTome and HXP illumination using 63×/1.4 oil immersion objective operated by ZEN blue software. *z*-stacks of the spindles were acquired by laser-scanning confocal microscopy using a Zeiss LSM 780 with 40×/1.2 W objective. Volume reconstructions used Arivis Vision 4D software. Length measurements of the spindle tubulin filaments were carried out using the Fiji Simple Neurite Tracer plugin (Longair et al., 2011). Images of histological sections were collected using an Olympus WF upright microscope.

Reverse transcription and quantitative PCR (qRT-PCR) analysis

Total RNA from oocytes was isolated using Trizol reagent (Ambion) and reverse transcribed using the SuperScript III First-Strand Synthesis System (Invitrogen). Real-time quantitative PCR was performed with Go Taq qPCR Master Mix (Promega) by Mx3000P multiplex PCR (Agilent). Ct values were normalized against *Rpl19*. Primer sequences and the lengths of the amplified products are given in Table S1. Fold differences in expression levels were calculated according to the $2^{-\Delta Ct}$ method (Livak and Schmittgen, 2001).

RNA isolation, amplification and sequencing

For single oocyte RNA profiling *Setd1b*^{Gdf9} cKO and control mice were superovulated and four oocytes collected for each genotype. Cells were lysed, RNA reverse transcribed and amplified according to the protocol of SMARTer Ultra Low RNA Kit for Illumina Sequencing (vers. 1, Clontech). Subsequently, samples were subjected to NEBNext Ultra DNA library preparation for Illumina using indexed adaptors (NEB). Resulting libraries were pooled in equimolar quantities for 75-bp single-read sequencing on Illumina HiSeq 2000, resulting in about 20 (16-23) million reads per sample.

The short reads were aligned to the mm10 transcriptome with GSNAP (2014-10-14; Wu and Watanabe, 2005; Wu and Nacu, 2010) and a table of read counts per gene was created based on the overlap of the uniquely mapped reads with the Ensembl Gene annotation version 69, using featureCounts (version 1.4.4; Liao et al., 2014). The raw read counts were further processed with the DESeq2 R package (version 1.3.79; Anders and Huber, 2010). Sample-to-sample correlation was computed by Euclidean distance between samples based on the normalized counts. Differential gene expression analysis was performed with DESeq2, for which a maximum of 1% false discoveries (1% FDR) was accepted. To identify enrichment for particular biological processes associated with the differentially expressed genes (DEGs), the DAVID GO/BP/FAT database (Huang et al., 2009) was used. Gene set enrichment analysis was performed using GSEA software from the Broad Institute (Subramanian et al., 2005).

Acknowledgements

We thank Stefanie Weidlich and Mandy Obst for excellent technical assistance. We thank the Biomedical Services of the Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, for excellent service. We are also grateful to Konstantin Lepikhov for helpful advice on monitoring replication in zygotes by BrdU incorporation.

Competing interests

The authors declare no competing or financial interests.

Author contributions

Conceptualization: K.A., A.F.S., A.K.; Methodology: D.B., H.H., A.K.; Validation: A.K.; Formal analysis: Q.Z., J.H., L.G., G.K., A.F.S., A.K.; Investigation: D.B., Q.Z., S.R., H.H., K.S., N.G., J.H., A.K.; Data curation: D.B., S.R., A.D.; Writing - original draft: D.B., A.F.S., A.K.; Writing - review & editing: A.F.S., A.K.; Visualization: Q.Z., H.H., L.G., A.K.; Supervision: A.F.S., A.K.; Funding acquisition: A.F.S., A.K.

Funding

This work was supported by funding from the Deutsche Forschungsgemeinschaft (KR2154/3-1 to A.K., STE903/4-1 to A.F.S.), the Deutsche Krebshilfe (110560) and the Else Kröner-Fresenius-Stiftung (2012_A300).

Data availability

RNA sequencing data have been deposited in Gene Expression Omnibus under accession number GSE85360.

