Dynamic CpG island methylation landscape in oocytes and preimplantation embryos

Sébastien A. Smallwood[1], **Shin-ichi Tomizawa**[1],

methylation of known maternal germline differentially methylated regions (DMRs) at imprinted loci (Supplementary Fig. 2).

CpG methylation overall, and in CGI and repetitive element contexts, showed a dynamic profile during oocyte growth: 0.5% of all CpGs assessed by RRBS were highly methylated in d5 oocytes (80% methylation), 11.3% in d20 germinal vesicle (GV) and 15.3% in ovulated metaphase II (MII) oocytes. CpG methylation was lower overall in mature oocytes than sperm (24.9% of CpGs highly methylated in sperm), consistent with previous observations on repetitive elements[10]; methylation in a CGI context, irrespective of location with respect to genes, was markedly lower in sperm (Fig. 1a, Supplementary Fig. 3a-b & 4). Using a threshold for scoring CGIs that reads should cover 10% of the CpGs per CGI (see Methods for a full account), we obtained information on ~15,000 (~65%) of the extended set of CGIs recently identified by CAP-Seq[11], and identified 1062 methylated CGIs (75% methylation) in mature oocytes (Fig 1b-c; Supplementary Table 1). By extrapolation, there may be ~1600 fully methylated CGIs in mature oocytes. Of interest, we found that the CGIs associated with the major promoters of *Dnmt3b* and *Dnmt1* (*Dnmt1s*) were methylated (Fig. 1d,e, Supplementary Fig. 9). Eighty-nine CGIs identified as methylated in MII oocytes were not fully methylated in GV oocytes, demonstrating that CGIs acquire methylation at different rates during oocyte growth, as reported for germline DMRs[12]-[13] (Supplementary Table 2). In sperm, we identified 185 fully methylated CGIs, 58 of which were methylated exclusively in sperm and 100 were also methylated in mature oocytes (27 of the CGIs methylated in sperm were not informative in mature oocyte datasets) (Fig 1b-c;

coverage in *Dnmt3a*

we performed RRBS on blastocysts (E3.5). This was validated by the expected degree of methylation at twelve known maternal germline DMRs (range 45.2%-58.7%). Consistent with genome-wide erasure, there was a substantial reduction in the proportion of methylated CpGs (60%) across the genome or within CGIs compared with gametes (Fig. 3b, Supplementary Fig. 8a). Crucially, a minority of CGIs methylated in germ cells showed complete protection from demethylation: only ~15% of CGIs methylated in oocytes retained 40% methylation in blastocysts (Fig. 3b, c). This substantial post-fertilisation reprogramming suggests that most CGI methylation in oocytes and sperm is unrelated to imprinting, and argues that maintenance of methylation in preimplantation embryos is a decisive factor in imprinting.

However, we observed that most CGIs methylated in oocytes displayed greater levels of methylation in blastocysts than expected if they were fully subject to passive demethylation, by which methylation should be <2% by the 32-cell stage. This was striking, as very few CGIs are methylated in blastocysts (Fig. 3b, c, Supplementary Fig 8b, c). To examine the degree to which gametic methylation is a factor in CGI methylation in preimplantation embryos, we looked at the dependence of methylation in blastocysts on prior methylation in gametes. Of 280 CGIs displaying intermediate methylation levels (25-40%) in blastocysts, the vast majority (234; 83%; $p<0.001$, ^{2} test) were fully methylated in MII oocytes (including 27 CGIs methylated in both oocyte and sperm) (Supplementary Fig. 8d). In contrast, less than 0.5% of CGIs unmethylated in both gametes are methylated 25% in blastocysts (Fig. 3c, Supplementary Table 1). To investigate whether CGI sequence influences the likelihood of maintaining methylation, we checked how the properties of CGIs highly methylated in MII oocytes (75%) differed according to methylation level in blastocysts. For most parameters the differences were minor, but there was a tendency for CGIs retaining higher levels of methylation to be shorter and to be intragenically located (Supplementary Fig. 8e, f). To validate CGI methylation allele-specifically, we examined a selection of CGIs in C57BL/6JxCast/Ei hybrid embryos by conventional bisulphite sequencing. As exemplified by the *Syt2* locus, the CGI is fully methylated in oocytes and the maternal allele partially retains methylation in blastocysts (Fig. 3d, Supplementary Fig. 9). For CGIs specifically methylated in sperm, there was less evidence for substantial maintenance of methylation in blastocysts (Fig. 3c, Supplementary Fig. 9). These findings extend observations of Borgel *et al.* who, from MeDIP-chip analysis of promoter methylation in preimplantation embryos, identified some non-imprinted sequences that resist demethylation in preimplantation development[25]. Thus, CGI methylation status in gametes strongly predisposes towards methylation in blastocysts, either by incomplete post-fertilisation demethylation of methylated CGIs, or because some legacy of gametic methylation instructs their re-methylation in a subpopulation of cells. By either mechanism, mosaicism of CGI methylation patterns between blastomeres is predicted to arise. This does not exclude a contribution of *de novo* methylation, as some CGIs unmethylated in gametes have become methylated in blastocysts (Fig. 3c, Supplementary Fig. 8d, Supplementary Table 1), including genes involved in trophectoderm development[26].

