Published in final edited form as:
Cell Rep. 2013 July 11; 4(1): 205–219. doi:10.1016/j.celrep.2013.06.004.

De Novo DNA Methylation in the Male Germ Line Occurs by Default but is Excluded at Sites of H3K4 Methylation

Purnima Singh[1], **Arthur X. Li**[2,&], **Diana A. Tran**[1,4,&], **Nathan Oates**[1], **Eun-Rim Kang**[1], **Xiwei Wu**[3], and **Piroska E. Szabó**[1,*]

[1]Department of Molecular and Cellular Biology, City of Hope National Medical Center, Duarte, CA 91010 U.S.A

[2]Department of Information Science, City of Hope National Medical Center, Duarte, CA 91010 U.S.A

[3]Department of Bioinformatics, City of Hope National Medical Center, Duarte, CA 91010 U.S.A

[4]Irell and Manella Graduate School of Biological Sciences of the Beckman Research Institute, City of Hope National Medical Center, Duarte, CA 91010 U.S.A

SUMMARY

To understand what dictates the emerging patterns of de novo DNA methylation in the male germline, we mapped DNA methylation, chromatin, and transcription changes in purified fetal mouse germ cells by using methylated CpG island recovery assay (MIRA)-chip, chromatin immunoprecipitation (ChIP)-chip, and strand-specific RNA deep sequencing, respectively. Global de novo methylation occurred by default in prospermatogonia without any apparent trigger from preexisting repressive chromatin marks but was preceded by broad, low-level transcription along the chromosomes, including the four known paternally imprinted differentially methylated regions (DMRs). Default methylation was excluded only at precisely aligned constitutive or emerging peaks of H3K4me2, including most CpG islands and some intracisternal A particles (IAPs). Similarly, each maternally imprinted DMR was protected from default DNA methylation among highly methylated DNA by an H3K4me2 peak and transcription initiation at least in one strand. Our results suggest that the pattern of de novo DNA methylation in prospermatogonia is dictated by opposing actions of broad, low-level transcription and dynamic patterns of active chromatin.

INTRODUCTION

Mammalian development involves two global waves of epigenetic remodeling events, related to the soma-germ line transition in primordial germ cells and the germ line-soma transition after fertilization (Sasaki and Matsui, 2008; Seisenberger et al., 2013). Primordial germ cells (PGCs) undergo global erasure of DNA methylation by mid-gestation with complex kinetics (Guibert et al., 2012; Hajkova et al., 2002; Kobayashi et al., 2013; Seisenberger et al., 2012). Gamete-specific DNA methylation is subsequently established in

*Correspondence: Piroska E. Szabó, pszabo@coh.org, (626) 301-8484.
&Contributed equally

ACCESSION NUMBERS

Primary data associated with this manuscript has been deposited to GEO database with SuperSeries accession number GSE46954.

fetal prospermatogonia in the male and after birth in growing oocytes in the female. Some fascinating questions remain, such as how the DNMTs find their targets in the germ lines. Why do certain CpG islands (CGIs) become methylated in sperm (Kobayashi et al., 2012; Smallwood et al., 2011) and why some IAP repeats remain partially unmethylated in sperm (Rakyan et al., 2003)? Why are the methylation patterns of the sperm and the egg different (Borgel et al., 2010; Kobayashi et al., 2012; Smallwood et al., 2011; Smith et al., 2012) even though the same enzyme, DNMT3A, and its cofactor, DNMT3L, are mainly responsible for de novo methylation in both germ lines (Bourc'his et al., 2001; Hata et al., 2002; Kaneda et al., 2004; Kobayashi et al., 2012; Okano et al., 1999; Smallwood et al., 2011)? We hypothesized that preexisting chromatin composition has a patterning role, because DNA and histone methylation are structurally and functionally linked (Cheng and Blumenthal, 2010). Additionally, de novo methylation in prospermatogonia coincides with global chromatin changes (Abe et al., 2011; Yoshioka et al., 2009). For example, global H3K9ac, H3K4me3, HK9me3, H3K27me3, H3K79me2 and H3K79me3 levels greatly increase in prospermatogonia by 15.5 dpc, before global DNA methylation occurs between 15.5–17.5 dpc, but not in female fetal germ cells which do not undergo global de novo DNA methylation (Abe et al., 2011).

Most of the gametic methylation differences are removed in the zygote or during embryo development (Borgel et al., 2010; Smallwood et al., 2011; Smith et al., 2012). Notable exemptions are the germ line differentially methylated regions (gDMRs) of imprinted genes (Ferguson-Smith, 2011). Here, methylation inherited from the sperm or egg, is maintained in the paternal (PAT) or maternal (MAT) chromosomes and often regulates the allele-specific expression of the associated imprinted genes in the soma. Imprint erasure and establishment in germ cells coincides with global epigenetic remodeling events (Reik et al., 2001; Sasaki and Matsui, 2008). In the mouse, DMRs become demethylated in PGCs between 10.5 and 12.5 dpc (Hajkova et al., 2002), during global DNA methylation erasure. Remethylation of PAT DMRs occurs at the fetal stages in prospermatogonia and of MAT DMRs after birth in growing oocytes (Hiura et al., 2006; Hiura et al., 2010; Kato et al., 2007; Lee et al., 2010; Lucifero et al., 2004). These time windows correspond to global waves of de novo methylation in the male (Kobayashi et al., 2013; Seisenberger et al., 2012) and female (Smallwood et al., 2011) germ lines. It has been long anticipated that a specific mechanism marks these DMRs for targeted de novo methylation in prospermatogonia and growing oocytes. Little is known about how transcription and chromatin affect the establishment of DNA methylation at DMRs. In the female germ line, KDM1B-dependent H3K4 demethylation is required for the establishment of methylation imprints at certain MAT DMRs in growing oocytes (Ciccone et al., 2009). Similarly, H3K4me3 allelic bias delays de novo methylation of the maternally inherited allele of the *H19-Igf2* PAT DMR in prospermatogonia (Lee et al., 2010). These studies suggest that H3K4 methylation modulates de novo methylation at specific loci. However, it is not known if the presence of any repressive histone mark is required in the germ line for global de novo methylation or for imprint establishment.

Transcription along gene bodies correlates with high levels of methylation (Rauch et al., 2009). According to high-throughput bisulfite sequencing results, most de novo methylation occurs along gene bodies in the growing oocyte (Kobayashi et al., 2012; Smallwood et al., 2011). Methylated CGIs reside downstream of alternative upstream promoters in growing oocytes at the time of de novo methylation (Smallwood et al., 2011). At the *Gnas-Nespas* imprinted domain, transcription run-through is required for de novo methylation of the domain's MAT DMRs in growing oocytes (Chotalia et al., 2009). Based on these mapping and genetic studies, transcription is considered as a trigger for de novo methylation in growing oocytes. Transcription run-through was also described at two PAT DMRs in prospermatogonia (Henckel et al., 2012), but it is not known whether these transcripts are

required for de novo methylation of the respective PAT DMRs or whether transcription has an effect on global de novo methylation, for example on the broad intergenic methylation found in sperm.

To find out whether transcription and/or chromatin guide the methylation machinery in prospermatogonia, it is necessary to align the distribution of transcripts and histone covalent modifications with emerging DNA methylation patterns. However, such mapping analysis has not been considered feasible because of the difficulty in isolating sufficient number of fetal germ cells for in vivo chromatin analysis (Abramowitz and Bartolomei, 2012; Smallwood and Kelsey, 2012). Here we report an in-depth mapping of CpG methylation, strand-specific transcription and histone covalent modifications in fetal mouse germ cells at the time when global DNA methylation and paternal imprint establishment occurs in prospermatogonia. Our study provides a new perspective on epigenetic remodeling in the male germ line. We suggest that broad, low-level transcription in prospermatogonia serves as general trigger for default de novo DNA methylation, which is prevented only at dynamic peaks of H3K4 methylation. CGIs, DMRs of imprinted genes and IAPs collectively follow these general rules of methylation establishment.

RESULTS

Timing of global de novo methylation in the male germ line

To map dynamic CpG methylation during fetal mouse germ cells development we purified GFP-positive male and female germ cells (MGC and FGC) by flow sorting based on *Pou5f1* (*Oct4*) promoter-driven EGFP expression (Szabó et al., 2002). We collected germ cells from CF1XOG2 embryo and fetus gonads (Figure 1A) as described previously (Lee et al., 2010) at different days of gestation, 11.5, 12.5, 13.5, 15.5 and 17.5 days post coitum (dpc). We isolated DNA from these cells and from control sperm and enriched the methylated fraction by MIRA (Rauch and Pfeifer, 2010). Unlike bisulfite-based methods, MIRA results cannot be confounded by hydroxymethylation (Jin et al., 2010). MIRA-chip is especially well suited for analyzing CpG methylation at CpG-rich genomic regions, such as CpG islands (CGIs) where DNA methylation levels are relevant for gene regulation. We visualized the MIRA signals on a custom NimbleGen tiling array that included all known PAT and MAT DMRs, all known imprinted genes (Williamson et al., 2013), and also selected control regions.

Global dynamics of de novo methylation are illustrated in Figure 1B at a region in chr11 that shows a typical methylation pattern and along the *Hoxc* cluster on chr15 with relatively low-level of DNA methylation. Regions of dense MIRA peaks in sperm and 17.5 dpc MGC aligned with long regions of high methylation in sperm DNA and partial methylation in 16.5 dpc MGC as measured by WGBS (Kobayashi et al., 2012; Seisenberger et al., 2012). Short methylation-free valleys in sperm WGBS aligned with MIRA valleys (Figure 1B). These results collectively suggest that de novo DNA methylation is the default path for fetal MGC DNA and only small valleys escape DNA methylation among highly methylated regions.