Supplementary information

Supplementary information available online at http://dev.biologists.org/lookup/doi/10.1242/dev.143347.supplemental

References

Abe, K., Yamamoto, R., Franke, V., Cao, M., Suzuki, Y., Suzuki, M. G., Vlahovicek, K., Svoboda, P., Schultz, R. M. and Aoki, F. (2015). The first murine zygotic transcription is promiscuous and uncoupled from splicing and 3′ processing. *EMBO J.* **34**, 1523-1537.

Allegrucci, C., Thurston, A., Lucas, E. and Young, L. (2005). Epigenetics and the germline. *Reproduction* **129**, 137-149.

Anders, S. and Huber, W. (2010). Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106.

Andreu-Vieyra, C. V., Chen, R., Agno, J. E., Glaser, S., Anastassiadis, K., Stewart, A. F. and Matzuk, M. M. (2010). MLL2 is required in oocytes for bulk histone 3 lysine 4 trimethylation and transcriptional silencing. *PLoS Biol.* **8**, e1000453.

Aoshima, K., Inoue, E., Sawa, H. and Okada, Y. (2015). Paternal H3K4 methylation is required for minor zygotic gene activation and early mouse embryonic development. *EMBO Rep.* **16**, 803-812.

Ayyanathan, K., Lechner, M. S., Bell, P., Maul, G. G., Schultz, D. C., Yamada, Y., Tanaka, K., Torigoe, K. and Rauscher, F. J.III. (2003). Regulated recruitment of HP1 to a euchromatic gene induces mitotically heritable, epigenetic gene silencing: a mammalian cell culture model of gene variegation. *Genes Dev.* **17**, 1855-1869.

Bannister, A. J. and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. *Cell Res.* **21**, 381-395.

Bledau, A. S., Schmidt, K., Neumann, K., Hill, U., Ciotta, G., Gupta, A., Torres, D. C., Fu, J., Kranz, A., Stewart, A. F. et al. (2014). The H3K4 methyltransferase Setd1a is first required at the epiblast stage, whereas Setd1b becomes essential after gastrulation. *Development* **141**, 1022-1035.

Chang, C. P., Jacobs, Y., Nakamura, T., Jenkins, N. A., Copeland, N. G. and Cleary, M. L. (1997). Meis proteins are major in vivo DNA binding partners for wild-type but not chimeric Pbx proteins. *Mol. Cell. Biol.* **17**, 5679-5687.

Chauhan, C., Zraly, C. B., Parilla, M., Diaz, M. O. and Dingwall, A. K. (2012). Histone recognition and nuclear receptor co-activator functions of Drosophila cara mitad, a homolog of the N-terminal portion of mammalian MLL2 and MLL3. *Development* **139**, 1997-2008.

Choi, Y., Qin, Y., Berger, M. F., Ballow, D. J., Bulyk, M. L. and Rajkovic, A. (2007). Microarray analyses of newborn mouse ovaries lacking Nobox. *Biol. Reprod.* **77**, 312-319.

Clift, D. and Schuh, M. (2013). Restarting life: fertilization and the transition from meiosis to mitosis. *Nat. Rev. Mol. Cell Biol.* **14**, 549-562.

Dahl, J. A., Jung, I., Aanes, H., Greggains, G. D., Manaf, A., Lerdrup, M., Li, G., Kuan, S., Li, B., Lee, A. Y. et al. (2016). Broad histone H3K4me3 domains in mouse oocytes modulate maternal-to-zygotic transition. *Nature* **537**, 548-552.

Dambacher, S., Hahn, M. and Schotta, G. (2010). Epigenetic regulation of development by histone lysine methylation. *Heredity* **105**, 24-37.

Dehe, P.-M., Dichtl, B., Schaft, D., Roguev, A., Pamblanco, M., Lebrun, R., Rodriguez-Gil, A., Mkandawire, M., Landsberg, K., Shevchenko, A. et al. (2006). Protein interactions within the Set1 complex and their roles in the regulation of histone 3 lysine 4 methylation. *J. Biol. Chem.* **281**, 35404-35412.

