Genetics & Epigenetics

REVIEW

DNA Demethylation Pathways: Recent Insights

Cong-jun Li

USDA, Agricultural Research Service, Bovine Functional Genomics Laboratory, Beltsville, MD, USA.
Corresponding author email: congjun.li@ars.usda.gov

Abstract: DNA methylation is a major epigenetic regulatory mechanism for gene expression and cell differentiation. Until recently, it was still unclear how unmethylated regions in mammalian genomes are protected from de novo methylation and whether or not active demethylating activity is involved. Even the role of molecules and the mechanisms underlying the processes of active demethylation itself is blurred. Emerging sequencing technologies have led to recent insights into the dynamic distribution of DNA methylation during development and the role of this epigenetic mark within a distinct genome context, such as the promoters, exons, or imprinted control regions. This review summarizes recent insights on the dynamic nature of DNA methylation and demethylation, as well as the mechanisms regulating active DNA demethylation in mammalian cells, which have been fundamental research interests in the field of epigenomics.

Keywords: AID/APOBEC proteins, BER glycosylase, DNA demethylation, epigenetic regulation, TED dioxygenases

Genetics & Epigenetics 2013:5 43–49

doi: 10.4137/GEG.S12143

This article is available from http://www.la-press.com.

Introduction

One of the most remarkable properties of complex genomes is their capacity to generate a range of different cell types with a set of identical genome in a highly ordered and reproducible manner. How this happens has intrigued geneticists and developmental biologists alike and has helped spur recent advances in epigenomics/epigenetics.[1] Whereas the term epigenomics describes the analysis of epigenetic changes across many genes in a cell or throughout an entire organism, epigenetics centers on processes that regulate how and when specific genes are turned on and turned off. The regulation of gene expression in many biological processes involves several different types of epigenetic mechanisms. These mechanisms, such as DNA methylation and histone modification, have been recognized for a long time and they are intricately interconnected. Some of these processes, such as the formation of microRNA, have only recently been discovered. Genomic imprinting, gene silencing, X chromosome inactivation, position effects, reprogramming, and the progress of carcinogenesis are all known epigenetic processes. By definition, RNA splicing, RNA editing, and prions also can be included as epigenetic mechanisms for gene regulation.

These regulatory mechanisms for modulation of gene function are multifaceted and complex. This complexity is presented by the dynamic nature of chromatin structure, and the nature of functional information in the genome. Epigenomic remodeling modulates the access to DNA by the functional elements of cells such as transcription and DNA replication machinery in response to different upstream signals. These epigenomic processes are critical for transcription and other DNA-related processes such as DNA replication and repair. The intrinsic worth of epigenetic regulation of gene expression is twofold. First, this type of regulation determines up- or down-regulation and the scope of gene responses to the activation of different signaling pathways. Second, epigenetic mechanisms contribute to stable, cell-type-specific patterns of gene activities (silencing or activation).[2] Despite the fact that 'epigenomes' of humans[3] and of other species have been extensively studied, we still know relatively little about how different gene expression patterns initially segregate in the developing embryo or how they are stably transmitted through cell division. In particular, the molecular details of the template mechanisms that duplicate epigenetic marks through DNA replication remain uncertain.

DNA methylation is one of the most studied epigenetic modifications. Methylation of DNA at position 5 of the cytosine ring (5-methylcytosine, 5mC) occurs at most CpG dinucleotides in the mammalian genome. Approximately 70 to 80% of cytosine in CpG dyads is methylated on both strands in human somatic cells. The patterns of DNA methylation are non-random, well regulated and tissue-specific. Epigenetic information encoded by 5-methylcytosine (5mC) has an overwhelming impact on mammalian development and human diseases.[4] The functional importance of DNA methylation is consistent with the pattern of DNA methylation.[5] Genome-wide analyses of the relationship between development or cell functions and epigenomic landscapes, as well as the interactions among different epigenomic mechanisms, have been exploring large volumes of new information about the functional implications of epigenetic processes. Emerging sequencing technologies have led to recent insights into the dynamic distribution of DNA methylation during development and the role of this epigenetic mark within a distinct genome context, such as the promoters, exons or imprinted control regions. However, one of the most fundamental areas of recent research interest is the active demethylation of 5mC in mammalian cells.