Global de novo methylation was lowest at 13.5 dpc, indicating completion of DNA methylation erasure (Figure 1B). Novel MIRA peaks, i.e. those not present at 12.5 dpc, generally did not emerge in FGC between 13.5–17.5 dpc. Sperm-specific MIRA peaks globally emerged in MGC in a short window of time between 15.5 and 17.5 dpc. These results are in agreement with immunostaining data along fetal stages (Abe et al., 2011) and recent whole genome bisulfite sequencing (WGBS) detecting partial global DNA methylation in 16.5 dpc MGC (Kobayashi et al., 2013; Seisenberger et al., 2012). The presence of global de novo methylation in MGC but not FGC is consistent with high transcription levels of *Dnmt3a* and *Dnmt3l* in MGC at 15.5 dpc (Figure 1D). Similarly, any

preexisting mechanism that has a patterning role in de novo DNA methylation must be in place in MGC at 15.5 dpc.

Differentially methylated regions of imprinted genes in the context of global remodeling

The dynamics of de novo methylation at PAT and MAT imprinted DMRs are illustrated in Figure 1C at the *Dlk1-Gtl2* (IG-DMR) and the *Slc38a4* DMR, respectively. The other three known PAT DMRs and three additional MAT DMRs are shown in Figure S1. We confirmed the MIRA-chip results at selected peaks by manual bisulfite sequencing (not shown). We detected MIRA peaks at the MAT and PAT DMRs in FGC and MGC at 11.5 dpc. These peaks disappeared by 13.5 dpc, confirming the erasure of gametic imprints. They remained largely absent in FGC. This is consistent with the lack of global de novo methylation in female germ cells at all fetal stages. MIRA peaks emerged at PAT DMRs in MGC between 15.5–17.5 dpc, confirming the establishment of paternal imprints. Importantly, methylation at PAT DMRs emerged in MGC simultaneously with numerous additional strong methylation peaks along highly methylated regions, suggesting that they follow the default path of global de novo DNA methylation. MAT DMRs, however, appeared as valleys of methylation among emerging MIRA peaks in MGC. This was consistent with an earlier observation obtained by manual bisulfite sequencing analysis: MAT DMRs in sperm DNA resided in small, unmethylated “islands” inside long stretches of highly methylated regions (Tomizawa et al., 2011). Indeed, MAT DMRs appeared similar to the many other small, unmethylated islands in MGC and sperm DNA, suggesting that they may be protected from default global de novo methylation by the same mechanism.

Broad, low-level transcription in fetal male germ cells precedes the wave of de novo DNA methylation

To elucidate what may trigger de novo DNA methylation in prospermatogonia, we performed strand-specific RNA deep sequencing at 15.5 dpc from purified MGC and control FGC. In addition we also purified GFP-negative male somatic cells (MSC) from the fetal testis and GFP-negative female somatic cells (FSC) from the fetal ovary. One representative view is shown for a four-megabase segment from chr9 in Figure 2A. Methylation analysis of MIRA-chip in 17.5 dpc MGC and sperm and WGBS data of sperm (Kobayashi et al., 2012) was aligned with strand-specific RNA-seq results obtained at 15.5 dpc. The FGC, MSC and FSC cells had clear transcriptionally silent intergenic regions between transcripts. MGC, however, displayed low-levels of broad, low-level transcription along the entire segment in one of the strands or both, and no clear gaps were apparent. Two additional segments are shown in Figure S2 from chr12 and chr15. These regions harbor the IG-DMR and *Slc38a4* DMR, respectively. At this magnification there is no difference in the overall methylation pattern or transcription pattern between the two regions: one harboring a PAT and the other harboring a MAT DMR: both are highly methylated in 17.5 dpc MGC and sperm and there appear to be no gaps in the transcription coverage of these chromosomal segments in MGC. The distribution of unique transcript reads slightly shifted toward intergenic regions in MGC (29.4%) as compared to FGC (24.5%), MSC (22.1%) and FSC (26.7%) (Figure 2B and Table S2). To determine the extent of the low-level intergenic transcription in MGC, we used chr7 as an example and plotted the RNA-seq read coverage on all CpG sites, excluding known transcripts. We randomly selected 30 million reads from each sample to make them comparable, then counted the number of CpG sites covered with reads. MGC had much more CpG sites with very low coverage (1–3 reads) than FGC, MSC and FSC, (Figure 2C), but higher coverage did not show difference. 13 % of CpGs were covered in chr7 by at least one read in MGC while 5% of CpGs were covered in FGC, MSC and FSC at the current sequencing depth. Most of the genome is methylated in sperm (Kobayashi et al., 2012), including intergenic regions. This cannot be explained by gene body methylation along known transcripts. However, the extensive transcription we detected in MGC at 15.5 dpc,

even at low-levels may be necessary and/or sufficient to trigger de novo DNA methylation along the entire genome.

Broad, low-level transcription runs through paternally methylated imprinted DMRs in prospermatogonia

Transcription run-through may be required for methylation imprint establishment at PAT DMRs, and it would be expected to occur at 15.5 dpc, when de novo methylation is just beginning in prospermatogonia. Closer inspection of the strand-specific transcription at the four known PAT DMRs (Figure 3) revealed that transcription run-through occurs across each of the PAT DMRs. The IG-DMR (Figure 3A) is traversed by an MGC-specific 5′-extension of *Gtl2* and by antisense transcription (a) in the reverse strand. The composite *Zdbf2* DMR (Figure 3B), consisting of three separate DMRs, is covered by germ cell-specific (MGC and FGC) broad, low-level transcription (b). The *H19-Igf2* DMR (Figure 3C) is crossed by MGC-specific transcription (c) in the forward direction. The *Rasgrf1* DMR (Figure 3D) is traversed by weak broad, low-level transcription in the forward, and MGC-specific transcription (d) in the reverse direction. Our findings at the DMRs are consistent with earlier studies that reported low-level transcription across the IG-DMR and *H19-Igf2* DMR in prospermatogonia (Henckel et al., 2012). We now show that each of the four known PAT DMRs display transcription run-through in MGC at 15.5 dpc. Importantly, we also show that these transcripts are not isolated or special features at PAT DMRs, but are characteristic of fetal MGC along the entire transcriptome. Broad, low-level transcripts in MGC saturate the entire DNA segments at display. These results suggest that the establishment of DNA methylation at PAT DMRs follows a default mechanism, and may be triggered by prospermatogonia-specific genome-wide transcription.

Mono- or bidirectional transcription initiates from maternally methylated DMRs in prospermatogonia

MAT DMRs remain unmethylated in the male germ line, in the same epigenetic environment where most of the genome and the PAT DMRs become methylated. We asked if MAT DMRs are exempt from the broad, low-level transcription in prospermatogonia. We found that transcription run-through occurs at the *Slc38a4* DMR by a weak extended form of the *Slc38a4* transcript in the reverse strand in the germ cells (Figure S3A). The regular initiation site at the DMR may also be active. In addition, a MGC-specific transcript initiates from the DMR in the forward direction (a). The *Grb10* transcript crosses the *Grb10* DMR in the reverse strand. In the forward direction a novel transcript initiates from the *Grb10* promoter in each cell type. This novel transcript appears to have a weak extension, but only in MGC (b) and this initiates from the DMR. The *Gnas* RNA crosses both the *Nespas* and *Gnas*1A DMRs (Figure S3C) in the forward strand while the *Nespas* RNA (c) and the *Gnas1*A RNA (d) initiate from their respective DMRs in the reverse or forward strands, respectively. The *Igf2r* RNA crosses the *Airn* DMR in the reverse direction. This RNA has a 5′ weak extension specifically in MGC. The *Airn* transcript initiates from its DMR in the forward strand. In summary, we found that at MAT DMRs are not exempt from transcription run-through in prospermatogonia. However, transcription initiation occurs from each the MAT DMRs at least in one strand and it often initiates in both strands (See Table S3). These findings are consistent with previous studies that described transcription initiation at two MAT DMRs in prospermatogonia (Henckel et al., 2012) and the protective role of promoters in oocytes (Kobayashi et al., 2012; Smallwood et al., 2011). Because MAT DMRs are not exempt from transcription run-through in prospermatogonia, transcription is not sufficient to trigger methylation of MAT DMRs in this methylation-permissive environment. Therefore, promoter-mediated protection must be dominant over elongation-mediated DNA methylation.

Chromatin analysis in fetal germ cells

To elucidate how chromatin is involved in patterning de novo methylation, we performed ChIP-chip analysis using fetal germ cells. We first optimized the ChIP-chip conditions for small numbers of cells (Lee et al., 2010; Singh and Szabó, 2012; Singh et al., 2011) and determined that chromatin from 400,000 germ cells was necessary per immunoprecipitation to achieve highly reproducible ChIP-chip data. We also found that LM-PCR amplification was much superior to whole genome amplification (WGA) using commercially available kits. We performed ChIP-chip analysis using MGC and FGC at 15.5 dpc with antibodies recognizing H3K4me1, H3K4me2, H3K4me3, H3K9ac, H3K27me3, H3K9me2, H3K9me3, H3K36me3, H3K79me2 and H3K79me3. We also analyzed MSC and FSC at 15.5 dpc using the H3K4me2, H3K9ac, H3K9me2, H3K9me3, H3K27me3, H3K79me2 and H3K79me3 antibodies. Additionally, ChIP-chip was performed at 13.5 dpc for MGC and FGC with the H3K4me2 and H3K9ac antibodies and at 16.5 dpc for MGC with the H3K4me2 antibody. The antibody-captured fraction was LM-PCR- amplified, labeled and hybridized to custom NimbleGen imprinting arrays. Control loci are shown in Figure S4.

Global de novo methylation is patterned by a negative mold of H3K4 methylation

To visualize the effect of active chromatin on global de novo DNA methylation, we aligned sperm DNA methylation with ChIP-chip results of dynamically changing H3K4me2 patterns in MGC at 13.5, 15.5 and 16.5 dpc. We show one region in chr11 that exhibits generally high methylation in sperm and at one region in chr15 that attains little methylation (Figure 4A) according to MIRA-chip and WGBS (Kobayashi et al., 2012) results. In both cases DNA methylation valleys precisely lined up with H3K4me2 peaks that are constitutive or growing between 13.5 and 16.5 dpc, suggesting that H3K4 methylation protects these valleys from default de novo methylation in prospermatogonia.