Dou, Y., Milne, T. A., Ruthenburg, A. J., Lee, S., Lee, J. W., Verdine, G. L., Allis, C. D. and Roeder, R. G. (2006). Regulation of MLL1 H3K4 methyltransferase activity by its core components. *Nat. Struct. Mol. Biol.* **13**, 713-719.

Edson, M. A., Nagaraja, A. K. and Matzuk, M. M. (2009). The mammalian ovary from genesis to revelation. *Endocr. Rev.* **30**, 624-712.

Elling, U., Klasen, C., Eisenberger, T., Anlag, K. and Treier, M. (2006). Murine inner cell mass-derived lineages depend on Sall4 function. *Proc. Natl. Acad. Sci. USA* **103**, 16319-16324.

Ernst, P. and Vakoc, C. R. (2012). WRAD: enabler of the SET1-family of H3K4 methyltransferases. *Brief. Funct. Genomics* **11**, 217-226.

Ernst, P., Mabon, M., Davidson, A. J., Zon, L. I. and Korsmeyer, S. J. (2004). An Mll-dependent Hox program drives hematopoietic progenitor expansion. *Curr. Biol.* **14**, 2063-2069.

Francis, N. J. and Kingston, R. E. (2001). Mechanisms of transcriptional memory. *Nat. Rev. Mol. Cell Biol.* **2**, 409-421.

Gahurova, L., Tomizawa, S. I., Smallwood, S. A., Stewart-Morgan, K. R., Saadeh, H., Kim, J., Andrews, S. R., Chen, T. and Kelsey, G. (2017). Transcription and chromatin determinants of de novo DNA methylation timing in oocytes. *Epigenetics Chromatin* **10**, 25.

Glaser, S., Schaft, J., Lubitz, S., Vintersten, K., van der Hoeven, F., Tufteland, K. R., Aasland, R., Anastassiadis, K., Ang, S.-L. and Stewart, A. F. (2006). Multiple epigenetic maintenance factors implicated by the loss of Mll2 in mouse development. *Development* **133**, 1423-1432.

Glaser, S., Lubitz, S., Loveland, K. L., Ohbo, K., Robb, L., Schwenk, F., Seibler, J., Roellig, D., Kranz, A., Anastassiadis, K. et al. (2009). The histone 3 lysine 4 methyltransferase, Mll2, is only required briefly in development and spermatogenesis. *Epigenet. Chromatin* **2**, 5.

Gu, B. and Lee, M. G. (2013). Histone H3 lysine 4 methyltransferases and demethylases in self-renewal and differentiation of stem cells. *Cell Biosci.* **3**, 39.

Guo, C., Chen, L. H., Huang, Y., Chang, C.-C., Wang, P., Pirozzi, C. J., Qin, X., Bao, X., Greer, P. K., McLendon, R. E. et al. (2013). KMT2D maintains neoplastic cell proliferation and global histone H3 lysine 4 monomethylation. *Oncotarget* **4**, 2144-2153.

Herz, H.-M., Mohan, M., Garruss, A. S., Liang, K., Takahashi, Y. H., Mickey, K., Voets, O., Verrijzer, C. P. and Shilatifard, A. (2012). Enhancer-associated H3K4 monomethylation by Trithorax-related, the Drosophila homolog of mammalian Mll3/Mll4. *Genes Dev.* **26**, 2604-2620.

Hodges, C. A. and Hunt, P. A. (2002). Simultaneous analysis of chromosomes and chromosome-associated proteins in mammalian oocytes and embryos. *Chromosoma* **111**, 165-169.

Hu, D., Gao, X., Morgan, M. A., Herz, H.-M., Smith, E. R. and Shilatifard, A. (2013). The MLL3/MLL4 branches of the COMPASS family function as major histone H3K4 monomethylases at enhancers. *Mol. Cell. Biol.* **33**, 4745-4754.

Huang, D. W., Sherman, B. T. and Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44-57.

Kageyama, S.-I., Liu, H., Kaneko, N., Ooga, M., Nagata, M. and Aoki, F. (2007). Alterations in epigenetic modifications during oocyte growth in mice. *Reproduction* **133**, 85-94.

