

# Emerging roles of TET proteins and 5-Hydroxymethylcytosines in active DNA demethylation and beyond

Junjie U. Guo,<sup>1,2</sup> Yijing Su,<sup>1,3</sup> Chun Zhong,<sup>1,3</sup> Guo-li Ming<sup>1-3</sup> and Hongjun Song<sup>1-3,\*</sup>

<sup>1</sup>Institute for Cell Engineering; <sup>2</sup>The Solomon H. Snyder Department of Neuroscience; <sup>3</sup>Department of Neurology; Johns Hopkins University School of Medicine; Baltimore, MD USA

**Key words:** TET1, 5-hydroxymethylcytosine, active DNA demethylation, epigenetic, DNA methylation, hippocampus, electroconvulsive stimulation, Gadd45b, BER

**Abbreviations:** TET, ten-eleven translocated oncogene; C, cytosine; 5mC, 5-methylcytosine; 5hmC, 5-hydroxymethylcytosine; Gadd45, growth arrest and DNA damage induced gene 45; BER, base excision repair; NER, nucleotide excision repair; AID, activation-induced cytidine deaminase; APOBEC, apolipoprotein B mRNA editing enzyme catalytic polypeptide; DNMT, DNA methyltransferase; MBD, methylated DNA binding domain-containing protein

Cytosine methylation is the major epigenetic modification of metazoan DNA. Although there is strong evidence that active DNA demethylation occurs in animal cells, the molecular details of this process are unknown. The recent discovery of the TET protein family (TET1–3) 5-methylcytosine hydroxylases has provided a new entry point to reveal the identity of the long-sought DNA demethylase. Here, we review the recent progress in understanding the function of TET proteins and 5-hydroxymethylcytosine (5hmC) through various biochemical and genomic approaches, the current evidence for a role of 5hmC as an early intermediate in active DNA demethylation and the potential functions of TET proteins and 5hmC beyond active DNA demethylation. We also discuss how future studies can extend our knowledge of this novel epigenetic modification.

## History of Active DNA Demethylation

Methylation at the C-5 position of cytosine (C) bases has long been considered the only biologically functional epigenetic covalent modification of the animal genomic DNA. In mammals, 5-methylcytosines (5mCs) are almost exclusively found in CpG dinucleotides, with the exception of non-CpG methylation found in pluripotent stem cells.<sup>1,2</sup> CpG methylation plays critical roles in transcriptional silencing of genes and retrotransposons, gene imprinting and X chromosome inactivation.<sup>3</sup> Deficiency of DNA methyltransferases (DNMTs), enzymes that add methyl groups onto Cs, leads to profound developmental defects.<sup>4</sup> Substantial evidence supports that the CpG methylation pattern across the genome can be replicated across cell division by the maintenance DNMT (DNMT1).<sup>5</sup> Indeed, DNMT1 has higher catalytic activity on hemimethylated DNA than on unmethylated DNA, supporting the notion that DNMT1 can replicate the DNA methylation on the parental DNA strand to the newly synthesized strand. It should be noted that inheritability and stability/reversibility of an epigenetic modification are two distinct and separable properties, although they both contribute to the overall duration of the modification.

According to the proposed replication mechanism, DNA methylation can be “diluted” by not replicating the methylated status to the newly synthesized strand. As a result, methylation at this specific locus will be lost in daughter cells upon further division. This process is termed “passive DNA demethylation.”<sup>6</sup> On the other hand, much effort has been made to understand the enzymatic removal of this covalent modification (active DNA demethylation) in mammalian cells. Such efforts have been partly encouraged by both the identification of histone demethylases<sup>7,8</sup> in mammals and the identification of 5mC glycosylases as DNA demethylases in plants.<sup>9</sup> More importantly, it has been shown in many different systems that active DNA demethylation does occur in mammalian cells.<sup>6</sup> The most well-known example is the rapid loss of 5mC immunoreactivity of the paternal pronucleus in the zygote soon after fertilization.<sup>10</sup> In primordial germ cells (PGCs), gene imprinting-related DNA methylation also needs to be re-established.<sup>10</sup> In contrast, active DNA methylation occurs in a highly locus-specific manner in differentiated somatic cells. For example, rapid demethylation of interleukin-2 (Il2) promoter in activated T cells has been shown to play a causal role in Il2 expression.<sup>11</sup>

Despite accumulating evidence for the presence of mammalian DNA demethylase, the identity of such enzymes has been

\*Correspondence to: Hongjun Song; Email: shongju1@jhmi.edu  
Submitted: 06/28/11; Accepted: 07/01/11  
DOI: 10.4161/cc.10.16.17093

controversial. Methylated DNA binding domain-containing protein MBD2 was reported to efficiently demethylate 5mCs and release methanol with no co-factor requirement.<sup>12</sup> However, no independent confirmation from other laboratories has been reported so far. Analogous to plant demethylases, mammalian DNA glycosylase TDG has been shown to possess weak 5mC glycosylase activity,<sup>13,14</sup> although it is unclear whether TDG is indeed responsible for active DNA demethylation observed *in vivo*. Other recent studies have pointed to an alternative mechanism that involves the sequential steps of deamination by AID/APOBEC deaminases<sup>15</sup> or DNMTs,<sup>16</sup> generation of a T:G mismatch, base excision repair (BER) initiated by T:G mismatch glycosylases, including TDG and MBD4, and refill of unmodified mononucleotides. Multiple lines of evidence have emerged in support of this BER-based mechanism, including extensive presence of DNA break markers, such as  $\gamma$ H2A.X and PARP-1, on the paternal pronuclei,<sup>17</sup> mechanistic requirements for BER enzymes<sup>18</sup> and a modest hypermethylation phenotype in AID-null PGCs.<sup>19</sup> Interestingly, demethylation of PGCs was still largely conserved in AID-null PGCs, suggesting that other pathways also play significant roles in the process.

Indirect/non-catalytic modulators have also been reported to participate in active DNA demethylation. In an overexpression screening, Gadd45a was found to promote global DNA demethylation through DNA repair.<sup>20</sup> Similar efforts have not only confirmed the roles of Gadd45 family proteins, but have also identified other proteins that might be involved in this process, including RING finger protein 4.<sup>21</sup> The detailed mechanisms by which these factors regulate active DNA demethylation remain to be determined.

