REPRODUCTION
REVIEW

Epigenetic reprogramming: is deamination key to active DNA demethylation?

Marta Teperek-Tkacz[1,2], Vincent Pasque[1,2], George Gentsch[3] and Anne C Ferguson-Smith[4]

[1]*Wellcome Trust/Cancer Research UK Gurdon Institute, Cambridge, CB2 1QN, UK,* [2]*Department of Zoology, University of Cambridge, Cambridge, CB2 2EJ, UK,* [3]*Division of Systems Biology, MRC National Institute for Medical Research, London, NW7 1AA, UK and* [4]*Department of Physiology, Development and Neuroscience, University of Cambridge, Cambridge, CB2 3EG, UK*

Correspondence should be addressed to M Teperek-Tkacz at Wellcome Trust/Cancer Research UK Gurdon Institute; Email: mt446@cam.ac.uk

Abstract

DNA demethylation processes are important for reproduction, being central in epigenetic reprogramming during embryonic and germ cell development. While the enzymes methylating DNA have been known for many years, identification of factors capable of mediating active DNA demethylation has been challenging. Recent findings suggest that cytidine deaminases may be key players in active DNA demethylation. One of the most investigated candidates is activation-induced cytidine deaminase (AID), best known for its role in generating secondary antibody diversity in B cells. We evaluate evidence for cytidine deaminases in DNA demethylation pathways in vertebrates and discuss possible models for their targeting and activity regulation. These findings are also considered along with alternative demethylation pathways involving hydroxymethylation.

Reproduction (2011) **142** 621–632

Introduction

The fifth carbon of cytosine in DNA can be either unmethylated or methylated to form 5-methylcytosine (5meC) and 5-hydroxymethylcytosine (5hmeC) on subsequent hydroxylation. This occurs on cytosines flanked by various nucleotides (Lister *et al*. 2009), but 5meC in a CpG dinucleotide context is the best characterised (Doerfler 2008). CpG dinucleotides are often enriched in promoters of genes and their methylation is associated with gene silencing (Chen *et al*. 2001, Kroft *et al*. 2001, Chan *et al*. 2004, Song *et al*. 2009). The opposite process, demethylation or replacement of methylated cytosines with unmethylated cytosines, can restore gene expression (Benvenuto *et al*. 1996, Papageorgis *et al*. 2010, Stengel *et al*. 2010). DNA methylation can therefore be used to establish correct gene expression patterns during development and differentiation (Maatouk *et al*. 2006, Song *et al*. 2009). Demethylation can be achieved by both passive and active mechanisms. Passive demethylation relies on DNA replication in the absence of DNA methyltransferase (DNMT) maintenance activity so that unmethylated cytosines are incorporated into new DNA strands. During active demethylation, methylated cytosines are replaced with unmethylated cytosines by an enzymatic process independent of DNA replication.

Active DNA demethylation is believed to occur on a global scale twice during mouse embryogenesis. First, the paternal genome is actively demethylated before the first cell division of the mouse zygote (Mayer *et al*. 2000, Oswald *et al*. 2000, Santos *et al*. 2002, Okada *et al*. 2010, Wossidlo *et al*. 2010). A second wave of global DNA demethylation occurs in primordial germ cells (PGCs) between embryonic days 11.25 and 13.5 (Hajkova *et al*. 2002, Feng *et al*. 2010, Surani & Hajkova 2010). Both events are likely to be involved in re-setting the genome for early development. Indeed, locus-specific DNA demethylation is required for reactivation of pluripotency genes during cell reprogramming (Simonsson & Gurdon 2004, Bhutani *et al*. 2010). It is also known that cloned embryos have defects in DNA methylation (Dean *et al*. 2001, Kang *et al*. 2001), which may result from aberrant gene reprogramming and lead to developmental abnormalities. Furthermore, demethylation of oncogenes is often associated with cancers (Nishigaki *et al*. 2005). Hence, DNA methylation dynamics are at the core of many developmentally regulated processes and their misregulation can lead to developmental defects and disease.

Potential DNA demethylases in vertebrates

The establishment of DNA methylation is achieved by DNMT enzymes that are well characterised in plants and animals. The mammalian DNMTs and their activity, specificity and regulation have been extensively

ISSN 1470–1626 (paper) 1741–7899 (online)
DOI: 10.1530/REP-11-0148
Online version via www.reproduction-online.org

reviewed (Bestor 2000, Hermann *et al*. 2004, Turek-Plewa & Jagodzinski 2005, Cheng & Blumenthal 2008). In contrast, enzymes involved in DNA demethylation have been identified in plants but their mammalian equivalents have been the subject of controversy. The bifunctional DNA glycosylases repressor of silencing 1 (ROS1) and Demeter (DME) are known to be the first enzymes in the demethylation pathway in plants (Fig. 1). They first recognise and bind methylated cytosine and then excise it from DNA through hydrolytic cleavage. This creates an abasic site that can be filled with an unmethylated cytosine by the DNA repair machinery (Zhu 2009). The mammalian glycosylases thymine DNA glycosylase (TDG) and methyl-CpG-binding domain protein 4 (MBD4) can both efficiently hydrolyse the N-glycosidic bond between thymine and deoxyribose, eventually leading to the base removal (Neddermann & Jiricny 1994, Hendrich *et al*. 1999), and they are capable of cleaving the bond between 5meC and deoxyribose *in vitro*; however, their activity in 5meCs is 30–40 times lower than in thymines (Zhu *et al*. 2000*a*, 2000*b*, Kim *et al*. 2009). *Tdg* deficiency in mice leads to embryonic lethality and aberrant *de novo* DNA methylation of developmentally regulated genes; however, DNA methylation levels at fertilisation and in PGCs have not been assessed in these mutant animals (Cortazar *et al*. 2011, Cortellino *et al*. 2011) and 5meC levels are not altered in *Mbd4*-deficient mice (Millar *et al*. 2002, Wong *et al*. 2002). Other proteins have been suggested to possess DNA demethylation activity, including DNMTs (Metivier *et al*. 2008), MBD2 (Bhattacharya *et al*. 1999, Detich *et al*. 2002), MBD3 (Brown *et al*. 2008) and growth arrest and DNA damage-inducible protein alpha (GADD45A; Barreto *et al*. 2007). GADD45A has been shown to contribute to active demethylation of plasmid DNA injected into frog oocytes (Barreto *et al*. 2007, Schafer *et al*. 2010) and locus-specific promoter demethylation in cultured mammalian cells (Schmitz *et al*. 2009, Schafer *et al*. 2010). However, *Gadd45a* mutant mice do not show defects in DNA methylation (Engel *et al*. 2009); hence, its role in DNA demethylation has also been questioned (Jin *et al*. 2008).