Repeat elements globally attain high levels of CpG methylation in the germ line. Repeats are excluded from microarrays, but we included 1 kilobase of unique flanking sequences for each IAP in our custom microarray expecting that the antibodies or MIRA proteins would enrich for the IAP regions and the specific flanking sequences would hybridize to the microarray. A heat map along the IAP- flanking regions depicts the relationship between H3K4me2 and DNA methylation (Figure 4B). IAPs with the highest level of H3K4me2 at 16.5 dpc exhibited the lowest level of DNA methylation levels in 16.5 dpc MGC as measured by WGBS (Seisenberger et al., 2012). Whereas most IAPs are 100% methylated in sperm, a small fraction of IAPs remain partially methylated (Kobayashi et al., 2012). This small fraction of partially methylated IAPs had the highest level of H3K4me2 in MGC at 16.5 dpc (Figure 4B) suggesting that active chromatin in prospermatogonia may protect certain IAP repeats from full de novo methylation.

The majority of the sperm genome is highly methylated. Important exemptions to this rule are CGIs, most of which remain unmethylated in sperm. However, a number of CGIs are methylated in the sperm or in the oocyte or in both gametes (Kobayashi et al., 2012; Smallwood et al., 2011). Our microarray contains a representative number of CGIs in the different methylation classes (Table S4). To elucidate the significance of H3K4 methylation in de novo methylation patterns at CGIs, we crosschecked the MIRA methylation profile of CGI methylation classes in FGC and MGC (Figure 4C) with their H3K4me1, H3K4me2 and H3K4me3 ChIP-chip profiles (Figure 4D). A strong MIRA methylation peak was detected in 17.5 dpc MGC and sperm at the sperm-methylated CGI class (80–100% methylation in sperm and 0–80% methylation in oocyte) (Figure 4C). This peak did not correspond to a strong H3K4me peak (Figure 4D). A small MIRA peak was detected in 11.5 dpc FGC and MGC at the oocyte-methylated CGIs (80–100% methylated in oocytes and 0–80% methylated in sperm). The MIRA peak diminished around 12.5–13.5 dpc FGC and MGC

and became a valley in 15.5–17.5 dpc MGC, but returned to become a small peak in 15.5 and 17.5 dpc FGC. This is consistent with the erasure of global DNA methylation by 13.5 dpc (Kobayashi et al., 2013; Seisenberger et al., 2012) and supports the novel finding that weak initiation of de novo methylation occurs in fetal FGC (Kobayashi et al., 2013). At the unmethylated class of CGIs (0–80% CpG methylation in sperm and oocyte), the MIRA peaks were almost nonexistent (Figure 4C). Both the oocyte-methylated and the unmethylated CGIs precisely aligned with sharp peaks of each H3K4 methylation mark, H3K4me2 being the most profound (Figure 4D).

A heat map of 15.5 dpc H3K4me2 and sperm bisulfite methylation (Kobayashi et al., 2012) provides a higher resolution analysis at CGI methylation categories (See Table S4) in Figure 4E. CGIs in the oocyte-methylated gDMR and the unmethylated categories (each with 0–20% methylation sperm) always corresponded to a strong enrichment of H3K4me2. However, CGIs in the sperm-methylated categories and the oocyte- and sperm methylated class (each with 8–100% methylation in sperm) displayed very little H3K4me2 enrichment. These results suggest that de novo DNA methylation of CGIs is inversely patterned by H3K4 methylation in prospermatogonia.

The H3K4me2 peaks underwent dynamic changes in MGC between 13.5 and 16.5 dpc (Figure 4A). To reveal if these changes affect de novo DNA methylation, we first determined which H3K4me2 peaks increased or decreased by at least three-fold between 13.5 and 16.5 dpc and then plotted H3K4me2 levels and the corresponding DNA methylation levels in 16.5 MGC and sperm at these peaks with 5 kilobase flanking regions (Figure 4F). We found that H3K4me2 peaks that emerged between 13.5 and 16.5 dpc corresponded to a lower level of DNA methylation in 16.5 dpc MGC and a clear window of unmethylated DNA in sperm among highly methylated DNA. This was not true for the diminishing H3K4me2 peaks. In summary, dynamically growing but not diminishing peaks of H3K4 methylation protect the underlying sequences from de novo methylation in prospermatogonia.

H3K4 methylation peaks distinguish maternally versus paternally methylated imprinted DMRs in prospermatogonia

To elucidate the patterning effect of active chromatin at imprinted DMRs of imprinted genes, we aligned the H3K4me2 peaks in 13.5, 15.5 and 16.5 dpc MGC and in control 13.5 and 15.5 dpc FGC with the MIRA methylation peaks and sperm bisulfite data. At the *H19-Igf2* and *Dlk1-Gtl2* PAT DMRs, 17.5 dpc MIRA peaks overlapped with H3K4me2 peaks at 13.5 dpc and these chromatin peaks diminished in MGC between 13.5 and 16.5 dpc, but did not diminish in FGC between 13.5 dpc and 15.5 dpc (Figure 5A and S5A). The *Rasgrf1* DMR and the three peaks of the *Zdbf2* composite DMR precisely resided in valleys between H3K4me2 peaks (Figure S5A). DNA methylation peaks of PAT DMRs precisely resided in H3K4me2 valleys at constitutive or diminishing H3K4me2 peaks in prospermatogonia. However, each MAT DMR (Figure 5A) precisely resided in DNA methylation valley and displayed a strong peak of H3K4me2, which remained high in MGC and FGC between 13.5 and 15.5 dpc (Figure 5A and S5B).

Gamete-specific methylation is not erased at imprinted gDMRs during the global wave of epigenetic erasure in the zygote/early embryo. CGIs inside known MAT and PAT DMRs belong to the oocyte-methylated resistant and the sperm-methylated resistant gDMR CGI category, respectively, each maintaining 20–100% methylation in blastocysts (Kobayashi et al., 2012) (Table S4). We analyzed the dynamic DNA methylation and H3K4 profiles of the “resistant” CGIs. Composite profiles revealed a strong peak of H3K4me2 at oocyte-methylated resistant and non-resistant gDMR CGIs in fetal germ cells, suggesting a role of this mark in protecting them from DNA methylation. However, H3K4me2 peak was not

detected at sperm-methylated gDMR CGIs in MGC between 13.5–16.5 dpc, where strong MIRA peaks emerged between 15.5 dpc and 17.5 dpc. At sperm methylated resistant gDMRs a H3K4me2 peak was detected in MGC and FGC at 13.5 dpc and this greatly decreased between 13.5 and 15.5 dpc, followed by the emergence of a very strong MIRA peak between 15.5 and 17.5 dpc. This finding supports the hypothesis that H3K4 has to be actively demethylated in prospermatogonia at some sperm-methylated resistant gDMRs to allow for de novo DNA methylation.

Fetal MGC and FGC have their own arsenal of epigenetic modifiers, providing very different epigenetic environment for these cell types (Lefevre and Mann, 2008). To search for candidate epigenetic modifiers that may pattern de novo DNA methylation in MGC we analyzed our RNA-seq data obtained at 15.5 dpc (Table S1). We found high transcript levels in MGC for the *Kdm1a*, *Kdm1b* and *Kdm5a* genes that encode H3K4 demethylases. One or more of these enzymes may play a role in facilitating CpG methylation in prospermatogonia by removing H3K4me2 peaks (Figure 4F) for example at the imprinted PAT DMRs (Figure 5A and C). In addition, H3K4 methyltransferases must play an important role in MGC, by building up the emerging H3K4me peaks (Figure 4F). *Ash2l* and *Mll3* are likely candidates for this function, because these genes are highly transcribed in MGC at 15.5 dpc (Figure 5D).

De novo methylation occurs by default without clues from repressive chromatin

We considered the possibility that reciprocally to active chromatin, repressive chromatin might trigger de novo methylation in prospermatogonia. SUV39H1-dependent H3K9 methylation is required for DNA methylation at least at the pericentric heterochromatin regions (Lehnertz et al., 2003). H3K27me3-dependent repression may become solidified by subsequent CpG methylation (Gal-Yam et al., 2008). Whereas H3K79me2 is biased toward the unmethylated allele of imprinted DMRs, H3K79me3 is biased toward the DNA-methylated allele (Singh et al., 2010). H3K36me3 and DNA methylation coexist along gene bodies and H3K36me3 may guide DNA methylation by recruiting DNMT3A to gene bodies via its PWWP domain (Dhayalan et al., 2010). We investigated the patterns of H3K9me2, H3K9me3, H3K27me3, H3K79me3 and H3K36me3 in the fetal gonad at 15.5 dpc by ChIP-chip. We derived composite profiles from the ChIP-chip data at the different CGI methylation classes (Table S4) +/−10 kb (Figure 6). CGIs in the unmethylated and oocyte-methylated classes (0–80% DNA methylation in sperm) had weak H3K79me2 and H3K27me3 peaks in fetal germ cells, but were slightly depleted in H3K79me3 and strongly depleted in H3K36me3. The control active chromatin marks H3K4me2 and H3K9ac had strong peaks at these CGIs. We detected a slight H3K27me3, H3K9me3 and H3K79me3 enrichment at the oocyte-methylated class in the control somatic cells, FSC and MSC. Importantly, the sperm-methylated CGIs class (80–100% methylation in sperm) did not display any of the repressive chromatin marks in MGC or in the other cell types (Figure 6). These results suggest that de novo methylation of CGIs in prospermatogonia does not require clues from repressive chromatin marks.

