EXCLI Journal 2014;13:592-610 – ISSN 1611-2156
Received: March 03, 2014, accepted: May 08, 2014, published: May 27, 2014

Review article:

HYDROXYMETHYLATION OF DNA: AN EPIGENETIC MARKER

Richa, Rajneesh and Rajeshwar P. Sinha*

Laboratory of Photobiology and Molecular Microbiology, Centre of Advanced Study in Botany, Banaras Hindu University, Varanasi-221005, India
* Corresponding author: e-mail: r.p.sinha@gmx.net

ABSTRACT

DNA methylation, an epigenetic mechanism is claimed to play essential roles in development, aging and disease over the past few decades. Cytosines (C) were known to exist in two functional states: unmethylated or methylated (5mC) in the mammalian genome for a very long time. However, the mechanisms controlling 5mC dynamics remain undefined. Recent studies of genomic DNA on human and mouse brain, neurons and from mouse embryonic stem cells have shown that 2-oxoglutarate and Fe(II)-dependent oxygenases of the ten-eleven translocation (Tet) proteins can catalyze the oxidation of 5mC at cpG dinucleotides to form 5-hydroxymethylcytosine (5-hmC). The exhilarating discovery of these novel 5-hmC has begun to focus on the dynamic nature of 5mC. The prevailing evidence has shown that Tet family proteins and 5-hmC are involved in the normal development as well as in many diseases. This review presents an overview of the role of Tet family proteins and 5-hmC. It also discusses their role as an epigenetic marker and the techniques used for their analysis.

Keywords: methylation, hydroxymethylation, TET-proteins, DNA methyltranferases

Abbreviations:
5caC : 5-carboxycytosine; Dnmt : DNA methyltransferase; ESCs : embryonic stem cells; 5fC : 5-formylcytosine; 5-hmC : 5-hydroxymethylcytosine; 5mC : 5-methylcytosine; PGCs : Primordial germ cells

INTRODUCTION

DNA methylation, i.e., methylation of cytosine at CpG sequences, and, to a lesser extent, in CpHpG or CpHpH (where H is A, C or T) sequences, is an epigenetic modification linked to gene expression, regulation and developmental processes in various eukaryotes (Bird, 2002; Feng et al., 2010; He et al., 2011a; Law and Jacobsen, 2010; Lister et al., 2009; Ramsahoye et al., 2000; Suzuki and Bird, 2008; Tomizawa et al., 2011; Zemach et al., 2010). It is a highly dynamic process during germ cell and pre-implantation development but is relatively static during development of somatic tissues in case of mammals. 5mC has been found to be the sole methylation product in higher eukaryotes, where the methylation occurs not only in sequence-specific but also in a locus-specific manner. In mammalian genomes (especially in somatic cells), methylation of cytosine predominantly occurs at CpG dinucleotides, whereas, a fraction of 5mC is associated with non-CG contexts in embryonic stem cells (Lister et al., 2009). CpG motifs are under-represented in mammals and tend to cluster in the high density regions, referred to as CpG islands. Studies related to 5mC distribution have shown that the majority (70–80 %) of CpGs are methylated,

except for those localized in CpG islands. Traditionally, 5mC when localized at CpG islands are important transcriptional silencers at gene promoters. The cytosine methylation typically leads to strong and heritable gene silencing. The epigenetic role of 5-methylcytosine has been widely studied and was found to be crucial in a variety of cellular processes including development gene regulation, differentiation, genomic imprinting, silencing of transposable elements, X-chromosome inactivation and others (Bird, 2002). Because of its significance and heritability, 5mC is often called the fifth base of DNA. The occurrence of 5mC containing photoproducts in UV irradiated DNA is still controversial (Rastogi et al., 2010). However, on the basis of HPLC and Mass spectrometry analyses a new photoproduct of 5mC has also been characterized (Su et al., 2010). It has been assumed by various group of workers that 5mC display a certain level of plasticity during development (Bernstein et al., 2007; Mohn and Schubeler, 2009), therefore, they needs to be rapidly removed by DNA demethylation processes at certain developmental stages. However, the mechanisms of DNA demethylation have always remained controversial (Ooi and Bestor, 2008; Wu and Zhang, 2010).

DNA hydroxymethylation, is a recently identified type of DNA modification in which the hydrogen atom at the C5-position in cytosine is replaced by a hydroxymethyl group, and whose, importance to biology and its role as an epigenetic marker have only been appreciated during the past few years. The presence of 5-hydroxymethylcytosine (5-hmC) has been observed in the genomes of many different organisms, and in mammals its levels are the highest in the brain and other tissues of the central nervous system (Globisch et al., 2010; Kinney et al., 2011; Kriaucionis and Heintz, 2009; Michaeli et al., 2013). 5-hmC is generated through oxidation of 5mC by the Ten-Eleven Translocation (TET) family of methylcytosine dioxygenases (Guo et al., 2011; Hashimoto et al., 2012; Wu and Zhang, 2011a;) and is considered to be a possible intermediate step in a replication-independent DNA demethylation pathway (Guo et al., 2011; Tahiliani et al., 2009; Wu and Zhang, 2010).

The discovery of a hydroxylated form of 5mC (5-hmC) (Kriaucionis and Heintz, 2009) and the TET family of enzymes required for its conversion (Tahiliani et al., 2009) has now offered insight into DNA methylation pathway involved in many biological processes including those pertaining to nervous system function (Kass et al., 2013). These modified bases may then function as DNA demethylation intermediates subject to deamination, glycosylase-dependent excision, and repair resulting in a reversion back to unmodified cytosine (Bhutani et al., 2011; Branco et al., 2012). However, recently, the DNA methyltransferases (DNMT) enzymes (Tan and Shi, 2012) has also been ascribed a role in the 5-hmC production, suggesting that this species may not be exclusively considered an intermediate of the 5mC demethylation process, but also an important epigenetic marker enriched within gene bodies, promoters, and transcription factor binding sites, where it may influence gene expression (Hahn et al., 2013; Mellén et al., 2012; Szulwach et al., 2011a) and thus, regulates the pluripotency of stem cells, cellular development, aging, and carcinogenesis (Liutkeviciute et al., 2009). Three major types of DNA methyltransferases are active on mammalian genomes (Figure 1A). DNMT1, the maintenance methyltransferase, maintains the genome-wide methylation patterns during replication as it preferentially methylates hemi-methylated cytosines in CpG sequences (Bestor, 1992; Li et al., 1992). DNMT3A and DNMT3B can methylate unmethylated CpG sequences and hence function as *de novo* methyltransferases (Okano et al., 1998 a). DNMT3L has no catalytic activity but recruits DNMT3A and DNMT3B to their targets by recognizing nucleosomes that carry unmethylated histone H3 lysine 4 (H3K4) (Aapola et al., 2000; Bourc'his and Bestor, 2004; Bourc'his et al., 2001; Hata et al., 2002; Ooi et al., 2007).

EXCLI Journal 2014;13:592-610 – ISSN 1611-2156
Received: March 03, 2014, accepted: May 08, 2014, published: May 27, 2014