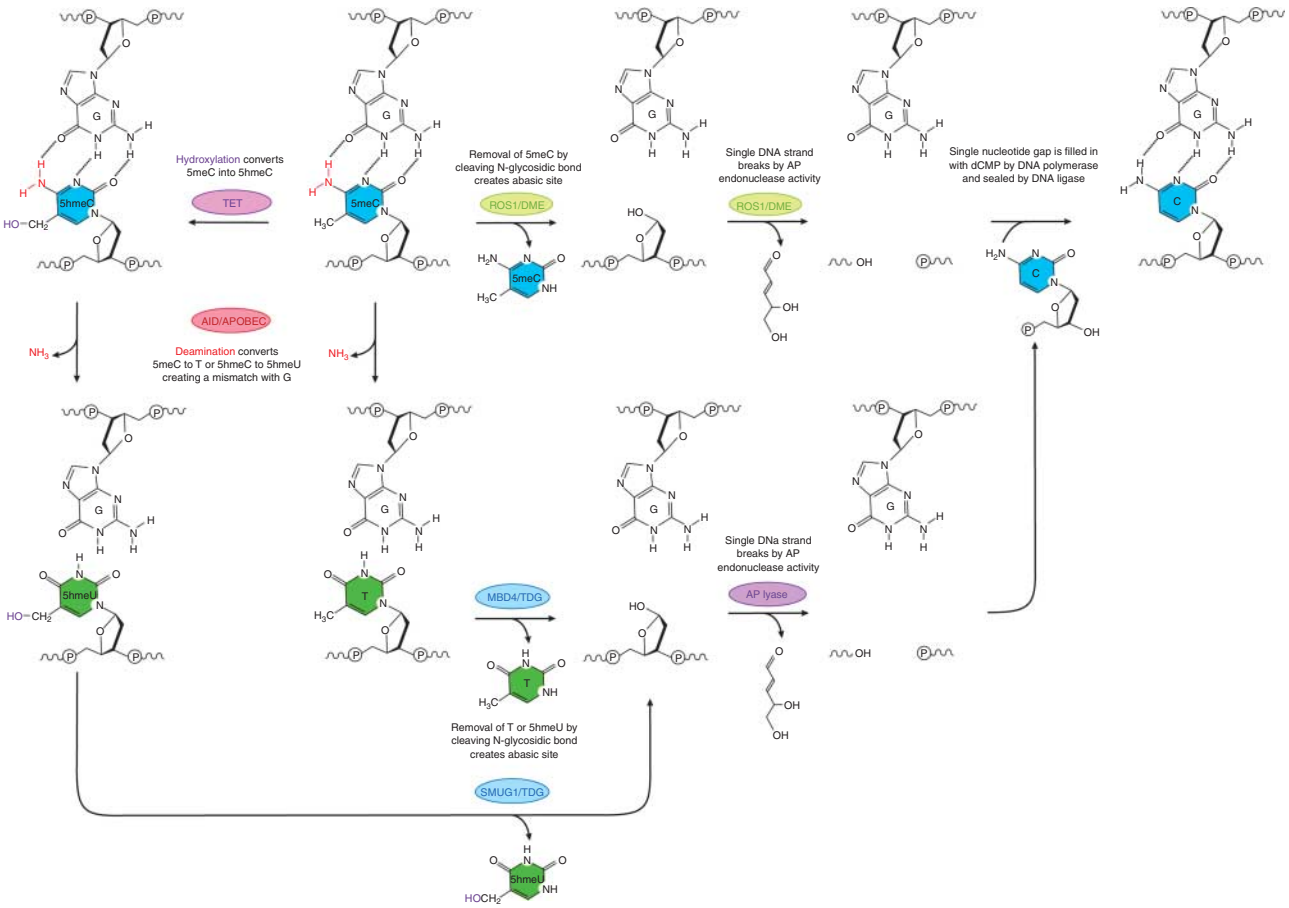

Figure 1 Possible mechanisms of active DNA demethylation. In plants, 5meC can be directly converted to unmethylated C by bifunctional DNA glycosylases ROS1 and DME, whereas in animals 5meC can be first hydroxylated to 5hmeC and subsequently deaminated by AID/APOBEC or directly deaminated by AID/APOBEC. These deamination products (5hmeU or T respectively) are further processed by the BER pathway.

Hydrolytic deamination catalysed by AID/APOBEC family members

Recent reports suggest that some members of the vertebrate-specific activation-induced cytidine deaminase (AID)/apolipoprotein B mRNA editing enzyme, catalytic polypeptide (APOBEC) family such as AID (Morgan *et al*. 2004, Rai *et al*. 2008, Bhutani *et al*. 2010, Popp *et al*. 2010), APOBEC1 (Morgan *et al*. 2004, Guo *et al*. 2011), APOBEC2 (Rai *et al*. 2008, Guo *et al*. 2011) and some proteins of the APOBEC3 branch (Guo *et al*. 2011) could play a key role in active DNA demethylation. AID and APOBEC proteins are zinc-dependent cytidine deaminases acting on single-stranded polynucleotides and deaminating cytosines in different contexts (Navaratnam *et al*. 1993, Teng *et al*. 1993, Bransteitter *et al*. 2003, Chelico *et al*. 2006). The zinc ion acting with the deaminase is coordinated by three amino acids: one histidine and two cysteines or three cysteines (Conticello *et al*. 2005). These residues are conserved within the motif signatures [H/C]xE and $PCx_{2-4}C$ (Gerber & Keller 1999). The latter motif is a zinc finger-like feature, a key to several proteins regulating chromatin modifications including DNA and histone methylation (Blackledge *et al*. 2010). This catalytic zinc site receives its activity by a fourth ligand, a water molecule, which is coordinated by the carboxylate ion of glutamate in [H/C]xE. The carboxylate group facilitates proton shuttling, which converts a water molecule – once it is trapped within the zinc coordination sphere – into a reactive hydroxide ion (Betts *et al*. 1994). Hydrolytic deamination of cytosines occurs through a nucleophilic attack of the zinc hydroxide on the pyrimidine carbon 4 carrying the amine group (Betts *et al*. 1994).

Among the AID/APOBEC family, *Aid* and *Apobec1* are expressed in mammalian oocytes and embryos at stages when global DNA demethylation occurs (Morgan *et al*. 2004). They can deaminate 5meC to thymine *in vitro* which, followed by G-T mismatch repair, could lead to DNA demethylation (Morgan *et al*. 2004). APOBEC1 was originally found to convert cytosine into uracil in the apolipoprotein B transcript (Navaratnam *et al*. 1993, Teng *et al*. 1993), whereas AID catalyses the same base conversion repetitively and preferentially on single-stranded DNA along immunoglobulin loci (Muramatsu *et al*. 1999, 2000, Pham *et al*. 2003). Interestingly, experiments by Rai *et al*. (2008) suggest that coupling of Aid and Apobec2 along with the glycosylase Mbd4 can lead to active DNA demethylation in zebrafish embryos.

Aid/Apobec-driven DNA demethylation in zebrafish

Demethylation of *in vitro* methylated circular or linearised DNA occurs when it is injected into zebrafish embryos at the one-cell stage. This is followed by its remethylation several hours later (Collas 1998, Rai *et al*. 2008). Recent overexpression and knockdown studies suggest that the presence of Aid/Apobec2 together with the Mbd4 DNA glycosylase and Gadd45 is responsible for the demethylation of plasmid DNA and concomitant demethylation of the embryonic genome (Rai *et al*. 2008). Knockdown of Aid or Mbd4 results in locus-specific hypermethylation and aberrant expression of genes important for neurogenesis. The repair of G-T mismatches is meant to be immediately initiated by the glycosylase Mbd4. Gadd45 may increase the DNA demethylation efficiency by promoting the Aid/Apobec2 physical interaction with Mbd4 (Rai *et al*. 2008). These results suggest that DNA glycosylases, known for their roles in DNA demethylation in plants, can also contribute to this process in vertebrates. However, according to this model, DNA glycosylases do not initiate the removal of 5meC as in plants, but only cut the N-glycosidic bond leading to the removal of thymines from G-T mismatches (Fig. 1). This may explain the lack of change in 5meC levels in *Mbd4* knockout mice, which instead have a higher frequency of G-T mutations (Millar *et al*. 2002, Wong *et al*. 2002).

While the model proposed by Rai *et al*. is appealing, caveats remain in the mechanistic understanding of Aid/Apobec2-induced demethylation. First, using cytidine deaminases in DNA demethylation seems risky. Cytidine deamination is mutagenic: besides creating G-T mismatches, the activity of Aid is primarily directed towards unmethylated cytosines, which converts them to uracils producing G-U mismatches, for example as occurs in B cells allowing secondary antibody diversification. Mismatches, if not repaired before DNA replication, will create permanent mutations. Secondly, these experiments have been mostly based on the artificial introduction of methylated plasmid DNA into zebrafish embryos. It is not known why this also leads to partial demethylation of the zebrafish genome itself. It is possible that the up-regulation of Aid/Apobec2 and Gadd45 expression and the stimulation of DNA demethylation is a type of immunological response to the presence of methylated exogenous DNA. Whether such a mechanism is common to other organisms is not yet known. Thirdly, it is not clear how symmetric demethylation of CpG in double-stranded DNA occurs since simultaneous 5meC deamination and subsequent nucleotide removal on both DNA strands would create mutagenic double-stranded breaks (Jiricny & Menigatti 2008). However, it is known that the genomes of vertebrate species such as fugu and zebrafish undergo CpG to TpG transitions over evolutionary time (Bird 1980, Glass *et al*. 2007). It cannot be excluded that these transitions occur through deamination of 5meC in CpG dinucleotides and subsequent erroneous repair. This is substantiated indirectly by the fact that CpG islands, generally protected from DNA methylation, do not show accumulation of C-to-T transitions.

Additionally, the role of Apobec2 remains controversial. The expression of APOBEC2 is mainly confined to

muscle and heart tissue (Liao *et al*. 1999). *Apobec2* knockout mice show a significant loss of body weight, a myofibre-type shift from fast to slow and centronuclear-like myopathy with age (Mikl *et al*. 2005, Sato *et al*. 2010). Recently, Apobec2 has been further implicated in left–right axis determination during early embryogenesis in *Xenopus* through an inhibition of the transforming growth factor β signalling pathway (Vonica *et al*. 2011). It is not clear how Apobec2 contributes to such a variety of developmental processes. Furthermore, its role as a potential DNA deaminase has been questioned due to lack of crucial residues (such as the positively charged amino acids found in AID) at the N-terminus facilitating binding to single-stranded DNA, and a tryptophan, which is located in the vicinity of the PCxxC motif and is required for APOBEC3G-catalysed DNA deamination (Pham *et al*. 2003, Chen *et al*. 2007, Sato *et al*. 2010).

Evidence for AID as a mammalian DNA demethylase

Two recent studies support a role for AID in DNA demethylation in mammalian systems (Bhutani *et al*. 2010, Popp *et al*. 2010). Bhutani *et al*. have shown that the use of siRNA against *Aid* interferes with reprogramming, demethylation and reactivation of the pluripotency genes *POU5F1* (*OCT4*) and *NANOG* in heterokaryons (fused cells containing multiple, genetically different nuclei) of human fibroblasts and mouse embryonic stem cells. Chromatin immunoprecipitation experiments implied that AID is bound to silent promoters of human somatic and mouse ES cells, but not to active, already unmethylated promoters in ES cell nuclei (Bhutani *et al*. 2010). These findings suggest a role for AID in promoter DNA demethylation, but leave the question open how AID associates with chromatin without causing immediate deamination in silent promoters. Moreover, the ES cells used for the fusions were actively dividing. Active divisions imply DNA replication that can therefore lead to passive DNA demethylation. It would be interesting to test whether AID is also important for active DNA demethylation in heterokaryons generated with cell cycle synchronised ES cells. In order to further confirm an involvement of AID in active DNA demethylation in heterokaryons, it would be worth performing the experiment using *Aid* mutant cells.

