# *DNA demethylation by DNA repair*

# DNA Demethylation by DNA Repair

Mary Gehring[1], Wolf Reik[2], and Steven Henikoff[1]

1 Howard Hughes Medical Institute
Fred Hutchinson Cancer Research Center
1100 Fairview Ave N
Seattle, WA 98109

2 Laboratory of Developmental Genetics and Imprinting
The Babraham Institute
Cambridge CB22 3AT, United Kingdom

*Corresponding author*: Henikoff, S (steveh@fhcrc.org)

**ABSTRACT**

Active DNA demethylation underlies key facets of reproduction in flowering plants and mammals and serves a general genome housekeeping function in plants. A family of 5-methylcytosine DNA glycosylases catalyzes plant demethylation via the well-known DNA base excision repair process. Although the existence of active demethylation has been known for a longer time in mammals, the means of achieving it remain murky, and mammals lack counterparts to the plant demethylases. Several intriguing experiments have suggested, but not conclusively proven, that DNA repair is also a plausible mechanism for animal demethylation. Here we examine what is known from flowering plants about the pathways and function of enzymatic demethylation and discuss possible mechanisms whereby DNA repair might also underlie global demethylation in mammals.

## DNA methylation and demethylation

5-methylcytosine is the only known epigenetic mark that can be inherited mitotically and in some instances meiotically. DNA methylation is associated with common and critical processes in both flowering plants and mammals, including transposon silencing and **genomic imprinting**. But DNA methylation can also be removed, a process that is far less well understood. The study of DNA demethylation has been led by research in plants, where it is clear that a DNA repair pathway has been adopted for removing 5-methylcytosine from DNA. The plant pathway, strongly supported biologically and biochemically, provide a framework to critically examine possible mechanisms of demethylation in mammals.

The enzymes responsible for DNA methylation are conserved between plants and animals, and their mechanisms of action are well understood [1]. DNA methylation is established

by *de novo* methyltransferases, which in plants are guided to their targets by small RNAs. Symmetric methylation patterns are maintained after DNA replication by maintenance methyltransferases, which methylate the new DNA strand based on the pattern found on the parent strand.Methylation can be lost passively when the maintenance methylation that usually follows DNA replication is inhibited, or by a more active process when 5-methylcytosine is enzymatically removed. Active demethylation has emerged as an important mechanism in the genomes of flowering plants for shaping methylation patterns. In this process, DNA glycosylases, which are normally associated with DNA repair, recognize and remove 5-methylcytosine from DNA, leading to its replacement with cytosine. DNA demethylation protects genic regions of the genome from a robust methylation-based transposon defense system and activates the expressed alleles of some imprinted genes. Extensive changes in DNA methylation levels occur during the genome-wide epigenetic reprogramming associated with mammalian reproduction [2](**Figure 1**). Indeed, erasure of methylation acquired over the lifetime of the organism, along with other chromatin marks, might be needed to establish **totipotency**. The rapid loss of DNA methylation that occurs within the period of a single cell cycle suggests the presence of enzymes that actively remove 5-methylcytosine from DNA. The responsible enzyme(s) remains unknown, and the plant 5-methylcytosine DNA glycosylases do not appear to have animal homologues. Until the basic biochemical mechanisms underlying demethylation are known, fundamental processes that involve changes in DNA methylation, including reproductive development, cancer progression, and epigenetic reprogramming[3], will remain insufficiently understood.

**When and where does demethylation occur?**

In mammals genome-wide active demethylation events occur at two points in reproductive development: in the male **pronucleus** of the zygote immediately after fertilization and possibly in the **primordial germ cells** (PGCs) of 11.5-12.5 day old embryos **(Figure 1)**. By contrast, there is no evidence that plants undergo genome-wide reprogramming events as a natural part of the life cycle. However, demethylation does occur somatically to counteract the establishment of potentially detrimental methylation. Gene imprinting is a key feature of reproduction in both classes of organisms, and active demethylation appears to play a role in each.

*Global DNA demethylation in the male pronucleus*

Studies of methylation levels at different stages of reproductive development show that sperm DNA is more methylated than oocyte DNA[4, 5]. At fertilization, the mature sperm arrives in an oocyte that is arrested in metaphase II of meiosis. The maternal and paternal nuclei differ markedly in terms of genome organization. In particular, sperm DNA is tightly packed by **protamines**, with some histones remaining[6]. After fertilization, the DNA is repackaged with nucleosomes; male chromatin decondensation is apparent within a few hours after fertilization [6] and correlates with methylation loss. DNA methylation within zygotes has been best studied in mice, both globally and at specific loci **(Box 1)**. Immunofluorescence studies using anti-5-methylcytosine antibodies show that although both the female and male pronuclei stain strongly for DNA methylation 3 hours after fertilization, between 4-8 hours after fertilization the signal from the male pronucleus is almost entirely lost[7, 8]. This change occurs before the first S-phase and so must represent an active demethylation mechanism. Furthermore, treatment of 6-hour-old zygotes with the DNA polymerase inhibitor aphidocolin does not prevent loss of the methylation signal[7]. The zygotic male pronuclei of rat, pigs, and cows also undergo demethylation[9]. Rabbits

and sheep have been reported to lack this event[10] but more recent data suggests that these discrepancies could stem from differences in timing between species[11] or the relative strength of the demethylating activity. Indeed, mouse sperm are demethylated when injected into sheep oocytes, albeit to a lesser extent than in mouse oocytes, and the reverse is also true. Therefore sheep oocytes possess a demethylating activity and their sperm has the capacity to be demethylated[12].

*Demethylation in Primordial Germ Cells*

Active demethylation in the male pronucleus is followed by passive demethylation of both genomes during pre-implantation development, except at imprinted loci[13]. After implantation, methylation levels increase in the blastocyst inner cell mass, the progenitor of the embryo proper. Parent-of-origin-specific imprinting marks must be removed in PGCs and later established according to the sex of the individual. Evidence indicates that this demethylation might also be active, although accessing the cells, their temporal asynchrony, and the fact that they are actively dividing, makes experiments challenging.

PGCs enter the genital ridge between 10.5 and 11.5 days post coitum (dpc), proliferate until day 13.5, and then enter either meiotic prophase (females) or mitotic arrest (males). It is in this 2-3 day time period that imprints are erased along with other methylation; the greatest loss of methylation probably occurs between 11.5 and 12.5 dpc. Embryos generated from PGC cells cloned by nuclear transfer at 11.5 dpc survive longer than those cloned from 12.5 dpc cells, which suffer from complete abolition of genomic imprinting[14]. Clones from day 11.5 have a mixture of imprinting status and methylation patterns at particular genes. A **bisulfite sequencing** study comparing day 12.5 to day 11.5 PGCs found a reduction in methylation at differentially

methylated regions of 5 examined imprinted genes as well as loss of methylation at non-imprinted sequences[15]. IAP (intracisternal A particle) and LINE (long interspersed nuclear elements) transposable elements resist demethylation to a variable extent[15, 16]. Because the doubling time of PGCs is 16 hours[17], the almost complete loss of methylation observed for single copy genes within a 24 hr time period suggests an active, rather than passive, demethylation process. Genome-wide characterization of methylation in these cells could provide important insights into the types of sequences subject to or protected from demethylation.