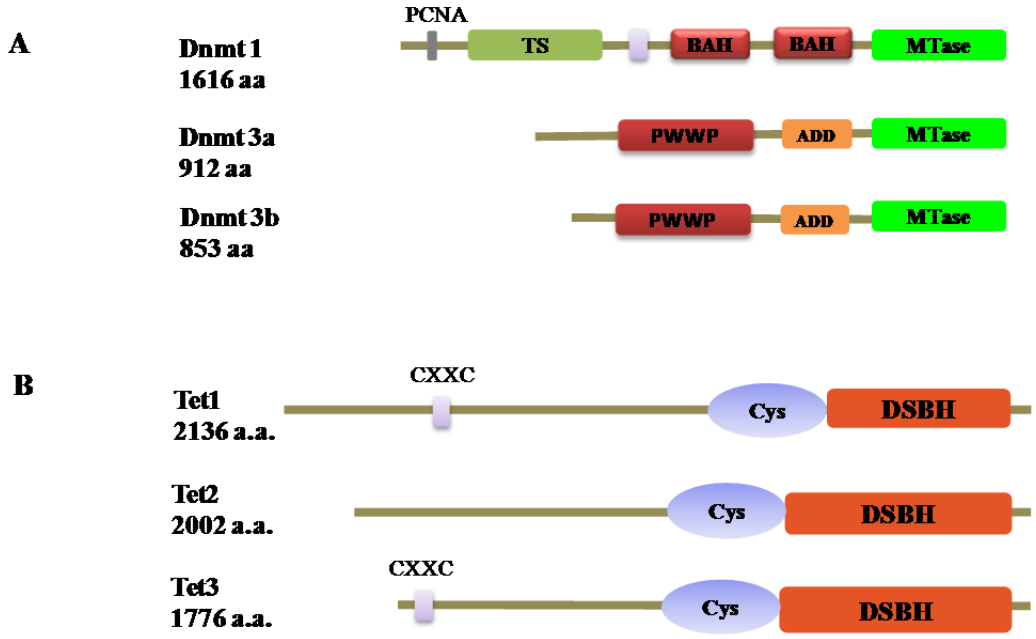

Figure 1: Schematic representation of the human DNA methyltransferase **(A)** and Tet family proteins **(B).** Dnmts share a conserved catalytic domain (MTase), but differ in their N-terminal regulatory regions. Dnmt1 contains a proliferating cell nuclear antigen (PCNA) binding domain, a pericentric heterochromatin targeting sequence (TS), a CXXC domain, and two bromo adjacent homology domains (BAH). Dnmt3A and 3B comprise a PWWP domain named after a conserved Pro-Trp-Trp-Pro motif and an ATRX-Dnmt3-Dnmt3L (ADD) domain also known as the PHD (plant homeodomain) domain. Tet 1, 2, 3 proteins contain a cysteine-rich (Cys) region and a double-stranded β-helix (DSBH) fold responsible for catalytic activity (modified from Kriukiene et al., 2012).

DOMAIN STRUCTURE AND DISTRIBUTION OF TET-FAMILY PROTEINS

TET proteins are mammalian homologs of the trypanosome base J binding proteins, JBP1 and JBP2, that contain catalytic motifs typical of Fe(II)- and 2-oxoglutarate (2OG)-dependent dioxygenases and require α-ketoglutarate as a co-substrate for enzymatic activity (Figure 1B) (Iyer et al., 2009; Tahiliani et al., 2009). In the N-terminal part, the Tet proteins possess a CXXC domain, a binuclear Zn-chelating domain found in certain chromatin-associated proteins such as Dnmt1 methyltransferase, and other elements which mediate interactions with multiple components in the cell (Xu et al., 2011b; Zhang et al., 2010). Unlike the CXXC domains in other proteins, such as, DNMT1, myeloid/lymphoid or mixed-lineage leukemia (MLL) and CXXC finger protein 1 (CFP1, or CXXC1), which are known to bind unmethylated CpG dinucleotides (Cierpicki et al., 2010; Song et al., 2011b; Xu et al., 2011a), the function of this domain in Tet1 and Tet3 is largely unknown.Various group of workers have reported that CXXC domain of TET1 recognizes not only unmodified cytosine but also 5mC and 5-hmc, and it prefers to bind to regions in the genome of high CpG content (Xu et al., 2011b; Zhang et al., 2010). Depending on this feature, genome-wide mapping of Tet1 binding by ChIP-seq approaches revealed its enrichment around transcription start sites (TSSs) in mouse ES cells (Figure 2) (Xu et al., 2011b; Williams et al., 2011; Wu et al., 2011). In addition, to the functional domains of Tet protein action, a spacer region, common to all TET family members but with varying length bridges the two parts of the disconnected DSBH enzymatic domain. The functional significance of this spacer region is currently unknown. However, recently, Upadhyay et al. (2011) have revealed that the spacer region of Tet1

EXCLI Journal 2014;13:592-610 – ISSN 1611-2156
Received: March 03, 2014, accepted: May 08, 2014, published: May 27, 2014

has significant sequence similarity to the C-terminal domain (CTD) of RNA polymerase II of *Saccharomyces cerevisiae*. All three members of the TETs (TET1–TET3) have been shown to specifically catalyze the conversion of 5mC to 5-hmC as well as its further oxidation into 5-formylcytosine (5-fC) and 5-carboxylcytosine (5-caC), in the presence of ATP respectively (Figure 3) (Ito et al., 2010; 2011; He et al., 2011b).

The distribution of TET1/2/3 proteins vary in different tissue (Cimmino et al., 2011; Szwagierczak et al., 2010; Tahiliani et al., 2009) (Table 1). 5-hmC is abundant in the brain (~40 % and ~13 % as abundant as 5mC in Purkinje and granule cells, respectively), but is present at lower levels in other mouse tissues (Kriaucionis and Heintz, 2009). 5-hmc is detected in mouse ESCs (~7-10 % as abundant as 5mC) but is undetectable in human T cells and in mouse dendritic cells (Tahiliani et al., 2009).

TET1 was first reported to have a role in maintaining the pluripotent state of embryonic stem cells (ESCs) (Ito et al., 2010), and together with TET2, it also regulates the cell lineage commitment of ESCs (Koh et al., 2011). TET2 modulates the balance between self-renewal and differentiation in hematopoietic stem cells, making them critical for normal myelopoiesis; TET2 mutations are seen in multiple types of leukemia and also

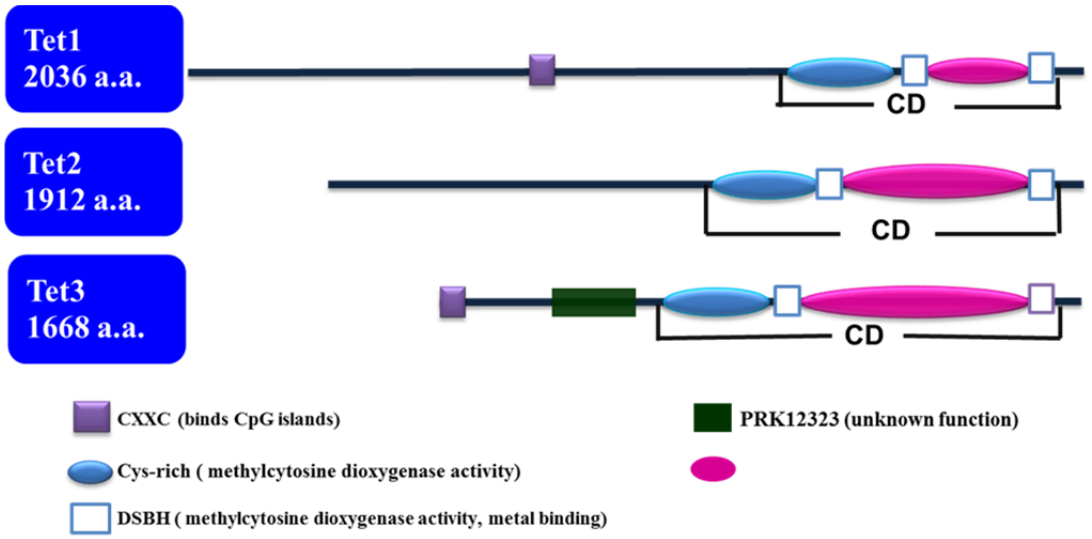

Figure 2: Structure and function of Tet family proteins in mouse.The Tet protein family contains three members: Tet1, Tet2 and Tet3. All Tet proteins contains a catalytic C-terminal CD domain having cysteine-rich and a double-stranded-β-helix, DSBH regions) that exhibits 2-oxoglutarate (2-OG)- and iron (II)-dependent dioxygenase activity. A spacer region of varying length is also present in the CD domain. The N-termini of Tet1 and Tet3, but not Tet2, contain a CXXC domain, which mediates their direct DNA-binding ability. There is also an unique PRK12323 (DNA polymerase III subunits gamma and tau; provisional) domain in Tet3, whose function is still not clear. a.a.: amino acids (modified from Tan and Shi, 2012).

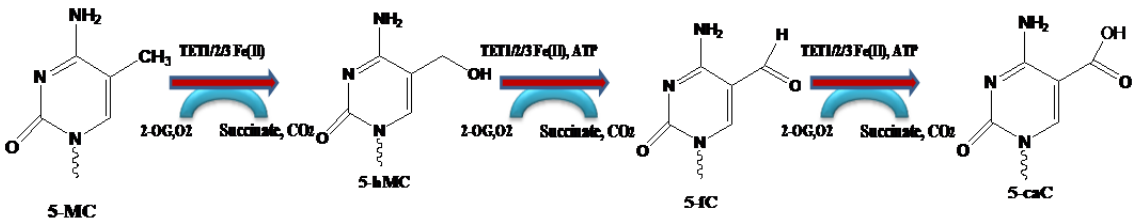

Figure 3: TET proteim mediated oxidation reaction.TET family proteins oxidize 5mC into 5hmC in the presence of 2-OG and iron. 5mC and 5hmC can also be further oxidized mainly to 5caC by Tet proteins in the presence of ATP.