Genetic evidence for the involvement of AID in DNA demethylation has been obtained from examining the DNA methylation level in PGCs from *Aid* knockout mice (Popp *et al*. 2010). Although the data reporting *Aid* expression in PGCs at the time of global DNA demethylation has been challenged (Morgan *et al*. 2004, Hajkova *et al*. 2010), in $Aid^{-/-}$ mice, the erasure of DNA methylation marks in PGCs is up to three times lower compared with wild-type controls. Intriguingly, considerable DNA demethylation still occurs in $Aid^{-/-}$ PGCs (Popp *et al*. 2010). This suggests that residual DNA demethylation results from the activity of other deaminases such as the co-expressed *Apobec1* (Morgan *et al*. 2004, Hajkova *et al*. 2010), other mechanisms that do not require deamination, or a combination of both. Similarly, reduced levels of DNA demethylation in zebrafish have only been observed after simultaneous knockdown of Aid and Apobec2 (Rai *et al*. 2008), suggesting redundancy among members of the AID/APOBEC family.

A consensus mechanism of deaminase-mediated DNA demethylation

The results of most of the experiments described above provide an emerging consensus as to how cytidine deaminases can initiate the removal of 5meC from DNA (Fig. 1). The hydrolytic deamination of 5meC converts the base to thymine. It is important to note, however, that cytidine deaminases preferentially act on unmethylated cytosines in DNA. Both activities are considered mutagenic. If cytidine deaminases act on cytosine, the resultant uracil is recognised and repaired by uracil DNA glycosylases (Talpaert-Borle *et al*. 1982, Olsen *et al*. 1989). Thymine (T), resulting from deamination of 5meC, is a true base in genomic DNA, so mismatch repair proteins need to distinguish Ts in a mismatch with guanosines from correctly paired Ts in DNA. Thymine DNA glycosylases (TDG/MBD4) are able to selectively recognise such mismatches as they interact not only with the T but also with the opposing base pair (Barrett *et al*. 1998, Yoon *et al*. 2003, Maiti *et al*. 2008). Additionally, MBD4 not only recognises G-T mismatches but also has a methyl-binding domain (Hendrich *et al*. 1999, Wu *et al*. 2003). This domain could target MBD4 to 5meC and mark it as a potential site for deamination. Moreover, GADD45 may couple the action of MBD4 with AID/APOBEC2 (Rai *et al*. 2008). Physical interaction of deaminases with glycosylases could be critical for immediate recognition of mismatches generated by deamination of 5meC, therefore preventing the mutagenic activity of deaminases. Thymine DNA glycosylases (TDG/MBD4) are the first enzymes of the base excision repair (BER) pathway, which triggers the removal of the mispaired base T from the DNA (Fig. 1). This is further processed by endonucleases, such as AP-endonuclease 1 (APE1). In a recent screen for factors promoting DNA demethylation, the RING finger protein 4 (RNF4) has been identified. RNF4 has been shown to enhance DNA demethylation by coupling TDG and APE1 (Hu *et al*. 2010). After APE1-mediated hydrolytic cleavage of the phosphodiester DNA backbone, the DNA polymerase β (POLB) removes the remaining deoxyribose moiety and fills in the nascent single nucleotide gap with an unmethylated cytosine. Eventually, the break is sealed by a DNA ligase (Dalhus *et al*. 2009, Kunz *et al*. 2009). Additionally, non-enzymatic proteins such

as X-ray repair cross-complementing proteins 1 and 2 (XRCC1 and XRCC2) create a scaffold for the mismatch repair machinery (Tebbs *et al*. 1999, Adam *et al*. 2007).

Incorrectly paired nucleotides can also be excised by the nucleotide excision repair (NER) pathway. In contrast to the BER pathway, the NER pathway removes an ~29 bp long single-stranded DNA fragment including the incorrectly paired nucleotide(s). Subsequently, new DNA is synthesised using the undamaged strand as a template and DNA ligase creates covalent phosphodiester bonds (reviewed in Niehrs (2009)). It cannot be excluded that at least some mispaired nucleotides resulting from 5meC deamination can be repaired using NER enzymes. Indeed, knockdown of components of the NER pathway (Gadd45a, XPA, XPG and XPF) inhibits DNA demethylation in *Xenopus* oocytes (Barreto *et al*. 2007, Schmitz *et al*. 2009). DNA demethylation is also inhibited by treatment with chemicals specifically blocking the NER pathway, but not the BER pathway, both in *Xenopus* oocytes and cultured HEK293 cells (Schafer *et al*. 2010). Furthermore, it has been shown that DNA demethylation in mouse zygotes leads to the creation of DNA breaks. Interestingly, aphidicolin treatment blocks repair of the breaks (Wossidlo *et al*. 2010). Aphidicolin blocks DNA synthesis but has no inhibitory effects on the BER-specific POLB (the major polymerase of the BER pathway). This could suggest that repair pathways other than BER may be involved in DNA demethylation in mouse zygotes.

Importance of BER enzymes in mouse embryogenesis

If DNA demethylation was to be initiated by deaminases and resultant mismatches processed by DNA repair pathways, then the enzymes from these pathways should also be expressed at times when DNA demethylation occurs. Indeed, it has been reported that MBD4 together with other components of the BER pathway, including *Ape1*, *Polb* and DNA ligase III, are expressed at all stages of mouse preimplantation development (Ruddock-D'Cruz *et al*. 2008, May *et al*. 2009). Moreover, BER enzymes are present in the paternal pronucleus in the zygote and in PGCs at the time of global DNA demethylation (Hajkova *et al*. 2010). *Tdg*, *Polb* as well as *Xrcc1* and *Xrcc2* knockout mice are lethal either at the embryonic (Gu *et al*. 1994, Tebbs *et al*. 1999, Deans *et al*. 2000, Adam *et al*. 2007, Cortazar *et al*. 2007, 2011, Cortellino *et al*. 2011) or at the neonatal stage (Sugo *et al*. 2000), suggesting that the BER pathway is important for embryonic development. The spatial and temporal co-expression of these components and the reported interactions among them (Bennett *et al*. 1997, Vidal *et al*. 2001, Dianova *et al*. 2004, Parsons *et al*. 2005, Fitzgerald & Drohat 2008, Hu *et al*. 2010) suggest that they may form functional deaminase–BER complexes. The existence of such a complex would ensure that any deaminated 5meC is immediately recognised, repaired and replaced with an unmethylated cytosine, hence diminishing the risk of mutating the genome through deamination.

How is site-specific activity of deaminases achieved?

Since the activity of AID is mutagenic, cells maintain a tight control on its nuclear localisation. AID is kept away from DNA by using a strong cytoplasmic retention signal, as well as a strong nuclear export signal (Patenaude *et al*. 2009). AID is also actively imported to the nucleus, and its concentration is thought to be regulated by the proteasome; on proteasome inhibition, a ubiquitinated nuclear form is found (Aoufouchi *et al*. 2008). The mechanisms used to target cytidine deamination to particular sites are not clear. AID/APOBECs could interact with other targeting proteins, such as MBD4 that has a 5meC-binding domain. However, this would result in binding of deaminases to all methylated cytosines in DNA. Intriguingly, Bhutani *et al*. (2010) have detected AID bound to silent promoters both in human fibroblasts and in ES cells, which are not demethylated. It may be that AID binding is not sufficient to trigger its activity; hence, other cues are required. It has been reported that AID can only induce deamination in the context of single-stranded DNA (Larijani & Martin 2007, Larijani *et al*. 2007, Brar *et al*. 2008). It is not known whether melting of double-stranded DNA *in vivo* is sufficient to induce AID activity; however, the presence of single-stranded DNA could represent such a signal. But how could it become activated, if, as in most cases, the DNA in eukaryotic cells is double-stranded? We present three possible models for AID activation and targeting.

Model 1: activation of deaminases by active transcription

It has been suggested that DNA demethylation of silenced genes cannot occur without histone acetylation-induced transcription (D'Alessio *et al*. 2007). Transcription leads to a transient formation of single-stranded DNA (Leibovitch & Harel 1978, Leibovitch *et al*. 1979), and it has been hypothesised that transcription may be needed for AID targeting (Chaudhuri *et al*. 2003, Shen *et al*. 2009). When overexpressed in NIH 3T3 cells, AID more efficiently edits a *GFP* reporter gene transcribed at higher levels, suggesting that transcription may stimulate its activity (Yoshikawa *et al*. 2002). Recently, elongator complex protein 3 (ELP3) and three other proteins from the ELP family comprising the elongator complex have been reported to be necessary for paternal DNA demethylation in mouse zygotes (Okada *et al*. 2010). The elongator complex has been previously found to be associated with RNA polymerase II and involved in

transcriptional elongation (Otero *et al*. 1999). Interestingly, it has been shown that interactions with the elongator complex and transcriptional elongation factor SPT5 may direct AID to transcribed targets (Besmer *et al*. 2006, Pavri *et al*. 2010). SPT5 facilitates transcriptional targeting by delivering AID to stalled RNA polymerase II, which was suggested to occur more frequently in the presence of R loop secondary structures (Pavri *et al*. 2010). The importance of transcription for AID targeting is further substantiated by recent findings that the RNA exosome complex recruits AID to both strands of transcribed DNA to ensure simultaneous deamination of template and non-template DNA (Basu *et al*. 2011). The RNA exosome is meant to remove nascent RNA from template DNA to expose it to AID for binding. Subsequently, protein kinase A may stabilise single-stranded DNA and support the recruitment of the repair machinery (Vuong *et al*. 2009). The notion that transcription may be a prerequisite for active DNA demethylation is challenged by findings in zebrafish and mouse embryos. Active demethylation of plasmid DNA in zebrafish occurs in the absence of transcription (Collas 1998). Additionally, the onset of transcription in mouse zygotes at the one-cell stage occurs after the erasure of methyl marks from the paternal genome (Bouniol *et al*. 1995, Aoki *et al*. 1997). To conclude, even though many experiments point towards a central role for transcription in the activity of deaminases making this model an attractive one, it is possible that in some cases deaminases may be regulated using alternative pathways, as discussed below.