In conclusion, we reveal the extent and dynamics of CGI methylation in oocytes; this provides an important reference by which to judge future studies on mechanisms of *de novo* methylation in germ cells. A comprehensive account of the differential CGI methylation in male and female gametes is also a prerequisite for defining the full repertoire of imprinted genes and the mechanistic basis of parent-of-origin expression effects in somatic tissues. We also describe an unexpectedly complex fate of gamete-derived methylation after fertilisation. Rather than a binary choice, with DMRs characterised by absolute maintenance and other gametic methylation comprehensively lost through active demethylation or lack of maintenance during the first cleavage divisions, our analysis suggests a greater diversity of methylation choices. This diversity might lead to the establishment of epigenetic mosaicism

within the early embryo, which might have the potential to influence first lineage specification[27]

ChIP-Seq

H3K4me3 immunoprecipitation (39159, Active Motif) was performed as described elsewhere[29] with minor modifications. ChIP was performed in duplicate from 3200 oocytes. Illumina Libraries were generated (input and IP) using an NEBNext kit (Set 1, NEB), except that adapter ligation was performed as for RRBS. Sequences were aligned using an ungapped Eland alignment with default stringency parameters. Owing to high background results from the limited starting material, reads from duplicates were combined. Technical assessment was made by comparison with ES cell H3K4me3 ChIP-Seq datasets (GSM594581 and GSM535982)[11],[30].

Direct Bisulphite Sequencing

DNA was purified by proteinase K digestion and phenol-chloroform extraction, spiked with Lambda DNA and bisulphite treated (Zymo). Each PCR comprised a minimum of 50 oocytes or 2-3 blastocyst equivalents. Cloning and analysis were performed as described elsewhere[31], with removal of clones with identical patterns of conversion. Primers used for the amplification of specific CGIs from bisulphite modified DNA are given in Supplementary Table 4.2.

Statistical analysis

For categorical data, such as distribution of CpGs or CGIs methylation, χ^2 tests were applied. For quantitative data, Mann-Whitney U tests (between 2 groups) and Kruskall-Wallis tests (between more than 2 groups), were applied.

Additional information

Dataset analysis was based on build NCBIM37/mm9 of the mouse genome and performed using Seqmonk (http://www.bioinformatics.bbsrc.ac.uk/projects). Promoter CGIs were defined as overlapping an annotated TSS (EnsEMBL, Refseq or UCSC); intragenic CGIs as overlapping an annotated gene without its TSS; intergenic CGIs were not overlapping annotated genes or promoters. Promoters were defined as the region 2kb upstream of annotated TSS. For repetitive element analysis, positions of individual instances of LINE, SINE, tandem repeats, long terminal repeats (LTR) and low complexity regions (LCR) were extracted from EnsEMBL. The overlap between full length CGIs and individual repeat types was determined as percentage of the CGI length using a custom Perl script. CpG periodicity was determined as the distribution of inter-CpG distance (from C to C) between all pairs of CpGs in each region, averaged over all of the regions in a particular grouping.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Bartolomei MS. Genomic imprinting: employing and avoiding epigenetic processes. Genes & Development. 2009; 23:2124–2133. [PubMed: 19759261]