*Demethylation in flowering plants*

Active DNA demethylation is involved in two processes in **angiosperms**: gene imprinting during reproduction and maintaining normal methylation patterns throughout the plant[18]. Small RNAs direct DNA methylation establishment in plants. These mainly arise from and target repetitive sequences such as transposable elements, which are generally highly methylated in comparison to genic sequences. In particular, the 5' and 3' regions of genes are generally depleted of methylation, where the accumulation of methylation can be detrimental to gene function[19]. Active DNA demethylation appears to remove marks that encroach on genic space[20, 21]. In this way plants enjoy a robust methylation defense system that silences transposable elements without negatively affecting nearby genes.

As in mammals, imprinted genes play a crucial role in reproductive development. Imprinting takes place in the endosperm, a tissue that supports embryo growth during seed development and seedling germination. Active DNA demethylation has been implicated in plant gene imprinting, although other chromatin-based mechanisms are also important[22]. Five genes are known to be imprinted in *Arabidopsis thaliana* endosperm: *FWA* (*FLOWERING WAGENINGEN*), *MEA* (*MEDEA)*, *FIS2* (*FERTILIZATION INDEPENDENT SEED 2*) and *MPC*

(*MATERNALLY EXPRESSED PAB C-TERMINAL*) are expressed maternally and silent paternally, whereas *PHE1* (*PHERES1*) is oppositely imprinted. Before fertilization, *FWA*, *MEA*, *FIS2*, and *MPC* are expressed in the central cell of the female gametophyte, which gives rise to the endosperm after fertilization **(Box 2)**. A maize gene, *FIE1* (*FERTILIZATION INDEPENDENT ENDOSPERM 1*), is imprinted in the endosperm in a similar manner[23, 24]. *FIE1* is less methylated in the central cell compared with the egg cell and sperm cells[23]. After fertilization, *FIE1* maternal endosperm alleles are hypomethylated compared with paternal alleles, although this difference is erased at later stages of seed development[23, 24]. Similarly, maternal *MEA* alleles are less methylated in the endosperm than paternal alleles and maternal and paternal embryo alleles[25]. *FWA* is also less methylated in the endosperm than in the embryo[26]. These data suggest that expressed maternal alleles of some imprinted genes are actively demethylated in the central cell before fertilization. The egg cell nucleus and one of the central cell nuclei are sisters, making it unlikely that passive loss of methylation due to replication can account for methylation differences between the egg and central cell, although recent data suggests it might also contribute to the process[27]. As the endosperm is a terminally differentiated tissue that does not contribute to the next generation, there is no need for methylation-resetting as there is in mammals.

For most of the Arabidopsis life cycle, active DNA demethylation, like DNA repair, primarily serves a genome "housekeeping" function. To date, DNA demethylation appears only to have a role in development with regard to the activation of imprinted genes essential for seed viability. Similarly, not all of the active demethylation observed in mammals is necessarily a developmental or essential event. **Round spermatids** can be successfully used for in vitro fertilization in mice even though their DNA is only transiently demethylated in the zygote and is

157 then aberrantly remethylated[28, 29]. Thus biological function might come from the process of
158 demethylation itself, rather than the final methylation status of the DNA.

159 **DNA repair as a mechanism for DNA demethylation**

160 The search for enzymes responsible for demethylation has produced varied candidates
161 and reaction mechanisms[30]. These fall into three general categories 1) direct removal of the
162 methyl group from the 5C position of cytosine[31] 2) base excision repair (BER) that leads to the
163 replacement of 5-methylcytosine with cytosine by either directly removing 5-methylcytosine or
164 through the directed deamination of 5-methylcytosine to thymine and 3) nucleotide excision
165 repair (NER) of DNA containing 5-methylcytosine. Here we focus on DNA repair-based
166 mechanisms. What are likely criteria for candidate demethylases? First, the demethylase must be
167 expressed in the cells in which demethylation occurs. Second, the reaction mechanism must be
168 fast enough to account for the observed rate of methylation changes. In mammals, the male
169 pronucleus is demethylated within just a few hours after fertilization. However, there has been no
170 genome-wide methylation profiling in any of the cells in which active demethylation occurs,
171 therefore the exact extent of demethylation is unclear. Finally, as all of the proposed DNA repair
172 mechanisms involve nicking the DNA backbone, a mechanism to ensure the prevention of
173 double stranded DNA breaks is required. It is with these criteria in mind that we consider the
174 evidence for candidate demethylases.

175

176 *Demethylation proceeds via a base excision repair mechanism in plants*

177 DNA glycosylases function in the first step of BER to cleave the "incorrect" or damaged
178 base from the sugar-phosphate backbone, leaving an abasic (AP) site that is repaired by other
179 enzymes **(Figure 2**). HhH-GPD (helix-hairpin-helix – Gly/Pro/Asp) DNA glycosylases

recognize diverse lesions in a wide range of organisms. There is strong genetic and biochemical evidence indicating that a family of four *Arabidopsis* HhH-GPD DNA glycosylases recognize and remove 5-methylcytosine from DNA, thereby acting as DNA demethylases. Two of the genes, *ROS1*(*REPRESSOR OF SILENCING 1*) and *DME* (*DEMETER*), were discovered in genetic screens that were not targeted at finding demethylases[32, 33]. *DML2* (*DEMETER-LIKE 2)* and *DML3* (*DEMETER-LIKE 3*) were identified based on homology to the founding members[32, 33].

*ROS1* is expressed broadly throughout plant development. It is required to demethylate endogenous loci and can also demethylate silenced transgenes[21, 33, 34]. Genomic methylation in whole plants has been profiled in *ros1 dml2 dml3* triple mutants using multiple methods. Methylation levels are not altered on a genome-wide scale, but these mutants do accumulate methylation at several hundred discrete regions, primarily near genes[20, 21]. In most instances examined, hypermethylation does not alter gene expression[21, 34], suggesting that the enzymes scan the genome and remove methylation whether or not it has functional consequences. This activity is similar to other DNA glycosylases involved in BER, which remove mutated or mismatched bases from DNA regardless of whether they are immediately damaging to cellular processes.

*DME* is expressed in the central cell of the female gametophyte[32] and is required for expression of the imprinted genes *FWA, MEA,* and to a lesser extent *FIS2* and *MPC* in the central cell before fertilization and in the endosperm after fertilization[25, 26, 32, 35, 36]. Hypomethylation of the maternal *MEA* allele in the endosperm is dependent on inheritance of a wild type maternal *DME* allele[25]. It is unknown whether demethylation in the central cell occurs at only a few imprinted loci, or whether the demethylation is more extensive. Maize endosperm is significantly

hypomethylated compared to other tissues, suggesting that many sequences might be subject to demethylation[37].

In biochemical assays, ROS1, DME, DML2, and DML3 can remove 5-methylcytosine from methylated olignucleotides, either in CG or non-CG contexts (both exist in plants). ROS1 and DME also excise thymine from T–G mismatches (the product of 5-methylcytosine deamination), albeit at a somewhat slower rate, but cannot excise U from U–G or U–A mismatches[21, 25, 38, 39]. Both DME and ROS1 are bifunctional DNA glycosylases/lyases, nicking the DNA backbone and producing the characteristic Schiff base intermediate and beta and delta elimination products[25, 38, 39]. The enzymes thus function in short patch BER, replacing only a single base at a time. None of the BER enzymes that further process the AP site have been identified, and *Arabidopsis* lacks a homologue to DNA polymerase β, which replaces the base in other organisms[40] (**Figure 2**).