Model 2: deaminase targeting and chromatin modification

Changes in chromatin state could be sufficient to target AID/APOBECs. Some histone modifications, like di- and tri-methylation of lysine 9 of histone H3 (H3K9me2/3), are associated with chromatin compaction, whereas others, like acetylated lysine 9 of histone H3 (H3K9ac) or trimethylated lysine 4 of histone H3 (H3K4me3), can make the chromatin more accessible (Jenuwein & Allis 2001). It has been suggested that the presence of H3K9ac, H3K14ac and H3K4me3 may be important for AID targeting (Wang *et al*. 2009). Interestingly, at the time of global DNA demethylation in the mouse zygote, the paternal genome is devoid of repressive H3K9me2/3 marks (Liu *et al*. 2004, Santos *et al*. 2005). Similarly, active DNA demethylation in PGCs occurs after a loss of repressive H3K9me2 (Hajkova *et al*. 2008). Furthermore, the presence of H3K4me2/me3 is associated with pluripotency gene reactivation during cell reprogramming (Murata *et al*. 2010). Thus, it is conceivable that accessibility and state of the chromatin may be important to recruit AID to the sites of deamination.

Model 3: RNA-mediated deaminase targeting

Another appealing mechanism for directing deaminases to specific sites in the genome could involve RNA-mediated targeting. Recently, it has been shown that non-coding RNA can bind to a complementary rDNA promoter region to form a triple helix. Formation of DNA:RNA triplexes facilitates recruitment of the DNMT3B methyltransferase to the rDNA promoter (Schmitz *et al*. 2010). Enzyme targeting by interaction with RNAs is attractive, as it ensures a high degree of site specificity. It could be that some of the reported requirements of transcription for AID activity might reflect a need for the generation of guiding non-coding RNA (Chaudhuri *et al*. 2003, Shen *et al*. 2009). However, so far, there is no evidence supporting this hypothesis.

With the current state of knowledge, it is difficult to decide which of the proposed models (if any) is true. It could be that all, or some combination of them, are utilised at different developmental stages or in different model systems. It should also be noted that most of our current mechanistic understanding about AID is based on its immunological role and may not be relevant for deaminase-mediated DNA demethylation. More work has to be done in order to shed light on 5meC deaminase targeting, regulation and the reciprocal relations between active DNA demethylation, chromatin changes, transcription and the requirement of single-stranded DNA for deaminase activity.

One universal mechanism or several independent ones?

Despite growing evidence for the involvement of AID and other cytidine deaminases in DNA demethylation, many questions still remain (Box 1). First, there is a need for more genetic evidence for the role of cytidine deaminases in DNA demethylation. Although *Aid* knockout mice have significantly decreased global DNA demethylation levels in PGCs, they are viable and residual demethylation in PGCs is still observed (Popp *et al*. 2010). Because of possible redundancy of AID with other cytidine deaminases, it is important to simultaneously knock out other cytidine deaminases to see whether residual DNA demethylation is still observed and whether paternal DNA demethylation after fertilisation in the zygote still occurs. Furthermore, to examine the involvement of other components of the proposed DNA demethylation pathways, they too will need to be systematically depleted from the examined cells/animals. This will be challenging as the list of potential candidates is large (Table 1) and possible redundancy has to be taken into account when analysing the roles of particular factors. As we cannot exclude that distinct active DNA demethylation mechanisms exist, the interpretation of results between different systems should also be conducted with caution. It is important to

Box 1 Outstanding questions

- How are deaminases targeted to specific promoters?
- Is transcription necessary for active DNA demethylation?
- Is the presence of ssDNA sufficient to trigger deaminase activity?
- How are symmetric 5meCs removed in deaminase-driven pathway?
- Are deaminases involved in the global DNA demethylation, e.g. of the paternal genome in fertilised mouse zygotes? What other components of the deaminase-mediated DNA demethylation pathway are necessary for its efficiency?
- How many pathways are used (and integrated) to mediate active DNA demethylation in mammals?

consider that mechanisms may distinguish between genome-wide and locus-specific DNA demethylation. Moreover, it is very likely that various other enzymes from distinct pathways can act synergistically (or independently) to achieve DNA demethylation.

Hydroxylation and deamination

The involvement of other pathways has received considerable recent attention. In particular, recognition of a contribution of TET proteins to DNA demethylation comes from studies on zygotic reprogramming (Hajkova *et al*. 2010, Wossidlo *et al*. 2011). TET proteins have the ability to convert 5meC to 5hmeC (Tahiliani *et al*. 2009, Ito *et al*. 2010). Interestingly, *Tet1* and *Tet2* are highly expressed in embryonic stem cells and in PGCs at the time of global DNA demethylation and are induced during reprogramming of fibroblasts to induced pluripotent stem cells (Hajkova *et al*. 2010, Koh *et al*. 2011). However, recent findings in stem cells with depletion of TET1 or TET2 demonstrate that the correlation of TET activity and DNA methylation pattern is complex (Ito *et al*. 2010, Ko *et al*. 2010, Koh *et al*. 2011) and TET1 protein, as well as being involved in DNA demethylation and sustaining transcriptional activity of several genes, may also be responsible for the silencing of others (Ficz *et al*. 2011, Williams *et al*. 2011, Wu *et al*. 2011). It is not known whether DNA demethylation in PGCs is impaired in *Tet*-deficient mice. In contrast, *Tet3* is expressed in mouse oocytes and is present in the early zygote. Absence of TET3 during the time of active demethylation results in failure to demethylate the paternal genome, suggesting that the conversion of 5meC to 5hmeC may constitute an intermediate step in the active demethylation process (Iqbal *et al*. 2011, Wossidlo *et al*. 2011).

It is possible that both enzymatic pathways involving deaminases and TET proteins could cooperatively lead to DNA demethylation. Indeed, it has been shown recently that co-expression of TET1 and AID or other cytidine deaminases can increase the efficiency of DNA demethylation of reporter plasmid DNA transfected into cultured cells (Guo *et al*. 2011). AID has been proposed to preferentially deaminate 5hmeC generated by TET1 to produce 5-hydroxymethyl uracil (5hmeU), which can then be processed by glycosylases from the BER pathway (Guo *et al*. 2011). Furthermore, it has been shown recently that TDG can initiate the removal of 5hmeU generated by AID-mediated deamination of 5hmeC (Cortellino *et al*. 2011; Fig. 1). This is an attractive model, because TET proteins, in contrast to cytidine deaminases, can efficiently act on 5meC in double-stranded DNA (Tahiliani *et al*. 2009, Ito *et al*. 2010), and because the product of 5hmeC deamination – 5hmeU – can be recognised by BER pathway uracil glycosylases. However, under at least some developmental circumstances, this orchestrated process of oxidation and deamination may not be required for active DNA demethylation. It has been shown that during DNA demethylation at fertilisation in mouse zygotes, some 5mC is converted to unmethylated cytosine without a 5hmeC intermediate,

Table 1 Example of proteins potentially involved in active DNA demethylation driven by cytidine deaminases.

Cytidine deaminases	Protein coupling deaminases, glycosylases and enzymes from the BER pathway	Proteins from the BER pathway
AID, APOBEC1, APOBEC2A, APOBEC2B, APOBEC3A, APOBEC3B, APOBEC3C, APOBEC3D (known as APOBEC3E), APOBEC3F, APOBEC3G, APOBEC3H, APOBEC4	GADD45A, GADD45B, GADD45G, RNF4	*Thymine DNA glycosylases:* MBD4, TDG *DNA endonucleases:* APE1, APE2 *DNA polymerases:* POLB *DNA ligases:* DNA ligase I, DNA ligase III *Protein coupling BER enzymes:* XRCC1

which would point towards the involvement of DNA demethylation mechanisms independent of TET proteins and possibly using cytidine deaminases (Wossidlo *et al*. 2011). One cannot rule out that various mechanisms and distinct enzymes might act at different developmental stages. For example, active demethylation pathways in the germline may be different from those operating in the zygote or at other developmental stages. Consistent with this, it is known that imprinting marks are not erased in the first wave of global demethylation after fertilisation in mouse zygotes (Mayer *et al*. 2000, Oswald *et al*. 2000), but that they are removed in the second wave occurring in mouse PGCs (Hajkova *et al*. 2002, Sato *et al*. 2003, Surani & Hajkova 2010). In addition, several distinct active DNA demethylation mechanisms may be active at the same time to target different regions of the genome. Clearly, our understanding of active DNA demethylation mechanisms is still limited.