### Discovery of the TET Proteins and 5hmCs

It was first reported in 1971 that 5hmCs were present in both rodent and frog brain DNA,<sup>22</sup> although the reported abundance was significantly higher than recent studies using more accurate methods.<sup>23-27</sup> This discovery did not attract significant attention until 2009, when two laboratories independently reported the groundbreaking (re)discovery of 5hmCs in mammalian genomic DNA. In one study, Kriaucionic and Heintz used an elegant genetic labeling approach to purify nuclei from two distinct neuronal subtypes from the mouse cerebellum.<sup>28</sup> Purkinje cells have large and generally euchromatic nuclei, whereas granule cells have much smaller and heterochromatic nuclei. When they used the nearest-neighbor assay to determine if the global methylation levels were different between the two subtypes, they observed an uncharacterized mononucleotide signal, more prominent in Purkinje cells than in granule cells. Further chemical characterization identified this unknown signal as 5hmC. They estimated the abundance of 5hmC in Purkinje cells was ~0.6% of all Cs, which translated to almost 40% of all 5mCs.

In contrast to the serendipity of the Heintz study, Rao and colleagues set out to look for mammalian enzymes that have the potential to modify DNA bases.<sup>29</sup> A homology search for the trypanosome thymine hydroxylases JBP1/2 led to three paralogous

human TET proteins, with their orthologs found throughout metazoan genomes. An elegant series of experiments showed that TET1 does not modify thymines but, rather, 5mCs both *in vitro* and in mammalian cells, generating 5hmCs. They further showed that the 5hmC level is higher in mouse embryonic stem cells (mESCs; 4.4% of all CCGGs) than in other cell types they measured, and both Tet1 expression and 5hmC levels are decreased upon differentiation of mESCs, linking this novel modification to pluripotency. Following these two studies, many other laboratories have confirmed and extended these results using various methods.<sup>24-27,30,31</sup>

### 5hmC as an Intermediate of Active DNA Demethylation: Indirect Evidence

The discovery of 5hmCs in mammalian DNA immediately led to wide speculation on its biological function, the dominant one being that 5hmCs may represent an intermediate product in the process of active DNA demethylation. The reason is that 5hmC satisfies a possible oxidative demethylation mechanism, analogous to reactions in the thymidine salvage pathway<sup>32</sup> and the direct repair of DNA alkylation damages by AlkB oxygenases.<sup>33,34</sup> In fact, the first piece of evidence in line with this hypothesis came from the very first paper that identified TET1 as a 5mC hydroxylase,<sup>29</sup> where the authors observed a slight but statistically significant increase in unmodified cytosine content upon TET1 overexpression in HEK293 cells.

Studies from other laboratories have provided more evidence supporting the oxidative demethylation hypothesis. Zhang and colleagues showed that Tet1 plays a role in mESC self-renewal and lineage commitment partly by maintaining the promoter of a pluripotency gene *Nanog* in an unmethylated state.<sup>30</sup> Tet1 knockdown caused *de novo* methylation of the *Nanog* promoter and gene silencing. Rao and colleagues also reported hypermethylation of a few gene promoters caused by Tet1/2 knockdown in mESCs, but hypomethylation at other promoters was also observed.<sup>35</sup> Shi and colleagues reported that TET1 overexpression led to active demethylation of artificially methylated plasmid DNA.<sup>31</sup>

The TET1 gene was originally identified through its translocation in acute myeloid leukemia (AML).<sup>36,37</sup> Later, TET2 was also found to be frequently mutated in various forms of myeloid malignancies.<sup>38</sup> A number of studies have provided fascinating links between TET2 malfunction, an oncogenic metabolite 2-hydroxyglutarate (2HG) and myeloid differentiation and malignancies.<sup>39-41</sup> One study showed a DNA hypermethylation phenotype (previously known in many AML cases) in TET2-mutated AML samples,<sup>40</sup> which is consistent with the oxidative demethylation hypothesis. Surprisingly, in another study, Rao and colleagues observed a DNA hypomethylation phenotype in TET2-mutated AML samples.<sup>39</sup> While the two studies used different methods for DNA methylation profiling, the exact reasons for the discrepancies remain unclear.

Multiple groups have asked the same question, whether 5hmC plays any role in zygotic paternal DNA demethylation.<sup>18,42,43</sup> Strikingly, 5hmC immunoreactive signals were found exclusively

on the paternal pronuclei, highly correlated with the disappearance of its 5mC immunoreactivity. The responsible hydroxylase appeared to be TET3. This observation resolved the apparent discrepancy between the global loss of 5mC immunoreactivity and the modest decrease in bisulfite sequencing-measured methylation levels of individual loci. It has been further shown that, although modest demethylation occurs, 5hmC signals on the paternal pronuclei persist upon multiple rounds of cell division, demarcating the paternally originated chromosomes.<sup>42,43</sup> These observations raise fascinating questions about how the imbalance of 5hmC levels between the two sets of chromosomes affects early development.

### Direct Evidence for 5hmC Demethylase Activity in Mammalian Cells

Although the abovementioned studies have significantly strengthened the oxidative demethylation hypothesis, it was still unproven whether 5hmCs generated by TET proteins are indeed converted to unmodified Cs. Changes in the methylation status of specific genomic loci could very well be indirect consequences of gain- or loss-of-function of TET proteins. This possibility may explain some of the differences in the methylation phenotypes in TET1/2 loss-of-function models. To directly test the hypothesis, an artificial DNA probe was designed consisting of a human Ubiquitin C promoter driving a GFP open reading frame (UbC-GFP).<sup>44</sup> This design could potentially allow one to measure any transcriptional regulatory effect of cytosine modifications. It was fortunate that most Taq DNA polymerases could incorporate pre-modified 5m-dCTP (5-methyl-2-deoxycytidine-5-triphosphate) and 5hm-dCTP (5-hydroxymethyl-2-deoxycytidine-5-triphosphate) just as efficiently as dCTP (2-deoxycytosine-5-triphosphate), making it possible to generate fully pre-modified probes by PCR. Fully hydroxymethylated UbC-GFP probes (5hmC-GFP) exhibit modest yet significant sensitivity to a methylation-sensitive enzyme HpaII at 48 h after transfection into HEK293 cells. More importantly, bisulfite sequencing, which also does not distinguish 5mC and 5hmC, shows the presence of unmodified Cs in the transfected 5hmC-GFP probes but not in untransfected 5hmC-GFP probes or transfected 5mC-GFP probes. These results provide the first direct evidence for a robust 5hmC-specific DNA demethylase activity in mammalian cells. Notably, such demethylation activity for 5hmCs is not limited to a CpG context.