EXCLI Journal 2014;13:592-610 – ISSN 1611-2156
Received: March 03, 2014, accepted: May 08, 2014, published: May 27, 2014

Table 1: Proteins involved in cytosine modification (modified from Saitou et al., 2012)

Protein	Function	Tissue Distribution	References
DNMT 1	Methylation of hemimethylated CpGs during DNA replication	Highly expressed in proliferating cells (ubiquitous expression)	Bestor et al., 1988; Lei et al., 1996
DNMT 2	Methylation of small RNA	High expression in heart, kidney and testis(ubiquitous expression)	Goll et al., 2006; Okano et al., 1998b
DNMT 3A	DNA methylation establishment in early development and germ cells (Involve in imprint establishment during gametogenesis with DNMT3L)	Highly expressed in embryonic tissue and undifferentiated ESCs (ubiquitous expression)	Kaneda et al., 2004; Okano et al., 1998a, 1999
DNMT 3B	DNA methylation establishment in early development and germ cells (crucial for methylation of pericentromeric major and minor satellite repeats)	Highly expressed in embryonic tissue and undifferentiated ESCs (ubiquitous expression)	Kaneda et al., 2004; Okano et al., 1998a,1999; Ueda et al., 2006
DNMT 3L	Non-catalytic activity; essential for the establishment of imprints in oocytes and for silencing of dispersed repeatedsequences in male germ cells	Specific expression in germ cells during gametogenesis and embryonic stages	Bourc'his and Bestor, 2004; Hata et al., 2002; Webster et al., 2005
NP95 (UHRF1)	Recruitment of DNMT1 into the replication foci	Highly expressed in proliferating cells (ubiquitous expression)	Bostick et al., 2007; Fujimori et al., 1998; Sharif et al., 2007
TET 1	Conversion of 5mC to 5hmC; limits accessibility of DNA to DNMTs by binding strongly to unmethylated CpG-rich regions via its CXXC domain	Highly expressed in ESCs	Dawlaty et al., 2011; Ito et al., 2010; Koh et al., 2011; Szwagierczak et al., 2010; Tahiliani et al., 2009; Xu et al., 2011b
TET 2	Conversion of 5mC to 5hmC	Expressed in ESCs and hematopoietic cells(expression in almost all tissues)	Szwagierczak et al., 2010; Figueroa et al., 2010; Ito et al., 2010; Ko et al., 2010; Koh et al., 2011; Langemeijer et al., 2009; Moran-Crusio et al., 2011; Quivoron et al., 2011
TET 3	Conversion of 5mC to 5hmC	Highly expressed in oocyte and zygote (expression in NSCs, lung, spleen and pancreas, etc.)	Gu et al., 2011; Iqbal et al., 2011; Wossidlo et al., 2011; Ito et al., 2010; Szwagierczak et al., 2010
AID	Deamination of cytosine to uracil in ssDNA, of 5mC to thymine and of 5hmC to 5hmU	Expressed in activated B cells and testis; expressed in oocytes, ESCs and PGCs (low level)	Muramatsu et al., 2000; Morgan et al., 2004; Bhutani et al., 2010; Popp et al., 2010; Guo et al., 2011
APOBEC1	Deamination of cytosine to uracil in RNA and DNA, of 5mC to thymine and of 5hmC to 5hmU	Expressed in small intestine	Hirano et al., 1996; Morrison et al., 1996; Navaratnam et al., 1993; Teng et al.,1993; Morgan et al., 2004; Guo et al., 2011

5hmC - 5-hydroxymethylcytosine, 5hmU - 5-hydroxymethyluracil, 5mC - 5-methylcytosine, AID - activation-induced deaminase, AP - apurinic/apyrimidinic, APOBEC1 - apolipoprotein B mRNA editing enzyme, catalytic polypeptides, DNMT - DNA methyltransferase, ESC - embryonic stem cells, IAP - intracisternal A particle, IgM - immunoglobulin M, LINE1 - long interspersed nuclear element 1, NSC - neural stem cell, PGC - primordial germ cell, siRNA - short interfering RNA, ssDNA - single stranded DNA, TET - ten-eleven translocation, UHRF1 - ubiquitin-like, containing PHD and RING finger domains 1.

in patients with diverse myeloid malignancies (Ko et al., 2010, 2011; Li et al., 2011; Moran-Crusio et al., 2011; Quivoron et al., 2011; Shide et al., 2012). Shide et al. (2012) reported the critical role of TET2 in survival and HSC homeostasis after *in vivo* analysis of its function in Ayu17-449 (TET2 trap) mice. TET3 contributes to the global DNA methylation erasure during the zygote stage of embryonic development (Gu et al., 2011; Inoue and Zhang, 2011). These studies together, demonstrates the critical role of TET-mediated 5-hmC modification in developmental processes and the possibility that dysregulation of 5-hmC may be associated with disease.

In mouse ESCs, Tet1 and Tet2 are highly expressed but are rapidly downregulated upon differentiation (Ito et al., 2010; Koh et al., 2011; Saitou et al., 2012; Tahiliani et al., 2009), whereas Tet3 is expressed at very low levels in ES cells and upregulated during ES cell differentiation. Dawlaty et al. (2011) demonstrated that Tet1 knockout ESCs showed a ~35 % reduction in 5-hmc levels, and exhibit only subtle changes in gene expression, are pluripotent and support full-term mouse development in the Tetraploid complementation assay. Koh et al. (2011) reported that both Tet1 and Tet2 are downstream targets of key pluripotency factors.

Wang et al. (2012) have reported that 5-hmC-mediated epigenetic pathways play evolutionarily conserved roles in the mammalian brain. In neuronal cells, 5-hmC is not enriched at enhancers but associates preferentially with gene bodies of activated neuronal function-related genes. Within these genes, gain of 5-hmC is often accompanied by loss of H3K27me3. Enrichment of 5-hmC is not associated with substantial DNA demethylation suggesting the potential of 5-hmc as a stable epigenetic mark. Functional perturbation of the H3K27 methyltransferase Ezh2 or of Tet2 and Tet3 leads to defects in neuronal differentiation suggesting that formation of 5-hmC and loss of H3K27me3 cooperate to promote brain development (Hahn et al., 2013).

The presence of 5-hydroxymethylcytosine in DNA was first observed in T-even bacteriophages (Wyatt and Cohen, 1953). In T-even phages, 5-hydroxymethylated cytosine is incorporated into the genome during DNA synthesis, and is subsequently modified by phage α- and β-glucosyltransferases, creating a highly glucosylated DNA containing 5-glucosyloxymethylcytosine (glc-hmC) residues. The precursor nucleotide, 2'-deoxy-5-hydroxymethylcytidine- 5'-monophosphate (dhmCMP), is produced by enzymatic addition of a methylene group at the 5-position of dCMP. This reaction is mechanistically similar to DNA C5-methylation as it involves the formation of a covalent intermediate between a cysteine residue from an enzyme and the C6-position of cytosine (Warren, 1980). A similar modified base J (5-(β-D-glucosyl)-oxymethyluracil) is present in DNA of flagellated protozoa of the order of Kinetoplastida and unicellular alga *Euglena gracilis* (Borst and Sabatini, 2008). However, discovery of physiologically relevant DNA modification in mammals, such as in mouse neurons and ESCs (Kriaucionis and Heintz, 2009; Tahiliani et al., 2009), is a more recent novel finding.

The chemical oxidation of 5mC and thymine is considered to be the major source of these hydroxymethylated nucleobases in higher eukaryotes. 5-hmC can be formed at 5mC sites in response to oxidative stress *in vitro* (Wagner and Cadet, 2010). Oxidation of thymine residues in DNA to hmU was considered to be an important source of endogenous oxidative DNA damage. 5mC is slightly more reactive towards hydroxyl radical attack than is thymine and it was predicted that in human cells, ~20 5mC residues would be oxidized to hmC per cell per day (Planechaud et al., 1997).

Valinluck et al. (2004) have reported that 5-hmC may also be a biological end-product of demethylation, as methyl-CpG binding proteins have a significantly lower affinity for 5-hmC. The ability of DNMT1 to recognize 5-hmC under *in vitro* condition is very poor (Valinluck and Sowers, 2007); thus, 5-

EXCLI Journal 2014;13:592-610 – ISSN 1611-2156
Received: March 03, 2014, accepted: May 08, 2014, published: May 27, 2014

hmc might passively convert into cytosine during replication. However, recently, it has been worked out by Frauer et al. (2011) that Np95, which recruits DNMT1 to replication foci, recognizes 5-hmC as efficiently as it does 5mC, raising the possibility that 5-hmC may have the same capacity as 5-mC for 5-mC propagation during replication.