Conclusions

In summary, despite several lines of evidence supporting the hypothesis that DNA demethylation can start with 5meC deamination, it is likely that, *in vivo*, an AID-driven pathway is not the only one leading to active DNA demethylation. It is crucial to get further insights into key mechanisms of DNA demethylation control. Determining the mechanisms of DNA demethylation at a molecular level will be important for understanding the erasure and establishment of the normal epigenetic programme in the germline and in the zygote and will provide insights into the epigenetic perturbations implicated in assisted reproductive technologies (Maher 2005, Laprise 2009). Ultimately, it may contribute to the design of treatments for diseases associated with aberrant DNA methylation, such as imprinting disorders or cancer. For example, even though direct links to genome-wide levels of 5mC have not been established, it has been shown that up-regulation of AID-mediated DNA demethylation pathway is associated with human colon cancers (Rai *et al*. 2010). In addition, deciphering mechanisms leading to DNA demethylation could help circumvent problems with inefficient DNA demethylation accompanying induced pluripotent stem cell derivation (Takahashi & Yamanaka 2006, Lister *et al*. 2011). We are getting closer to an understanding of how active DNA demethylation is achieved on a molecular level, and how cytidine deaminases contribute to this process where repair is desired rather than rejected as observed at immunoglobulin loci to cause antibody diversity in B cells.

Declaration of interest

The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

Funding

This work was supported by the grants from the MRC, Wellcome Trust [089613], [081277], [081562] and a Wallonia-Brussels International Excellence Grant.

Acknowledgements

We would like to thank R Shields, K Miyamoto, J Jullien and J Green for critically reading and commenting on the manuscript.

References

Adam J, Deans B & Thacker J 2007 A role for Xrcc2 in the early stages of mouse development. *DNA Repair* **6** 224–234. (doi:10.1016/j.dnarep.2006.10.024)

Aoki F, Worrad DM & Schultz RM 1997 Regulation of transcriptional activity during the first and second cell cycles in the preimplantation mouse embryo. *Developmental Biology* **181** 296–307. (doi:10.1006/dbio.1996.8466)

Aoufouchi S, Faili A, Zober C, D'Orlando O, Weller S, Weill JC & Reynaud CA 2008 Proteasomal degradation restricts the nuclear lifespan of AID. *Journal of Experimental Medicine* **205** 1357–1368. (doi:10.1084/jem.20070950)

Barreto G, Schafer A, Marhold J, Stach D, Swaminathan SK, Handa V, Doderlein G, Maltry N, Wu W, Lyko F *et al.* 2007 Gadd45a promotes epigenetic gene activation by repair-mediated DNA demethylation. *Nature* **445** 671–675. (doi:10.1038/nature05515)

Barrett TE, Savva R, Panayotou G, Barlow T, Brown T, Jiricny J & Pearl LH 1998 Crystal structure of a G:T/U mismatch-specific DNA glycosylase: mismatch recognition by complementary-strand interactions. *Cell* **92** 117–129. (doi:10.1016/S0092-8674(00)80904-6)

Basu U, Meng FL, Keim C, Grinstein V, Pefanis E, Eccleston J, Zhang T, Myers D, Wasserman CR, Wesemann DR *et al.* 2011 The RNA exosome targets the AID cytidine deaminase to both strands of transcribed duplex DNA substrates. *Cell* **144** 353–363. (doi:10.1016/j.cell.2011.01.001)

Bennett RA, Wilson DM III, Wong D & Demple B 1997 Interaction of human apurinic endonuclease and DNA polymerase beta in the base excision repair pathway. *PNAS* **94** 7166–7169. (doi:10.1073/pnas.94.14.7166)

Benvenuto G, Carpentieri ML, Salvatore P, Cindolo L, Bruni CB & Chiariotti L 1996 Cell-specific transcriptional regulation and reactivation of galectin-1 gene expression are controlled by DNA methylation of the promoter region. *Molecular and Cellular Biology* **16** 2736–2743.

Besmer E, Market E & Papavasiliou FN 2006 The transcription elongation complex directs activation-induced cytidine deaminase-mediated DNA deamination. *Molecular and Cellular Biology* **26** 4378–4385. (doi:10.1128/MCB.02375-05)

Bestor TH 2000 The DNA methyltransferases of mammals. *Human Molecular Genetics* **9** 2395–2402. (doi:10.1093/hmg/9.16.2395)

Betts L, Xiang S, Short SA, Wolfenden R & Carter CW Jr 1994 Cytidine deaminase. The 2.3 A crystal structure of an enzyme: transition-state analog complex. *Journal of Molecular Biology* **235** 635–656. (doi:10.1006/jmbi.1994.1018)

Bhattacharya SK, Ramchandani S, Cervoni N & Szyf M 1999 A mammalian protein with specific demethylase activity for mCpG DNA. *Nature* **397** 579–583. (doi:10.1038/17533)

Bhutani N, Brady JJ, Damian M, Sacco A, Corbel SY & Blau HM 2010 Reprogramming towards pluripotency requires AID-dependent DNA demethylation. *Nature* **463** 1042–1047. (doi:10.1038/nature08752)

Bird AP 1980 DNA methylation and the frequency of CpG in animal DNA. *Nucleic Acids Research* **8** 1499–1504. (doi:10.1093/nar/8.7.1499)

Blackledge NP, Zhou JC, Tolstorukov MY, Farcas AM, Park PJ & Klose RJ 2010 CpG islands recruit a histone H3 lysine 36 demethylase. *Molecular Cell* **2** 179–190. (doi:10.1016/j.molcel.2010.04.009)

Bouniol C, Nguyen E & Debey P 1995 Endogenous transcription occurs at the 1-cell stage in the mouse embryo. *Experimental Cell Research* **218** 57–62. (doi:10.1006/excr.1995.1130)

Bransteitter R, Pham P, Scharff MD & Goodman MF 2003 Activation-induced cytidine deaminase deaminates deoxycytidine on single-stranded DNA but requires the action of RNase. *PNAS* **100** 4102–4107. (doi:10.1073/pnas.0730835100)

Brar SS, Sacho EJ, Tessmer I, Croteau DL, Erie DA & Diaz M 2008 Activation-induced deaminase, AID, is catalytically active as a monomer on single-stranded DNA. *DNA Repair* **7** 77–87. (doi:10.1016/j.dnarep.2007.08.002)

Brown SE, Suderman MJ, Hallett M & Szyf M 2008 DNA demethylation induced by the methyl-CpG-binding domain protein MBD3. *Gene* **420** 99–106. (doi:10.1016/j.gene.2008.05.009)

Chan Y, Fish JE, D'Abreo C, Lin S, Robb GB, Teichert AM, Karantzoulis-Fegaras F, Keightley A, Steer BM & Marsden PA 2004 The cell-specific expression of endothelial nitric-oxide synthase: a role for DNA methylation. *Journal of Biological Chemistry* **279** 35087–35100. (doi:10.1074/jbc.M405063200)

Chaudhuri J, Tian M, Khuong C, Chua K, Pinaud E & Alt FW 2003 Transcription-targeted DNA deamination by the AID antibody diversification enzyme. *Nature* **422** 726–730. (doi:10.1038/nature01574)

Chelico L, Pham P, Calabrese P & Goodman MF 2006 APOBEC3G DNA deaminase acts processively 3′→5′ on single-stranded DNA. *Nature Structural & Molecular Biology* **13** 392–399. (doi:10.1038/nsmb1086)

Chen C, Yang MC & Yang TP 2001 Evidence that silencing of the HPRT promoter by DNA methylation is mediated by critical CpG sites. *Journal of Biological Chemistry* **276** 320–328. (doi:10.1074/jbc.M007096200)

Chen KM, Martemyanova N, Lu Y, Shindo K, Matsuo H & Harris RS 2007 Extensive mutagenesis experiments corroborate a structural model for the DNA deaminase domain of APOBEC3G. *FEBS Letters* **581** 4761–4766. (doi:10.1016/j.febslet.2007.08.076)

Cheng X & Blumenthal RM 2008 Mammalian DNA methyltransferases: a structural perspective. *Structure* **16** 341–350. (doi:10.1016/j.str.2008.01.004)

Collas P 1998 Modulation of plasmid DNA methylation and expression in zebrafish embryos. *Nucleic Acids Research* **26** 4454–4461. (doi:10.1093/nar/26.19.4454)

Conticello SG, Thomas CJ, Petersen-Mahrt SK & Neuberger MS 2005 Evolution of the AID/APOBEC family of polynucleotide (deoxy)cytidine deaminases. *Molecular Biology and Evolution* **22** 367–377. (doi:10.1093/molbev/msi026)

Cortazar D, Kunz C, Saito Y, Steinacher R & Schar P 2007 The enigmatic thymine DNA glycosylase. *DNA Repair* **6** 489–504. (doi:10.1016/j.dnarep.2006.10.013)

Cortazar D, Kunz C, Selfridge J, Lettieri T, Saito Y, MacDougall E, Wirz A, Schuermann D, Jacobs AL, Siegrist F *et al.* 2011 Embryonic lethal phenotype reveals a function of TDG in maintaining epigenetic stability. *Nature* **470** 419–423. (doi:10.1038/nature09672)