This simple assay using an easily accessible cell line, such as HEK293 cells, greatly facilitated the understanding of its molecular details (Fig. 1).<sup>44</sup> First, pharmacological inhibition of key BER enzymes could partially block 5hmC demethylation, consistent with a requirement for these enzymes in zygotic demethylation.<sup>18</sup> Second, overexpression of each of the 12 known human DNA glycosylases does not have significant impact on global 5hmC levels or demethylation efficiency of 5hmC-GFP, suggesting that BER might be initiated by an indirect mechanism. Finally, overexpression of a few members of the AID/APOBEC deaminase family could enhance demethylation of 5hmC-GFP probes, suggesting that these enzymes might have a direct or indirect role in

this process. It will be important for future studies to determine directly the capacity of these deaminases to use 5hmC as their substrates to generate 5-hydroxymethyluracil (5hmU).

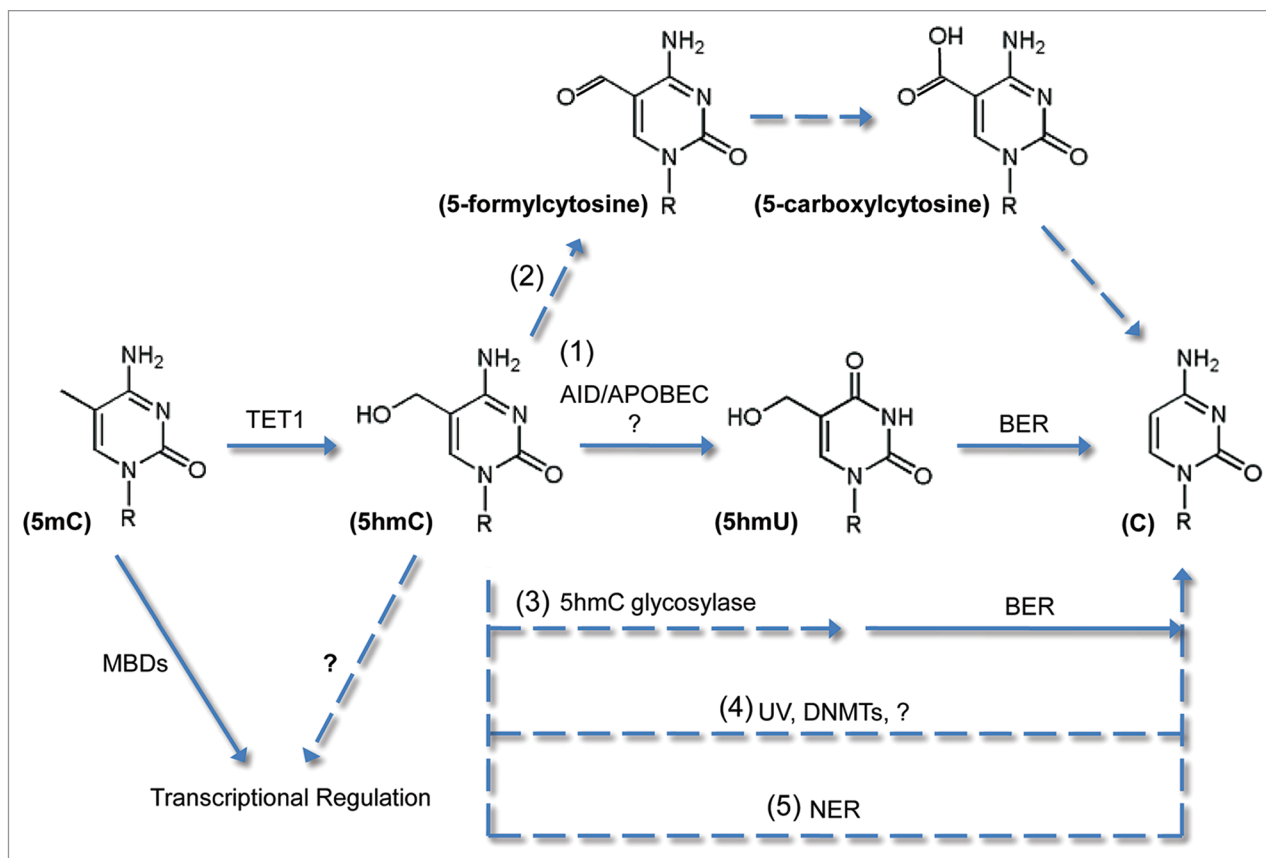
The single-strand, single-nucleotide resolution of bisulfite sequencing revealed more mechanistic properties of the mammalian 5hmC demethylase activity.<sup>44</sup> First, 5hmC demethylation appears to be highly processive, with a few DNA strands being extensively demethylated, while others barely showed any demethylated 5hmCs. Within each cluster of demethylated 5hmCs, very few 5hmCs are skipped (unpublished observations), implying that a “hopping” mechanism is unlikely. Second, transcription plays a critical role in the process. Analysis of promoter-truncated transcriptionally inactive 5hmC-GFP probes showed much less demethylation. It is possible that the 5hmC demethylase targets single-strand DNA, or that the 5hmC demethylase is recruited by specific components of the transcription initiation/elongation machineries. Finally, of the two DNA strands in the transcribed region, 5hmC demethylation is much more efficient on the non-transcribed (sense) strand than on the transcribed (antisense) strand. All these mechanistic aspects resemble some of the known properties of AID and somatic hypermutation,<sup>45,46</sup> making deamination a more promising candidate step for 5hmC demethylation.