Szulwach et al. (2011b) reported that 5-hmC-mediated epigenetic modification is critical in neurodevelopment and diseases both in mice and human. The kinetics study of hydroxymethylation during the first cell stage have shown that hydroxymethylation remains very low in the maternal pronucleus while it increases in the paternal one, in parallel to demethylation (Inoue and Zhang, 2011; Iqbal et al., 2011; Wossidlo et al., 2011; Ruzov et al., 2011; Gu et al., 2011). Based on their work on mouse zygote Salviang et al. (2012) hypothesized that 5-hydroxymethylcytosine is not a simple intermediate in an active demethylation process and could play a role of its own during early development.

The presence of 5-hmC within eukaryotic genomes was unknown mainly because standard methods for detection of 5-methylcytosine (5-mC) cannot distinguish between this base and 5-hmC (Huang et al., 2010; Münzel et al., 2010a; Jin et al., 2010). A newly developed method for single-nucleotide resolution 5-hmC sequencing, based on oxidative conversion of 5-hmC to 5-formyl-cytosine and its further bisulfite conversion to uracil was given by Booth et al. (2012). Overall, 5-hmC content might also be determined *via* selective chemical labeling with β-glucosyltransferase of 5-hmC by a glucose moiety containing an azide group modified, for example, by biotin (Pastor et al., 2011; Song et al., 2011a). Wanunu et al. (2010) reported another approach using electric signal for detection on a DNA solution flown through nanopores. However, to understand the principles of recognition and binding of 5-hmC in biological context requires its atomic structure. The interaction of 5-hmC-modified DNA with the SET and RING-associated domain of the Uhrf1 protein that belongs to the big family of 5-mC binding proteins was studied by crystal structure and molecular dynamics analyses (Frauer et al., 2011) and it was found that 5-hmC base is flipped out of the DNA double helix and that the 5-hmC hydroxyl group participates in hydrogen bonding, which helps to stabilize the interaction. Expansion of studies on 5-hmC was enabled by the development of suitable phosphoramidites (Münzel et al., 2011). These authors found that both 5-hmC and 5-mC were not causing any insertion of wrong nucleotides even by low-fidelity polymerases, and therefore, concluded that position 5 of cytosine is the ideal place to store epigenetic information. By using X-ray diffraction Renciuk et al. (2013) reported that the cytosine modification (5mC and 5-hmC) neither influences the helical structure of B-DNA nor the thermodynamics.

ROLES OF 5-HMC

(a) DNA methylation and gene expression

Multiple pathways have been suggested by various workers for 5-hmc mediated regulation of the DNA methylation dynamics and gene transcription. 5-hmC was found to be more sensitive to deamination than 5mC by activation-induced deaminase (AID), and the deamination product, 5-hydroxymethyluracil (5hmU), activates base-excision repair (BER) pathway-mediated demethylation (Guo et al., 2011), suggesting one possible way to acheive DNA demethylation. Recently, it has been shown that Tet proteins can oxidize 5mC not only to 5-hmc but also to 5fC and 5caC, which can subsequently be recognized and excised by thymine DNA glycosylase (TDG) *in vitro* and *in vivo* (He et al., 2011b; Ito et al., 2011; Pfaffeneder et al., 2011; Zhang et al., 2012). The gradual loss of 5-hmC, 5fC and 5caC, associated with the paternal genome in the zygote during preimplantation development suggested a dominant mechanism of passive DNA demethylation at this developmental stage (Inoue and Zhang, 2011; Inoue et al., 2011; Oda et al. 2009). 5-hmC is important for proper gene

transcription as it directs the dynamic remodeling and organization of the chromatin structures (Tan and Shi, 2012). Jin et al. (2010) reported that many methyl-CpG-binding domain (MBD) proteins, such as MBD1, MBD2, methyl CpG-binding protein 2 (MECP2) and MBD4, do not bind 5-hmC *in vitro*. However, recently it has been proposed that Tet1 is associated with the Mbd3/NuRD complex in mouse ES cells and that Mbd3 can bind 5-hmC and unmodified cytosine, but not 5mC, *in vitro* (Yildirim et al., 2011). Hashimoto et al. (2012) based on his studies on DNA methyltranferases reported the high intrinsic activity of DNMT1 towards the hemi-methylated CpG substrate but a greatly reduced activity for the hemi-hydroxymethylated CpG substrate in methyl transfer assays *in vitro*, whereas the DNMT3A and DNMT3B exhibited equal activities on unmodified, hemimethylated and hemi-hydroxymethylated CpG substrates. Based on these studies it can be considered that 5-hmC acts as a stable epigenetic marker that directly influence genome structure and function, instead of being simply an intermediate of DNA demethylation.

(b) Mammalian development

(i) Mouse primordial germ cells and gametes

Gene knockout strategies have been used to investigate the role of DNA methylation and demethylation involving TET family proteins and 5-hmc in mammalian development. The mouse oocyte and sperm develop from primordial germ cells (PGCs) which are specified from epiblast cells and migrate into the gonad region at embryonic day (E) 7.25. The genome-wide DNA demethylation, erasure of genomic imprints, and large-scale chromatin remodeling takes place at E11.5 (Hajkova et al., 2008; Kota and Feil, 2011). Both Tet1 and Tet2 are expressed in E11.5 and E12.5 PGCs but less so in oocytes and zygotes, whereas Tet3 is highly expressed in oocytes but not in PGCs (Hajkova et al., 2010; Surani and Hajkova, 2010; Gu et al., 2011). The development of PGCs, and subsequently of spermatocytes or oocytes, is normal in *Tet1* or *Tet2* knockout mice (Dawlaty et al., 2011; Ko et al., 2011; Li et al., 2011; Moran-Crusio et al., 2011; Quivoron et al., 2011), whereas some workers established PGC-conditional *Tet3* knockout mice and found that oogenesis and spermatogenesis are normal when Tet3 is absent in PGCs (Gu et al., 2011).

(ii) Mouse zygote

When DNA methylation ends up in parental genome rapidly after fertilization, then DNA demethylation takes place (Oswald et al., 2000). Recently, it has been reported that the oxidation of 5mc to 5-hmc takes place only in paternal but not in the maternal pronucleus (Iqbal et al., 2011; Wossidlo et al., 2011). In contrast to Tet1 and Tet2, TET 3 are highly expressed in oocytes and fertilized zygotes but disappears rapidly during cleavage (Gu et al., 2011; Iqbal et al., 2011; Wossidlo et al., 2011). Gu et al. (2011) showed maternal Tet3 to be the key enzyme catalyzing 5mC into 5-hmC in the paternal pronuclei of developing zygotes. Nakamura et al. (2007) reported the presence of DNA demethylation and 5-hmc in the maternal pronuclei of *Stella* (*Dppa3*) null zygotes. Gu et al. (2011) reported the requirement of oxidation of 5mc to 5-hmc in oocytes during somatic cell nuclear transfer.

(iii) Mouse pre- and post-implantation embryos and in ES cells

High levels of Tet1 and 5-hmC are found in mouse ES cells, inner cell mass of the mouse blastocyst, and in cells of the early mouse epiblast (Ito et al., 2010; Koh et al., 2011; Ruzov et al., 2011). Recently, it has been found that the shRNA-mediated knockdown of *Tet1*, but not *Tet2* or *Tet3*, led to the spontaneous differentiation of mouse ES cells even in the presence of LIF (Ito et al., 2010). However, others believe that Tet1 is not required for maintaining the self-renewal of mouse ES cells (Koh et al., 2011; Williams et al., 2011; Xu et al., 2011b). The role of knockdown and knockout of *Tet1* in forcing the transdifferentiation of mouse ES cells into trophoblasts has not been detected *in vi-*

vo (Dawlaty et al., 2011; Ito et al., 2010), and *Tet1–/–* and *Tet2–/–* mice appeared to develop normally, appear healthy through adulthood and were found to be fertile (Dawlaty et al., 2011; Li et al., 2011; Moran-Crusio et al., 2011; Quivoron et al., 2011). Gu et al. (2011) showed that homozygous mutation of *Tet3* led to neonatal lethality, highlighting its importance in the development of multiple organs during embryogenesis. Thus, Tet1/Tet2 and 5-hmC plays an important but redundant role in the regulation of postimplantation embryonic development.