Cortellino S, Xu J, Sannai M, Moore R, Caretti E, Cigliano A, Le Coz M, Devarajan K, Wessels A, Soprano D *et al.* 2011 Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. *Cell* **146** 67–79. (doi:10.1016/j.cell.2011.06.020)

D'Alessio AC, Weaver IC & Szyf M 2007 Acetylation-induced transcription is required for active DNA demethylation in methylation-silenced genes. *Molecular and Cellular Biology* **27** 7462–7474. (doi:10.1128/MCB.01120-07)

Dalhus B, Laerdahl JK, Backe PH & Bjoras M 2009 DNA base repair – recognition and initiation of catalysis. *FEMS Microbiology Reviews* **33** 1044–1078. (doi:10.1111/j.1574-6976.2009.00188.x)

Dean W, Santos F, Stojkovic M, Zakhartchenko V, Walter J, Wolf E & Reik W 2001 Conservation of methylation reprogramming in mammalian development: aberrant reprogramming in cloned embryos. *PNAS* **98** 13734–13738. (doi:10.1073/pnas.241522698)

Deans B, Griffin CS, Maconochie M & Thacker J 2000 Xrcc2 is required for genetic stability, embryonic neurogenesis and viability in mice. *EMBO Journal* **19** 6675–6685. (doi:10.1093/emboj/19.24.6675)

Detich N, Theberge J & Szyf M 2002 Promoter-specific activation and demethylation by MBD2/demethylase. *Journal of Biological Chemistry* **277** 35791–35794. (doi:10.1074/jbc.C200408200)

Dianova II, Sleeth KM, Allinson SL, Parsons JL, Breslin C, Caldecott KW & Dianov GL 2004 XRCC1-DNA polymerase beta interaction is required for efficient base excision repair. *Nucleic Acids Research* **32** 2550–2555. (doi:10.1093/nar/gkh567)

Doerfler W 2008 In pursuit of the first recognized epigenetic signal – DNA methylation: a 1976 to 2008 synopsis. *Epigenetics* **3** 125–133. (doi:10.4161/epi.3.3.6249)

Engel N, Tront JS, Erinle T, Nguyen N, Latham KE, Sapienza C, Hoffman B & Liebermann DA 2009 Conserved DNA methylation in Gadd45a(−/−) mice. *Epigenetics* **4** 98–99. (doi:10.4161/epi.4.2.7858)

Feng S, Jacobsen SE & Reik W 2010 Epigenetic reprogramming in plant and animal development. *Science* **330** 622–627. (doi:10.1126/science.1190614)

Ficz G, Branco MR, Seisenberger S, Santos F, Krueger F, Hore TA, Marques CJ, Andrews S & Reik W 2011 Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. *Nature* **473** 398–402. (doi:10.1038/nature10008)

Fitzgerald ME & Drohat AC 2008 Coordinating the initial steps of base excision repair. Apurinic/apyrimidinic endonuclease 1 actively stimulates thymine DNA glycosylase by disrupting the product complex. *Journal of Biological Chemistry* **283** 32680–32690. (doi:10.1074/jbc.M805504200)

Gerber AP & Keller W 1999 An adenosine deaminase that generates inosine at the wobble position of tRNAs. *Science* **286** 1146–1149. (doi:10.1126/science.286.5442.1146)

Glass JL, Thompson RF, Khulan B, Figueroa ME, Olivier EN, Oakley EJ, Van Zant G, Bouhassira EE, Melnick A, Golden A *et al.* 2007 CG dinucleotide clustering is a species-specific property of the genome. *Nucleic Acids Research* **35** 6798–6807. (doi:10.1093/nar/gkm489)

Gu H, Marth JD, Orban PC, Mossmann H & Rajewsky K 1994 Deletion of a DNA polymerase beta gene segment in T cells using cell type-specific gene targeting. *Science* **265** 103–106. (doi:10.1126/science.8016642)

Guo JU, Su Y, Zhong C, Ming GL & Song H 2011 Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. *Cell* **145** 423–434. (doi:10.1016/j.cell.2011.03.022)

Hajkova P, Erhardt S, Lane N, Haaf T, El-Maarri O, Reik W, Walter J & Surani MA 2002 Epigenetic reprogramming in mouse primordial germ cells. *Mechanisms of Development* **117** 15–23. (doi:10.1016/S0925-4773(02)00181-8)

Hajkova P, Ancelin K, Waldmann T, Lacoste N, Lange UC, Cesari F, Lee C, Almouzni G, Schneider R & Surani MA 2008 Chromatin dynamics during epigenetic reprogramming in the mouse germ line. *Nature* **452** 877–881. (doi:10.1038/nature06714)

Hajkova P, Jeffries SJ, Lee C, Miller N, Jackson SP & Surani MA 2010 Genome-wide reprogramming in the mouse germ line entails the base excision repair pathway. *Science* **329** 78–82. (doi:10.1126/science.1187945)

Hendrich B, Hardeland U, Ng HH, Jiricny J & Bird A 1999 The thymine glycosylase MBD4 can bind to the product of deamination at methylated CpG sites. *Nature* **401** 301–304. (doi:10.1038/45843)

Hermann A, Gowher H & Jeltsch A 2004 Biochemistry and biology of mammalian DNA methyltransferases. *Cellular and Molecular Life Sciences* **61** 2571–2587. (doi:10.1007/s00018-004-4201-1)

Hu XV, Rodrigues TM, Tao H, Baker RK, Miraglia L, Orth AP, Lyons GE, Schultz PG & Wu X 2010 Identification of RING finger protein 4 (RNF4) as a modulator of DNA demethylation through a functional genomics screen. *PNAS* **107** 15087–15092. (doi:10.1073/pnas.1009025107)

Iqbal K, Jin SG, Pfeifer GP & Szabo PE 2011 Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine. *PNAS* **108** 3642–3647. (doi:10.1073/pnas.1014033108)

Ito S, D'Alessio AC, Taranova OV, Hong K, Sowers LC & Zhang Y 2010 Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* **466** 1129–1133. (doi:10.1038/nature09303)

Jenuwein T & Allis CD 2001 Translating the histone code. *Science* **293** 1074–1080. (doi:10.1126/science.1063127)

Jin SG, Guo C & Pfeifer GP 2008 GADD45A does not promote DNA demethylation. *PLoS Genetics* **4** e1000013. (doi:10.1371/journal.pgen.1000013)

Jiricny J & Menigatti M 2008 DNA cytosine demethylation: are we getting close? *Cell* **135** 1167–1169. (doi:10.1016/j.cell.2008.12.008)

Kang YK, Koo DB, Park JS, Choi YH, Chung AS, Lee KK & Han YM 2001 Aberrant methylation of donor genome in cloned bovine embryos. *Nature Genetics* **28** 173–177. (doi:10.1038/88903)

Kim MS, Kondo T, Takada I, Youn MY, Yamamoto Y, Takahashi S, Matsumoto T, Fujiyama S, Shirode Y, Yamaoka I *et al.* 2009 DNA demethylation in hormone-induced transcriptional derepression. *Nature* **461** 1007–1012. (doi:10.1038/nature08456)

Ko M, Huang Y, Jankowska AM, Pape UJ, Tahiliani M, Bandukwala HS, An J, Lamperti ED, Koh KP, Ganetzky R *et al.* 2010 Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. *Nature* **468** 839–843. (doi:10.1038/nature09586)

Koh KP, Yabuuchi A, Rao S, Huang Y, Cunniff K, Nardone J, Laiho A, Tahiliani M, Sommer CA, Mostoslavsky G *et al.* 2011 Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. *Cell Stem Cell* **8** 200–213. (doi:10.1016/j.stem.2011.01.008)

Kroft TL, Jethanandani P, McLean DJ & Goldberg E 2001 Methylation of CpG dinucleotides alters binding and silences testis-specific transcription directed by the mouse lactate dehydrogenase C promoter. *Biology of Reproduction* **65** 1522–1527. (doi:10.1095/biolreprod65.5.1522)

Kunz C, Saito Y & Schar P 2009 DNA repair in mammalian cells: mismatched repair: variations on a theme. *Cellular and Molecular Life Sciences* **66** 1021–1038. (doi:10.1007/s00018-009-8739-9)

Laprise SL 2009 Implications of epigenetics and genomic imprinting in assisted reproductive technologies. *Molecular Reproduction and Development* **76** 1006–1018. (doi:10.1002/mrd.21058)

Larijani M & Martin A 2007 Single-stranded DNA structure and positional context of the target cytidine determine the enzymatic efficiency of AID. *Molecular and Cellular Biology* **27** 8038–8048. (doi:10.1128/MCB.01046-07)

Larijani M, Petrov AP, Kolenchenko O, Berru M, Krylov SN & Martin A 2007 AID associates with single-stranded DNA with high affinity and a long complex half-life in a sequence-independent manner. *Molecular and Cellular Biology* **27** 20–30. (doi:10.1128/MCB.00824-06)

Leibovitch SA & Harel J 1978 Active DNA transcription sites released from the genome of normal embryonic chicken cells. *Nucleic Acids Research* **5** 777–787. (doi:10.1093/nar/5.3.777)