2. Morgan HD, Santos F, Green K, Dean W, Reik W. Epigenetic reprogramming in mammals. Human Molecular Genetics. 2005; 14:R47–R58. [PubMed: 15809273]

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

4. Bourc'his D, Xu G-L, Lin C-S, Bollman B, Bestor TH. Dnmt3L and the Establishment of Maternal Genomic Imprints. Science. 2001; 294:2536–2539. [PubMed: 11719692]

5. Kaneda M, et al. Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. Nature. 2004; 429:900–903. [PubMed: 15215868]

6. Jia D, Jurkowska RZ, Zhang X, Jeltsch A, Cheng X. Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation. Nature. 2007; 449:248–251. [PubMed: 17713477]

7. Schaefer CB, Ooi SKT, Bestor TH, Bourc'his D. Epigenetic Decisions in Mammalian Germ Cells. Science. 2007; 316:398–399. [PubMed: 17446388]

8. Meissner A, et al. Genome-scale DNA methylation maps of pluripotent and differentiated cells. Nature. 2008; 454:766–770. [PubMed: 18600261]

9. Smith ZD, Gu H, Bock C, Gnirke A, Meissner A. High-throughput bisulfite sequencing in mammalian genomes. Methods. 2009; 48:226–232. [PubMed: 19442738]

10. Howlett SK, Reik W. Methylation levels of maternal and paternal genomes during preimplantation development. Development. 1991; 113:119–127. [PubMed: 1764989]

11. Illingworth RS, et al. Orphan CpG Islands Identify Numerous Conserved Promoters in the Mammalian Genome. PLoS Genet. 2010; 6:e1001134. [PubMed: 20885785]

12. Chotalia M, et al. Transcription is required for establishment of germline methylation marks at imprinted genes. Genes & Development. 2009; 23:105–117. [PubMed: 19136628]

13. Lucifero D, Mann MRW, Bartolomei MS, Trasler JM. Gene-specific timing and epigenetic memory in oocyte imprinting. Human Molecular Genetics. 2004; 13:839–849. [PubMed: 14998934]

14. Illingworth R, et al. A Novel CpG Island Set Identifies Tissue-Specific Methylation at Developmental Gene Loci. PLoS Biol. 2008; 6:e22. [PubMed: 18232738]

15. Maunakea AK, et al. Conserved role of intragenic DNA methylation in regulating alternative promoters. Nature. 2010; 466:253–257. [PubMed: 20613842]

16. Reinhart B, Paoloni-Giacobino A, Chaillet JR. Specific Differentially Methylated Domain Sequences Direct the Maintenance of Methylation at Imprinted Genes. Mol. Cell. Biol. 2006; 26:8347–8356. [PubMed: 16954379]

17. Bock C, Halachev K, Buch J, Lengauer T. EpiGRAPH: user-friendly software for statistical analysis and prediction of (epi)genomic data. Genome Biology. 2009; 10:R14. [PubMed: 19208250]

18. 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]

19. Popp C, et al. Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. Nature. 2010; 463:1101–1105. [PubMed: 20098412]

20. Ciccone DN, et al. KDM1B is a histone H3K4 demethylase required to establish maternal genomic imprints. Nature. 2009; 461:415–418. [PubMed: 19727073]

21. Fang R, et al. Human LSD2/KDM1b/AOF1 Regulates Gene Transcription by Modulating Intragenic H3K4me2 Methylation. Molecular Cell. 2010; 39:222–233. [PubMed: 20670891]

22. Ooi SKT, et al. DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. Nature. 2007; 448:714–717. [PubMed: 17687327]

23. Zhang Y, 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 Research. 2010; 38:4246–4253. [PubMed: 20223770]

24. Dhayalan A, et al. The Dnmt3a PWWP Domain Reads Histone 3 Lysine 36 Trimethylation and Guides DNA Methylation. Journal of Biological Chemistry. 2010; 285:26114–26120. [PubMed: 20547484]