Klymenko, T. and Müller, J. (2004). The histone methyltransferases Trithorax and Ash1 prevent transcriptional silencing by Polycomb group proteins. *EMBO Rep.* **5**, 373-377.

Kranz, A., Fu, J., Duerschke, K., Weidlich, S., Naumann, R., Stewart, A. F. and Anastassiadis, K. (2010). An improved Flp deleter mouse in C57Bl/6 based on Flpo recombinase. *Genesis* **48**, 512-520.

Lan, Z.-J., Xu, X. and Cooney, A. J. (2004). Differential oocyte-specific expression of Cre recombinase activity in GDF-9-iCre, Zp3cre, and Msx2Cre transgenic mice. *Biol. Reprod.* **71**, 1469-1474.

Lee, J.-H. and Skalnik, D. G. (2012). Rbm15-Mkl1 interacts with the Setd1b histone H3-Lys4 methyltransferase via a SPOC domain that is required for cytokine-independent proliferation. *PLoS ONE* **7**, e42965.

Lee, J.-H., Tate, C. M., You, J.-S. and Skalnik, D. G. (2007). Identification and characterization of the human Set1B histone H3-Lys4 methyltransferase complex. *J. Biol. Chem.* **282**, 13419-13428.

Lee, J. E., Wang, C., Xu, S., Cho, Y. W., Wang, L., Feng, X., Baldridge, A., Sartorelli, V., Zhuang, L., Peng, W. et al. (2013). H3K4 mono- and di-methyltransferase MLL4 is required for enhancer activation during cell differentiation. *Elife* **2**, e01503.

Lepikhov, K. and Walter, J. (2004). Differential dynamics of histone H3 methylation at positions K4 and K9 in the mouse zygote. *BMC Dev. Biol.* **4**, 12.

Li, B., Howe, L., Anderson, S., Yates, J. R., III and Workman, J. L. (2003). The Set2 histone methyltransferase functions through the phosphorylated carboxyl-terminal domain of RNA polymerase II. *J. Biol. Chem.* **278**, 8897-8903.

Liao, Y., Smyth, G. K. and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930.

Liu, X., Wang, C., Liu, W., Li, J., Li, C., Kou, X., Chen, J., Zhao, Y., Gao, H., Wang, H. et al. (2016). Distinct features of H3K4me3 and H3K27me3 chromatin domains in pre-implantation embryos. *Nature* **537**, 558-562.

Livak, K. J. and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**, 402-408.

Longair, M. H., Baker, D. A. and Armstrong, J. D. (2011). Simple Neurite Tracer: open source software for reconstruction, visualization and analysis of neuronal processes. *Bioinformatics* **27**, 2453-2454.

Lupo, A., Cesaro, E., Montano, G., Zurlo, D., Izzo, P. and Costanzo, P. (2013). KRAB-zinc finger proteins: a repressor family displaying multiple biological functions. *Curr. Genomics* **14**, 268-278.

Moens, C. B. and Selleri, L. (2006). Hox cofactors in vertebrate development. *Dev. Biol.* **291**, 193-206.

Park, S.-J., Komata, M., Inoue, F., Yamada, K., Nakai, K., Ohsugi, M. and Shirahige, K. (2013). Inferring the choreography of parental genomes during fertilization from ultralarge-scale whole-transcriptome analysis. *Genes Dev.* **27**, 2736-2748.

Peters, A. H. F. M., O'Carroll, D., Scherthan, H., Mechtler, K., Sauer, S., Schöfer, C., Weipoltshammer, K., Pagani, M., Lachner, M., Kohlmaier, A. et al. (2001). Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. *Cell* **107**, 323-337.

Pfeiffer, M. J., Siatkowski, M., Paudel, Y., Balbach, S. T., Baeumer, N., Crosetto, N., Drexler, H. C. A., Fuellen, G. and Boiani, M. (2011). Proteomic analysis of mouse oocytes reveals 28 candidate factors of the "reprogrammome". *J. Proteome Res.* **10**, 2140-2153.