Leibovitch SA, Leibovitch MP, Kruh J & Harel J 1979 Relationship between single-stranded DNA isolated from cultured muscular cells during differentiation and the transcription of messenger RNA. *European Journal of Biochemistry* **97** 327–333. (doi:10.1111/j.1432-1033.1979.tb13118.x)

Liao W, Hong SH, Chan BH, Rudolph FB, Clark SC & Chan L 1999 APOBEC-2, a cardiac- and skeletal muscle-specific member of the cytidine deaminase supergene family. *Biochemical and Biophysical Research Communications* **2** 398–404. (doi:10.1006/bbrc.1999.0925)

Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, Nery JR, Lee L, Ye Z, Ngo QM *et al.* 2009 Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **462** 315–322. (doi:10.1038/nature08514)

Lister R, Pelizzola M, Kida YS, Hawkins RD, Nery JR, Hon G, Antosiewicz-Bourget J, O'Malley R, Castanon R, Klugman S *et al.* 2011 Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. *Nature* **471** 68–73. (doi:10.1038/nature09798)

Liu H, Kim JM & Aoki F 2004 Regulation of histone H3 lysine 9 methylation in oocytes and early pre-implantation embryos. *Development* **131** 2269–2280. (doi:10.1242/dev.01116)

Maatouk DM, Kellam LD, Mann MR, Lei H, Li E, Bartolomei MS & Resnick JL 2006 DNA methylation is a primary mechanism for silencing postmigratory primordial germ cell genes in both germ cell and somatic cell lineages. *Development* **133** 3411–3418. (doi:10.1242/dev.02500)

Maher ER 2005 Imprinting and assisted reproductive technology. *Human Molecular Genetics* **14** R133–R138. (doi:10.1093/hmg/ddi107)

Maiti A, Morgan MT, Pozharski E & Drohat AC 2008 Crystal structure of human thymine DNA glycosylase bound to DNA elucidates sequence-specific mismatch recognition. *PNAS* **105** 8890–8895. (doi:10.1073/pnas.0711061105)

May A, Kirchner R, Muller H, Hartmann P, El Hajj N, Tresch A, Zechner U, Mann W & Haaf T 2009 Multiplex rt-PCR expression analysis of developmentally important genes in individual mouse preimplantation embryos and blastomeres. *Biology of Reproduction* **80** 194–202. (doi:10.1095/biolreprod.107.064691)

Mayer W, Niveleau A, Walter J, Fundele R & Haaf T 2000 Demethylation of the zygotic paternal genome. *Nature* **403** 501–502. (doi:10.1038/35000656)

Metivier R, Gallais R, Tiffoche C, Le Peron C, Jurkowska RZ, Carmouche RP, Ibberson D, Barath P, Demay F, Reid G *et al.* 2008 Cyclical DNA methylation of a transcriptionally active promoter. *Nature* **452** 45–50. (doi:10.1038/nature06544)

Mikl MC, Watt IN, Lu M, Reik W, Davies SL, Neuberger MS & Rada C 2005 Mice deficient in APOBEC2 and APOBEC3. *Molecular and Cellular Biology* **16** 7270–7277. (doi:10.1128/MCB.25.16.7270-7277.2005)

Millar CB, Guy J, Sansom OJ, Selfridge J, MacDougall E, Hendrich B, Keightley PD, Bishop SM, Clarke AR & Bird A 2002 Enhanced CpG mutability and tumorigenesis in MBD4-deficient mice. *Science* **297** 403–405. (doi:10.1126/science.1073354)

Morgan HD, Dean W, Coker HA, Reik W & Petersen-Mahrt SK 2004 Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: implications for epigenetic reprogramming. *Journal of Biological Chemistry* **279** 52353–52360. (doi:10.1074/jbc.M407695200)

Muramatsu M, Sankaranand VS, Anant S, Sugai M, Kinoshita K, Davidson NO & Honjo T 1999 Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells. *Journal of Biological Chemistry* **274** 18470–18476. (doi:10.1074/jbc.274.26.18470)

Muramatsu M, Kinoshita K, Fagarasan S, Yamada S, Shinkai Y & Honjo T 2000 Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell* **102** 553–563. (doi:10.1016/S0092-8674(00)00078-7)

Murata K, Kouzarides T, Bannister AJ & Gurdon JB 2010 Histone H3 lysine 4 methylation is associated with the transcriptional reprogramming efficiency of somatic nuclei by oocytes. *Epigenetics & Chromatin* **3** 4. (doi:10.1186/1756-8935-3-4)

Navaratnam N, Morrison JR, Bhattacharya S, Patel D, Funahashi T, Giannoni F, Teng BB, Davidson NO & Scott J 1993 The p27 catalytic subunit of the apolipoprotein B mRNA editing enzyme is a cytidine deaminase. *Journal of Biological Chemistry* **28** 20709–20712.

Neddermann P & Jiricny J 1994 Efficient removal of uracil from G.U mispairs by the mismatch-specific thymine DNA glycosylase from HeLa cells. *PNAS* **91** 1642–1646. (doi:10.1073/pnas.91.5.1642)

Niehrs C 2009 Active DNA demethylation and DNA repair. *Differentiation* **77** 1–11. (doi:10.1016/j.diff.2008.09.004)

Nishigaki M, Aoyagi K, Danjoh I, Fukaya M, Yanagihara K, Sakamoto H, Yoshida T & Sasaki H 2005 Discovery of aberrant expression of R-RAS by cancer-linked DNA hypomethylation in gastric cancer using microarrays. *Cancer Research* **65** 2115–2124. (doi:10.1158/0008-5472.CAN-04-3340)

Okada Y, Yamagata K, Hong K, Wakayama T & Zhang Y 2010 A role for the elongator complex in zygotic paternal genome demethylation. *Nature* **463** 554–558. (doi:10.1038/nature08732)

Olsen LC, Aasland R, Wittwer CU, Krokan HE & Helland DE 1989 Molecular cloning of human uracil-DNA glycosylase, a highly conserved DNA repair enzyme. *EMBO Journal* **8** 3121–3125.

Oswald J, Engemann S, Lane N, Mayer W, Olek A, Fundele R, Dean W, Reik W & Walter J 2000 Active demethylation of the paternal genome in the mouse zygote. *Current Biology* **10** 475–478. (doi:10.1016/S0960-9822(00)00448-6)

Otero G, Fellows J, Li Y, de Bizemont T, Dirac AM, Gustafsson CM, Erdjument-Bromage H, Tempst P & Svejstrup JQ 1999 Elongator, a multisubunit component of a novel RNA polymerase II holoenzyme for transcriptional elongation. *Molecular Cell* **3** 109–118. (doi:10.1016/S1097-2765(00)80179-3)

Papageorgis P, Lambert AW, Ozturk S, Gao F, Pan H, Manne U, Alekseyev YO, Thiagalingam A, Abdolmaleky HM, Lenburg M *et al.* 2010 Smad signaling is required to maintain epigenetic silencing during breast cancer progression. *Cancer Research* **70** 968–978. (doi:10.1158/0008-5472.CAN-09-1872)

Parsons JL, Dianova II & Dianov GL 2005 APE1-dependent repair of DNA single-strand breaks containing 3′-end 8-oxoguanine. *Nucleic Acids Research* **33** 2204–2209. (doi:10.1093/nar/gki518)

Patenaude AM, Orthwein A, Hu Y, Campo VA, Kavli B, Buschiazzo A & Di Noia JM 2009 Active nuclear import and cytoplasmic retention of activation-induced deaminase. *Nature Structural & Molecular Biology* **16** 517–527. (doi:10.1038/nsmb.1598)

Pavri R, Gazumyan A, Jankovic M, Di Virgilio M, Klein I, Ansarah-Sobrinho C, Resch W, Yamane A, Reina San-Martin B,

Barreto V *et al.* 2010 Activation-induced cytidine deaminase targets DNA at sites of RNA polymerase II stalling by interaction with Spt5. *Cell* **143** 122–133. (doi:10.1016/j.cell.2010.09.017)

Pham P, Bransteitter R, Petruska J & Goodman MF 2003 Processive AID-catalysed cytosine deamination on single-stranded DNA simulates somatic hypermutation. *Nature* **424** 103–107. (doi:10.1038/nature01760)

Popp C, Dean W, Feng S, Cokus SJ, Andrews S, Pellegrini M, Jacobsen SE & Reik W 2010 Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. *Nature* **463** 1101–1105. (doi:10.1038/nature08829)

Rai K, Huggins IJ, James SR, Karpf AR, Jones DA & Cairns BR 2008 DNA demethylation in zebrafish involves the coupling of a deaminase, a glycosylase, and gadd45. *Cell* **135** 1201–1212. (doi:10.1016/j.cell.2008.11.042)

Rai K, Sarkar S, Broadbent TJ, Voas M, Grossmann KF, Nadauld LD, Dehghanizadeh S, Hagos FT, Li Y, Toth RK *et al.* 2010 DNA demethylase activity maintains intestinal cells in an undifferentiated state following loss of APC. *Cell* **142** 930–942. (doi:10.1016/j.cell.2010.08.030)

Ruddock-D'Cruz NT, Xue J, Wilson KJ, Heffernan C, Prashadkumar S, Cooney MA, Sanchez-Partida LG, French AJ & Holland MK 2008 Dynamic changes in the localization of five members of the methyl binding domain (MBD) gene family during murine and bovine preimplantation embryo development. *Molecular Reproduction and Development* **75** 48–59. (doi:10.1002/mrd.20712)