Passive and Active DNA Demethylation

DNA demethylation can be achieved either passively, by simply not methylating the new DNA strand after replication, or actively, by a replication-independent process. Passive demethylation most likely occurs during mammalian development and cell differentiation, mainly in the maternal genome during pre-implantation growth. In mice, remarkable reprogramming with waves of demethylation and then remethylation occur in germ cells and early embryos.[6] After fertilization, the maternal genome undergoes passive, replication-dependent demethylation during subsequent cleavage divisions, whereas most of the paternal genome is rapidly demethylated before DNA replication begins, signifying active enzymatic demethylation.[7] This active demethylation of the paternal genome probably is associated with epigenetic remodeling of sperm chromatin, so as to establish

parent-specific developmental programs during early embryogenesis.[7] After implantation, global de novo methylation re-establishes the DNA methylation patterns that will be maintained in somatic tissues. In addition to these genome-wide changes, gene-specific de novo methylation and demethylation also take place during lineage-specific differentiation.[8] It has been established that DNA methyltransferases 3A (DNMT3A) and 3B (DNMT3B) are responsible for de novo methylation in early development,[9] and once a DNA methylation pattern is established, DNMT1 is responsible for faithful maintenance of the DNA methylation pattern through cell divisions.[9]

Imprinted genes are one class of genes that critically depend upon epigenetic modifications for correct expression, and many imprinted genes have roles in controlling fetal growth as well as neonatal and adult metabolism. Imprinted genes have been considered a potential target or mediator of programming events.[10,11] Because epigenetic marks such as DNA methylation or histone tail modifications could provide a persistent memory of earlier nutritional states, it has been intensively studied whether expression of imprinted genes is altered in a paradigm of developmental programming and whether this is associated with altered methylation of their differentially methylated regions (DMRs). Using an established mouse model of developmental programming that employs protein restriction of maternal diets during gestation or lactation, one study found that although changes in the expression level of imprinted genes can be detected in the livers of offspring from dietary-restricted female mice, DMR methylation appears to be robust.[10] However, in contrast, another study in sheep[12] found that expression analysis of 9 imprinted genes and 3 DNA methyltransferase (DNMTs) genes showed significant effects from the different maternal diets on the expression of these genes, as well as the methylation levels of CpG islands of both IGF2R and H19 in fetuses of both males and females. Apparently, the amount of methylation could be caused by a combination of undermethylation of DNA, or could be the result of highly dynamic and sometimes opposing demethylation or de novo methylation processes in parental genome.[13] This de novo DNA methylation and demethylation machinery is tightly regulated in a dynamic manner. Once established, the patterns of methylation can be maintained in a stable manner over the lifetime of the organism. However, during aging and as part of the disease process, DNA methylation modulation can occur and is tightly regulated using complex machinery.[14]

Early studies indicated that demethylation is an active event, which does not take place passively as a result of DNA replication in the absence of methylation maintenance. It was initially established that DNA hypomethylation can arise without replication by analysis of methylation changes in the delta-crystallin genes of the chicken lens. During embryonic development, a large fraction of cells in the lens stops dividing as part of the differentiation process. Shortly after this stage, the delta-crystallin genes in samples of the whole lens become hypomethylated, suggesting the possibility that this process might be occurring in the subset of cells that is no longer dividing. Since hypomethylation of these genes does occur in post-mitotic lens cells, this result implicated an excision mechanism in this tissue.[15] For many years, the biochemical mechanism of active demethylation was unknown. Only over the past few years has the enzymology of demethylation has been clarified. With all these developments, we have come to realize that DNA methylation is a dynamic equilibrium of methylation and demethylation and not just a stable DNA mark.