One concern with a BER-based demethylation mechanism is that the generation of single-stranded breaks on complementary strands, for example when a symmetrical CpG dinucleotide is demethylated, could lead to the formation of double strand DNA breaks (DSBs). BER-dependent DSB formation occurs in *E. coli* with clustered DNA damage[41]. Plant and mammalian BER pathways contain some mechanisms to safeguard against this occurrence. Human AP endonuclease is inhibited from incising an AP site if another AP site or a β-δ single strand break (the product of bifunctional DNA glycosylases) is located at the –1, -3 or +1 position on the opposite strand[42]. DME inefficiently removes 5-methylcytosine if an opposing AP site is present; this inhibition is reduced as the AP site is placed farther away from the 5-methylcytosine[25].

A persistent challenge in the field of DNA repair is to understand how DNA glycosylases recognize their targets among the vast excess of normal bases [43]. This might be a particular problem for 5-methylcytosine DNA glycosylases, as 5-methylcytosine is not a damaged base, and is correctly paired with guanine. This question can probably only be addressed by determining the crystal structure of one of these glycosylases in contact with its substrate. Additionally, how these enzymes work within the context of chromatin has not been explored **(Box 3)**.

*Base Excision Repair and Demethylation in Mammals*

Several different DNA repair based mechanisms have been suggested for animal demethylation. These include processes initiated by DNA glycosylases, DNA methyltransferases, and DNA deaminases.

***Direct Removal of 5-methylcytosine***

Long before the plant 5-methylcytosine DNA glycosylases were discovered, work in animals suggested the existence of similar enzymes. Initially, 5-methylcytosine DNA glycosylase activity was purified from chicken embryos along with T–G mismatch glycosylase activity. The activity, which is RNase-sensitive, preferentially cleaves hemi-methylated double-stranded oligos compared to fully methylated oligos[44]. Subsequent mass spectrometric analysis indicated that the active complex contained a homologue of human thymine DNA glycosylases (TDG)[45]. Purified recombinant protein produced from the chicken TDG can remove 5-methylcytosine, but has 30-40 fold higher activity against T–G mismatches[45]. It appears that cytosine can be flipped into the human TDG (hTDG) active site, but the enzyme does not have enough catalytic power to break the N-glycosidic bond[46]. Another enzyme, MBD4 (methyl-CpG binding domain protein 4), has also been suggested to posses 5-methylcytosine DNA glycosylase

activity in addition to thymine DNA glycosylase activity[47]. MBD4 contains both a N-terminal methyl-binding domain and a C-terminal HhH-GPD DNA glycosylase domain[48]. The human enzyme removes T and U from T–G and U–G mismatches, with greater activity when the mismatches are in a CpG context[48]. Initial characterization of the enzymatic activity did not identify any activity against fully or hemimethylated CpG sites[48]. Zhu *et al*. did detect 5-methylcytosine DNA glycosylase activity from recombinant human protein, but at levels, again, 30-40x lower than T–G mismatch activity[47]. To date, there is no biological evidence to support a role for either MBD4 or TDG functioning as 5-methylcytosine DNA glycosylases *in vivo*. Indeed, active demethylation in the paternal pronucleus appears to take place normally in *Mbd4*-knockout mouse zygotes.

***Enzymatic deamination of 5-methylcytosine followed by BER***

In addition to direct removal of 5-methylcytosine by a DNA glycosylase, demethylation might also be achieved by enzymatic deamination of 5-methylcytosine to thymine, followed by T–G mismatch repair that specifically replaces thymine with cytosine. Two different types of enzymes have been proposed to carry out the first step in this process: cytosine deaminases and DNA methyltransferases.

Cytosine DNA deaminases, which convert cytosine to uracil in nucleic acids, are well known from their roles in RNA editing, viral defense and antibody affinity maturation[49]. Activation-induced deaminase (AID) is responsible for both somatic hypermutation and immunoglobin locus class switch recombination in B lymphocytes[50]. AID and the related deaminase APOBEC1 (apolipoprotein B mRNA editing enzyme, catalytic polypeptide 1) were identified in a screen for cytosine deaminases expressed in mouse oocytes[51]. They are also

expressed in embryonic stem cells, and AID is detected in primordial germ cells[51], making them potential candidates for performing global demethylation **(Box 4)**. *In vitro*, both enzymes have robust 5-methylcytosine deaminase activity, resulting in thymine and hence T–G mismatches in DNA, which can be effectively repaired through the BER pathway[51].

Cytosine and 5-methylcytosine can also be enzymatically deaminated by DNA methyltransferases. DNA methyltransferases are primarily known as enzymes that transfer a methyl group to the C-5 position of cytosine from the methyl donor S-adenosylmethionine (SAM), generating S-adenosylhomocysteine (SAH) as a byproduct. If SAM levels are low or nonexistent, the bacterial methylases M.HpaII, M.EcoRII, and to a lesser extent M.MspI, can deaminate cytosine, generating uracil[52-55] **(Box 4)**. Recent work in mammalian cell lines suggests that deamination by the Dnmt3a and Dnmt3b DNA methyltransferases is a means of achieving fast, active, DNA demethylation at promoters undergoing transcriptional cycling. A subset of the CpG sites in the promoter of an estrogen responsive gene undergo cycles of methylation and demethylation that correlate with transcriptional cycling[56, 57]. Coincident with loss of methylation is the recruitment of Dnmt3a, Dnmt3b, TDG, and other BER enzymes. The authors propose that Dnmt3a and Dnmt3b deaminate 5-methylcytosine, generating thymine, which is repaired by TDG and other enzymes. Dnmt3a and Dnmt3b are known to associate with TDG, although previously this interaction has been interpreted as a means of ensuring remethylation of new cytosines incorporated after spontaneous 5-methylcytosine deamination and BER[58, 59]. TDG interacts with various nuclear receptors, including estrogen receptor alpha (ERα), and can either coactivate or repress transcription[60]; however, TDG-mediated ERα stimulation does not require DNA glycosylase activity[61].

The suggestion that DNA methyltransferases can demethylate DNA raises further questions. Why would the deamination reaction be favored over the methylation reaction, unless SAM is completely absent? This type of demethylation mechanism would require that SAM levels also rapidly cycle *in vivo*, on the order of tens of minutes, without producing inhibitory concentrations of SAH. Considering the crucial importance of SAM in a wide variety of biochemical reactions, it is difficult to understand how these requirements would be fulfilled. Based on their expression patterns, neither Dnmt3a nor Dnmt3b is a particularly good candidate for a demethylase associated with reproduction. Dnmt3b is completely absent from the oocyte before and after fertilization; maternally supplied Dnmt3a is found in the oocyte nucleus before fertilization and in the pronuclei after fertilization, but maternal and paternal pronuclei have not been distinguished [62]. Moreover, Dnmt3a is absent from PGCs during the time period of active demethylation, and Dnmt3b is restricted to the cytoplasm[15, 63].

***Nucleotide Excision Repair and Demethylation***

Another DNA repair pathway, NER, has also been implicated in active DNA demethylation[64]. NER differs from BER in several respects. It is responsible for removing helix-distorting lesions that can stall replication or transcription, such as those induced by UV damage or carcinogens. DNA distortion is recognized by the XPC protein (named for xeroderma pigmentosum), which facilitates formation of a pre-incision complex made up a variety of helicases, DNA binding proteins, and endonucleases. DNA around the lesion is unwound and cleaved on either side by the NER nucleases XPG and XPF, removing a 25-30 nt stretch of DNA that is filled in by a DNA polymerase and sealed by a DNA ligase. Base–base mismatches do not appear to be substrates for NER.