We further investigated the pattern of chromatin marks at the CGI categories (Table S4) into which the known imprinted DMRs belong (Figure S6A). Even though we observed strong peaks for H3K79me3 and H3K9me3 at resistant sperm gDMR CGIs in FSC and MSC, these were missing in FGC and MGC. Resistant oocyte gDMRs were depleted in H3K79me3 and H3K36me3 in fetal germ cells. Examples of ChIP-chip results are shown in Figure S6B and C. We found that in FGC and MGC imprinted DMRs lacked the H3K9me3 and H3K79me3 peaks that were present in FSC and MSC. Strong H3K9me3 peaks did occur in germ cells at other sequences, for example at the *Chrac1* gene (Figure S6B). PAT DMRs lacked H3K36me3 and the surrounding sequences also showed very weak association with this mark. This is likely due to the relatively low-level of transcription across these sequences.

MAT DMRs displayed a valley in the H3K36me3 mark (Figures S6B) similarly to the unmethylated and oocyte-methylated CGIs (Figure 6). The H3K79me2 peaks that occurred at DMRs were weaker in germ cells than in somatic cells. MAT DMRs exhibited slightly more H3K79me2 in FGC than in MGC (Figure S6C). Unlike H3K4me2, the active H3K79me2 mark is, therefore, less likely to protect MAT DMRs from de novo methylation in fetal MGC. These results suggest that H3K9me2, H3K9me3, H3K79me3, H3K27me3 or H3K36me3 histone marks do not trigger de novo methylation of PAT DMRs in fetal MGC.

DISCUSSION

De novo DNA methylation occurs by default in prospermatogonia

We found that sperm-specific DNA methylation broadly emerges in prospermatogonia in the very short window of time, between 15.5 and 17.5 dpc. Based on our chromatin mapping results, repressive chromatin is not involved in this process. However, broad, low-level transcription precedes the wave of remethylation at 15.5 dpc along the chromosomes including intergenic regions and it may serve as a trigger or facilitator for global de novo DNA methylation in the male germline. Similarly, transcription along gene bodies correlates with high level of DNA methylation (Rauch et al., 2009) and de novo methylation occurs along gene bodies in the growing oocyte (Kobayashi et al., 2012; Smallwood et al., 2011). DNMTs are linked to transcribed loci (Jin et al., 2012) during de novo DNA methylation in human tumor cells. Our results provide the framework for testing the role of broad low-level intergenic transcripts in male germline de novo DNA methylation. We suggest that global de novo DNA methylation occurs in prospermatogonia by default in response to broad, low-level transcription, with its pattern being dictated by a reciprocal mold of H3K4 methylation (Figure 7).

We showed that each of the four known imprinted PAT DMRs are traversed by MGC-specific weak transcripts in the male germline. It will be important to genetically test the role of weak MGC-specific transcripts in the establishment of DNA methylation at the PAT DMRs. Indeed, at the *Gnas-Nespas* imprinted domain, transcription run-through is required for de novo methylation of the domain's MAT DMRs in growing oocytes (Chotalia et al., 2009) and the human SNRPN DMR requires transcription run through in the mouse oocyte in a transgenic experimental setting (Smith et al., 2011). In addition, our data revealed that broad, low-level transcripts are general features of MGC at 15.5 dpc, suggesting that the deposition of CpG methylation at a PAT DMRs in response to those transcripts follows the default path in prospermatogonia rather than being a specifically targeted process.

H3K4 methylation peaks exclude de novo DNA methylation in prospermatogonia

The methylation valleys among highly methylated sperm DNA precisely aligned with sharp peaks of H3K4me2 in prospermatogonia, suggesting that default de novo DNA methylation is prevented at these sequences by H3K4 methylation. This protection can be explained by the fact that methylation of H3K4 blocks the access of the ADD domains of both DNMT3A and DNMT3B/3L complex to histone H3 tails (Ooi et al., 2007; Zhang et al., 2010). While CGIs in the unmethylated and oocyte-methylated classes were highly enriched in H3K4 methylation, they were strongly depleted in H3K36me3. This is expected, because the major H3K36 demethylase, KDM2A is specifically targeted to unmethylated CGIs by its CXXC domain, resulting in H3K36 demethylation (Blackledge et al., 2010).

IAPs repeat elements are known to retain residual CpG methylation in PGCs during germ line epigenetic remodeling and can mediate transgenerational epigenetic inheritance at metastable epialleles (Guibert et al., 2012; Lane et al., 2003; Morgan et al., 1999; Seisenberger et al., 2012). Whereas the silent state of the epiallele corresponds to full

methylation, the penetrant state corresponds to hypomethylation of the IAP-LTR (Rakyan et al., 2003). If the phenotype depends on DNA methylation differences, in order to transmit the penetrant state sperm DNA has to be hypomethylated at the IAP and it must escape the default de novo methylation process in prospermatogonia. Our data suggests that active chromatin provides partial protection to certain IAP repeats in the male germ line. We propose that H3K4 methylation-dependent partial protection of IAPs in the germ line from default de novo methylation may be an essential component of transgenerational epigenetic inheritance.

We revealed that H3K4me2 undergoes remodeling in MGC between 13.5–16.5 dpc. Importantly, dynamically growing H3K4 methylation peaks protect the underlying sequences in prospermatogonia from default de novo DNA methylation while other H3K4me peaks are actively removed to allow DNA methylation. These changes must entail the dynamic action of histone methyltransferases and demethylases. Our results are consistent with a previous study that proved the importance of KDM1B H3K4 demethylase in the establishment of de novo DNA methylation at four out of seven MAT DMRs in the growing oocyte (Ciccone et al., 2009). It is also supported by a genetic study, where we showed that the erasure of active chromatin in PGCs lags behind the erasure of DNA methylation at the *H19-Igf2* PAT DMR and delays de novo methylation in the maternally inherited allele in prospermatogonia (Lee et al., 2010). It is likely that dynamic changes in H3K4 methylation play similar patterning role in de novo methylation of growing oocytes between day 5 and day 20. Indeed, H3K4me3 enrichment inversely correlated with CGI methylation in day 15 oocytes that are on their way to global methylation (Smallwood et al., 2011). Conditional inactivation will be necessary to reveal the exact roles and specificities of epigenetic modifiers, including KDM1A and KDM1B histone demethylases and the ASH2L and MLL3 histone methyltransferases in patterning de novo DNA methylation globally and at specific genomic sites, such as CGIs, IAPs and imprinted DMRs in prospermatogonia.

Differential protection from global de novo methylation and global demethylation are key to genomic imprinting

It has been a long quest to determine what cis-elements and trans-factors target gametic DNA methylation differences to DMRs of imprinted genes. In order to generate a gDMR, one germ line has to fully methylate the same sequence, which the reciprocal germ line has to keep unmethylated. We found that PAT DMRs follow the default path of de novo methylation in prospermatogonia. In the light of our results, it is perhaps not so much the methylation targeting in prospermatogonia that specifies a sperm-methylated gDMR, but rather keeping the same sequences unmethylated in growing oocytes. Similarly, de novo methylation of MAT DMRs may be viewed as part of the general processes in growing oocytes (Kelsey and Feil, 2013). Indeed, we showed here that each known MAT DMR is specifically kept unmethylated in prospermatogonia by precisely aligned sharp peaks of H3K4me2 together with transcription initiation.

The second important key to genomic imprints is the maintenance of gametic DNA methylation at imprinted DMRs during the wave of global DNA demethylation in the zygote and early embryo. For example, PGC7 protects methylation imprints at certain MAT DMRs and the *H19-Igf2* and *Rasgrf1* PAT DMRs in the zygote (Nakamura et al., 2007; Nakamura et al., 2012) by organizing TET3 dioxygenase-repellent higher-order chromatin. ZFP57 transcription factor is required for the maintenance of methylation imprints at certain MAT DMRs and the PAT IG-DMR, during early development (Li et al., 2008). Maternal expression of an epigenetic modifier, TRIM28, is required for maintaining genomic imprints, including methylation of the *H19-Igf2* DMR at the time of oocyte-embryo transition (Messerschmidt et al., 2012). In the light of these results, we suggest that the

establishment of genomic imprints is the result of two differential protection mechanisms: first, protection from default de novo DNA methylation during gametogenesis and second, protection from default DNA demethylation after fertilization.

EXPERIMENTAL PROCEDURES

The experiments involving mice had been approved by the IACUC of the City of Hope. Housing and care of the animals has been consistent with the Public Health Service Policy, the NIH "Guide for the Care and Use of Laboratory Animals" and the Animal Welfare Act.

Purification of germ cells

GFP-positive germ cells and GFP-negative gonadal somatic cells were flow sorted based on *Pou5f1* promoter-driven EGFP expression in germ cells (Szabó et al., 2002) from CF1XOG2 embryo and fetus gonads as described previously (Lee et al., 2010) using a MoFlo or Aria II flow cytometer. Germ cell purity was 96–100% based on double staining with DDX4 antibody. Motile spermatozoa were extracted from the cauda epididymis of 8 weeks old CF1XOG2 male.

Methylated CpG island recovery assay (MIRA) and MIRA-chip

The methylated fraction of sonicated GFP-positive germ cell genomic DNA was captured using recombinant MBD2b and MBD3L1 proteins as described earlier (Rauch and Pfeifer, 2010). Ligation-mediated PCR (LM-PCR) was performed to amplify ChIP- and MIRA-enriched DNA as previously described (Kim et al., 2007) with minor modifications (Singh et al., 2011).

RNA isolation and RNA-Seq

15.5 dpc gonadal germ cells or somatic cells were sorted as described. The cell pellets (~10^6 cells) were frozen in 100 μl RNA Bee (Tel Test Inc.) before isolating total RNA as per the manufacturer's instructions. Contaminating DNA was removed with the DNA*free* Kit (Ambion). The amount of RNA was monitored using a Nanodrop spectrophotometer followed by a quality check using the Bioanalyzer. All RNA samples had a RIN>9.0. 2μg total RNA samples were depleted for rRNA using the RiboZero Kit (Epicenter, Illumina). Paired end sequencing was performed at the UCLA Microarray core using the IntegenX RNA kit (Illumina) with read length/coverage of 2x100.