Pierre, A., Gautier, M., Callebaut, I., Bontoux, M., Jeanpierre, E., Pontarotti, P. and Monget, P. (2007). Atypical structure and phylogenomic evolution of the new eutherian oocyte- and embryo-expressed KHDC1/DPPA5/ECAT1/OOEP gene family. *Genomics* **90**, 583-594.

Rajkovic, A., Pangas, S. A., Ballow, D., Suzumori, N. and Matzuk, M. M. (2004). NOBOX deficiency disrupts early folliculogenesis and oocyte-specific gene expression. *Science* **305**, 1157-1159.

Reik, W. (2007). Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature* **447**, 425-432.

Robert, V. J., Mercier, M. G., Bedet, C., Janczarski, S., Merlet, J., Garvis, S., Ciosk, R. and Palladino, F. (2014). The SET-2/SET1 histone H3K4 methyltransferase maintains pluripotency in the Caenorhabditis elegans germline. *Cell Rep.* **9**, 443-450.

Roguev, A., Schaft, D., Shevchenko, A., Pijnappel, W. W., Wilm, M., Aasland, R. and Stewart, A. F. (2001). The Saccharomyces cerevisiae Set1 complex includes an Ash2 homologue and methylates histone 3 lysine 4. *EMBO J.* **20**, 7137-7148.

Ruthenburg, A. J., Allis, C. D. and Wysocka, J. (2007). Methylation of lysine 4 on histone H3: intricacy of writing and reading a single epigenetic mark. *Mol. Cell* **25**, 15-30.

Santos-Rosa, H., Schneider, R., Bannister, A. J., Sherriff, J., Bernstein, B. E., Emre, N. C. T., Schreiber, S. L., Mellor, J. and Kouzarides, T. (2002). Active genes are tri-methylated at K4 of histone H3. *Nature* **419**, 407-411.

Schaft, D., Roguev, A., Kotovic, K. M., Shevchenko, A., Sarov, M., Shevchenko, A., Neugebauer, K. M. and Stewart, A. F. (2003). The histone 3 lysine 36 methyltransferase, SET2, is involved in transcriptional elongation. *Nucleic Acids Res.* **31**, 2475-2482.

Seibler, J., Zevnik, B., Kuter-Luks, B., Andreas, S., Kern, H., Hennek, T., Rode, A., Heimann, C., Faust, N., Kauselmann, G. et al. (2003). Rapid generation of inducible mouse mutants. *Nucleic Acids Res.* **31**, e12.

Shen, W. F., Rozenfeld, S., Kwong, A., Kömüves, L. G., Lawrence, H. J. and Largman, C. (1999). HOXA9 forms triple complexes with PBX2 and MEIS1 in myeloid cells. *Mol. Cell. Biol.* **19**, 3051-3061.

Shilatifard, A. (2012). The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis. *Annu. Rev. Biochem.* **81**, 65-95.

Sieber, M. H., Thomsen, M. B. and Spradling, A. C. (2016). Electron transport chain remodeling by GSK3 during oogenesis connects nutrient state to reproduction. *Cell* **164**, 420-432.

Staessen, C. and Van Steirteghem, A. C. (1997). The chromosomal constitution of embryos developing from abnormally fertilized oocytes after intracytoplasmic sperm injection and conventional *in-vitro* fertilization. *Hum. Reprod.* **12**, 321-327.

Stewart, K. R., Veselovska, L., Kim, J., Huang, J., Saadeh, H., Tomizawa, S.-I., Smallwood, S. A., Chen, T. and Kelsey, G. (2015). Dynamic changes in histone modifications precede de novo DNA methylation in oocytes. *Genes Dev.* **29**, 2449-2462.