A screen for *Xenopus laevis* cDNAs that activate expression of a methylated reporter plasmid transfected into a human embryonic kidney cell line identified *Gadd45a* (growth arrest and DNA-damage-inducible alpha)[64]. *Gadd45a* is a p53-inducible gene involved in a multitude of cellular processes, including NER, although this involvement appears to be indirect[65,66]. Gadd45a-mediated reporter activation is inhibited by XPG knock-down[64]. Ectopic expression of Gadd45a also correlates with a partial reduction in methylation at the endogenous *Oct4* promoter and a reduction in total cellular 5-methylcytosine content. Conversely, Gadd45□ or XPG knockdown increase 5-methylcytosine content. However, in very similar assays using human *Gadd45a*, activation of methylated reporter plasmids was not detected, nor was the *Oct4* promoter demethylated in an endogenous or reporter plasmid context[67].

Removing 30-nt stretches of DNA by NER could potentially lead to loss of methylation by replacing methylated cytosines with cytosines during fill-in by DNA polymerase. Alternatively, the involvement of XPG and Gadd45□ might reflect roles in BER. XPG stimulates BER of oxidative damage by the bifunctional DNA glycosylase/lyase Nth *in vitro*[68,69], independent of XPG's nuclease activity. Furthermore, DNA damage induced by methyl methanesulfonate, which is repaired exclusively by BER, is repaired more slowly in Gadd45α-null mouse cell lines than in wild type[70]. Given the evidence for involvement of XPG and Gadd45α in both NER and BER, and the known *in vivo* targets of these two pathways, BER is the more plausible mechanism.

**Concluding remarks**

Key questions remain in our understanding of demethylation and demethylases in both flowering plants and mammals (Text Box 5). The conservation of *de novo* and maintenance DNA methylation pathways, and the harnessing of DNA methylation for genomic imprinting in

both plants and animals, leads to the expectation that common mechanisms might also underlie active demethylation. In plants, direct genetic and biochemical evidence demonstrates that demethylation results from BER. However, the lack of orthologous glycosylases, and the absence of genetic evidence tying BER enzymes to global demethylation has impeded progress in understanding whether BER is also responsible for methylation removal in mammals. This situation is likely to change in the near future. Advances in knockdown technologies make possible surrogate forward-genetic screens for functional demethylases, and we expect these will be hotly pursued. More routine use of high-resolution methylation mapping [71] coupled with techniques that can distinguish maternal and paternal genomes[72] will be important for testing at the genomic level global demethylation events observed thus far primarily cytologically. More precisely defining the sequences subject to demethylation in mammals , as has been done in plants[20, 21], is vital to understanding its function. Our knowledge of what is demethylated is still extremely limited – consisting of total methylation content and precise methylation patterns at a few loci in only some of the relevant cells. Unbiased methylation maps that encompass all regions of the genome will be crucial for guiding researchers. Developing methods for isolating precisely staged female or male nuclei from the zygote is also key. Alignment of methylation patterns with high-resolution chromatin profiles generated from wild type and mutant organisms is needed to understand the contribution of demethylation to changes in chromatin structure and vice versa (Text Box 4). These are likely to be fruitful areas of research in both plants and animals. We look forward to the end of an era in which mammalian DNA demethylation is regarded as merely 'colorful'[30], and to the beginning of one in which basic mechanistic insights will emerge.

## Glossary

**Genomic Imprinting**: differential expression of alleles depending on the parent-of-origin. Genomic imprinting is often associated with differential methylation of DNA.

**Totipotency**: the ability to differentiate into any cell type

**Pronucleus**: the nuclei from sperm and egg after fertilization but before fusion.

**Primordial Germ Cells**: diploid germ cell precursors.

**Protamines**: small basic proteins that replace histones during sperm maturation, allowing compaction of DNA into the sperm head.

**Bisulfite Sequencing**: a method to determine methylation at individual cytosines. Treatment of DNA with sodium bisulfite converts cytosine to uracil but does not affect 5-methylcytosine. After conversion, amplification of a region of interest by PCR and subsequent sequencing reveals methylation patterns.

**Angiosperms**: flowering plants. Double fertilization produces the embryo and endosperm, a nutritive tissue that supports the embryo during seed development.

**Round Spermatids**: spermatogenic cells that have not undergone the histone-protamine transistion

**Figure 1: Mammalian Methylation Cycle**

DNA is actively demethylated genome-wide at two points during mammalian development. i) DNA of the paternal pronucleus (blue) is demethylated after fertilization but before the first zygotic division, whereas the female pronucleus (pink) remains unaffected. ii)Both genomes are passively demethylated, except at imprinted genes and some transposons, for several rounds of cell division thereafter. The genome is *de novo* methylated around the blastocyst stage,

beginning in the inner cell mass (ICM). All embryonic lineages differentiate from the ICM. iii) At E11.5-12.5, PGCs are drastically demethylated, erasing parent-of-origin specific imprinting marks. Methylation patterns are reestablished at later stages during gametogenesis by *de novo* methyltransferases. Figure adapted, with permission from Oxford University Press, from Ref [3].

**Figure 2: Short Patch Base Excision Repair by Bifunctional DNA Glycosylase/lyases**

i) Bifunctional DNA glycosylases/lyases recognize the lesion base (purple box), flipping it out of the DNA helix into the active site and cleaving the N-glycosidic bond between the sugar-phosphate backbone and the deoxyribose. The enzyme (green circle) is covalently bound to the opened sugar ring via a nucleophillic lysine (K). ii) The intrinsic AP lyase activity of the enzyme then nicks the DNA backbone. iii) AP endonuclease cleans up the nick, leaving a 3' hydroxyl from which iv) DNA polymerase adds the correct base. DNA ligase seals the nick. The plant enzymes that repair the abasic site after 5-methylcytosine removal are unknown. Plants lack homologues to DNA polymerase β and DNA ligase III, which perform these functions in animals. DNA polymerase λ might substitute for DNA polymerase β.

**Box 1: DNA demethylation at specific mouse loci**

A bisulfite sequencing study of four single copy mouse genes found that they became extensively demethylated after fertilization independent of DNA replication[73]. For two loci where parental origin could be assayed, this change occurred specifically on the paternal allele. A recent study of mouse promoter methylation in mature sperm and embryonic stem and germ cells found that, somewhat surprisingly, the sperm promoter methylome closely resembles that of pluripotent cells except at a few key pluripotency genes[71]. This result underscores the need for more precise studies of DNA methylation at various stages of development. As the vast majority

of methylation occurs at repetitive sequences, assays that measure total methylation levels cannot accurately represent methylation at discrete regulatory sequences.

Not all sequences in the paternal pronucleus are subject to demethylation. Imprinted genes and some retrotransposons resist demethylation. Stella (also called PGC7), a non-specific DNA binding protein expressed in PGCs, oocytes, and embryonic stem cells, provides partial protection against demethylation in both the maternal and paternal pronucleus[74]. Stella accumulates in the cytoplasm of unfertilized eggs, but translocates to both pronuclei after fertilization. In Stella mutant females both the paternal and the maternal genomes lose methylation. DNA methylation is specifically reduced at some, but not all, maternally and paternally methylated imprinted genes. How might Stella protect paternal methylation? This could be achieved if histones persist at imprinted genes in sperm and Stella is directed to nucleosomal DNA.