Other possibilities exist for the bona fide 5hmC demethylase (Fig. 1). First, it is not known whether TET proteins or other enzymes can further oxidize the hydroxymethyl group to generate formyl and carboxyl groups, which may serve as better leaving groups. Second, 5hmC glycosylase activity was found in calf thymus extract over 20 y ago,<sup>47</sup> although the identity of this glycosylase still remains unknown. Third, direct reversal of 5hmCs to unmodified Cs has been reported to occur under certain conditions or to be performed by bacterial methyltransferase HhaI.<sup>48</sup> Finally, similar to its potential role in 5mC demethylation,<sup>20,49</sup> nucleotide excision repair (NER), which does not involve modification-specific recognition, may be utilized to re-synthesize the DNA strand that contains 5hmCs. On the other hand, 5hmC in certain genomic contexts may remain stable and resistant to the 5hmC demethylase. It may reverse the silencing effect of 5mCs mediated by MBD proteins, which do not seem to recognize 5hmC-containing DNA,<sup>50</sup> or it may have its own protein “readers” to achieve transcriptional regulations.

### Active DNA Demethylation in the Mammalian Brain

An advantageous experimental system to study active DNA demethylation would be non-dividing cells. Post-mitotic neurons in the mammalian brain, in particular, have attracted much attention because of their remarkable ability to alter gene expression profiles in response to external stimuli.<sup>51-53</sup> Active DNA demethylation has been shown to occur both in cultured neurons<sup>54</sup> and in various brain regions in vivo.<sup>55-57</sup> Genetic<sup>58,59</sup> and pharmacological<sup>57,60</sup> studies have further indicated important roles for DNA methylation/demethylation dynamics in regulating neuronal plasticity and animal behaviors.

The adult dentate gyrus provides a relative homogenous population of post-mitotic neurons that can be activated synchronously



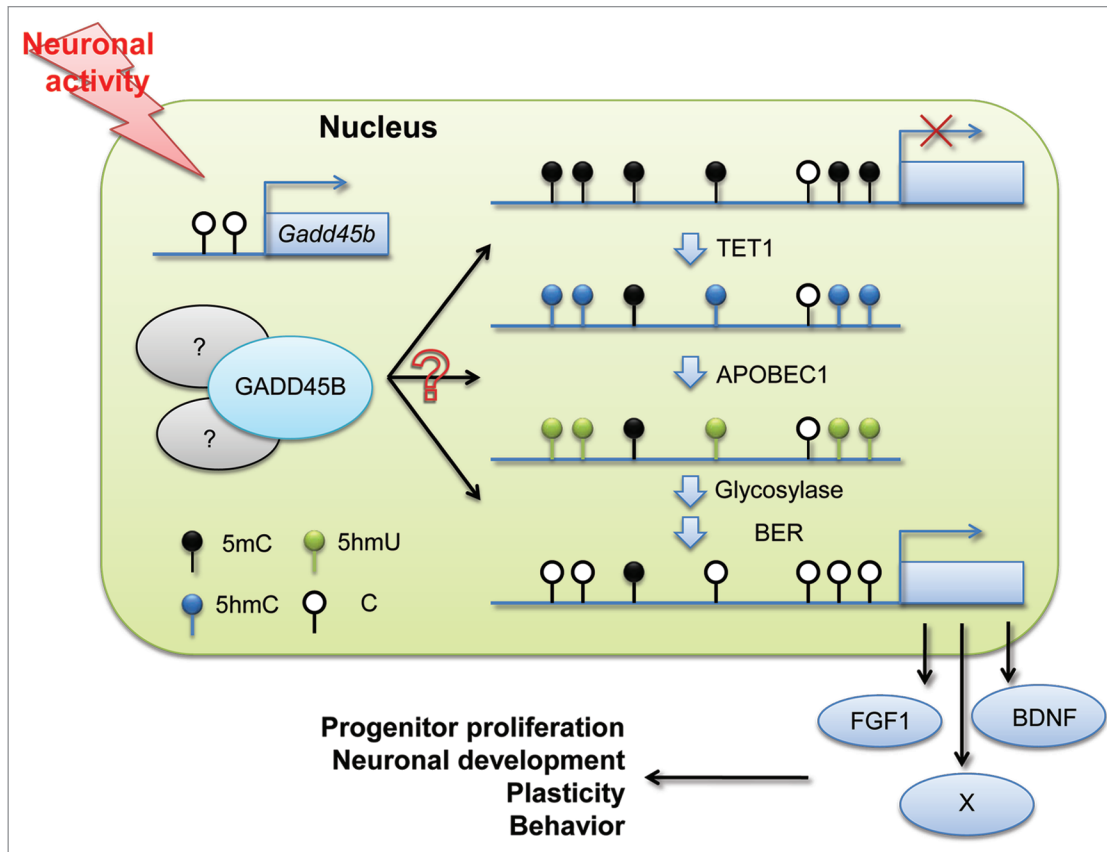
**Figure 1.** Potential mechanisms for active DNA demethylation through 5hmC. (1) Evidence suggests that 5hmC may be converted to 5hmU by AID/APOBEC deaminases and repaired by the base excision repair (BER) pathway. (2) 5hmC may also be further oxidized to 5-carboxylcytosine and decarboxylated/repaired to C. (3) A 5hmC glycosylase may excise 5hmC and directly initiate BER. (4) Direct conversion from 5hmC to C may also occur. (5) Nuclear excision repair (NER) pathway may excise 5hmC-containing DNA strands. Besides promoting DNA demethylation, 5hmC may regulate transcription through unidentified 5hmC-binding proteins. Solid arrows indicate catalytic steps that are supported by experimental evidence, whereas dashed arrows indicate speculative processes.