Chromatin Immunoprecipitation and ChIP-chip

For chromatin immunoprecipitation, cells were crosslinked before sorting. ChIP was performed as described previously (Lee et al., 2010; Singh and Szabó, 2012). Chromatin from 400,000 cells was used for one ChIP using various antibodies listed in the Extended Experimental Procedures. Custom-designed tiling arrays (110228_MM9_PS_ChIP), manufactured by Roche/NimbleGen, were used for the histone modification profile analysis. Amplified ChIP DNA fractions were compared with amplified input DNA. Data were extracted from scanned images by using NimbleScan 2.3 extraction software (NimbleGen Systems). Bioinformatics analysis is described in the Extended Experimental Procedures.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are grateful to Alexander Spalla and Lucy Brown (Analytical Cytometry Core, City of Hope) for FACS sorting of germ cells and the Animal Resource Centre of the City of Hope for providing mouse care. We thank Xinmin Li, Jamie Zhou and Jose Solis (UCLA Microarray Core) for Nimblegen array hybridizations and RNA deep sequencing. We thank Dong-Hoon Lee for purified MBD3L1 and MBD2b proteins, Angela Bai (Eugene and Ruth Roberts Academy) for assistance in bisulfite analysis and Raymund Stefancsik for tabulating the IAP flanking regions. We thank Wolf Reik and Felix Krueger for providing the fetal WGBS analysis for comparison. We thank Gerd Pfeifer for critically reading the manuscript. This work was supported by grant RO1GM064378 from the NIH and an Excellence Award from the City of Hope to PES.

References

Abe M, Tsai SY, Jin SG, Pfeifer GP, Szabó PE. Sex-specific dynamics of global chromatin changes in fetal mouse germ cells. PLoS One. 2011; 6:e23848. [PubMed: 21886830]

Abramowitz LK, Bartolomei MS. Genomic imprinting: recognition and marking of imprinted loci. Curr Opin Genet Dev. 2012; 22:72–78. [PubMed: 22195775]

Blackledge NP, Zhou JC, Tolstorukov MY, Farcas AM, Park PJ, Klose RJ. CpG islands recruit a histone H3 lysine 36 demethylase. Mol Cell. 2010; 38:179–190. [PubMed: 20417597]

Borgel J, Guibert S, Li Y, Chiba H, Schubeler D, Sasaki H, Forne T, Weber M. Targets and dynamics of promoter DNA methylation during early mouse development. Nat Genet. 2010; 42:1093–1100. [PubMed: 21057502]

Bourc'his D, Xu GL, Lin CS, Bollman B, Bestor TH. Dnmt3L and the establishment of maternal genomic imprints. Science. 2001; 294:2536–2539. [PubMed: 11719692]

Cheng X, Blumenthal RM. Coordinated chromatin control: structural and functional linkage of DNA and histone methylation. Biochemistry. 2010; 49:2999–3008. [PubMed: 20210320]

Chotalia M, Smallwood SA, Ruf N, Dawson C, Lucifero D, Frontera M, James K, Dean W, Kelsey G. Transcription is required for establishment of germline methylation marks at imprinted genes. Genes & development. 2009; 23:105–117. [PubMed: 19136628]

Ciccone DN, Su H, Hevi S, Gay F, Lei H, Bajko J, Xu G, Li E, Chen T. KDM1B is a histone H3K4 demethylase required to establish maternal genomic imprints. Nature. 2009; 461:415–418. [PubMed: 19727073]

Dhayalan A, Rajavelu A, Rathert P, Tamas R, Jurkowska RZ, Ragozin S, Jeltsch A. The Dnmt3a PWWP domain reads histone 3 lysine 36 trimethylation and guides DNA methylation. J Biol Chem. 2010; 285:26114–26120. [PubMed: 20547484]

Ferguson-Smith AC. Genomic imprinting: the emergence of an epigenetic paradigm. Nat Rev Genet. 2011; 12:565–575. [PubMed: 21765458]

Gal-Yam EN, Egger G, Iniguez L, Holster H, Einarsson S, Zhang X, Lin JC, Liang G, Jones PA, Tanay A. Frequent switching of Polycomb repressive marks and DNA hypermethylation in the PC3 prostate cancer cell line. Proc Natl Acad Sci U S A. 2008; 105:12979–12984. [PubMed: 18753622]

Guibert S, Forne T, Weber M. Global profiling of DNA methylation erasure in mouse primordial germ cells. Genome Res. 2012; 22:633–641. [PubMed: 22357612]

Hajkova P, Erhardt S, Lane N, Haaf T, El-Maarri O, Reik W, Walter J, Surani MA. Epigenetic reprogramming in mouse primordial germ cells. Mech Dev. 2002; 117:15–23. [PubMed: 12204247]

Hata K, Okano M, Lei H, Li E. Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. Development. 2002; 129:1983–1993. [PubMed: 11934864]

Henckel A, Chebli K, Kota SK, Arnaud P, Feil R. Transcription and histone methylation changes correlate with imprint acquisition in male germ cells. EMBO J. 2012; 31:606–615. [PubMed: 22117218]

Hiura H, Obata Y, Komiyama J, Shirai M, Kono T. Oocyte growth-dependent progression of maternal imprinting in mice. Genes Cells. 2006; 11:353–361. [PubMed: 16611239]

Hiura H, Sugawara A, Ogawa H, John RM, Miyauchi N, Miyanari Y, Horiike T, Li Y, Yaegashi N, Sasaki H, et al. A tripartite paternally methylated region within the Gpr1-Zdbf2 imprinted domain on mouse chromosome 1 identified by meDIP-on-chip. Nucleic Acids Res. 2010; 38:4929–4945. [PubMed: 20385583]

Jin B, Ernst J, Tiedemann RL, Xu H, Sureshchandra S, Kellis M, Dalton S, Liu C, Choi JH, Robertson KD. Linking DNA methyltransferases to epigenetic marks and nucleosome structure genome-wide in human tumor cells. Cell Rep. 2012; 2:1411–1424. [PubMed: 23177624]

Jin SG, Kadam S, Pfeifer GP. Examination of the specificity of DNA methylation profiling techniques towards 5-methylcytosine and 5-hydroxymethylcytosine. Nucleic Acids Res. 2010; 38:e125. [PubMed: 20371518]

Kaneda M, Okano M, Hata K, Sado T, Tsujimoto N, Li E, Sasaki H. Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. Nature. 2004; 429:900–903. [PubMed: 15215868]

Kato Y, Kaneda M, Hata K, Kumaki K, Hisano M, Kohara Y, Okano M, Li E, Nozaki M, Sasaki H. Role of the Dnmt3 family in de novo methylation of imprinted and repetitive sequences during male germ cell development in the mouse. Hum Mol Genet. 2007; 16:2272–2280. [PubMed: 17616512]

Kelsey G, Feil R. New insights into establishment and maintenance of DNA methylation imprints in mammals. Philos Trans R Soc Lond B Biol Sci. 2013; 368:20110336. [PubMed: 23166397]

Kim TH, Barrera LO, Ren B. ChIP-chip for genome-wide analysis of protein binding in mammalian cells. Curr Protoc Mol Biol. 2007; Chapter 21(Unit 21):13. [PubMed: 18265397]

Kobayashi H, Sakurai T, Imai M, Takahashi N, Fukuda A, Yayoi O, Sato S, Nakabayashi K, Hata K, Sotomaru Y, et al. Contribution of intragenic DNA methylation in mouse gametic DNA methylomes to establish oocyte-specific heritable marks. PLoS Genet. 2012; 8:e1002440. [PubMed: 22242016]

Kobayashi H, Sakurai T, Miura F, Imai M, Mochiduki K, Yanagisawa E, Sakashita A, Wakai T, Suzuki Y, Ito T, et al. High-resolution DNA methylome analysis of primordial germ cells identifies gender-specific reprogramming in mice. Genome Res. 2013

Lane N, Dean W, Erhardt S, Hajkova P, Surani A, Walter J, Reik W. Resistance of IAPs to methylation reprogramming may provide a mechanism for epigenetic inheritance in the mouse. Genesis. 2003; 35:88–93. [PubMed: 12533790]

Lee DH, Singh P, Tsai SY, Oates N, Spalla A, Spalla C, Brown L, Rivas G, Larson G, Rauch TA, et al. CTCF-dependent chromatin bias constitutes transient epigenetic memory of the mother at the H19-Igf2 imprinting control region in prospermatogonia. PLoS Genet. 2010; 6:e1001224. [PubMed: 21124827]

Lefevre C, Mann JR. RNA expression microarray analysis in mouse prospermatogonia: identification of candidate epigenetic modifiers. Dev Dyn. 2008; 237:1082–1089. [PubMed: 18330932]

Lehnertz B, Ueda Y, Derijck AA, Braunschweig U, Perez-Burgos L, Kubicek S, Chen T, Li E, Jenuwein T, Peters AH. Suv39h-mediated histone h3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. Curr Biol. 2003; 13:1192–1200. [PubMed: 12867029]

Li X, Ito M, Zhou F, Youngson N, Zuo X, Leder P, Ferguson-Smith AC. A maternal-zygotic effect gene, Zfp57, maintains both maternal and paternal imprints. Dev Cell. 2008; 15:547–557. [PubMed: 18854139]

Lucifero D, Mann MR, Bartolomei MS, Trasler JM. Gene-specific timing and epigenetic memory in oocyte imprinting. Hum Mol Genet. 2004; 13:839–849. [PubMed: 14998934]

Messerschmidt DM, de Vries W, Ito M, Solter D, Ferguson-Smith A, Knowles BB. Trim28 is required for epigenetic stability during mouse oocyte to embryo transition. Science. 2012; 335:1499–1502. [PubMed: 22442485]