**Box 2: Angiosperm gametogenesis and fertilization**

Unlike animals, plants do not set aside a germ line during embryogenesis. Instead, reproductive organs such as ovules and anthers, which produce the female and male gametes, develop from floral meristems of mature plants. In ovules, a megaspore mother cell undergoes meiosis to give rise to four haploid gametes, 3 of which die. The remaining cell divides mitotically three times to generate a female gametophyte (or embryo sac) with 7 cells (and 8 nuclei) that resides completely within diploid maternal tissue. These seven cells consist of 3 antipodal cells, a diploid central cell, which contains two nuclei that fuse either before or at the time of fertilization, two synergid cells, and an egg cell (**Figure I**). In the anthers, meiotic division of pollen mother cells produces four haploid spores. A subsequent mitotic division generates a vegetative nucleus (VN) and a generative nucleus, which will divide again to create two sperm

cells (SC) (**Figure I**). One sperm fertilizes the egg cell (EC), giving rise to the diploid embryo, and the other fertilizes the central cell (CC), generating a triploid endosperm. The endosperm is terminally differentiated tissue that nourishes the embryo during later stages of seed development, loosely analogous to the mammalian placenta.

**Box 2, Figure I: Haploid gametophyte formation and double fertilization**

Mitosis following meiosis leads to the formation of female gametophytes and male gametophytes (pollen). One haploid sperm cell (SC) fertilizes the haploid egg cell (EC). This develops into the diploid embryo. The other sperm fertilizes the diploid central cell (CC), generating the triploid endosperm.

**Box 3: DNA repair in the context of chromatin**

The substrate for most biochemical characterizations of DNA glycosylases and BER reactions is naked DNA. A few studies have examined DNA glycosylase activity against mononucleosomal DNA[75]. In this context, excision of target bases is reduced, but not eliminated. SMUG1 (*single-strand-selective monofunctionl uracil-DNA glycosylase 1)*, which removes uracil from U–A and U–G base pairs, removes uracil from a mononucleosomal particle at a rate 9-fold less than from naked DNA[76]. Notably, this does not cause disruption or sliding of the nucleosome, even when U is removed from core DNA. A study of thymine glycol removal by the human bifunctional DNA glycosylase NTH1 found that lesions facing away from the nucleosome particle were excised almost as efficiently as in naked DNA, but that inward facing lesions were excised 10 times less efficiently[77]. This difference was eliminated at physiological concentrations of the enzyme, leading the authors to favor a model whereby damaged DNA is

455 accessed through transient DNA unwrapping, which can be taken advantage of if the enzyme
456 concentration is high[77]. For some reactions, active chromatin remodeling is probably necessary
457 for base excision to occur[78]. The efficiency of the uracil DNA glycosylase UDG has been
458 investigated on folded oligonucleosome arrays. UDG and AP endonuclease removed lesions in
459 linker DNA and in the core only 2-3 fold more slowly than on naked DNA, and did so without
460 disrupting or sliding nucleosomes[79]. This study and others found that DNA polymerase β is
461 strongly inhibited by nucleosomal substrates and required the addition of the yeast chromatin
462 remodellers ISWI and ISW2 in order to synthesize DNA[78, 79]. Thus, it is possible that DNA
463 demethylation by DNA glycosylases can begin *in vivo* without a requirement for extensive
464 chromatin remodeling or disruption, but that completing the repair process leads to disruption of
465 nucleosome–DNA contacts or causes nucleosome sliding. Such activity might facilitate other
466 aspects of epigenetic reprogramming. Around the time of active demethylation in PGCs, several
467 dramatic chromatin changes take place[63, 80, 81]. The linker histone H1 is lost from the DNA, tri-
468 methylated H3K9 and H3K27, as well as other histone modifications, disappear, and the nucleus
469 enlarges to accommodate less condensed chromatin. These changes are transient, and the
470 histones and associated modifications return by late E12.5. Large-scale histone replacement
471 might stem from DNA demethylation by a DNA repair mechanism, or could precede and
472 facilitate it. Further experimentation is required to finely dissect the timing of these two
473 processes.
474
475 **Box 4: Enzymes capable of deaminating 5-methylcytosine**
476
477 AID deaminates 5-methylcytosine in an *E. coli* assay and in *in vitro* oligonucleotide
478 assays, where it acts on single-stranded DNA substrates and prefers 5-methylcytosine in an A/T-

G/A-C sequence context[51]. Apobec1 also deaminates single-stranded methylated oligonucleotides *in vitro*. Although *Aid* knockout mice display the expected immunological phenotypes, no reproductive phenotypes have been reported[50]. If AID does act as a 5-methylcytosine DNA deaminase in oocytes or PGCs, it raises a tricky question: how does a genome maintain integrity if deamination is rampant, especially if both 5-methylcytosine and cytosine are targets? In B cells, change in DNA sequence is the adaptive outcome at the immunoglobin locus, but recent evidence indicates that AID action broadly affects other transcribed genes[82]. This property enhances the candidacy of AID as a DNA demethylase, as a lack of specific targeting could result in global genome-wide demethylation. Most genes deaminated by AID in B cells are subject to high fidelity base-excision and mismatch repair, but other genes, which are often found mutated in B cell tumors are repaired, like the immunoglobin genes, in an error-prone manner[82]. Mutations might accumulate not because the mutational load is too high, but because error-prone repair dominates[82]. Clearly, the accumulation of genic mutations would not be acceptable in the zygotic or PGC genome, and high fidelity repair would have to exclusively dominate.

A DNA methyltransferase-like gene, *rid*, is essential for repeat-induced point mutation in *Neurospora crassa*, a process whereby C-G to T-A mutations accumulate in repeated sequences during the sexual cycle[1]. However, it has not been shown that RID functions as a deaminase during this process. Bacterial DNA methyltransferases do have deaminase activity *in vitro* and in *in vivo* bacterial reversion assays. Like the methylation reaction, deamination is inhibited by SAH. Overall the rate of deamination, even when SAM is absent, is much lower than the rate of methylation[53]. Additionally**,** deamination by M.HpaII and M.EcoRII is negligible at around 0.3 μM SAM[53, 54]. At mammalian physiological concentrations of SAM and SAH, these reactions

would not proceed. In rats the concentration of SAM in various tissues is between 20 and 70 μM, whereas SAH ranges from 3-45 μM[83]. 5-methylcytosine is also a target for deamination by M.EcoRII[84]. The rate of this reaction is lower still than the rate of cytosine deamination but can proceed at higher levels of SAM (up to 50 μM) and SAH (10 μM) than can the deamination reaction against cytosine.

**Text Box 5: Key Questions**

Does base excision repair have a role in mammalian demethylation?

Are 5-methylcytosine DNA glycosylases targeted to their substrate? Are they targeted by RNAs[85]?

Do demethylases act as part of complexes with other proteins? What are they?

What chromatin changes precede or accompany demethylation and how do they facilitate the process?

**ACKNOWLEGDEMENTS**

The authors thank their colleagues for discussions. WR also thanks BBSRC, MRC, EU NoE The Epigenome, CellCentric, and TSB for funding. MG is a Howard Hughes Medical Institute Fellow of the Life Sciences Research Foundation.