in vivo and made readily accessible in large quantities.<sup>54</sup> Such properties are particularly important for epigenetic analysis, because each diploid cell displays only two locus-specific modifications. In a search for epigenetic factors that may mediate neuronal activation-induced changes in gene expression, Gadd45b (growth arrest and DNA damage induced gene 45  $\beta$ ) was found to be dramatically induced by electroconvulsive stimulation (ECS), a treatment currently used in clinics for depression disorders.<sup>54</sup> As discussed above, Gadd45 family proteins have been shown to be enhancing factors in DNA demethylation.<sup>15,20,21,49,61</sup> Importantly, Gadd45b plays a critical role in ECS-induced demethylation of *Bdnf* (brain-derived neurotrophic factor) and *Fgf1* (fibroblast growth factor 1) promoters in the adult mouse dentate gyrus (Fig. 2). Activity-induced expression of these neurogenic niche factors and activity-enhanced neurogenesis in the dentate gyrus are both greatly attenuated in Gadd45b-deficient mice, suggesting an important role for DNA demethylation in these processes.<sup>51,62</sup> Moreover, endogenously expressed Tet1 and AID homolog Apobec1 also appear to be involved in neuronal activity-induced demethylation of *Bdnf* and *Fgf1* promoters.<sup>44</sup> Consistent with this dynamic DNA methylation regulation in neurons, cultured mouse hippocampal neurons can

also demethylate the fully modified 5hmC-GFP probes with high processivity.<sup>44</sup> It remains to be determined whether or not and how Gadd45b and TET1 may interact to regulate active DNA demethylation in neurons in vivo (Fig. 2). The impact of Gadd45b or TET1 deletion in mature neurons on neuronal function and animal behaviors also needs to be determined. It is interesting to note that Gadd45b has been shown to be abnormally expressed in mental retardation patients carrying mutations in UPF3B, a member of the nonsense-mediated mRNA decay complex,<sup>63</sup> and in brain tissue of autistic patients,<sup>64,65</sup> supporting a critical role of epigenetic neuronal DNA demethylation in normal brain function.

### Genomic Distribution of TET Proteins and 5hmCs

The past decade has witnessed a revolution of genomics incited by massively parallel DNA sequencing technologies. Soon after the discovery of TET proteins and 5hmCs, significant efforts have been made to characterize TET protein binding and 5hmC deposition across the mammalian genome, resulting in a flurry of papers recently published from a number of laboratories.<sup>66-74</sup> The first published 5hmC profile was that of mouse cerebellum,<sup>71</sup>



**Figure 2.** Neuronal activity-induced DNA demethylation in the adult brain. Neuronal activation in the dentate gyrus induces expression of Gadd45b, which plays an essential role in demethylation of Bdnf and Fgf1 promoters. This process also involves Tet1 and Apobec1. Gadd45b-dependent DNA demethylation leads to expression of important genes, such as Bdnf and Fgf1, and can potentially have a broad impact in the mature nervous system.

where 5hmCs were present in high abundance. He, Jin and colleagues ingeniously leveraged on the known biology of 5hmC in T-even bacteriophage and used T4  $\beta$ -glucosyltransferase to tag 5hmCs with glucose-like moieties that could be further biotinylated and affinity-enriched.<sup>71</sup> Unlike 5mC, which is abundant all over the genome, 5hmC signals appeared to be enriched around genes. Furthermore, 5hmC levels in gene bodies positively correlated with gene expression. Similar findings were reported later using an hMeDIP (hydroxymethylated DNA immunoprecipitation) method.<sup>72</sup>

Several studies have focused on mESCs and have used high quality TET1 antibodies to generate Tet1 ChIP-seq profiles.<sup>67,69,73</sup> Surprisingly, Tet1 shows strong binding to predominantly CpG islands, which are known to be largely resistant to DNA methylation. Tet1-bound promoters show significantly lower 5mC levels, reflecting the enzymatic action of Tet1. Bisulfite sequencing further showed that aberrant DNA methylation occurs on these normally unmethylated CpG islands upon Tet1 knockdown,<sup>69</sup> establishing a role for Tet1 in maintaining the unmethylated states of these critical regulatory regions. This finding has been confirmed on a genome-wide scale by Shi and colleagues.<sup>73</sup> It will be interesting to determine whether this function of Tet1 is mediated by its CXXC domain, which usually interacts with unmethylated CpGs. Surprisingly, Tet1 binds to both active and

repressed (bivalent) genes in mESCs. Tet1 knockdown led to a loss of H3K27me3 modification and transcriptional derepression of many bivalent genes, placing Tet1 in an upstream position in Polycomb repressive complex-mediated gene silencing in mESCs.<sup>69</sup>

Unlike 5mCs, 5hmCs are enriched at gene transcription start sites (TSS).<sup>71</sup> Loss-of-function experiments show that Tet1 binds to 30–50% of genes marked by 5hmC,<sup>67,68,70</sup> suggesting that other Tet proteins may function in parallel. Multiple studies have made the unexpected observation that Tet1 knockdown leads to more de-repressed genes than deactivated genes. Helin and colleagues showed an even more surprising finding that such transcriptional effects were conserved in DNMT triple-knockout mESCs,<sup>67</sup> suggesting that Tet1 can regulate gene expression independently of its 5mC hydroxylase activity. They further reported the physical interaction and similar genome distribution between Tet1 and SIN3A repressive complex, linking this well-studied transcriptional regulator to the novel functions of Tet1.

### Concluding Remarks

DNA methylation has been traditionally perceived as a binary switch in the genomic DNA. As a majority of CpGs in the

mammalian genome are methylated, its impact on genome functions has been a major topic in epigenetics. With the discovery of TET proteins and 5hmC, we now know that Cs in the genome can exist in more than just two forms. Technological barriers need to be overcome to adapt to such new knowledge. Future methods that can simultaneously distinguish C, 5mC and 5hmC,<sup>75,76</sup> preferably compatible with high-throughput sequencing, will certainly be the new driving force for the field.

5hmC, recently regarded as “the sixth base” in the genome, greatly expands the epigenetic plasticity of the genome and

requires researchers to revisit virtually all known biology of DNA methylation. Our current understanding of this novel epigenetic modification is only the tip of the iceberg.

#### Acknowledgments

We thank K. Christian for comments. Supported by Dr. Miriam and Sheldon G. Adelson Medical Research Foundation, NIH (NS048271, HD069184) to G.L.M., NIH (AG024984, NS047344) to H.J.S., The FARMS Fellowship to J.U.G., and by postdoctoral fellowship from MSCRF to Y.S. and C.Z.