Morgan HD, Sutherland HG, Martin DI, Whitelaw E. Epigenetic inheritance at the agouti locus in the mouse. Nat Genet. 1999; 23:314–318. [PubMed: 10545949]

Nakamura T, Arai Y, Umehara H, Masuhara M, Kimura T, Taniguchi H, Sekimoto T, Ikawa M, Yoneda Y, Okabe M, et al. PGC7/Stella protects against DNA demethylation in early embryogenesis. Nat Cell Biol. 2007; 9:64–71. [PubMed: 17143267]

Nakamura T, Liu YJ, Nakashima H, Umehara H, Inoue K, Matoba S, Tachibana M, Ogura A, Shinkai Y, Nakano T. PGC7 binds histone H3K9me2 to protect against conversion of 5mC to 5hmC in early embryos. Nature. 2012; 486:415–419. [PubMed: 22722204]

Okano M, Bell DW, Haber DA, Li E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell. 1999; 99:247–257. [PubMed: 10555141]

Ooi SK, Qiu C, Bernstein E, Li K, Jia D, Yang Z, Erdjument-Bromage H, Tempst P, Lin SP, Allis CD, et al. DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. Nature. 2007; 448:714–717. [PubMed: 17687327]

Rakyan VK, Chong S, Champ ME, Cuthbert PC, Morgan HD, Luu KV, Whitelaw E. Transgenerational inheritance of epigenetic states at the murine Axin(Fu) allele occurs after maternal and paternal transmission. Proc Natl Acad Sci U S A. 2003; 100:2538–2543. [PubMed: 12601169]

Rauch TA, Pfeifer GP. DNA methylation profiling using the methylated-CpG island recovery assay (MIRA). Methods. 2010; 52:213–217. [PubMed: 20304072]

Rauch TA, Wu X, Zhong X, Riggs AD, Pfeifer GP. A human B cell methylome at 100-base pair resolution. Proc Natl Acad Sci U S A. 2009; 106:671–678. [PubMed: 19139413]

Reik W, Dean W, Walter J. Epigenetic reprogramming in mammalian development. Science. 2001; 293:1089–1093. [PubMed: 11498579]

Sasaki H, Matsui Y. Epigenetic events in mammalian germ-cell development: reprogramming and beyond. Nat Rev Genet. 2008; 9:129–140. [PubMed: 18197165]

Seisenberger S, Andrews S, Krueger F, Arand J, Walter J, Santos F, Popp C, Thienpont B, Dean W, Reik W. The dynamics of genome-wide DNA methylation reprogramming in mouse primordial germ cells. Mol Cell. 2012; 48:849–862. [PubMed: 23219530]

Seisenberger S, Peat JR, Hore TA, Santos F, Dean W, Reik W. Reprogramming DNA methylation in the mammalian life cycle: building and breaking epigenetic barriers. Philos Trans R Soc Lond B Biol Sci. 2013; 368:20110330. [PubMed: 23166394]

Singh P, Han L, Rivas GE, Lee DH, Nicholson TB, Larson GP, Chen T, Szabó PE. Allele-specific H3K79 Di- versus trimethylation distinguishes opposite parental alleles at imprinted regions. Mol Cell Biol. 2010; 30:2693–2707. [PubMed: 20351169]

Singh P, Szabó PE. Chromatin immunoprecipitation to characterize the epigenetic profiles of imprinted domains. Methods Mol Biol. 2012; 925:159–172. [PubMed: 22907496]

Singh P, Wu X, Lee DH, Li AX, Rauch TA, Pfeifer GP, Mann JR, Szabó PE. Chromosome-wide analysis of parental allele-specific chromatin and DNA methylation. Mol Cell Biol. 2011; 31:1757–1770. [PubMed: 21321082]

Smallwood SA, Kelsey G. De novo DNA methylation: a germ cell perspective. Trends Genet. 2012; 28:33–42. [PubMed: 22019337]

Smallwood SA, Tomizawa S, Krueger F, Ruf N, Carli N, Segonds-Pichon A, Sato S, Hata K, Andrews SR, Kelsey G. Dynamic CpG island methylation landscape in oocytes and preimplantation embryos. Nat Genet. 2011; 43:811–814. [PubMed: 21706000]

Smith EY, Futtner CR, Chamberlain SJ, Johnstone KA, Resnick JL. Transcription is required to establish maternal imprinting at the Prader-Willi syndrome and Angelman syndrome locus. PLoS Genet. 2011; 7:e1002422. [PubMed: 22242001]

Smith ZD, Chan MM, Mikkelsen TS, Gu H, Gnirke A, Regev A, Meissner A. A unique regulatory phase of DNA methylation in the early mammalian embryo. Nature. 2012; 484:339–344. [PubMed: 22456710]

Szabó PE, Hubner K, Scholer H, Mann JR. Allele-specific expression of imprinted genes in mouse migratory primordial germ cells. Mech Dev. 2002; 115:157–160. [PubMed: 12049782]

Tomizawa S, Kobayashi H, Watanabe T, Andrews S, Hata K, Kelsey G, Sasaki H. Dynamic stage-specific changes in imprinted differentially methylated regions during early mammalian development and prevalence of non-CpG methylation in oocytes. Development. 2011; 138:811–820. [PubMed: 21247965]

Williamson, CM.; Blake, A.; Thomas, S.; Beechey, CV.; Hancock, J.; Cattanach, BM.; Peters, J. MRC Harwell; Oxfordshire: 2013. World Wide Web Site - Mouse Imprinting Data and References - http://www.har.mrc.ac.uk/research/genomic_imprinting/

Yoshioka H, McCarrey JR, Yamazaki Y. Dynamic nuclear organization of constitutive heterochromatin during fetal male germ cell development in mice. Biol Reprod. 2009; 80:804–812. [PubMed: 19129513]

Zhang Y, Jurkowska R, Soeroes S, Rajavelu A, Dhayalan A, Bock I, Rathert P, Brandt O, Reinhardt R, Fischle W, et al. Chromatin methylation activity of Dnmt3a and Dnmt3a/3L is guided by interaction of the ADD domain with the histone H3 tail. Nucleic Acids Res. 2010; 38:4246–4253. [PubMed: 20223770]

HIGHLIGHTS

1. Germ cell de novo DNA methylation occurs without clues from repressive chromatin
2. Methylation is preceded by low-level genome-wide transcription in prospermatogonia
3. Dynamic H3K4 methylation is the “negative molding form” for global DNA methylation
4. CGIs, gDMRs, IAPs follow the default pattern: unmethylated if protected by H3K4me2

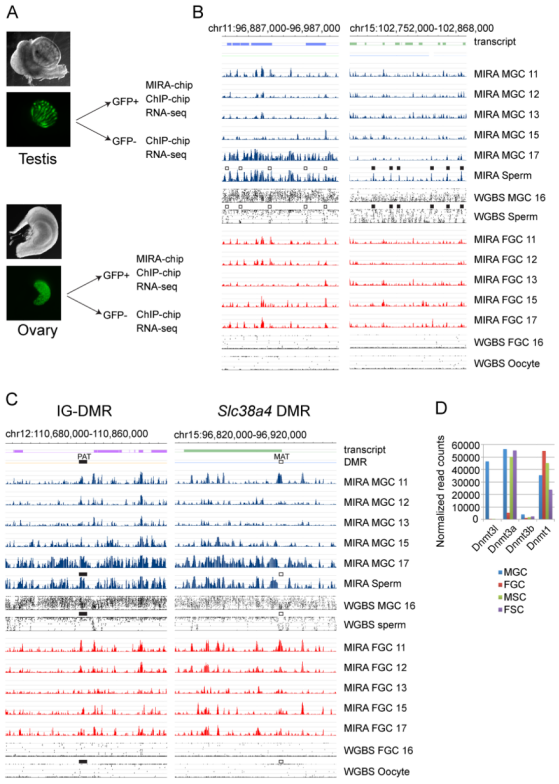

Figure 1. Global DNA methylation erasure is complete by 13.5 dpc and remethylation occurs by default between 15 and 17 dpc in prospermatogonia

(A) Experimental system. Male and female GFP+ germ cells (MGC and FGC) and GFP- somatic cells (MSC and FSC) were purified by flow sorting from fetal testes and ovaries. (B) One highly methylated region depicts the general methylation pattern to the left. One region with relatively less methylation, the *Hoxc* cluster depicts an exemption to the right. (C) One paternally (PAT) and one maternally (MAT) methylated DMR is shown. (B–C) The DNA methylation signals of MIRA versus input DNA were plotted as $-\log_{10}$ p-value scores ranging from 0–8.4 for FGC (red) and MGC (blue) at gestational stages 11.5–17.5 (11–17) dpc, and in sperm DNA as indicated to the right. The average % methylation levels at each CpG as determined by whole genome bisulfite sequencing (WGBS) are shown in comparison from MGC and FGC at 16.5 dpc (Seisenberger et al., 2012), sperm, and oocyte (Kobayashi et al., 2012). The location of transcripts and known DMRs is indicated. Note that DNA methylation valleys among highly methylated regions (open squares) and the MIRA peaks among poorly methylated regions (closed squares) precisely line up between MIRA and WGBS. PAT DMRs (closed rectangle) are found in one of the emerging MIRA peaks and the MAT DMRs (open rectangle) reside in a methylation valley inside broadly methylated regions in MGC. See additional DMRs in Figure S1 (D) Transcription levels of *Dnmts* in the cells of the fetal gonad. RNA-seq results are shown. See also Table S1.