Santos F, Hendrich B, Reik W & Dean W 2002 Dynamic reprogramming of DNA methylation in the early mouse embryo. *Developmental Biology* **241** 172–182. (doi:10.1006/dbio.2001.0501)

Santos F, Peters AH, Otte AP, Reik W & Dean W 2005 Dynamic chromatin modifications characterise the first cell cycle in mouse embryos. *Developmental Biology* **280** 225–236. (doi:10.1016/j.ydbio.2005.01.025)

Sato S, Yoshimizu T, Sato E & Matsui Y 2003 Erasure of methylation imprinting of Igf2r during mouse primordial germ-cell development. *Molecular Reproduction and Development* **65** 41–50. (doi:10.1002/mrd.10264)

Sato Y, Probst HC, Tatsumi R, Ikeuchi Y, Neuberger MS & Rada C 2010 Deficiency in APOBEC2 leads to a shift in muscle fiber-type, diminished body mass and myopathy. *Journal of Biological Chemistry* **285** 7111–7118. (doi:10.1074/jbc.M109.052977)

Schafer A, Schomacher L, Barreto G, Doderlein G & Niehrs C 2010 Gemcitabine functions epigenetically by inhibiting repair mediated DNA demethylation. *PLoS ONE* **5** e14060. (doi:10.1371/journal.pone.0014060)

Schmitz KM, Schmitt N, Hoffmann-Rohrer U, Schafer A, Grummt I & Mayer C 2009 TAF12 recruits Gadd45a and the nucleotide excision repair complex to the promoter of rRNA genes leading to active DNA demethylation. *Molecular Cell* **33** 344–353. (doi:10.1016/j.molcel.2009.01.015)

Schmitz KM, Mayer C, Postepska A & Grummt I 2010 Interaction of noncoding RNA with the rDNA promoter mediates recruitment of DNMT3b and silencing of rRNA genes. *Genes and Development* **24** 2264–2269. (doi:10.1101/gad.590910)

Shen HM, Poirier MG, Allen MJ, North J, Lal R, Widom J & Storb U 2009 The activation-induced cytidine deaminase (AID) efficiently targets DNA in nucleosomes but only during transcription. *Journal of Experimental Medicine* **206** 1057–1071. (doi:10.1084/jem.20082678)

Simonsson S & Gurdon J 2004 DNA demethylation is necessary for the epigenetic reprogramming of somatic cell nuclei. *Nature Cell Biology* **6** 984–990. (doi:10.1038/ncb1176)

Song F, Mahmood S, Ghosh S, Liang P, Smiraglia DJ, Nagase H & Held WA 2009 Tissue specific differentially methylated regions (TDMR): changes in DNA methylation during development. *Genomics* **93** 130–139. (doi:10.1016/j.ygeno.2008.09.003)

Stengel S, Fiebig U, Kurth R & Denner J 2010 Regulation of human endogenous retrovirus-K expression in melanomas by CpG methylation. *Genes, Chromosomes & Cancer* **49** 401–411. (doi:10.1002/gcc.20751)

Sugo N, Aratani Y, Nagashima Y, Kubota Y & Koyama H 2000 Neonatal lethality with abnormal neurogenesis in mice deficient in DNA polymerase beta. *EMBO Journal* **19** 1397–1404. (doi:10.1093/emboj/19.6.1397)

Surani MA & Hajkova P 2010 Epigenetic reprogramming of mouse germ cells toward totipotency. *Cold Spring Harbor Symposia on Quantitative Biology* **75** 211–218. (doi:10.1101/sqb.2010.75.010)

Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L *et al.* 2009 Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324** 930–935. (doi:10.1126/science.1170116)

Takahashi K & Yamanaka S 2006 Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126** 663–676. (doi:10.1016/j.cell.2006.07.024)

Talpaert-Borle M, Campagnari F & Creissen DM 1982 Properties of purified uracil-DNA glycosylase from calf thymus. An *in vitro* study using synthetic DNA-like substrates. *Journal of Biological Chemistry* **257** 1208–1214.

Tebbs RS, Flannery ML, Meneses JJ, Hartmann A, Tucker JD, Thompson LH, Cleaver JE & Pedersen RA 1999 Requirement for the Xrcc1 DNA base excision repair gene during early mouse development. *Developmental Biology* **208** 513–529. (doi:10.1006/dbio.1999.9232)

Teng B, Burant CF & Davidson NO 1993 Molecular cloning of an apolipoprotein B messenger RNA editing protein. *Science* **260** 1816–1819. (doi:10.1126/science.8511591)

Turek-Plewa J & Jagodzinski PP 2005 The role of mammalian DNA methyltransferases in the regulation of gene expression. *Cellular & Molecular Biology Letters* **10** 631–647.

Vidal AE, Boiteux S, Hickson ID & Radicella JP 2001 XRCC1 coordinates the initial and late stages of DNA abasic site repair through protein–protein interactions. *EMBO Journal* **20** 6530–6539. (doi:10.1093/emboj/20.22.6530)

Vonica A, Rosa A, Arduini BL & Brivanlou AH 2011 APOBEC2, a selective inhibitor of TGFbeta signaling, regulates left–right axis specification during early embryogenesis. *Developmental Biology* **350** 13–23. (doi:10.1016/j.ydbio.2010.09.016)

Vuong BQ, Lee M, Kabir S, Irimia C, Macchiarulo S, McKnight GS & Chaudhuri J 2009 Specific recruitment of protein kinase A to the immunoglobulin locus regulates class-switch recombination. *Nature Immunology* **10** 420–426. (doi:10.1038/ni.1708)

Wang L, Wuerffel R, Feldman S, Khamlichi AA & Kenter AL 2009 S region sequence, RNA polymerase II, and histone modifications create chromatin accessibility during class switch recombination. *Journal of Experimental Medicine* **206** 1817–1830. (doi:10.1084/jem.20081678)

Williams K, Christensen J, Pedersen MT, Johansen JV, Cloos PA, Rappsilber J & Helin K 2011 TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. *Nature* **473** 343–348. (doi:10.1038/nature10066)

Wong E, Yang K, Kuraguchi M, Werling U, Avdievich E, Fan K, Fazzari M, Jin B, Brown AM, Lipkin M *et al.* 2002 Mbd4 inactivation increases Cright-arrowT transition mutations and promotes gastrointestinal tumor formation. *PNAS* **99** 14937–14942. (doi:10.1073/pnas.232579299)

Wossidlo M, Arand J, Sebastiano V, Lepikhov K, Boiani M, Reinhardt R, Scholer H & Walter J 2010 Dynamic link of DNA demethylation, DNA strand breaks and repair in mouse zygotes. *EMBO Journal* **29** 1877–1888. (doi:10.1038/emboj.2010.80)

Wossidlo M, Nakamura T, Lepikhov K, Marques CJ, Zakhartchenko V, Boiani M, Arand J, Nakano T, Reik W & Walter J 2011 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. *Nature Communications* **2** 241. (doi:10.1038/ncomms1240)

Wu P, Qiu C, Sohail A, Zhang X, Bhagwat AS & Cheng X 2003 Mismatch repair in methylated DNA. Structure and activity of the mismatch-specific thymine glycosylase domain of methyl-CpG-binding protein MBD4. *Journal of Biological Chemistry* **278** 5285–5291. (doi:10.1074/jbc.M210884200)

Wu H, D'Alessio AC, Ito S, Xia K, Wang Z, Cui K, Zhao K, Eve Sun Y & Zhang Y 2011 Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. *Nature* **473** 389–393. (doi:10.1038/nature09934)

Yoon JH, Iwai S, O'Connor TR & Pfeifer GP 2003 Human thymine DNA glycosylase (TDG) and methyl-CpG-binding protein 4 (MBD4) excise thymine glycol (Tg) from a Tg:G mispair. *Nucleic Acids Research* **31** 5399–5404. (doi:10.1093/nar/gkg730)

Yoshikawa K, Okazaki IM, Eto T, Kinoshita K, Muramatsu M, Nagaoka H & Honjo T 2002 AID enzyme-induced hypermutation in an actively transcribed gene in fibroblasts. *Science* **296** 2033–2036. (doi:10.1126/science.1071556)

Zhu JK 2009 Active DNA demethylation mediated by DNA glycosylases. *Annual Review of Genetics* **43** 143–166. (doi:10.1146/annurev-genet-102108-134205)

Zhu B, Zheng Y, Angliker H, Schwarz S, Thiry S, Siegmann M & Jost JP 2000*a* 5-Methylcytosine DNA glycosylase activity is also present in the human MBD4 (G/T mismatch glycosylase) and in a related avian sequence. *Nucleic Acids Research* **28** 4157–4165. (doi:10.1093/nar/28.21.4157)

Zhu B, Zheng Y, Hess D, Angliker H, Schwarz S, Siegmann M, Thiry S & Jost JP 2000*b* 5-Methylcytosine-DNA glycosylase activity is present in a cloned G/T mismatch DNA glycosylase associated with the chicken embryo DNA demethylation complex. *PNAS* **97** 5135–5139. (doi:10.1073/pnas.100107597)

Received 3 May 2011
First decision 27 June 2011
Revised manuscript received 31 August 2011
Accepted 12 September 2011