(iv) Mouse postnatal development

The role of 5-hmC dynamics in postnatal neurodevelopment and brain function and their abnormality leading to neural disorders has been worked out by various workers. 5-hmc is found to be present in almost all organs and tissues of mice but with varying levels (Ito et al., 2010; Li et al., 2011). The genome-wide analysis of brain revealed that 5-hmC levels in the cerebellum and hippocampus of adult mice are higher than those observed in postnatal day (P) 7 mice (Szulwach et al., 2011b). MECP2, a methyl-CpG-binding protein plays an important role in neural development and brain function (Lewis et al., 1992). Amir et al. (1999) reported that humans suffer from Rett syndrome, a neurodevelopmental disorder due to the mutations in *MECP2*. Szulwach et al. (2011b) using *Mecp2* knockout and knockin mice found that the overall abundance of 5-hmC in the cerebellum was negatively correlated with Mecp2 dosage. Recent evidence suggests that Tet2 and 5-hmC are also associated with postnatal development of the hematopoietic lineage, and alteration of *Tet2* expression or the levels of 5-hmC may disrupt the homeostasis of cells and ultimately lead to abnormalities in hematopoietic cell proliferation and maturation (Albano et al., 2011; Chou et al., 2011; Weissmann et al., 2012). In TET knockout phenotype (*Tet2–/–)* of mice an abnormality in myeloid lineage differentiation and spontaneous leukemia but with a overall normal growth appeared (Ko et al., 2011; Li et al., 2011; Moran-Crusio et al., 2011; Quivoron et al., 2011). Gu et al. (2011) showed that although the *Tet3–/–* phenotype was neonatal lethal, the exact functions of Tet3 protein in postnatal development are largely unknown and need further investigation.

(v) Human cerebellum development

5-hmC is abundant in human ESCs (hESCs) and brain regions, and the genomic features of 5-hmC in hESCs has been well characterized by Szulwach et al. (2011a) and Stroud et al. (2011). Similar to hESCs, in cerebellum, 5-hmC displayed a bimodal distribution around transcription start sites (TSS) and was enriched in gene body regions. In comparison with hESCs, less 5-hmC was found at CpG islands and CpG shores, suggesting that CpG islands and shores may have a specific protective mechanism against 5-hmC modification in cerebellum. Wang et al. (2012) have reported a positive correlation between 5-hmC levels and cerebellum development in humans. Genome-wide profiling reveals that 5-hmC is highly enriched on specific gene regions including exons and particularly the untranslated regions (5'UTRs), but it is depleted on introns and intergenic regions in both the fetal and adult cerebellum samples. Furthermore, they have also identified fetus-specific and adult-specific differentially hydroxymethylated regions (DhMRs), most of which overlap with genes and CpG island shores. Surprisingly, during development, DhMRs are highly enriched in genes encoding mRNAs that can be regulated by fragile X mental retardation protein (FMRP), some of which are disrupted in autism, as well as in many known autism genes.

(c) Human diseases

The crucial role of 5-hmc for normal neuronal function being evident with the mutation of MeCP2 protein, causing Rett syndrome, an autism disorder characterized by severe deterioration of neuronal function after birth (Szulwach et al., 2011b). The role of Tet proteins and 5-hmC has also been studied in the context of haematopoiesis and

EXCLI Journal 2014;13:592-610 – ISSN 1611-2156
Received: March 03, 2014, accepted: May 08, 2014, published: May 27, 2014

cancer. Aberrant DNA methylation is a hallmark of cancer, and cancer cells often display global hypomethylation and promoter hypermethylation (Esteller, 2008). The Tet1 gene was originally identified through its translocation in acute myeloid leukemia (AML). Later, somatic Tet2 mutations were tified in patients with a variety of myeloid malignancies, including myelodysplastic syndromes (MDS), chronic myelomonocytic leukemia (CMML), acute myeloid leukemias and many others (Ko et al., 2010). Figueroa et al. (2010) demonstrated that Tet2 mutations in AML patients are predominantly associated with a DNA hypermethylation phenotype.

Among all the members of TET family, TET2 is the most studied genes in MDS and other types of leukemia and whose mutations or dysregulation might lead to lethality. In case of myeloid malignancies, it has been reported as one of the most frequently mutated genes (Albano et al., 2011; Chou et al., 2011; Weissmann et al., 2012). Some of the mutation sites in *TET2* affects the residues within the catalytic domain and could thus damage catalytic activity, whereas, many mutations appear unrelated to enzymatic activity (Ko et al., 2010). Recent studies on isocitrate dehydrogenase (IDH) gene which catalyzes the conversion of isocitrate to 2-oxoglutarate (2OG) also suggests the involvement of 5-hmC in cancer. The mutations in the isocitrate dehydrogenase gene *IDH1/2* in patients suffering from AML were found to stimulate the accumulation of 2-hydroxyglutarate (2-HG) in cells (Klose et al., 2006). This metabolite impairs catalytic activity of many 2-oxoglutarate-dependent enzymes, including Tet proteins, by competing with its co-substrate 2OG. Thus, these results suggest that the deficiencies in both *IDH* and *Tet* genes contribute to cancer development *via* a common disease mechanism that leads to altered 5mC and 5-hmC patterns. However, the role of TET1 in hematopoietic development and leukemogenesis still remains uncertain. The levels of 5-hmC have been found to be decreased in a broad range of solid tumors, including glioma, colon cancer, breast cancer and melanoma (Li and Liu, 2011; Jin et al., 2011a; Haffner et al., 2011).

(d) Transcriptional regulation

The abundance of Tet1 and 5-hmc at the promoter sites suggests a role for the Tet-mediated hydroxymethylation in transcriptional regulation (Wu and Zhang, 2011b). Tet1 and 5-hmc are enriched at 5' gene regulatory regions of various pluripotency factors, such as *Nanog*, *Esrrb* and *Tcl1* thereby promoting the transcription of the gene. Therefore, the depletion of *Tet1/2* has been reported to decrease the expression of a group of genes, including these pluripotency-related factors (Ficz et al., 2011; Ito et al., 2010; Wu et al., 2011). In undifferentiated mouse ES cells, Tet1, probably in conjunction with Tet2, are required for promoting transcription of a cohort of pluripotency factors by maintaining a hypomethylated state at their promoters. Williams et al. (2011) reported that the colocalization of TET1 the SIN3A co-repressor brings about the gene repression.

METHODS FOR *IN VITRO* PRODUCTION AND ANALYSIS OF 5-HMC IN DNA

The chemical synthesis of hmC containing oligonucleotides was first worked out by Planechaud et al. (1997), and is currently being used by the majority of commercial supplies. DNA polymerase-dependent processes (including PCR) utilising commercial 2′-deoxynucleoside-5′-triphosphate (dhmCTP) are being used for enzymatic incorporation of 5-hmC into DNA. Recently, chemoenzymatic reaction of DNA cytosine-5 methyltransferases has been developed for *in vitro* sequence specific incorporation of 5-hmC into DNA duplexes (Liutkevičiūtė et al., 2009).

Following the discovery of 5-hmc, all the existing DNA methylation investigation methods needed to be re-evaluated as they were designed to distinguish only the two

epigenetic states of cytosine: methylated versus unmodified. Therefore, there is a need to develop methods having the ability to discriminate between 5-hmC and 5mC in DNA. Previously, bisulfite sequencing (Huang et al., 2010) were widely used for DNA methylation studies, in which the unmodified cytosine readily forms a 5,6-dihydro-6-sulfonyl adduct in the presence of bisulfite, which undergoes hydrolytic deamination to uracil, and thus appears as T in DNA sequencing. 5mC is very stable to bisulfite-promoted deamination and is subsequently read as normal C. Bisulfite reacts predominantly with the 5-hydroxymethyl group yielding a hydrolytically stable 5-sulfonylmethylcytosine (smC, also called cytosine 5-methylenesulfonate), and thus hmC appears as C (Huang et al., 2010). In contrast, the products of hmC oxidation, caC and fC, are interpreted as unmodified C (Booth et al., 2012) owing to their transient bisulfite-induced conversion to C. Thus, it is evident that bisulfite conversion cannot discriminate between 5mC and hmC (Huang et al., 2010; Jin et al., 2010). Recently, two methods have been proposed that permit detection of hmC in bisulfite sequencing after chemical or enzymatic modifications of hmC (Booth et al., 2012; Yu et al., 2012).