Dynamics of DNA Demethylation and the DNA Demethylation Pathway

Although earlier studies implicated a rapid and active mechanism independent of cell division in the demethylation of DNA,[7,16] the mechanism(s) that underlie the active removal of methylation remain unidentified. The discovery of cytosine hydroxymethylation (5hmC) suggested a simple means of demethylating DNA.[17,18] However, detailed DNA demethylation pathways indicate otherwise. Recent evidence[19,20,21] has lead to the postulation that DNA methylation and demethylation can go in 2 directions, and it was suggested that 5hmC may serve as an intermediate for the removal of methylated cytosines. 3 enzyme families have been implicated in active DNA demethylation via DNA repair pathways.[22] The 10–11 translocation (TET) family, the activation-induced cytidine deaminase (AID)/apolipoprotein B mRNA-editing catalytic polypeptides (APOBEC) family and base excision repair (BER) family have been identified as having roles in active DNA demethylation.

The ten-eleven translocation (TET, TED dioxygenases) family

TET proteins were initially discovered through their involvement in myeloid leukemia in which the *TET1* gene, located on chromosome 10, can translocate with the H3K4 histone methyltransferase *MLL* gene on chromosome 11.[23] TET enzymes are members of the TET/J-binding protein (JBP) family of α-ketoglutarate- and iron (II)-dependent dioxygenases, closely related to the JBP1 and JBP2 proteins found in kinetoplastids such as trypanosomes and leishmanias. In mammals, the TET/JBP family is composed of the founding member *TET1* along with *TET2* and *TET3*. These 3 genes encode proteins sharing a double-stranded β-helix-fold and a cysteine-rich region within the catalytic domain.[24] The involvement of TETs in active DNA demethylation was overlooked until recently. It was first reported in 2009 that TET1 catalyzes the conversion of 5mC to 5-hydroxymethylcytosine (5hmC) in cultured cells and in vitro and suggested that 5-hmC might be an intermediate in the pathway to unmodified cytosine.[18] Later, it was demonstrated that TET proteins (TET1, TET2 and TET3), in addition to 5hmC, can generate 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) from 5mC in an enzymatic activity-dependent manner. Thus, 2 previously unknown cytosine derivatives in genomic DNA were identified as the products of TET proteins.[25]

One development related to the TET protein is TET-assisted bisulfite sequencing (TAB-Seq), based on the findings that TET proteins not only oxidize 5mC to 5hmC but also further oxidize 5hmC to 5caC, and that 5caC exhibits behavior similar to that of unmodified cytosine after bisulfite treatment. The study of 5-hydroxylmethylcytosines (5hmC) has been hampered by the lack of a method to map it at a single-base resolution on a genome-wide scale. Affinity purification-based methods cannot precisely locate 5hmC nor accurately determine its relative abundance at each modified site. A genome-wide approach, TET-assisted bisulfite sequencing (TAB-Seq) combined with traditional bisulfite sequencing can be used for mapping 5hmC at base resolution and quantifying the relative abundance of 5hmC as well as 5mC.[26,27]

The AID/APOBEC family

The AID/APOBEC proteins are found in vertebrates and share the ability to insert mutations in DNA and RNA by deaminating cytidine to uridine. The first family member to be identified and characterized was the apolipoprotein B editing complex 1 (APOBEC1), a protein involved in the editing of the apolipoprotein B (ApoB) pre-mRNA.[28,29] Further members were identified as DNA mutators. Activation-induced deaminase (AID) was revealed to be essential for the antigen-driven diversification of already rearranged immunoglobulin genes in the vertebrate adaptive immune system.[30] The pioneering work in plants provides strong evidence for a set of glycosylase/lyase enzymes (Demeter, ROS, DML2, DML3) in the removal of the 5-meC in various biological contexts.[31] The activation induced deaminase/apolipoprotein B RNA-editing catalytic component-1 (AID/APOBEC-1) family of RNA cytidine deaminases was reported to have 5-meC deaminase activities. If these deaminases are tightly and efficiently coupled to G/T mismatch repair systems, their activity could lead to DNA demethylation. However, this mechanism was only one of many proposed but unproven mechanisms for DNA demethylation in vertebrates until a recent study suggests that global active demethylation in zebrafish embryos can be achieved by the coupled action of AID and MBD4.[32] One of the important findings of this study is that overexpression of AID/APOBEC along with hMbd4, but not either alone, caused significant demethylation, providing evidence for a coupled mechanism of 5-meC demethylation, whereby AID deaminates 5-meC, followed by thymine base excision repair by Mbd4. Since then, more evidence has found that AID/APOBEC has a role in DNA demethylation.[19,20,33]