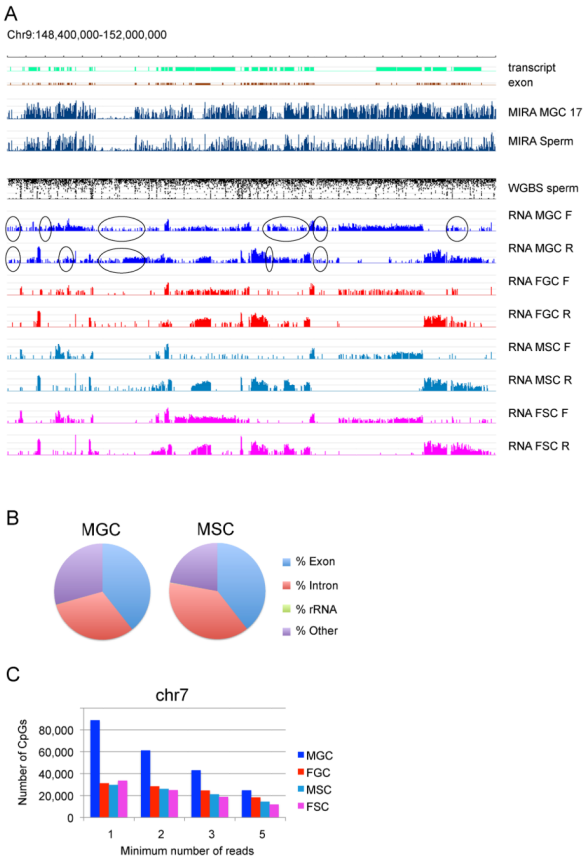

Figure 2. Genome-wide broad, low-level transcription preexist de novo methylation in prospermatogonia

(A) RNA-seq results are displayed for a region in chr9 in forward and reverse (F and R) direction in FGC, MGC, FSC, and MSC at 15.5 dpc. The total reads were evenly scaled to 50 M and are depicted in $\log_2$ scale ranging from 0–15. The location of known transcripts and exons is shown at the top. DNA methylation lanes by MIRA-chip and WGBS are included as in Figure 1. FGC, MSC and FSC have clear gaps between transcripts where MGC displays read-through (ovals) by low-level transcription. See also Figure S2. (B) Comparison of the relative distribution of transcription among genomic features between MGC and MSC. Unique pair-end reads were counted. See also Table S2. (C) Quantitation of broad, low-level intergenic transcription. The numbers of CpGs that occur outside of known transcripts and overlap RNA-seq reads were plotted in chr7.

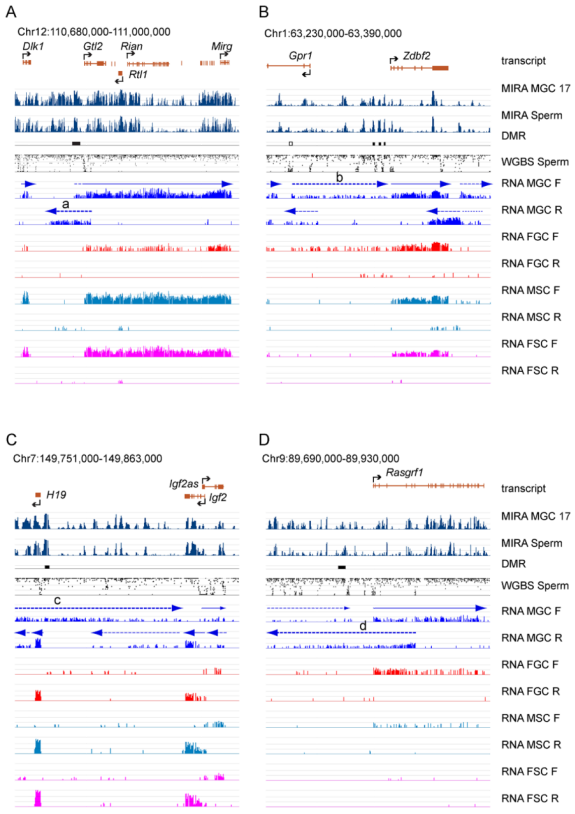

Figure 3. Broad, low-level transcripts traverse paternally methylated imprinted DMRs in fetal male germ cells

Transcription and DNA methylation is depicted with high resolution at the four known paternally methylated imprinted DMRs (closed rectangles). Labeling details are as in Figure 2. The DMR genomic coordinates are from (Tomizawa et al., 2011) and (Hiura et al., 2010). The location of known transcripts is shown at the top and marked above the plotted RNA-seq reads by solid arrows. Note the presence of novel transcripts in MGC (broken arrows). Some of these novel low-level transcripts (heavy broken arrows a–d) run across PAT DMRs. See transcription at imprinted MAT DMRs in Figure S3 and Table S3.

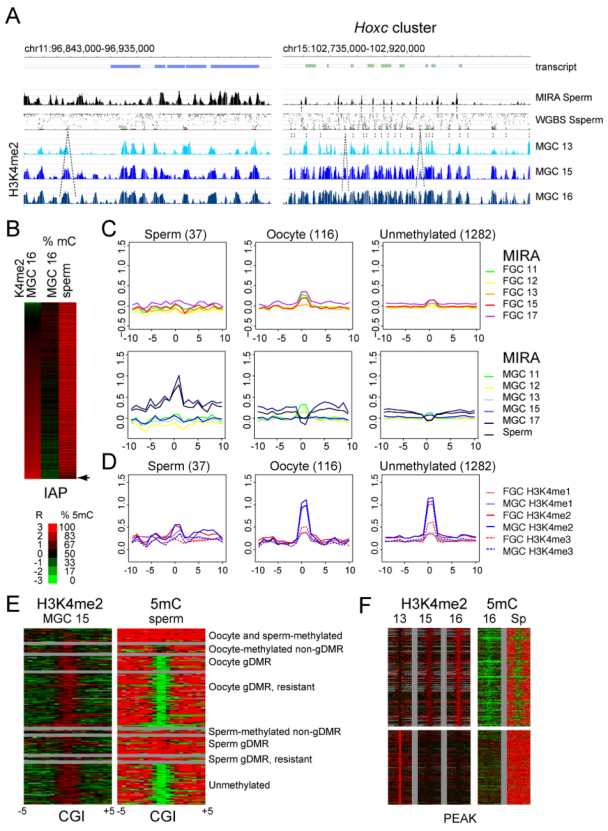

Figure 4. Default DNA methylation is patterned by the negative mold of dynamic H3K4me2 peaks in prospermatogonia

(A) Constitutive and emerging (highlighted with broken lines) H3K4me2 peaks in MGC are precisely aligned with unmethylated valleys of sperm DNA. H3K4me2 ChIP-chip results are shown at 13.5, 15.5 and 16.5 dpc of at a highly methylated region along chr11 (to the left), and one region with relatively less methylation (to the right). ChIP versus input DNA values were plotted as $-\log_{10}$ p-value scores in the range of 0–8.4. (B) IAPs can escape full methylation when protected by H3K4me2. Heat maps of depict WGBS DNA methylation levels (Kobayashi et al., 2012; Seisenberger et al., 2012) along the +/−1 kb IAP flanking regions in 16.5 dpc MGC and sperm according to increasing H3K4me2 levels in MGC at 16.5 dpc. An arrow points to the IAPs that remain partially methylated in sperm. (C) Dynamic methylation of CGIs in prospermatogonia. The average MIRA values are depicted at CGIs (number in parentheses) in the different methylation classes (See Table S4) (Kobayashi et al., 2012), +/− 10 kb in FGC (top) and MGC (bottom), as indicated by colors to the right, at 11.5, 12.5, 13.5, 15.5 and 17.5 dpc and in sperm. (D) Oocyte-methylated and unmethylated but not sperm-methylated CGIs precisely align with sharp peaks of H3K4 methylation in prospermatogonia. ChIP-chip composite profiles of H3K4me1, H3K4me2 and H3K4me3 are displayed in MGC and FGC at 15.5 dpc at the same CGIs as in Figure 4C. (E) Sharply defined unmethylated islands correspond to H3K4me2 at oocyte-methylated gDMRs and unmethylated CGIs. Heat map depicts the DNA methylation levels in sperm according to WGBS (Kobayashi et al., 2012) and the H3K4me2 levels in 15.5 dpc MGC at CGIs (+/− 5 kb) in the methylation categories (See Table S4). A truncated list is shown for unmethylated CGIs. (F) Emerging H3K4me2 peaks protect the underlying sequences from global de novo DNA methylation. Heat maps depict % DNA methylation in 16.5 dpc MGC (Seisenberger et al., 2012) and sperm (Kobayashi et al., 2012) and H3K4me2 ChIP-chip enrichment ratios in MGC at 13.5, 15.5 and 16.5 dpc. The heat maps interrogate +/− 5 kb

regions in the proximity of H3K4me2 peaks that emerge (top panel) or diminish (bottom panel) in MGC between 13.5 and 16.5 dpc. See Figure S4 for ChIP-chip at control regions.