25. Borgel J, et al. Targets and dynamics of promoter DNA methylation during early mouse development. Nat Genet. 2010; 42:1093–1100. [PubMed: 21057502]

26. Tartakover-Matalon S, et al. Impaired migration of trophoblast cells caused by simvastatin is associated with decreased membrane IGF-I receptor, MMP2 activity and HSP27 expression. Human Reproduction. 2007; 22:1161–1167. [PubMed: 17158816]

Methods References

27. Hemberger M, Dean W, Reik W. Epigenetic dynamics of stem cells and cell lineage commitment: digging Waddington's canal. Nat Rev Mol Cell Biol. 2009; 10:526–537. [PubMed: 19603040]

28. Krueger F, Andrews SR. Bismark: A flexible aligner and methylation caller for Bisulfite-Seq applications. Bioinformatics. 2011 doi:10.1093/bioinformatics/btr167.

29. Dahl JA, Collas P. A rapid micro chromatin immunoprecipitation assay (ChIP). Nat. Protocols. 2008; 3:1032–1045.

30. Creyghton MP, et al. Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proceedings of the National Academy of Sciences. 2010; 107:21931–21936.

31. Tomizawa S, et al. 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]
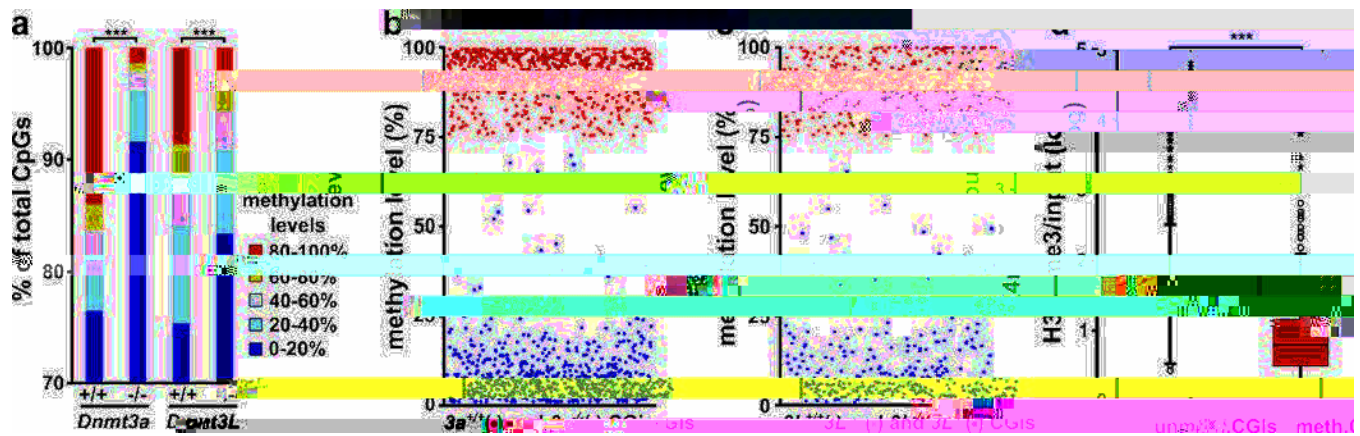
Figure 1. DNA methylation landscape in oocytes and sperm determined by RRBS
a-b, Distribution of CpG methylation levels across the genome (