#### References

- Ramsahoye BH, Binizskiewicz D, Lyko F, Clark V, Bird AP, Jaenisch R. Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proc Natl Acad Sci USA* 2000; 97:5237-42; DOI:10.1073/pnas.97.10.5237; PMID:10805783.
- Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, et al. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* 2009; 462:315-22; DOI:10.1038/nature08514; PMID:19829295.
- Bird A. DNA methylation patterns and epigenetic memory. *Genes Dev* 2002; 16:6-21; DOI:10.1101/gad.947102; PMID:11782440.
- Goll MG, Bestor TH. Eukaryotic cytosine methyltransferases. *Annu Rev Biochem* 2005; 74:481-514; DOI:10.1146/annurev.biochem.74.010904.153721; PMID:15952895.
- Jones PA, Liang G. Rethinking how DNA methylation patterns are maintained. *Nat Rev Genet* 2009; 10:805-11; DOI:10.1038/nrg2651; PMID:19789556.
- Wu SC, Zhang Y. Active DNA demethylation: many roads lead to Rome. *Nat Rev Mol Cell Biol* 2010; 11:607-20; DOI:10.1038/nrm2950; PMID:20683471.
- Shi Y, Lan F, Matson C, Mulligan P, Whetstone JR, Cole PA, et al. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* 2004; 119:941-53; DOI:10.1016/j.cell.2004.12.012; PMID:15620353.
- Klose RJ, Zhang Y. Regulation of histone methylation by demethylination and demethylation. *Nat Rev Mol Cell Biol* 2007; 8:307-18; DOI:10.1038/nrm2143; PMID:17342184.
- Zhu JK. Active DNA demethylation mediated by DNA glycosylases. *Annu Rev Genet* 2009; 43:143-66; DOI:10.1146/annurev-genet-102108-134205; PMID:19659441.
- Surani MA, Hajkova P. Epigenetic Reprogramming of Mouse Germ Cells toward Totipotency. *Cold Spring Harb Symp Quant Biol* 75:211-8.
- Bruniquel D, Schwartz RH. Selective, stable demethylation of the interleukin-2 gene enhances transcription by an active process. *Nat Immunol* 2003; 4:235-40; DOI:10.1038/ni887; PMID:12548284.
- Bhattacharya SK, Ramchandani S, Cervoni N, Szyf M. A mammalian protein with specific demethylase activity for mCpG DNA. *Nature* 1999; 397:579-83; DOI:10.1038/17533; PMID:10050851.
- Zhu B, Zheng Y, Hess D, Anglikar H, Schwarz S, Siegmann M, et al. 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 USA* 2000; 97:5135-9; DOI:10.1073/pnas.100107597; PMID:10779566.
- Hardeland U, Bentele M, Jiricny J, Schar P. The versatile thymine DNA-glycosylase: a comparative characterization of the human, *Drosophila* and fission yeast orthologs. *Nucleic Acids Res* 2003; 31:2261-71; DOI:10.1093/nar/gkg344; PMID:12711670.
- 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; 135:1201-12; DOI:10.1016/j.cell.2008.11.042; PMID:19109892.
- Métivier R, Gallais R, Tiffocche C, Le Peron C, Jurkowska RZ, Carmouche RP, et al. Cyclical DNA methylation of a transcriptionally active promoter. *Nature* 2008; 452:45-50; DOI:10.1038/nature06544; PMID:18322525.
- Wossidlo M, Arand J, Sebastiano V, Lepikhov K, Boiani M, Reinhardt R, et al. Dynamic link of DNA demethylation, DNA strand breaks and repair in mouse zygotes. *EMBO J* 2010; 29:1877-88; DOI:10.1038/emboj.2010.80; PMID:20442707.
- Hajkova P, Jeffries SJ, Lee C, Miller N, Jackson SP, Surani MA. Genome-wide reprogramming in the mouse germ line entails the base excision repair pathway. *Science* 2010; 329:78-82; DOI:10.1126/science.1187945; PMID:20595612.
- Popp C, Dean W, Feng S, Cokus SJ, Andrews S, Pellegrini M, et al. Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. *Nature* 2010; 463:1101-5; DOI:10.1038/nature08829; PMID:20098412.
- Barreto G, Schafer A, Marhold J, Stach D, Swaminathan SK, Handa V, et al. Gadd45a promotes epigenetic gene activation by repair-mediated DNA demethylation. *Nature* 2007; 445:671-5; DOI:10.1038/nature05515; PMID:17268471.
- Hu XV, Rodrigues TM, Tao H, Baker RK, Miraglia L, Orth AP, et al. Identification of RING finger protein 4 (RNF4) as a modulator of DNA demethylation through a functional genomics screen. *Proc Natl Acad Sci USA* 2010; 107:15087-92; DOI:10.1073/pnas.1009025107; PMID:20696907.
- Penn NW, Suwalski R, O'Riley C, Bojanowski K, Yura R. The presence of 5-hydroxymethylcytosine in animal deoxyribonucleic acid. *Biochem J* 1972; 126:781-90; PMID:4538516.
- Kinney SM, Chin HG, Vaisvila R, Bitinaite J, Zheng Y, Esteve PO, et al. Tissue specific distribution and dynamic changes of 5-hydroxymethylcytosine in mammalian genome. *J Biol Chem*.
- Krais AM, Park YJ, Plass C, Schmeiser HH. Determination of genomic 5-hydroxymethyl-2'-deoxycytidine in human DNA by capillary electrophoresis with laser induced fluorescence. *Epigenetics* 2011; 6:560-5; DOI:10.4161/epi.6.5.15678; PMID:21593596.
- Globisch D, Munzel M, Muller M, Michalakakis S, Wagner M, Koch S, et al. Tissue distribution of 5-hydroxymethylcytosine and search for active demethylation intermediates. *PLoS ONE* 2010; 5:15367; DOI:10.1371/journal.pone.0015367; PMID:21203455.
- Munzel M, Globisch D, Bruckl T, Wagner M, Welzmueller V, Michalakakis S, et al. Quantification of the sixth DNA base hydroxymethylcytosine in the brain. *Angew Chem Int Ed Engl* 2010; 49:5375-7; DOI:10.1002/anie.201002033; PMID:20583021.
- Szwagierczak A, Bultmann S, Schmidt CS, Spada F, Leonhardt H. Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA. *Nucleic Acids Res* 2010; 38:181; DOI:10.1093/nar/gkq684; PMID:20685817.
- Kriaucionis S, Heintz N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science* 2009; 324:929-30; DOI:10.1126/science.1169786; PMID:19372393.
- Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, et al. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* 2009; 324:930-5; DOI:10.1126/science.1170116; PMID:19372391.
- Ito S, D'Alessio AC, Taranova OV, Hong K, Sowers LC, Zhang Y. Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* 2010; 466:1129-33; DOI:10.1038/nature09303; PMID:20639862.
- Zhang H, Zhang X, Clark E, Mulcahey M, Huang S, Shi YG. TET1 is a DNA-binding protein that modulates DNA methylation and gene transcription via hydroxylation of 5-methylcytosine. *Cell Res* 2010; 20:1390-3; DOI:10.1038/cr.2010.156; PMID:21079648.
- Warn-Cramer BJ, Macrander LA, Abbott MT. Markedly different ascorbate dependencies of the sequential alpha-ketoglutarate dioxygenase reactions catalyzed by an essentially homogeneous thymine 7-hydroxylase from *Rhodotorula glutinis*. *J Biol Chem* 1983; 258:10551-7; PMID:6684117.
- Aas PA, Otterlei M, Falnes PO, Vagbo CB, Skorpen F, Akbari M, et al. Human and bacterial oxidative demethylases repair alkylation damage in both RNA and DNA. *Nature* 2003; 421:859-63; DOI:10.1038/nature01363; PMID:12594517.
- Falnes PØ, Johansen RF, Seiber E. AlkB-mediated oxidative demethylation reverses DNA damage in *Escherichia coli*. *Nature* 2002; 419:178-82; DOI:10.1038/nature01048; PMID:12226668.
- Koh KP, Yabuuchi A, Rao S, Huang Y, Cunniff K, Nardone J, et al. Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. *Cell Stem Cell* 2011; 8:200-13; DOI:10.1016/j.stem.2011.01.008; PMID:21295276.
- 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; 62:4075-80; PMID:12124344.
- Lorsbach RB, Moore J, Mathew S, Raimondi SC, Mukatira ST, Downing JR. TET1, a member of a novel protein family, is fused to MLL in acute myeloid leukemia containing the t(10;11)(q22;q23). *Leukemia* 2003; 17:637-41; DOI:10.1038/sj.leu.2402834; PMID:12646957.
- Bejar R, Levine R, Ebert BL. Unraveling the molecular pathophysiology of myelodysplastic syndromes. *J Clin Oncol* 2011; 29:504-15; DOI:10.1200/JCO.2010.31.1175; PMID:21220588.