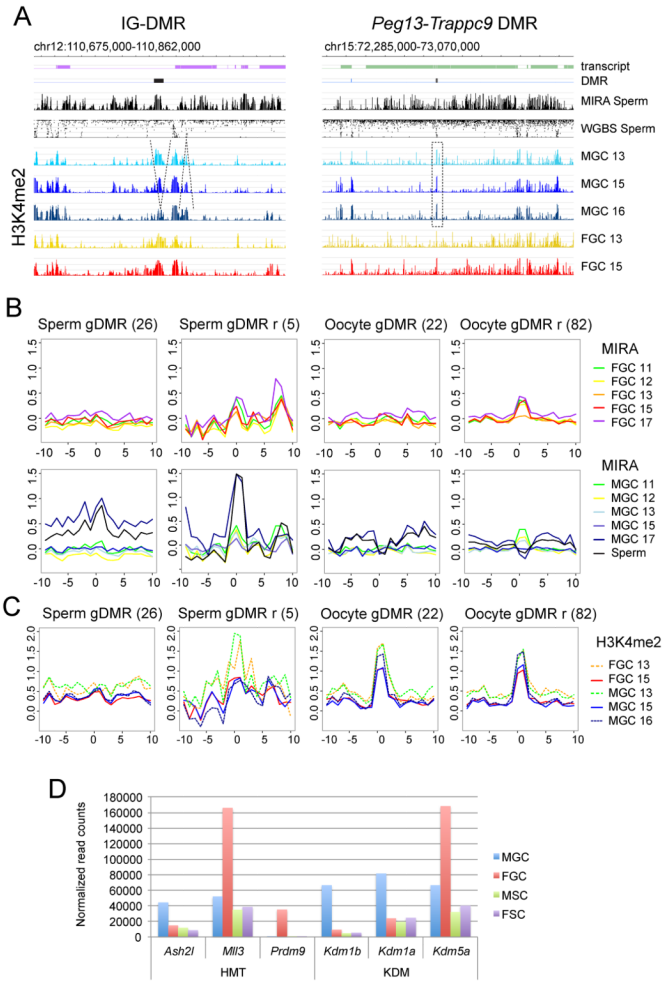

Figure 5. Imprinted DMR methylation patterns follow the general rules of default de novo methylation

(A) PAT DMRs correspond to one of the emerging methylation peaks at no or diminishing H3K4me2 but MAT DMRs occur in a DNA methylation valley at a sharp H3K4me2 peak. ChIP-chip H3K4me2 results are shown in MGC at 13.5, 15.5, 16.5 dpc and FGC at 13.5, 15.5 at one paternally and one maternally methylated DMR. DNA methylation patterns in sperm are also included for comparison (as in Figure 1). See additional DMRs in Figure S5. (B) Dynamic DNA methylation at gDMR CGIs (See Table S4). MIRA composite profiles of CGIs +/− 5 kb are depicted in FGC (top) and MGC (bottom), as indicated by colors to the right, at 11.5, 12.5, 13.5, 15.5 and 17.5 dpc and in sperm. (C) Developmental profile of H3K4me2 at gDMR CGIs. ChIP-chip composite profiles of H3K4me2 are displayed in 13.5, 15.5 and 16.5 dpc MGC and 13.5 and 15.5 dpc FGC at the CGI categories as in Figure 5B. Note that sperm-methylated gDMRs do not have distinct H3K4me peak, except the resistant category at 13.5 dpc. This peak diminishes by 15.5 dpc. (D) Putative epigenetic modifiers of the dynamic H3K4me2 patterns in fetal germ cells. Normalized read counts of H3K4 methyltransferase (HMT) and demethylase (KDM) transcripts are plotted in MGC, FGC, FSC and MSC from an RNA-seq experiment of the 15.5 dpc gonads (See also Table S1).

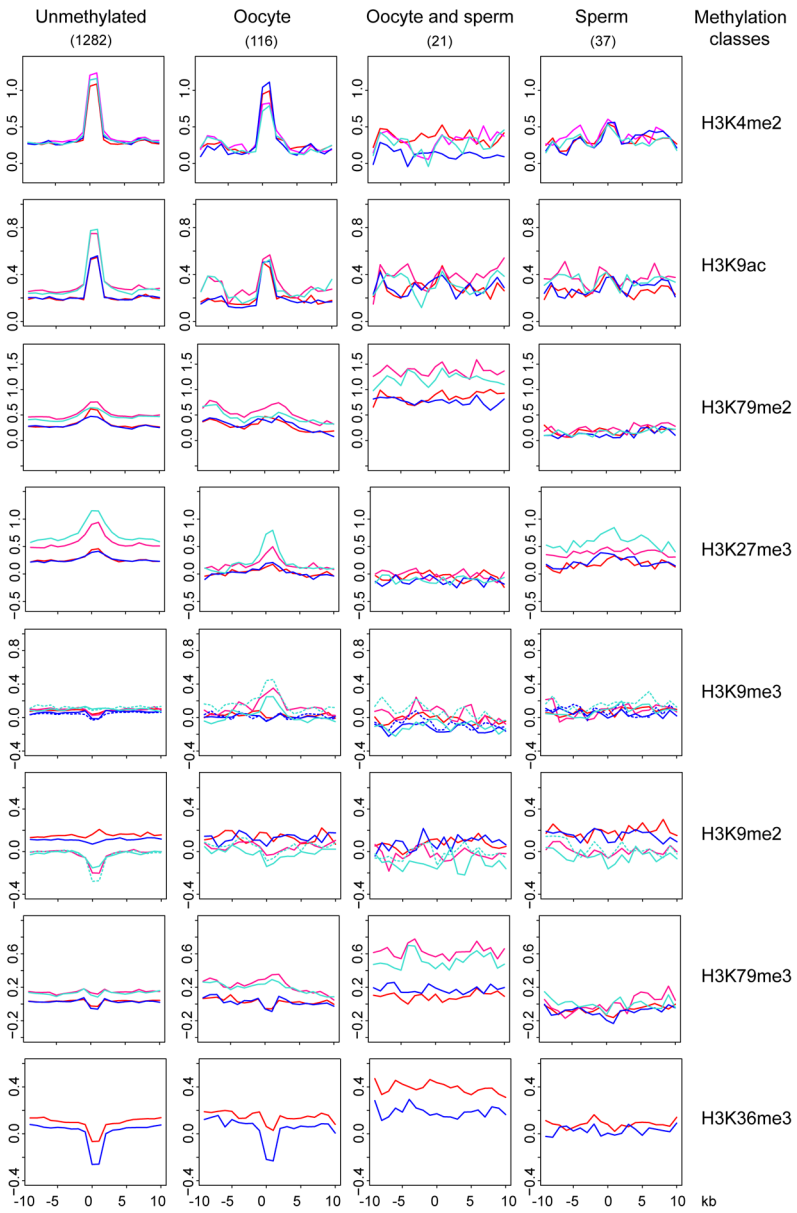

Figure 6. De novo methylation occurs at CGIs with no apparent clues from repressive chromatin

Chromatin profiles are shown at 15.5 dpc for the CGI methylation classes (Table S4). ChIP-chip composites are depicted at the CGI +/− 10 kb with the antibodies indicated to the right in the four cell types: FGC (red), MGC (blue), FSC (hot pink), and MSC (turquoise). The second antibody (ab2) for H3K9me2 and H3K9me3 is shown in broken lines. Note that unmethylated and oocyte-methylated CGIs are precisely marked in each cell type with H3K4me2 and H3K9ac peaks and only in somatic cells by diffused peaks of H3K79me2 and H3K27me3. FGC and MGC are depleted at these CGIs in H3K36me3 and somewhat in H3K79me3. CGIs that become methylated in sperm do not display enrichment or depletion of any chromatin marks in fetal germ cells. See also Figure S6 for gDMR CGI composites and examples of imprinted DMRs.

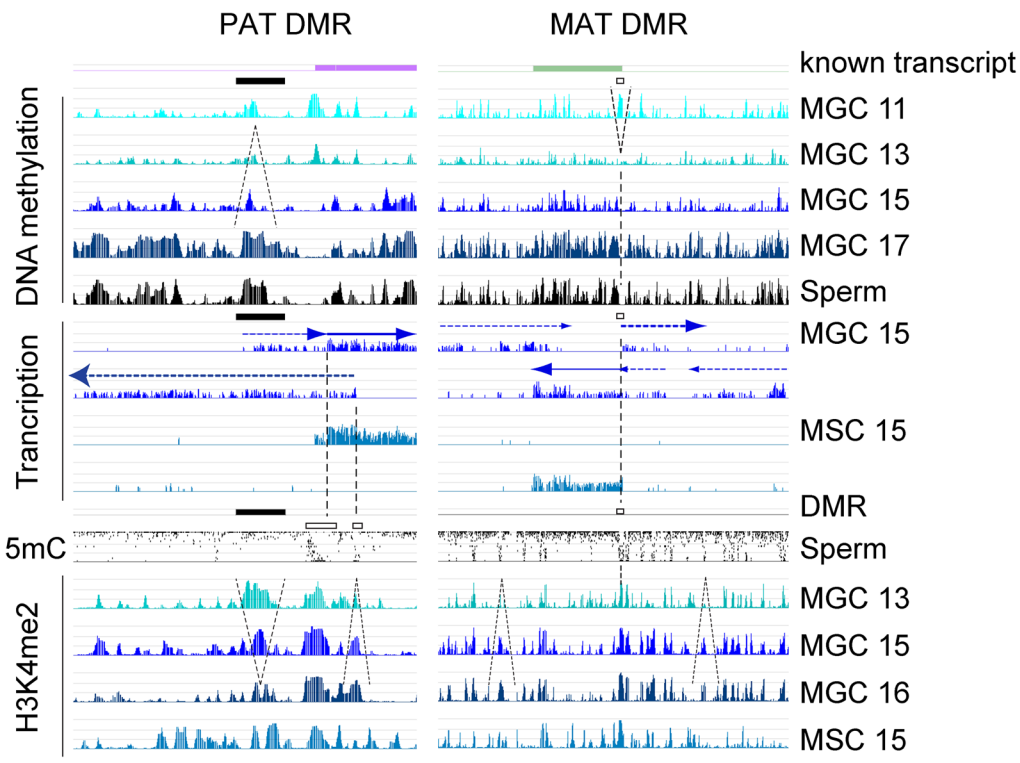

Figure 7. Summary

The pattern of de novo DNA methylation in prospermatogonia is dictated by opposing actions of broad, low-level transcription and dynamic patterns of active chromatin. The key findings are illustrated at the sequences that harbor the PAT IG-DMR and at the *Slc38a4* MAT DMR. 1, Default DNA methylation takes place in MGC between 15.5–17.5 dpc. 2, Broad low-level transcription occurs at the initiation of de novo DNA methylation in MGC at 15.5 dpc. 3, H3K4 methylation is dynamically remodeled between 13.5–16.5 dpc. 4. H3K4me2 peaks are the negative mold of emerging DNA methylation. 5. Imprinted DMRs, CGIs (and IAPs) follow these same clues in prospermatogonia.