Methyl-sensitive CG-specific restriction endonucleases that differentially cleave unmethylated and methylated target sites are also an important tool for analysis of DNA methylation. Most of the restriction enzymes do not clearly discriminate between 5mC and 5-hmC (Nestor et al., 2010) although some hmC-specific restriction endonucleases have recently been described (Wang et al., 2011). Radiolabeling strategies have also been used for analysis of hmC in DNA. In one such approach named ‘‘nearest neighbour analysis’’, single stranded DNA breaks are randomly introduced by DNaseI and then DNA polymerase and [α-^{33}P]- dGTP is used to incorporate the radioactively labeled G nucleotide at the DNA breaks in the case when C is in the opposite strand. Then DNA is enzymatically hydrolysed to yield 3′-nucleotides such that the radiolabeled phosphate group appears on a 5′-neighbouring nucleotide (Kriaučionis and Heintz, 2009). Alternatively, another method with higher sensitivity is also being used in which the modified cytosine in CG dinucleotides can be enriched by using R.*Msp*I or R.*Taq*I restriction endonucleases, that cleaves CCGG or TCGA sites, respectively, between the two pyrimidines, regardless of whether C, 5mC or hmC is present in the second position. The produced DNA fragments with 5′-terminal CG sites are ^{33}P-labeled and DNA is degraded to 5′-NMPs for thin layer chromatography (TLC) (Tahiliani et al., 2009).

HPLC can also be used for the analysis of 5-hmc. hmC 2′-deoxynucleoside elutes closely following, and is partially disguised by, a major peak of 2′-deoxycytidine. Similar elution times on reverse-phase columns and similar UV spectra of C and hmC may complicate detection of hmC in routine analyses of mammalian DNA using UV detection (Ito et al., 2011). But, this problem can be overcome by the use of synthetic stable-isotope labeled internal standards in mass spectrometry detectors for reliable quantification of nucleosides (Münzel et al., 2010b).

Genome-wide localization of hmC can also be detected by affinity enrichment based methods (Figure 4). Such methods rely on selective binding of short hmC containing DNA fragments (200-800 bp) to hmC-specific antibodies or other hmC-binding proteins permitting their physical extraction from the rest of DNA for analysis using quantitative PCR, DNA microarrays or sequencing. Polyclonal and monoclonal antibodies have been raised against hmC itself or the product of bisulfite treatment of hmC, 5-sulfonylmethylcytosine (smC) (Ficz et al., 2011; Jin et al., 2011b). During the last few years, methylated-DNA immunoprecipitation (MeDIP), hMeDIP (hydroxymethylated-DNA immunoprecipitation) are being widely used for 5-hmC profiling studies.

EXCLI Journal 2014;13:592-610 – ISSN 1611-2156
Received: March 03, 2014, accepted: May 08, 2014, published: May 27, 2014

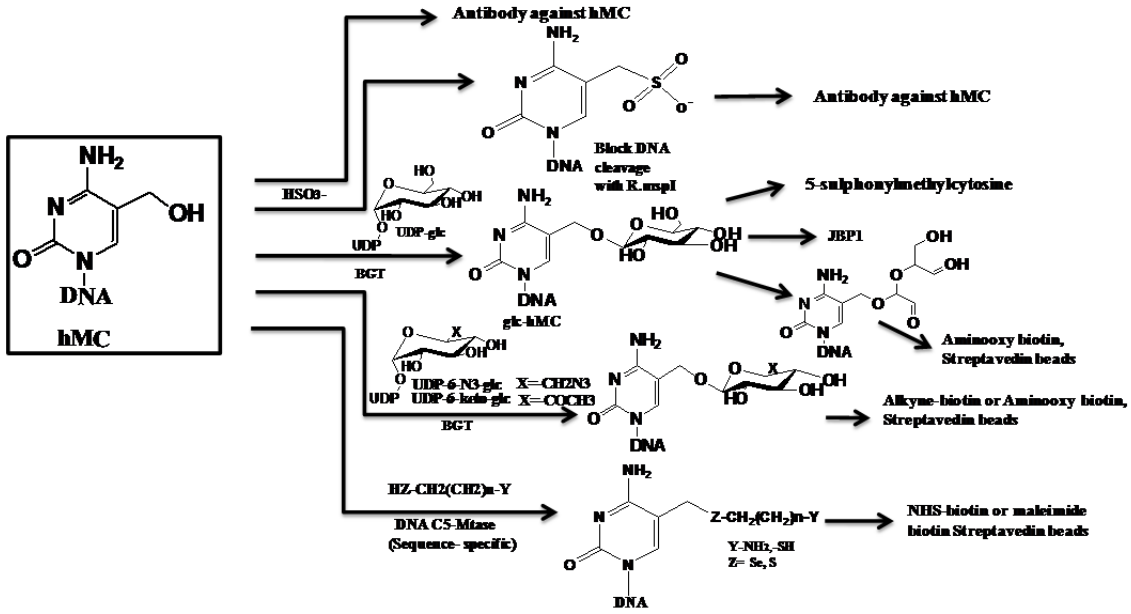

Figure 4: Analytical strategies for labelling and enrichment of hydroxymethylated DNA (modified from Kriukiene et al., 2012)

The enzymatic glucosylation of 5-hmC, using T4 beta-glucosyltransferase (BGT) and the uridine-5'-diphospho-D-glucose (UDP-glc) cofactor, to a much bulkier and distinctive residue, 5-glucosyloxymethylcytosine is highly specific and selective for the 5-hydroxymethyl group of hmC (or hmU). Such treatment can be used to attach tritium labeled glucose moieties from UDP-[^{3}H] glucose to DNA, permitting direct quantification of hmC by scintillation counting (Szwagierczak et al., 2010). JBP1 protein has been reported to selectively bind DNA fragments containing glucosylated hmC which can be used for selective isolation and analysis of hmC containing DNA (Robertson et al., 2011).

Single-nucleotide resolution mapping of hmC residues in the genome can also be done by combining glucosylation of hmC and MspI restriction endonuclease digestion. R.*Msp*I cleaves the CCGG target site if the second cytosine is unmethylated, methylated or hydroxymethylated, but glucosylation of hmC residues renders the sites resistant to *Msp*I cleavage. Thus, glc-hmC DNA can be enriched and analysed using qPCR (Davis and Vaisvila, 2011), microarrays or next-generation sequencing.

The enrichment of 5-hmc in DNA is associated with chemical capture of glucosylated hmC residues (Pastor et al., 2011). Oxidation of the glucose moiety with $NaIO_4$ creates two reactive aldehyde groups which can be subsequently modified by commercially available aldehyde reactive probes containing biotin (Kriaučionis and Heintz, 2009). Nanopore sequencing has also been shown to directly discriminate 5-hmC from 5mC without any further chemical modifications of the base (Wanunu et al., 2010). Zhao and Xue (2012) developed a highly sensitive fluorescence polarization method for genome-wide quantification of DNA methylation. In addition, to estimation of methylation level of submicrograms of lambda or human DNA, and of a 255-bp DNA segment containing a single *Hpa*II /*Msp*I restriction site, fluorescence polarization (FP) based measurement of DNA methylation (FPDM) was also applied to measure dose-dependent DNA hypomethylation in human embryonic kidney 293T cells treated with the DNA-methyltransferase inhibitor 5-aza-dC. Song et al. (2012) used single-molecule real-time

(SMRT) DNA sequencing for detection of 5-hydroxymethylcytosine (5-hmC) in genomic DNA.

CONCLUSION

5-Hydroxymethylcytosine (5-hmC) a newly discovered modified form of cytosine has been assumed to be an important epigenetic modification having important role in development ageing and diseases. 5-hmC is generated through oxidation of 5mC by the Ten-Eleven Translocation (TET) family of methylcytosine dioxygenases and is considered to be a possible intermediate step in a replication-independent DNA demethylation pathway. The role of Tet proteins and 5hmC in development and disease has been widely studied during the past few years. Moreover, further investigations are needed to distinguish the catalytic activity-dependent and independent functions of Tet proteins. Most studies of Tet proteins and 5hmC have been done on mice, rats and human models, therefore, investigations are needed to be expanded to other model systems (such as *Xenopus*, zebrafish and even non-vertebrates) in order to elucidate the evolutionally conserved functions of these proteins.

ACKNOWLEDGEMENTS

Richa and Rajneesh are thankful to the DST (Project No. SR/WOS-A/LS-140/2011) and DBT, Govt. of India, respectively, for the financial assistance in the form of fellowship.

REFERENCES

Aapola U, Kawasaki K, Scott HS, Ollila J, Vihinen M, Heino M et al. Isolation and initial characterization of a novel zinc finger gene, DNMT3L, on 21q22.3, related to the cytosine-5-methyltransferase 3 gene family. Genomics 2000;65:293-8.

Albano F, Anelli L, Zagaria A, Coccaro N, Minervini A, Rossi AR et al. Decreased TET2 gene expression during chronic myeloid leukemia progression. Leuk Res 2011;35:e220-e222.