39. Ko M, Huang Y, Jankowska AM, Pape UJ, Tahiliani M, Bandukwala HS, et al. Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. *Nature* 2010; 468:839-43; DOI:10.1038/nature09586; PMID:21057493.
40. Figueroa ME, Abdel-Wahab O, Lu C, Ward PS, Patel J, Shih A, et al. Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function and impair hematopoietic differentiation. *Cancer Cell* 2010; 18:553-67; DOI:10.1016/j.ccr.2010.11.015; PMID:21130701.
41. Xu W, Yang H, Liu Y, Yang Y, Wang P, Kim SH, et al. Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of alpha-ketoglutarate-dependent dioxygenases. *Cancer Cell* 2011; 19:17-30; DOI:10.1016/j.ccr.2010.12.014; PMID:21251613.
42. Wossidlo M, Nakamura T, Lepikhov K, Marques CJ, Zakhartchenko V, Boiani M, et al. 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. *Nat Commun* 2011; 2:241; PMID:21407207.
43. Iqbal K, Jin SG, Pfeifer GP, Szabo PE. Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine. *Proc Natl Acad Sci USA* 2011; 108:3642-7; DOI:10.1073/pnas.1014033108; PMID:21321204.
44. 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; 145:423-34; DOI:10.1016/j.cell.2011.03.022; PMID:21496894.
45. Pham P, Branstetter R, Petruska J, Goodman MF. Processive AID-catalysed cytosine deamination on single-stranded DNA simulates somatic hypermutation. *Nature* 2003; 424:103-7; DOI:10.1038/nature01760; PMID:12819663.
46. Chaudhuri J, Tian M, Khuong C, Chua K, Pinaud E, Alt FW. Transcription-targeted DNA deamination by the AID antibody diversification enzyme. *Nature* 2003; 422:726-30; DOI:10.1038/nature01574; PMID:12692563.
47. Cannon SV, Cummings A, Teebor GW. 5-Hydroxymethylcytosine DNA glycosylase activity in mammalian tissue. *Biochem Biophys Res Commun* 1988; 151:1173-9; DOI:10.1016/S0006-291X(88)80489-3; PMID:3355548.
48. Liutkeviciute Z, Lukinavicius G, Masevicius V, Daujotyte D, Klimasauskas S. Cytosine-5-methyltransferases add aldehydes to DNA. *Nat Chem Biol* 2009; 5:400-2; DOI:10.1038/nchembio.172; PMID:19430486.
49. Schmitz KM, Schmitt N, Hoffmann-Rohrer U, Schafer A, Grummt I, Mayer C. TAF12 recruits Gadd45a and the nucleotide excision repair complex to the promoter of rRNA genes leading to active DNA demethylation. *Mol Cell* 2009; 33:344-53; DOI:10.1016/j.molcel.2009.01.015; PMID:19217408.
50. Valinluck V, Tsai HH, Rogstad DK, Burdzy A, Bird A, Sowers LC. Oxidative damage to methyl-CpG sequences inhibits the binding of the methyl-CpG binding domain (MBD) of methyl-CpG binding protein 2 (MeCP2). *Nucleic Acids Res* 2004; 32:4100-8; DOI:10.1093/nar/gkh739; PMID:15302911.
51. Ma DK, Marchetto MC, Guo JU, Ming GL, Gage FH, Song H. Epigenetic choreographers of neurogenesis in the adult mammalian brain. *Nat Neurosci* 2010; 13:1338-44; DOI:10.1038/nn.2672; PMID:20975758.
52. Day JJ, Sweatt JD. DNA methylation and memory formation. *Nat Neurosci* 2010; 13:1319-23; DOI:10.1038/nn.2666; PMID:20975755.
53. Meaney MJ, Ferguson-Smith AC. Epigenetic regulation of the neural transcriptome: the meaning of the marks. *Nat Neurosci* 2010; 13:1313-8; DOI:10.1038/nn1110-1313; PMID:20975754.
54. Martinowich K, Hattori D, Wu H, Fouse S, He F, Hu Y, et al. DNA methylation-related chromatin remodeling in activity-dependent BDNF gene regulation. *Science* 2003; 302:890-3; DOI:10.1126/science.1090842; PMID:14593184.
55. Ma DK, Jang MH, Guo JU, Kitabatake Y, Chang ML, Pow-Anpongkul N, et al. Neuronal activity-induced Gadd45b promotes epigenetic DNA demethylation and adult neurogenesis. *Science* 2009; 323:1074-7; DOI:10.1126/science.1166859; PMID:19119186.
56. Elliott E, Ezra-Nevo G, Regev L, Neufeld-Cohen A, Chen A. Resilience to social stress coincides with functional DNA methylation of the Crf gene in adult mice. *Nat Neurosci* 2010; 13:1351-3; DOI:10.1038/nn.2642; PMID:20890295.
57. Miller CA, Sweatt JD. Covalent modification of DNA regulates memory formation. *Neuron* 2007; 53:857-69; DOI:10.1016/j.neuron.2007.02.022; PMID:17359920.