**REFERENCES**

1.Goll, M.G., and Bestor, T.H. (2005) Eukaryotic cytosine methyltransferases. *Annu Rev Biochem* 74, 481-514

2.Reik, W., *et al*. (2001) Epigenetic reprogramming in mammalian development. *Science* 293, 1089-1093

3.Morgan, H.D., *et al*. (2005) Epigenetic reprogramming in mammals. *Hum Mol Genet* 14 Spec No 1, R47-58

4.Howlett, S.K., and Reik, W. (1991) Methylation levels of maternal and paternal genomes during preimplantation development. *Development* 113, 119-127

5.Monk, M., *et al*. (1987) Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development. *Development* 99, 371-382

6.McLay, D.W., and Clarke, H.J. (2003) Remodelling the paternal chromatin at fertilization in mammals. *Reproduction* 125, 625-633

7.Mayer, W., *et al*. (2000) Demethylation of the zygotic paternal genome. *Nature* 403, 501-502

8.Santos, F., *et al*. (2002) Dynamic reprogramming of DNA methylation in the early mouse embryo. *Dev Biol* 241, 172-182

9. Dean, W., *et al*. (2001) Conservation of methylation reprogramming in mammalian development: aberrant reprogramming in cloned embryos. *Proc Natl Acad Sci U S A* 98, 13734-13738

10.Beaujean, N., *et al*. (2004) Non-conservation of mammalian preimplantation methylation dynamics. *Curr Biol* 14, R266-267

11.Lepikhov, K., Zakhartchenko, V., Hao, R., Yang, F., Wrenzycki, C., Niemann, H., Wolf, E., Walter, J. (2008) Evidence for conserved DNA and histone H3 methylation reprogramming in mouse, bovine, and rabbit zygotes. *Epigenetics & Chromatin* 1, 8

12.Beaujean, N., *et al*. (2004) The effect of interspecific oocytes on demethylation of sperm DNA. *Proc Natl Acad Sci U S A* 101, 7636-7640

13. Rougier, N., *et al*. (1998) Chromosome methylation patterns during mammalian preimplantation development. *Genes Dev* 12, 2108-2113

14.Lee, J., *et al*. (2002) Erasing genomic imprinting memory in mouse clone embryos produced from day 11.5 primordial germ cells. *Development* 129, 1807-1817

15.Hajkova, P., *et al*. (2002) Epigenetic reprogramming in mouse primordial germ cells. *Mech Dev* 117, 15-23

16.Lane, N., *et al*. (2003) Resistance of IAPs to methylation reprogramming may provide a mechanism for epigenetic inheritance in the mouse. *Genesis* 35, 88-93

17.Tam, P.P., and Snow, M.H. (1981) Proliferation and migration of primordial germ cells during compensatory growth in mouse embryos. *J Embryol Exp Morphol* 64, 133-147

18.Gehring, M., and Henikoff, S. (2008) DNA Methylation and Demethylation in Arabidopsis: May 23, 2008. *The Arabidopsis Book*, Rockville, MD: American Society of Plant Biologists, doi:10.1199/tab.0102, http://www.aspb.org/publications/arabidopsis.

19.Zilberman, D., *et al*. (2007) Genome-wide analysis of Arabidopsis thaliana DNA methylation uncovers an interdependence between methylation and transcription. *Nat Genet* 39, 61-69

20.Lister, R., *et al*. (2008) Highly integrated single-base resolution maps of the epigenome in Arabidopsis. *Cell* 133, 523-536

21.Penterman, J., *et al*. (2007) DNA demethylation in the Arabidopsis genome. *Proc Natl Acad Sci U S A* 104, 6752-6757

22.Huh, J.H., *et al*. (2008) Cellular programming of plant gene imprinting. *Cell* 132, 735-744

23.Gutierrez-Marcos, J.F., *et al*. (2006) Epigenetic asymmetry of imprinted genes in plant gametes. *Nat Genet* 38, 876-878

24.Hermon, P., *et al*. (2007) Activation of the imprinted Polycomb Group Fie1 gene in maize endosperm requires demethylation of the maternal allele. *Plant Mol Biol* 64, 387-395

25.Gehring, M., *et al*. (2006) DEMETER DNA glycosylase establishes MEDEA polycomb gene self-imprinting by allele-specific demethylation. *Cell* 124, 495-506

26.Kinoshita, T., *et al*. (2004) One-way control of FWA imprinting in Arabidopsis endosperm by DNA methylation. *Science* 303, 521-523

27.Jullien, P.E., *et al*. (2008) Retinoblastoma and its binding partner MSI1 control imprinting in Arabidopsis. *PLoS Biol* 6, e194

28.Kishigami, S., *et al*. (2006) Epigenetic abnormalities of the mouse paternal zygotic genome associated with microinsemination of round spermatids. *Dev Biol* 289, 195-205

29.Polanski, Z., *et al*. (2008) Hypomethylation of paternal DNA in the late mouse zygote is not essential for development. *Int J Dev Biol* 52, 295-298

30.Ooi, S.K., and Bestor, T.H. (2008) The colorful history of active DNA demethylation. *Cell* 133, 1145-1148

31.Bhattacharya, S.K., *et al*. (1999) A mammalian protein with specific demethylase activity for mCpG DNA. *Nature* 397, 579-583

32.Choi, Y., *et al*. (2002) DEMETER, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in arabidopsis. *Cell* 110, 33-42

33.Gong, Z., *et al*. (2002) ROS1, a repressor of transcriptional gene silencing in Arabidopsis, encodes a DNA glycosylase/lyase. *Cell* 111, 803-814

34. Zhu, J., *et al*. (2007) The DNA glycosylase/lyase ROS1 functions in pruning DNA methylation patterns in Arabidopsis. *Curr Biol* 17, 54-59

591 35.Jullien, P.E., *et al*. (2006) Maintenance of DNA methylation during the Arabidopsis life cycle
592 is essential for parental imprinting. *Plant Cell* 18, 1360-1372

593 36.Tiwari, S., *et al*. (2008) MATERNALLY EXPRESSED PAB C-TERMINAL, a Novel
594 Imprinted Gene in Arabidopsis, Encodes the Conserved C-Terminal Domain of Polyadenylate
595 Binding Proteins. *Plant Cell* 20, 2387-2398

596 37.Lauria, M., *et al*. (2004) Extensive maternal DNA hypomethylation in the endosperm of Zea
597 mays. *Plant Cell* 16, 510-522

598 38.Agius, F., *et al*. (2006) Role of the Arabidopsis DNA glycosylase/lyase ROS1 in active DNA
599 demethylation. *Proc Natl Acad Sci U S A* 103, 11796-11801

600 39.Morales-Ruiz, T., *et al*. (2006) DEMETER and REPRESSOR OF SILENCING 1 encode 5-
601 methylcytosine DNA glycosylases. *Proc Natl Acad Sci U S A* 103, 6853-6858

602 40.Britt, A. (2002) Repair of Damaged Bases: April 4, 2002. *The Arabidopsis Book*, Rockville,
603 MD: American Society of Plant Biologists, doi:10.119/tab.0005,
604 http://www.aspb.org/publications/arabidopsis.