Figure 2. Mechanism of DNA methylation establishment in oocytes
a, Distribution of CpG methylation levels across the genome in *Dnmt3a*–/– and *Dnmt3L*–/–
oocytes and their wild-type counterparts (+/+); the number of CpGs analysed is indicated in
Suppl. Fig 1b (***: p<0.001, χ^2 test). **b-c,** Methylation levels of CGIs in *Dnmt3a*–/– and
Dnmt3L–/– oocytes; only those CGIs for which methylation was ≥75% in the corresponding
wild-type oocytes are displayed. **d,** Overall correlation between H3K4me3 enrichment
determined in d15 oocytes by ChIP-seq and methylation status of CGIs (all CGIs
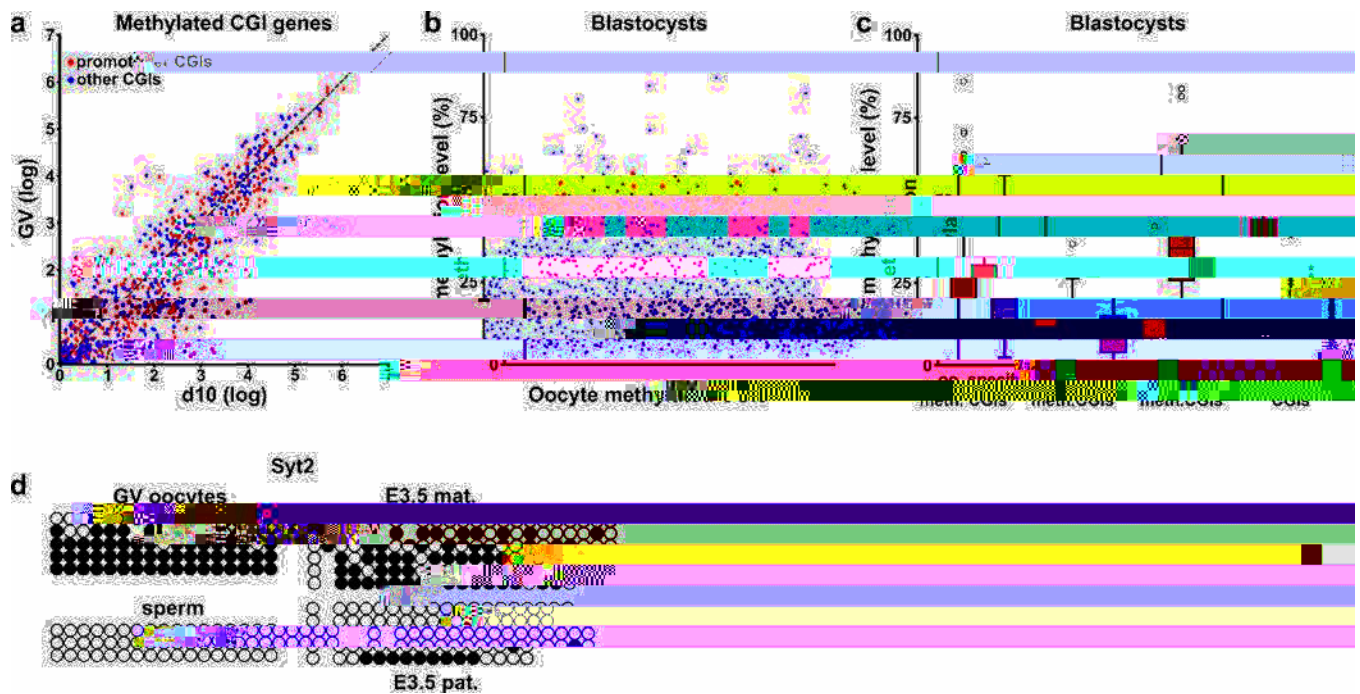irrespective of genomic location; ***: p<0.001, Mann-Whitney U test).

Figure 3. Biological significance and fate of CGI methylation in oocytes
a, mRNA expression levels in d10 and GV oocytes of the genes associated with methylated CGIs, either promoter (red, n=410) or intragenic (blue, n=555). **b,** Methylation levels in blastocysts of the CGIs identified as methylated in mature oocytes; twelve known germline DMRs with informative coverage are displayed in red (range 45.2%-58.7%). **c,** Range of methylation in blastocysts of the CGIs methylated specifically in oocytes (n=803) or sperm (n=51), methylated in both oocytes and sperm (n=86) and unmethylated in gametes (n=11512). **d,** Bisulphite sequencing in GV oocytes, sperm and C57BL/6JxCAST/Ei hybrid E3.5 blastocysts of the *Syt2* CGI. Bisulphite sequence profiles from the maternal (mat) and paternal (pat) alleles in blastocysts were discriminated by polymorphisms between C57BL/6J and CAST/Ei. Open circles represent unmethylated CpGs and filled circles methylated CpGs.