58. LaPlant Q, Vialou V, Covington HE, 3rd, Dumitriu D, Feng J, Warren BL, et al. Dnmt3a regulates emotional behavior and spine plasticity in the nucleus accumbens. *Nat Neurosci* 2010; 13:1137-43; DOI:10.1038/nn.2619; PMID:20729844.
59. Feng J, Zhou Y, Campbell SL, Le T, Li E, Sweatt JD, et al. Dnmt1 and Dnmt3a maintain DNA methylation and regulate synaptic function in adult forebrain neurons. *Nat Neurosci* 2010; 13:423-30; DOI:10.1038/nn.2514; PMID:20228804.
60. Weaver IC, Cervoni N, Champagne FA, D'Alessio AC, Sharma S, Seckl JR, et al. Epigenetic programming by maternal behavior. *Nat Neurosci* 2004; 7:847-54; DOI:10.1038/nn1276; PMID:15220929.
61. Sen GL, Reuter JA, Webster DE, Zhu L, Khavari PA. DNMT1 maintains progenitor function in self-renewing somatic tissue. *Nature* 2010; 463:563-7; DOI:10.1038/nature08683; PMID:20081831.
62. Ma DK, Guo JU, Ming GL, Song H. DNA excision repair proteins and Gadd45 as molecular players for active DNA demethylation. *Cell Cycle* 2009; 8:1526-31; DOI:10.4161/cc.8.10.8500; PMID:19377292.
63. Tarpey PS, Raymond FL, Nguyen LS, Rodriguez J, Hackett A, Vandeleur L, et al. Mutations in UPF3B, a member of the nonsense-mediated mRNA decay complex, cause syndromic and nonsyndromic mental retardation. *Nat Genet* 2007; 39:1127-33; DOI:10.1038/ng2100; PMID:17704778.
64. Garbett K, Ebert PJ, Mitchell A, Lintas C, Manzi B, Mirnics K, et al. Immune transcriptome alterations in the temporal cortex of subjects with autism. *Neurobiol Dis* 2008; 30:303-11; DOI:10.1016/j.nbd.2008.01.012; PMID:18378158.
65. Voineagu I, Wang X, Johnston P, Lowe JK, Tian Y, Horvath S, et al. Transcriptomic analysis of autistic brain reveals convergent molecular pathology. *Nature* 2011.
66. Pastor WA, Pape UJ, Huang Y, Henderson HR, Lister R, Ko M, et al. Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. *Nature* 2011; 473:394-7; DOI:10.1038/nature10102; PMID:21552279.
67. Williams K, Christensen J, Pedersen MT, Johansen JV, Cloos PA, Rappilber J, et al. TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. *Nature* 2011; 473:343-8; DOI:10.1038/nature10066; PMID:21490601.
68. Ficiz G, Branco MR, Seisenberger S, Santos F, Krueger F, Hore TA, et al. Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. *Nature* 2011; 473:398-402; DOI:10.1038/nature10008; PMID:21460836.
69. Wu H, D'Alessio AC, Ito S, Xia K, Wang Z, Cui K, et al. Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. *Nature* 2011; 473:389-93; DOI:10.1038/nature09934; PMID:21451524.
70. Wu H, D'Alessio AC, Ito S, Wang Z, Cui K, Zhao K, et al. Genome-wide analysis of 5-hydroxymethylcytosine distribution reveals its dual function in transcriptional regulation in mouse embryonic stem cells. *Genes Dev* 2011; 25:679-84; DOI:10.1101/gad.2036011; PMID:21460036.
71. Song CX, Szulwach KE, Fu Y, Dai Q, Yi C, Li X, et al. Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine. *Nat Biotechnol* 2011; 29:68-72; DOI:10.1038/nbt.1732; PMID:21151123.
72. Jin SG, Wu X, Li AX, Pfeifer GP. Genomic mapping of 5-hydroxymethylcytosine in the human brain. *Nucleic Acids Res*.
73. Xu Y, Wu F, Tan L, Kong L, Xiong L, Deng J, et al. Genome-wide Regulation of 5hmC, 5mC and Gene Expression by Tet1 Hydroxylase in Mouse Embryonic Stem Cells. *Mol Cell* 2011; 42:451-64; DOI:10.1016/j.molcel.2011.04.005; PMID:21514197.
74. Stroud H, Feng S, Morey Kinney S, Pradhan S, Jacobsen SE. 5-hydroxymethylcytosine is associated with enhancers and gene bodies in human embryonic stem cells. *Genome Biol* 2011; 12:54; DOI:10.1186/gb-2011-12-6-r54; PMID:21689397.
75. Wallace EV, Stoddard D, Heron AJ, Mikhailova E, Maglia G, Donohoe TJ, et al. Identification of epigenetic DNA modifications with a protein nanopore. *Chem Commun (Camb)* 2010; 46:8195-7; DOI:10.1039/c0cc02864a; PMID:20927439.
76. Flusberg BA, Webster DR, Lee JH, Travers KJ, Olivares EC, Clark TA, et al. Direct detection of DNA methylation during single-molecule, real-time sequencing. *Nat Methods* 2010; 7:461-5; DOI:10.1038/nmeth.1459; PMID:20453866.