The BER glycosylase family

Base excision repair (BER) corrects DNA damage from oxidation, deamination and alkylation. Such base lesions cause little distortion to the DNA helix structure. BER is initiated by a DNA glycosylase that recognizes and removes the damaged base, leaving a basic site that is further processed by short-patch repair or long-patch repair that largely uses different proteins to complete BER. At least 11 distinct mammalian DNA glycosylases are known, each recognizing a few related lesions, frequently with some overlap in specificities.[34] The accumulation data from the TED and AID/APPOBEC studies suggest that active demethylation involves cytosine

replacement via DNA repair. In plants, active DNA demethylation is a well-characterized process. BER glycosylases mediate the first step in the repair pathway by removing the methylated cytosine and creating a basic site, which is then further processed by other enzymes.[35] At the time, it was indicated from several intriguing experiments that DNA repair is also a plausible mechanism for animal demethylation. Glycosylase-dependent DNA demethylation was first proposed in animals.[36,37,38] The 5-methylcytosine glycosylase activity was initially detected in chicken embryo extracts, which contain thymine-DNA glycosylase (TDG). However, the glycosylase activity of TDG is much lower against 5-methylcytosine than against mismatched thymine,[37] indicating that TDG might require some other proteins to activate DNA demethylation pathways. Recently, it was demonstrated that either knockout or catalytic inactivation of the DNA repair enzyme thymine DNA glycosylase (TDG) leads to embryonic lethality in mice. It was also indicated that TDG is necessary for recruiting the transcription activator p300 to retinoic acid (RA)-regulated promoters, protection of CpG islands from hypermethylation, and active demethylation of tissue-specific developmentally and hormonally regulated promoters and enhancers.[33] It was also demonstrated that TDG interacts with the deaminase AID and the damage response protein GADD45a, suggesting a 2-step mechanism for DNA demethylation in mammals, whereby 5-methylcytosine and 5-hydroxymethylcytosine are first deaminated by AID to thymine and 5-hydroxymethyluracil, respectively, followed by TDG-mediated thymine and 5-hydroxymethyluracil excision repair.[33] Direct evidence was also reported that human cells possess a robust demethylating activity toward 5hmC-containing DNA, which is DNA replication independent and requires an intact BER pathway. Furthermore, AID/APOBEC cytidine deaminases promote 5hmC demethylation both in cultured human cells and in the adult mouse brain. Similar to deamination, 5hmC demethylation is processive, transcription dependent, and strand biased.[20] These studies demonstrated that 5-methylcytosine (5mC) hydroxylase TET1, by converting 5mCs to 5-hydroxymethylcytosines (5hmCs), promotes DNA demethylation in mammalian cells through a process that requires the base excision repair pathway.