605 41.Blaisdell, J.O., and Wallace, S.S. (2001) Abortive base-excision repair of radiation-induced
606 clustered DNA lesions in Escherichia coli. *Proc Natl Acad Sci U S A* 98, 7426-7430

607 42.David-Cordonnier, M.H., *et al*. (2002) Efficiency of incision of an AP site within clustered
608 DNA damage by the major human AP endonuclease. *Biochemistry* 41, 634-642

609 43.Fromme, J.C., *et al*. (2004) DNA glycosylase recognition and catalysis. *Curr Opin Struct Biol*
610 14, 43-49

611 44.Jost, J.P., *et al*. (1995) Mechanisms of DNA demethylation in chicken embryos. Purification
612 and properties of a 5-methylcytosine-DNA glycosylase. *J Biol Chem* 270, 9734-9739

613 45.Zhu, B., *et al*. (2000) 5-methylcytosine-DNA glycosylase activity is present in a cloned G/T
614 mismatch DNA glycosylase associated with the chicken embryo DNA demethylation complex.
615 *Proc Natl Acad Sci U S A* 97, 5135-5139

616 46.Bennett, M.T., *et al*. (2006) Specificity of human thymine DNA glycosylase depends on N-
617 glycosidic bond stability. *J Am Chem Soc* 128, 12510-12519

618 47.Zhu, B., *et al*. (2000) 5-Methylcytosine DNA glycosylase activity is also present in the human
619 MBD4 (G/T mismatch glycosylase) and in a related avian sequence. *Nucleic Acids Res* 28, 4157-
620 4165

621 48.Hendrich, B., *et al*. (1999) The thymine glycosylase MBD4 can bind to the product of
622 deamination at methylated CpG sites. *Nature* 401, 301-304

623 49.Conticello, S.G. (2008) The AID/APOBEC family of nucleic acid mutators. *Genome Biol* 9,
624 229

50.Muramatsu, M., *et al*. (2000) Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell* 102, 553-563

51. Morgan, H.D., *et al*. (2004) Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: implications for epigenetic reprogramming. *J Biol Chem* 279, 52353-52360

52.Bandaru, B., *et al*. (1995) HpaII methyltransferase is mutagenic in Escherichia coli. *J Bacteriol* 177, 2950-2952

53.Shen, J.C., *et al*. (1992) High frequency mutagenesis by a DNA methyltransferase. *Cell* 71, 1073-1080

54. Wyszynski, M., *et al*. (1994) Cytosine deaminations catalyzed by DNA cytosine methyltransferases are unlikely to be the major cause of mutational hot spots at sites of cytosine methylation in Escherichia coli. *Proc Natl Acad Sci U S A* 91, 1574-1578

55. Zingg, J.M., *et al*. (1998) Enzyme-mediated cytosine deamination by the bacterial methyltransferase M.MspI. *Biochem J* 332 (Pt 1), 223-230

56.Kangaspeska, S., *et al*. (2008) Transient cyclical methylation of promoter DNA. *Nature* 452, 112-115

57.Metivier, R., *et al*. (2008) Cyclical DNA methylation of a transcriptionally active promoter. *Nature* 452, 45-50

58.Boland, M.J., and Christman, J.K. (2008) Characterization of Dnmt3b:thymine-DNA glycosylase interaction and stimulation of thymine glycosylase-mediated repair by DNA methyltransferase(s) and RNA. *J Mol Biol* 379, 492-504

59.Li, Y.Q., *et al*. (2007) Association of Dnmt3a and thymine DNA glycosylase links DNA methylation with base-excision repair. *Nucleic Acids Res* 35, 390-400

60.Kovtun, I.V., and McMurray, C.T. (2007) Crosstalk of DNA glycosylases with pathways other than base excision repair. *DNA Repair (Amst)* 6, 517-529

61.Chen, D., *et al*. (2003) T:G mismatch-specific thymine-DNA glycosylase potentiates transcription of estrogen-regulated genes through direct interaction with estrogen receptor alpha. *J Biol Chem* 278, 38586-38592

62.Hirasawa, R., *et al*. (2008) Maternal and zygotic Dnmt1 are necessary and sufficient for the maintenance of DNA methylation imprints during preimplantation development. *Genes Dev* 22, 1607-1616

63.Seki, Y., *et al*. (2005) Extensive and orderly reprogramming of genome-wide chromatin modifications associated with specification and early development of germ cells in mice. *Dev Biol* 278, 440-458

64.Barreto, G., *et al*. (2007) Gadd45a promotes epigenetic gene activation by repair-mediated DNA demethylation. *Nature* 445, 671-675

65.Kazantsev, A., and Sancar, A. (1995) Does the p53 up-regulated Gadd45 protein have a role in excision repair? *Science* 270, 1003-1004; author reply 1005-1006

66.Kearsey, J.M., *et al*. (1995) Does the p53 up-regulated Gadd45 protein have a role in excision repair? *Science* 270, 1004-1005; author reply 1005-1006

67.Jin, S.G., *et al*. (2008) GADD45A does not promote DNA demethylation. *PLoS Genet* 4, e1000013

68.Bessho, T. (1999) Nucleotide excision repair 3' endonuclease XPG stimulates the activity of base excision repairenzyme thymine glycol DNA glycosylase. *Nucleic Acids Res* 27, 979-983

69.Klungland, A., *et al*. (1999) Base excision repair of oxidative DNA damage activated by XPG protein. *Mol Cell* 3, 33-42

70.Jung, H.J., *et al*. (2007) Base excision DNA repair defect in Gadd45a-deficient cells. *Oncogene* 26, 7517-7525

71.Farthing, C.R., *et al*. (2008) Global mapping of DNA methylation in mouse promoters reveals epigenetic reprogramming of pluripotency genes. *PLoS Genet* 4, e1000116

72.Babak, T., *et al*. (2008) Global survey of genomic imprinting by transcriptome sequencing. *Curr Biol* 18, 1735-1741.

73.Oswald, J., *et al*. (2000) Active demethylation of the paternal genome in the mouse zygote. *Curr Biol* 10, 475-478

74.Nakamura, T., *et al*. (2007) PGC7/Stella protects against DNA demethylation in early embryogenesis. *Nat Cell Biol* 9, 64-71

75.Jagannathan, I., *et al*. (2006) Base excision repair in nucleosome substrates. *Chromosome Res* 14, 27-37

76.Nilsen, H., *et al*. (2002) DNA base excision repair of uracil residues in reconstituted nucleosome core particles. *Embo J* 21, 5943-5952

77.Prasad, A., *et al*. (2007) Initiation of base excision repair of oxidative lesions in nucleosomes by the human, bifunctional DNA glycosylase NTH1. *Mol Cell Biol* 27, 8442-8453

78.Menoni, H., *et al*. (2007) ATP-dependent chromatin remodeling is required for base excision repair in conventional but not in variant H2A.Bbd nucleosomes. *Mol Cell Biol* 27, 5949-5956

79.Nakanishi, S., *et al*. (2007) Different structural states in oligonucleosomes are required for early versus late steps of base excision repair. *Nucleic Acids Res* 35, 4313-4321

80.Hajkova, P., *et al*. (2008) Chromatin dynamics during epigenetic reprogramming in the mouse germ line. *Nature* 452, 877-881

81.Seki, Y., *et al*. (2007) Cellular dynamics associated with the genome-wide epigenetic reprogramming in migrating primordial germ cells in mice. *Development* 134, 2627-2638

82.Liu, M., *et al*. (2008) Two levels of protection for the B cell genome during somatic hypermutation. *Nature* 451, 841-845

83. Eloranta, T.O. (1977) Tissue distribution of S-adenosylmethionine and S-adenosylhomocysteine in the rat. Effect of age, sex and methionine administration on the metabolism of S-adenosylmethionine, S-adenosylhomocysteine and polyamines. *Biochem J* 166, 521-529

84. Yebra, M.J., and Bhagwat, A.S. (1995) A cytosine methyltransferase converts 5-methylcytosine in DNA to thymine. *Biochemistry* 34, 14752-14757

85.Zheng, X., *et al*. (2008) ROS3 is an RNA-binding protein required for DNA demethylation in Arabidopsis. *Nature* 455, 1259-1262