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S. et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* **102**, 15545-15550.

van der Heijden, G. W., Dieker, J. W., Derijck, A. A. H. A., Muller, S., Berden, J. H. M., Braat, D. D. M., van der Vlag, J. and de Boer, P. (2005). Asymmetry in histone H3 variants and lysine methylation between paternal and maternal chromatin of the early mouse zygote. *Mech. Dev.* **122**, 1008-1022.

van Nuland, R., Smits, A. H., Pallaki, P., Jansen, P. W. T. C., Vermeulen, M. and Timmers, H. T. M. (2013). Quantitative dissection and stoichiometry determination of the human SET1/MLL histone methyltransferase complexes. *Mol. Cell. Biol.* **33**, 2067-2077.

Vermeulen, M., Mulder, K. W., Denissov, S., Pijnappel, W. W. M. P., van Schaik, F. M. A., Varier, R. A., Baltissen, M. P. A., Stunnenberg, H. G., Mann, M. and Timmers, H. T. M. (2007). Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. *Cell* **131**, 58-69.

Voigt, P., LeRoy, G., Drury, W. J., III, Zee, B. M., Son, J., Beck, D. B., Young, N. L., Garcia, B. A. and Reinberg, D. (2012). Asymmetrically modified nucleosomes. *Cell* **151**, 181-193.

Wang, L., Wang, Z.-B., Zhang, X., FitzHarris, G., Baltz, J. M., Sun, Q.-Y. and Liu, X. J. (2008). Brefeldin A disrupts asymmetric spindle positioning in mouse oocytes. *Dev. Biol.* **313**, 155-166.

Wang, B., Li, L., Ni, F., Song, J., Wang, J., Mu, Y., Ma, X. and Cao, Y. (2009). Mutational analysis of SAL-Like 4 (SALL4) in Han Chinese women with premature ovarian failure. *Mol. Hum. Reprod.* **15**, 557-562.

Wossidlo, M., Arand, J., Sebastiano, V., Lepikhov, K., Boiani, M., Reinhardt, R., Schöler, H. and Walter, J. (2010). Dynamic link of DNA demethylation, DNA strand breaks and repair in mouse zygotes. *EMBO J.* **29**, 1877-1888.

Wu, T. D. and Nacu, S. (2010). Fast and SNP-tolerant detection of complex variants and splicing in short reads. *Bioinformatics* **26**, 873-881.

Wu, T. D. and Watanabe, C. K. (2005). GMAP: a genomic mapping and alignment program for mRNA and EST sequences. *Bioinformatics* **21**, 1859-1875.

Wysocka, J., Swigut, T., Milne, T. A., Dou, Y., Zhang, X., Burlingame, A. L., Roeder, R. G., Brivanlou, A. H. and Allis, C. D. (2005). WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. *Cell* **121**, 859-872.

Xiao, Y., Bedet, C., Robert, V. J. P., Simonet, T., Dunkelbarger, S., Rakotomalala, C., Soete, G., Korswagen, H. C., Strome, S. and Palladino, F. (2011). Caenorhabditis elegans chromatin-associated proteins SET-2 and ASH-2 are differentially required for histone H3 Lys 4 methylation in embryos and adult germ cells. *Proc. Natl. Acad. Sci. USA* **108**, 8305-8310.

Xuan, T., Xin, T., He, J., Tan, J., Gao, Y., Feng, S., He, L., Zhao, G. and Li, M. (2013). dBre1/dSet1-dependent pathway for histone H3K4 trimethylation has essential roles in controlling germline stem cell maintenance and germ cell differentiation in the Drosophila ovary. *Dev. Biol.* **379**, 167-181.

Yan, D., Neumüller, R. A., Buckner, M., Ayers, K., Li, H., Hu, Y., Yang-Zhou, D., Pan, L., Wang, X., Kelley, C. et al. (2014). A regulatory network of Drosophila germline stem cell self-renewal. *Dev. Cell* **28**, 459-473.

Zhang, T., Cooper, S. and Brockdorff, N. (2015). The interplay of histone modifications - writers that read. *EMBO Rep.* **16**, 1467-1481.

Zhang, B., Zheng, H., Huang, B., Li, W., Xiang, Y., Peng, X., Ming, J., Wu, X., Zhang, Y., Xu, Q. et al. (2016). Allelic reprogramming of the histone modification H3K4me3 in early mammalian development. *Nature* **537**, 553-557.