The 2 most recent research reports exposed much of the details of active demethylation pathways and highlighted the dynamics of active DNA demethylation. These 2 new studies show that the reprogramming of 5-methylcytosine via TET- and TDG-family enzymes is both extensive throughout the genome and functionally significant. First, the final oxidation product of methylated cytosine was identified by enrichment and sequencing and reveals that DNA demethylation is common throughout the genome. TET proteins can further oxidize 5-hmC to 5-fC (5-formylcytosine) and 5-caC (5-carboxylcytosine).[39] The low abundance of these modifications of cytosine makes profiling their location in the genome difficult. Even though these modifications present in low abundance, they have been named as the 5th, 6th, 7th and 8th DNA bases (5mC, 5hmC, 5fC and 5caC).[40] Moreover, with genetic enrichment, Zhong's laboratory also confirmed that thymine-DNA glycosylase (TDG) replaces 5-fC and 5-caC with unmodified cytosine via base excision repair. In cells depleted of TDG, the accumulation of 5-fC and 5-caC showed the dynamic nature of demethylation in nonrepetitive and regulatory elements, indicating a role for active cycling in regulating gene expression. A DNA methylation-demethylation cycle run by DNA methylase, TET and TDG seems a very apropos name for this event.

The second report[41] reveals roles of active 5mC/5hmC oxidation and TDG-mediated demethylation in epigenetic tuning at regulatory elements. Genome-wide mapping of 5fC in mouse embryonic stem cells (mESCs) discovered that 5fC preferentially occurs at poised enhancers among other gene regulatory elements. The data also suggests that 5-fC production may coordinate with transcription factor p300 in remodeling epigenetic states of enhancers. This process, which is not influenced by 5hmC, appears to be associated with further oxidation of 5hmC and commitment to demethylation through 5fC. The data also suggest that many regions annotated as unmethylated are actually constantly demethylated.

Both reports revealed rich information by profiling demethylation intermediates. To accurately understand the functions of 5mC, 5hmC, 5fC and 5caC, it is important to develop systems to map these modifications. We also have to keep in mind that these modifications are not only active demethylation intermediates but may be epigenetic modifications by themselves.

The proteins involved in the active demethylation process are multi-functional. For example, TETs can regulate transcription activity independently of their catalytic activity. It remains a challenge to elucidate the exact pathway following each specific context. Furthermore, the TET-AID/APOBEC-BER Glycosylase pathway may not be the only pathway participating in active demethylation. Schiesser et al reported fascinating 5-caC-decarboxylation activity in stem cells (mESCs).[42] This appears to be a highly attractive alternative, because this mechanism allows exchange of mC by dC without formation of intermediate strand breaks. Strand breaks are DNA lesions that are known to cause genome instability without formation of potentially harmful side products, such as formaldehyde.

Summary and Perspective

5-methylcyosine is an important, well-known nucleobase modification that is essential to establish tissue-specific gene expression and involved in many biological processes such as gene expression and genomic imprinting. Our understanding about active demethylation sets the stage for the systematic study of the function of DNA methylation and demethylation dynamics. Discovery of 5fC, 5caC has lead to the proposal of an active DNA demethylation cycle relying on the initial oxidation of 5mC into 5hmC, and further oxidation to the 5fC and 5caC. However, in contrast to the abundant 5hmC, the 5fC and 5caC derivatives are present in a much lower abundance. This imbalance of derivatives or the scarcity of 5fC and 5caC has to be better explained experimentally. The discovery of these derivatives of 5mC also triggers a possibility that these derivatives may also serve as epigenomic markers. And finally, the most important impact of DNA methylation patterns is to maintain the cell and organismal fate. The intriguing phenomenon of how a stable differentiated state is maintained by regulatory pathways that are surprisingly dynamic and perturbable raises the tricky question of how cellular plasticity is kept in check to maintain cellular fate. One future goal and major challenge is to understand the mechanisms that regulate this contradictory phenomenon.

Funding

Author discloses no funding sources.

Author Contributions

Conceived the concepts and wrote the manuscript: CJL. The author reviewed and approved of the final manuscript.

Competing Interests

Author discloses no potential conflicts of interest.

Disclosures and Ethics

As a requirement of publication author(s) have provided to the publisher signed confirmation of compliance with legal and ethical obligations including but not limited to the following: authorship and contributorship, conflicts of interest, privacy and confidentiality and (where applicable) protection of human and animal research subjects. The authors have read and confirmed their agreement with the ICMJE authorship and conflict of interest criteria. The authors have also confirmed that this article is unique and not under consideration or published in any other publication, and that they have permission from rights holders to reproduce any copyrighted material. Any disclosures are made in this section. The external blind peer reviewers report no conflicts of interest. Provenance: the author was invited to submit this paper.

Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture. USDA is an equal opportunity provider and employer.

References

1. Cantone I, Fisher AG. Epigenetic programming and reprogramming during development. *Nat Struct Mol Biol.* 2013;20(3):282–289.
2. Tarakhovsky A. Tools and landscapes of epigenetics. *Nat Immunol.* 2010; 11(7):565–568.
3. Dunham I, Kundaje A, Aldred SF, et al. An integrated encyclopedia of DNA elements in the human genome. *Nature.* 2012;489(7414):57–74.
4. Klose RJ, Bird AP. Genomic DNA methylation: the mark and its mediators. *Trends Biochem Sci.* 2006;31(2):89–97.
5. Illingworth R, Kerr A, Desousa D, et al. (2008). A novel CpG island set identifies tissue-specific methylation at developmental gene loci. *PLoS Biol.* 2008;6(1):e22.
6. Morgan HD, Santos F, Green K, Dean W, Reik W. Epigenetic reprogramming in mammals. *Hum Mol Genet.* 2005;14(Spec No 1):R47–R58.
7. Mayer W, Niveleau A, Walter J, Fundele R, Haaf T. Demethylation of the zygotic paternal genome. *Nature.* 2000;403(6769):501–502.
8. Ji H, Ehrlich LI, Seita J, et al. Comprehensive methylome map of lineage commitment from haematopoietic progenitors. *Nature.* 2010;467(7313): 338–342.
9. Law JA, Jacobsen SE. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat Rev Genet.* 2010;11(3):204–220.