Figure 1

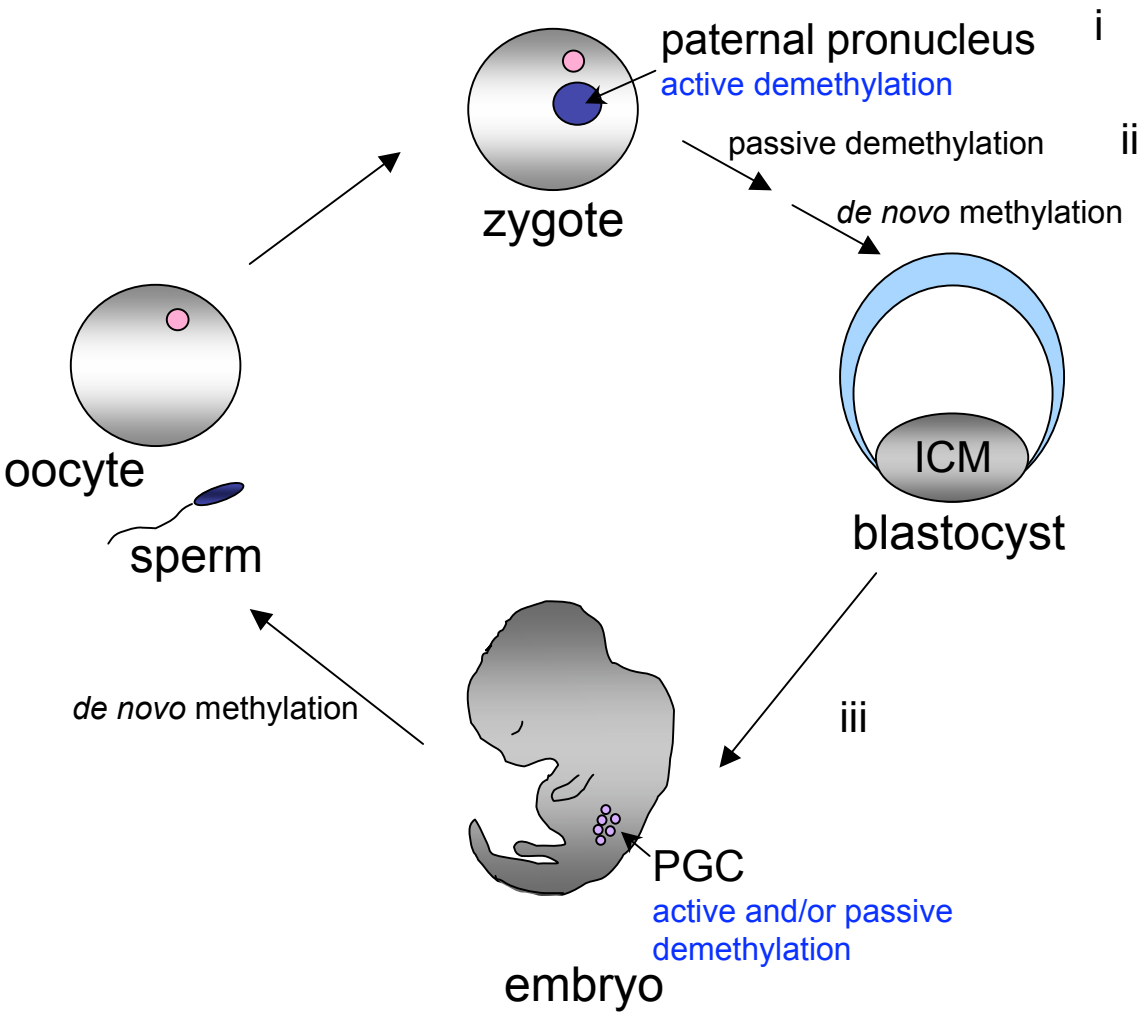
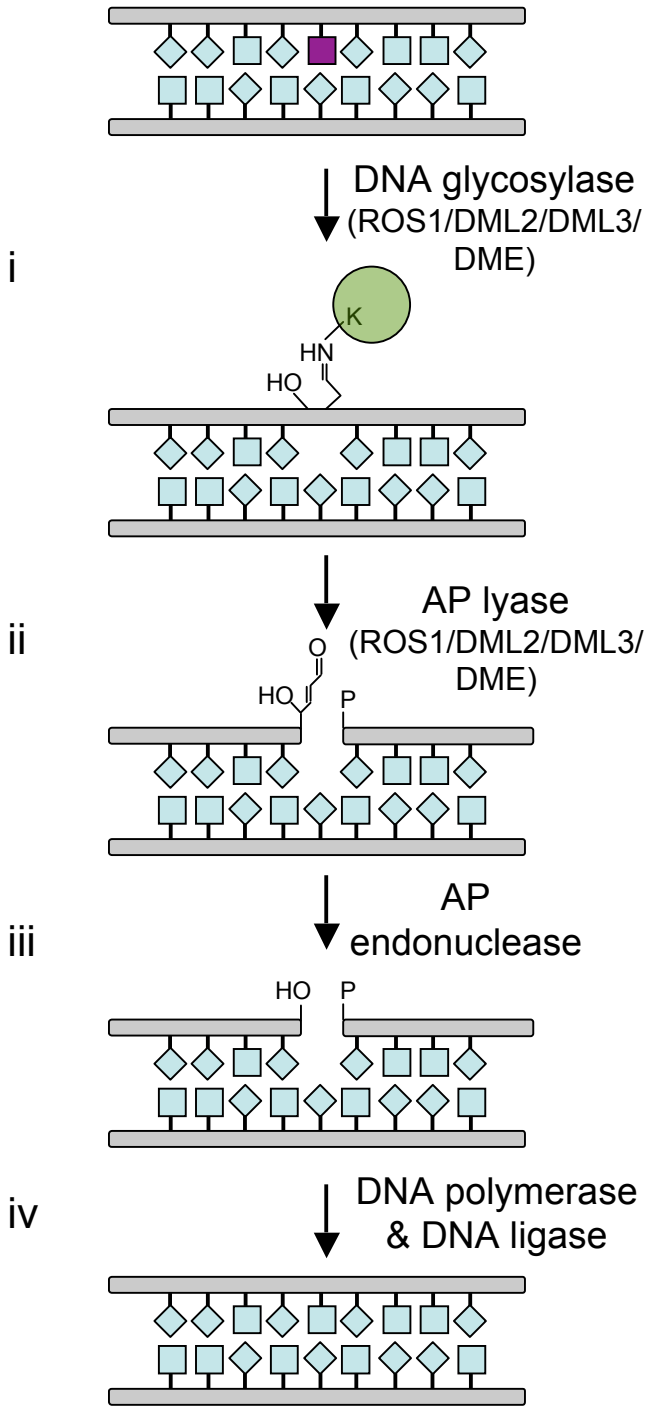

paternal pronucleus
i
active demethylation
passive demethylation
ii
de novo methylation
zygote
ICM
blastocyst
oocyte
sperm
de novo methylation
iii
PGC
active and/or passive
demethylation
embryo

Figure 2


DNA glycosylase
(ROS1/DML2/DML3/
DME)
i
K
HN
HO
AP lyase
(ROS1/DML2/DML3/
DME)
ii
O
HO
P
AP
endonuclease
iii
HO
P
DNA polymerase
& DNA ligase
iv

Box 2, Figure I