Amir RE, Van den Veyver IB, Wan M, Tran CQ, Francke U, Zoghbi HY. Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. Nat Genet 1999;23:185-8.

Bernstein BE, Meissner A, Lander ES. The mammalian epigenome. Cell 2007;128:669-81.

Bestor TH. Activation of mammalian DNA methyltransferase by cleavage of a Zn binding regulatory domain. EMBO J 1992;11:2611-7.

Bestor T, Laudano A, Mattaliano R, Ingram V. Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases. J Mol Biol 1988; 203:971-83.

Bhutani N, Brady JJ, Damian M, Sacco A, Corbel SY, Blau HM. Reprogramming towards pluripotency requires AID-dependent DNA demethylation. Nature 2010;463:1042-7.

Bhutani N, Burns DM, Blau HM. DNA demethylation dynamics. Cell 2011;146:866-72.

Bird A. DNA methylation patterns and epigenetic memory. Genes Dev 2002;16:6-21.

Booth MJ, Branco MR, Ficz G, Oxley D, Krueger F, Reik W et al. Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution. Science 2012;336:934-7.

Borst P, Sabatini R. Base J: Discovery, biosynthesis and possible functions. Annu Rev Microbiol 2008;62: 235-51.

Bostick M, Kim JK, Esteve PO, Clark A, Pradhan S, Jacobsen SE. UHRF1 plays a role in maintaining DNA methylation in mammalian cells. Science 2007; 317:1760-4.

Bourc'his D, Bestor TH. Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. Nature 2004;431:96-9.

Bourc'his D, Xu GL, Lin CS, Bollman B, Bestor TH. Dnmt3L and the establishment of maternal genomic imprints. Science 2001;294:2536-9.

Branco MR, Ficz G, Reik W. Uncovering the role of 5-hydroxymethylcytosine in the epigenome. Nat Rev Genet 2012;13:7-13.

Chou WC, Chou SC, Liu CY, Chen CY, Hou HA, Kuo YY et al. TET2 mutation is an unfavourable prognostic factor in acute myeloid leukemia patients with intermediate-risk cytogenetics. Blood 2011;118: 3803-10.

Cierpicki T, Risner LE, Grembecka J, Lukasik SM, Popovic R, Omonkowska M et al. Structure of the MLL CXXC domain-DNA complex and its functional role in MLL-AF9 leukemia. Nat Struct Mol Biol 2010;17:62-8.

Cimmino L, Abdel-Wahab O, Levine RL, Aifantis I. TET family proteins and their role in stem cell differentiation and transformation. Cell Stem Cell 2011;9: 193-204.

Davis T, Vaisvila R. High sensitivity 5-hydroxymethylcytosine detection in Balb/C brain Tissue. J Visualized Exp 2011;48:e2661.

Dawlaty MM, Ganz K, Powell BE, Hu YC, Markoulaki S, Cheng AW et al. Tet1 is dispensable for maintaining pluripotency and its loss is compatible with embryonic and postnatal development. Cell Stem Cell 2011;9:166-75.

Esteller M. Epigenetics in cancer. N Engl J Med 2008;358:1148-59.

Feng S, Cokus SJ, Zhang X, Chen PY, Bostick M, Goll MG et al. Conservation and divergence of methylation patterning in plants and animals. Proc Natl Acad Sci U S A 2010;107:8689-94.

Ficz 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.

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.

Frauer C, Hoffmann T, Bultmann S, Casa V, Cardoso MC, Antes I et al. Recognition of 5-hydroxymethylcytosine by the Uhrf1 SRA domain. PLoS One 2011; 6:e21306.

Fujimori A, Matsuda Y, Takemoto Y, Hashimoto Y, Kubo E, Araki R et al. Cloning and mapping of Np95 gene which encodes a novel nuclear protein associated with cell proliferation. Mamm. Genome 1998;9: 1032-5.

Globisch D, Münzel M, Müller M, Michalakis S, Wagner M, Koch S et al. Tissue distribution of 5-hydroxymethylcytosine and search for active demethylation intermediates. PLoS One 2010;5:e15367.

Goll MG, Kirpekar F, Maggert KA, Yoder JA, Hsieh CL, Zhang X et al. Methylation of tRNAA sp by the DNA methyltransferase homolog Dnmt2. Science 2006;311:395-8.

Gu TP, Guo F, Yang H, Wu HP, Xu GF, Liu W et al. The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. Nature 2011;477:606-10.

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.

Haffner MC, Chaux A, Meeker AK, Esopi DM, Gerber J, Pellakuru LG et al. Global 5-hydroxymethylcytosine content is significantly reduced in tissue stem/progenitor cell compartments and in human cancers. Oncotarget 2011;2:627-37.

Hahn MA, Qiu R, Wu X, Li AX, Zhang H, Wang J et al. Dynamics of 5-hydroxymethylcytosine and chromatin marks in mammalian neurogenesis. Cell Rep 2013;3:291-300.

Hajkova P, Ancelin K, Waldmann T, Lacoste N, Lange UC, Cesari F et al. Chromatin dynamics during epigenetic reprogramming in the mouse germ line. Nature 2008;452:877-81.

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.

Hashimoto H, Liu Y, Upadhyay AK, ChangY, Howerton SB, Vertino PM et al. Recognition and potential mechanisms for replication and erasure of cytosine hydroxymethylation. Nucleic Acids Res 2012;40: 4841-9.

Hata K, Okano M, Lei H, Li E. Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. Development 2002;129:1983-93.

He G, Elling AA, Deng XW. The epigenome and plant development. Annu Rev Plant Biol 2011a;62: 411-35.

He YF, Li BZ, Li Z, Liu P, Wang Y, Tang Q et al. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. Science 2011b;333:1303-7.

Hirano K, Young SG, Farese RV Jr, Ng J, Sande E, Warburton C et al. Targeted disruption of the mouse apobec-1 gene abolishes apolipoprotein B mRNA editing and eliminates apolipoprotein B48. J Biol Chem 1996;271:9887-90.

Huang Y, Pastor WA, Shen Y, Tahiliani M, Liu DR, Rao A. The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing. PLoS One 2010;5:e8888.

Inoue A, Shen L, Dai Q, He C, Zhang Y. Generation and replication-dependent dilution of 5fC and 5caC during mouse preimplantation development. Cell Res 2011;21:1670-6.

Inoue A, Zhang Y. Replication-dependent loss of 5-hydroxymethylcytosine in mouse preimplantation embryos. Science 2011;334:194.

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 U S A 2011;108:3642-7.

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.

Ito S, Shen L, Dai Q, Wu SC, Collins LB, Swenberg JA et al. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. Science 2011;333:1300-3.

Iyer LM, Tahiliani M, Rao A, Aravind L. Prediction of novel families of enzymes involved in oxidative and other complex modifications of bases in nucleic acids. Cell Cycle 2009;8:1698-710.

Jin SG, Kadam S, Pfeifer GP. Examination of the specificity of DNA methylation profiling techniques towards 5-methylcytosine and 5-hydroxymethylcytosine. Nucleic Acids Res 2010;38:e125.

Jin SG, Jiang Y, Qiu R, Rauch TA, Wang Y, Schackert G et al. 5-Hydroxymethylcytosine is strongly depleted in human cancers but its levels do not correlate with IDH1 mutations. Cancer Res 2011a;71:7360-5.

Jin SG, Wu X, Li AX, Pfeifer GP. Genomic mapping of 5-hydroxymethylcytosine in the human brain. Nucleic Acids Res 2011b;39:5015-24.

Kaas GA, Zhong C, Eason DE, Ross DL, Vachhani RV, Ming GL et al. TET1 Controls CNS 5-methylcytosine hydroxylation, active DNA demethylation, gene transcription, and memory formation. Neuron 2013;79:1086-93.

Kaneda M, Okano M, Hata K, Sado T, Tsujimoto N, Li E et al. Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. Nature 2004;429:900-3.

Kinney SM, Chin HG, Vaisvila R, Bitinaite J, Zheng Y, Estève PO et al. Tissue specific distribution and dynamic changes of 5-Hydroxymethylcytosine in mammalian genomes. J Biol Chem 2011;286:4685-93.

Klose RJ, Kallin EM, Zhang Y. JmjC-domain-containing proteins and histone demethylation. Nat Rev Genet 2006;7:715-27.

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.

Ko M, Bandukwala HS, An J, Lamperti ED, Thompson EC, Hastie R et al. Ten-Eleven-Translocation 2 (TET2) negatively regulates homeostasis and differentiation of hematopoietic stem cells in mice. Proc Natl Acad Sci U S A 2011;108:14566-71.