10. Ivanova E, Chen JH, Segonds-Pichon A, Ozanne SE, Kelsey G. DNA methylation at differentially methylated regions of imprinted genes is resistant to developmental programming by maternal nutrition. *Epigenetics.* 2012;7(10):1200–1210.
11. Smallwood SA, Kelsey G. De novo DNA methylation: a germ cell perspective. *Trends Genet*. 2012;28(1):33–42.
12. Lan X, Cretney EC, Kropp J, et al. Maternal Diet during Pregnancy Induces Gene Expression and DNA Methylation Changes in Fetal Tissues in Sheep. *Front Genet*. 2013;4:49.
13. Yoder JA, Walsh CP, Bestor TH. Cytosine methylation and the ecology of intragenomic parasites. *Trends Genet.* 1997;13(8):335–340.
14. Bergman Y, Cedar H. DNA methylation dynamics in health and disease. *Nat Struct Mol Biol.* 2008;20(3):274–281.
15. Sullivan CH, Grainger RM. Delta-crystallin genes become hypomethylated in postmitotic lens cells during chicken development. *Proc Natl Acad Sci U S A.* 1987;84(2):329–333.
16. Zhang F, Pomerantz JH, Sen G, Palermo AT, Blau HM. Active tissue-specific DNA demethylation conferred by somatic cell nuclei in stable heterokaryons. *Proc Natl Acad Sci U S A.* 2007;104(11):4395–4400.
17. Kriaucionis S, Heintz N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science*. 2009;324(5929): 929–930.
18. Tahiliani M, Koh KP, Shen Y, et al. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science.* 2009;324(5929):930–935.
19. Bhutani N, Brady JJ, Damian M, Sacco A, Corbel SY, Blau HM. Reprogramming towards pluripotency requires AID-dependent DNA demethylation. *Nature.* 2010;463(7284):1042–1047.
20. Guo JU, Su Y, Zhong C, Ming GL, Song H. Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. *Cell.* 2011;145(3):423–434.
21. He YF, Li BZ, Li Z, et al. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science.* 2011;333(6047): 1303–1307.
22. Bhutani N, Burns DM, Blau HM. DNA demethylation dynamics. *Cell.* 146(6):866–872.
23. Ono R, Taki T, Taketani T, Taniwaki M, Kobayashi H, Hayashi Y. LCX, leukemia-associated protein with a CXXC domain, is fused to MLL in acute myeloid leukemia with trilineage dysplasia having t(10;11)(q22;q23). *Cancer Res.* 2002;62(14):4075–4080.
24. Delatte B, Fuks F. TET proteins: on the frenetic hunt for new cytosine modifications. *Brief Funct Genomics.* 2013;12(3):191–204.
25. Ito S, Shen L, Dai Q, et al. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science.* 2011;333(6047):1300–1303.
26. Yu M, Hon GC, Szulwach KE, et al. Tet-assisted bisulfite sequencing of 5-hydroxymethylcytosine. *Nat Protoc.* 2012;7(12):2159–2170.
27. Yu M, Hon GC, Szulwach KE, et al. Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome. *Cell*. 2012;149(6): 1368–1380.
28. Navaratnam N, Morrison JR, Bhattacharya S, et al. The p27 catalytic subunit of the apolipoprotein B mRNA editing enzyme is a cytidine deaminase. *J Biol Chem.* 1993;268(28):20709–20712.
29. Teng B, Burant CF, Davidson NO. Molecular cloning of an apolipoprotein B messenger RNA editing protein. *Science.* 1993;260(5115):1816–1819.
30. Muramatsu M, Sankaranand VS, Anant S, et al. Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells. *J Biol Chem.* 1999;274(26):18470–18476.
31. Penterman J, Zilberman D, Huh JH, Ballinger T, Henikoff S, Fischer RL. DNA demethylation in the Arabidopsis genome. *Proc Natl Acad Sci U S A.* 2007;104(16):6752–6757.
32. Rai K, Huggins IJ, James SR, Karpf AR, Jones DA, Cairns BR. DNA demethylation in zebrafish involves the coupling of a deaminase, a glycosylase, and gadd45. *Cell.* 2008;135(7):1201–1212.
33. Cortellino S, Xu J, Sannai M, et al. Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. *Cell.* 2011;146(1):67–79.
34. Krokan HE, Bjørås M. Base excision repair. *Cold Spring Harb Perspect Biol.* 2013;5(4):a012583.
35. Gehring, Reik, et al. DNA demethylation by DNA repair. *Trends Genet.* 2009;25(2):82–90.
36. Jost JP, Siegmann M, Sun L, Leung R. Mechanisms of DNA demethylation in chicken embryos. Purification and properties of a 5-methylcytosine-DNA glycosylase. *J Biol Chem.* 1995;270(17):9734–9739.
37. Zhu B, Zheng Y, Hess D, et al. (2000). 5-methylcytosine-DNA glycosylase activity is present in a cloned G/T mismatch DNA glycosylase associated with the chicken embryo DNA demethylation complex. *Proc Natl Acad Sci U S A.* 2000;97(10):5135–5139.
38. Jost JP, Oakeley EJ, Zhu B, et al. 5-Methylcytosine DNA glycosylase participates in the genome-wide loss of DNA methylation occurring during mouse myoblast differentiation. *Nucleic Acids Research.* 2001;29(21):4452–4461.
39. Shen L, Wu H, Diep D, et al. Genome-wide analysis reveals TET- and TDG-dependent 5-methylcytosine oxidation dynamics. *Cell.* 2013;153(3): 692–706.
40. Thomson JP, Hunter JM, Meehan RR. Deep C diving: mapping the low-abundance modifications of the DNA demethylation pathway. *Genome Biology.* 2013;14(5):118.
41. Song CX, Szulwach KE, Dai Q, et al. Genome-wide profiling of 5-formylcytosine reveals its roles in epigenetic priming. *Cell.* 2013;153(3): 678–691.
42. Schiesser S, Hackner B, Pfaffeneder T, et al. Mechanism and stem-cell activity of 5-carboxycytosine decarboxylation determined by isotope tracing. *Angew Chem Int Ed Engl.* 2012;51(26):6516–6520.