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.

Kota SK, Feil R. Epigenetic transitions in germ cell development and meiosis. Dev Cell 2011;19:675-86.

Kriaučionis S, Heintz N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. Science 2009;324:929-30.

Kriukiene E., Liutkeviciute Z, Klimasauskas S. 5-Hydroxymethylcytosine - the elusive epigenetic mark in mammalian DNA. Chem Soc Rev 2012;41:6916-30.

Langemeijer SM, Kuiper RP, Berends M, Knops R, Aslanyan MG, Massop M et al. Acquired mutations in TET2 are common in myelodysplastic syndromes. Nat Genet 2009;41:838-42.

Law JA, Jacobsen SE. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. Nat Rev Genet 2010;11:204-20.

Lei H, Oh SP, Okano M, Juttermann R, Goss KA, Jaenisch R et al. De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. Development 1996;122:3195-205.

Lewis JD, Meehan RR, Henzel WJ, Maurer-Fogy I, Jeppesen P, Klein F et al. Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. Cell 1992;69:905-14.

Li E, Bestor TH, Jaenisch R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell 1992;69:915-26.

Li W, Liu M. Distribution of 5-hydroxymethylcytosine in different human tissues. J Nucleic Acids 2011; 2011:870726.

Li Z, Cai X, Cai CL, Wang J, Zhang W, Petersen BE et al. Deletion of Tet2 in mice leads to dysregulated hematopoietic stem cells and subsequent development of myeloid malignancies. Blood 2011;118:4509-18.

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.

Liutkeviciute Z, Lukinavicius G, Masevicius V, Daujotyte D, Klimasauskas S. Cytosine-5-methyltransferases add aldehydes to DNA. Nat Chem Biol 2009; 5:400-2.

Mellén M, Ayata P, Dewell S, Kriaucionis S, Heintz N. MeCP2 binds to 5-hmc enriched within active genes and accessible chromatin in the nervous system. Cell 2012;151:1417-30.

Michaeli Y, Shahal T, Torchinsky D, Grunwald A, Hoch R, Ebenstein Y. Optical detection of epigenetic marks: sensitive quantification and direct imaging of individual hydroxymethylcytosine bases. Chem Commun (Camb) 2013;49:8599-601.

Mohn F, Schubeler D. Genetics and epigenetics: stability and plasticity during cellular differentiation. Trends Genet 2009;25:129-36.

Moran-Crusio K, Reavie L, Shih A, Abdel-Wahab O, Ndiaye-Lobry D, Lobry C et al. Tet2 loss leads to increased hematopoietic stem cell self-renewal and myeloid transformation. Cancer Cell 2011;20:11-24.

Morgan HD, Dean W, Coker HA, Reik W, Petersen-Mahrt SK. Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: implications for epigenetic reprogramming. J Biol Chem 2004;279:52353-60.

Morrison JR, Paszty C, Stevens ME, Hughes SD, Forte T, Scott J, Rubin EM. Apolipoprotein B RNA editing enzyme-deficient mice are viable despite alterations in lipoprotein metabolism. Proc Natl Acad Sci U S A 1996;93:7154-9.

Münzel M, Lercher L, Müller M, Carell T. Chemical discrimination between dC and (5Me)dC via their hydroxylamine adducts. Nucleic Acids Res 2010a;38: e192.

Münzel M, Globisch D, Bruckl T, Wagner M, Welzmiller V, Michalakis S et al. Quantification of the sixth DNA base hydroxymethylcytosine in the brain.Angew Chem Int Ed 2010b;49:5375-77.

Münzel M, Lischke U, Stathis D, Pfaffeneder T, Gnerlich FA, Deiml CA et al. Improved synthesis and mutagenicity of oligonucleotides containing 5-hydroxymethylcytosine, 5-formylcytosine and 5-carboxylcytosine. Chem Eur J 2011;17:13782-8.

Muramatsu M, Kinoshita K, Fagarasan S, Yamada S, Shinkai Y, Honjo T. Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. Cell 2000;102:553-63.

Nakamura T, Arai Y, Umehara H, Masuhara M, Kimura T, Taniguchi H et al. PGC7/Stella protects against DNA demethylation in early embryogenesis. Nat Cell Biol 2007;9:64-71.

Navaratnam N, Morrison JR, Bhattacharya S, Patel D, Funahashi T, Giannoni F et al. The p27 catalytic subunit of the apolipoprotein B mRNA editing enzyme is a cytidine deaminase. J Biol Chem 1993;268:20709-12.

Nestor C, Ruzov A, Meehan RR, Dunican DS. Enzymatic approaches and bisulfite sequencing cannot distinguish between 5-methylcytosine and 5-hydroxymethylcytosine in DNA. Biotechniques 2010;48:317-9.

Oda M, Glass JL, Thompson RF, Mo Y, Olivier EN, Figueroa ME et al. High-resolution genome-wide cytosine methylation profiling with simultaneous copy number analysis and optimization for limited cell numbers. Nucleic Acids Res 2009;37:3829-39.

Okano M, Xie S, Li E. Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. Nat Genet 1998a;19:219-20.

Okano M, Xie S, Li E. Dnmt2 is not required for de novo and maintenance methylation of viral DNA in embryonic stem cells. Nucleic Acids Res 1998b;26: 2536-40.

Okano M, Bell DW, Haber DA, Li E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell 1999;99:247-57.

Ooi SK, Bestor TH. The colorful history of active DNA demethylation. Cell 2008;133:1145-8.

Ooi SK, Qiu C, Bernstein E, Li K, Jia D, Yang Z et al. DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. Nature 2007;448: 714-7.

Oswald J, Engemann S, Lane N, Mayer W, Olek A, Fundele R et al. Active demethylation of the paternal genome in the mouse zygote. Curr Biol 2000;10:475-8.

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.

Pfaffeneder T, Hackner B, Truss M, Münzel M, Müller M, Deiml CA et al. The discovery of 5-formylcytosine in embryonic stem cell DNA. Angew Chem Int Ed 2011;50:7008-12.

Planechaud ST, Fujimoto J, Lin SS, Sowers LC. Solid phase synthesis and restriction endonuclease cleavage of oligodeoxynucleotides containing 5-(hydroxylmethyl)-cytosine. Nucleic Acids Res 1997;25:553-9.

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.

Quivoron C, Couronne L, Della Valle V, Lopez CK, Plo I, Wagner-Ballon O et al. TET2 inactivation results in pleiotropic hematopoietic abnormalities in mouse and is a recurrent event during human lymphomagenesis. Cancer Cell 2011;20:25-38.

Ramsahoye BH, Biniszkiewicz 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 U S A 2000;97:5237-42.

Rastogi RP, Richa, Kumar A, Tyagi MB, Sinha RP. Molecular mechanisms of ultraviolet radiation-induced DNA damage and repair. J Nucleic Acids 2010;2010:592980.

Renciuk D, Blacque O, Vorlickova M, Spingler B. Crystal structures of B-DNA dodecamer containing the epigenetic modifications 5-hydroxymethylcytosine or 5-methylcytosine. Nucleic Acids Res 2013;41:9891-900.

Robertson AB, Dahl JA, Vagbo CB, Tripathi P, Krokan HE, Klungland A. A novel method for the efficient and selective identification of 5-hydroxymethylcytosine in genomic DNA. Nucleic Acids Res 2011;39:e55.

Ruzov A, Tsenkina Y, Serio A, Dudnakova T, Fletcher J, Bai Y et al. Lineage specific distribution of high levels of genomic 5-hydroxymethylcytosine in mammalian development. Cell Res 2011;21:1332-42.

Saitou M, Kagiwada S, Kurimoto K. Epigenetic reprogramming in mouse pre-implantation development and primordial germ cells. Development 2012;139: 15-31.

Salvaing J, Lavin TA, Boulesteix C, Lehmann G, Debey P, Beaujean N. 5-methylcytosine and 5-hydroxymethylcytosine spatiotemporal profiles in the mouse zygote. PloS One 2012;7(5):e38156.

Sharif J, Muto M, Takebayashi S, Suetake I, Iwamatsu A, Endo TA et al. The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. Nature 2007;450:908-12.

Shide K, Kameda T, Shimoda H, Yamaji T, Abe H, Kamiunten A et al. TET2 is essential for survival and hematopoietic stem cell homeostasis. Leukemia 2012; 26:2216-23.

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 2011a;29:68-72.

Song CX, Clark TA, Lu XY, Kislyuk A, Dai Q, Turner SW et al. Sensitive and specific single-molecule sequencing of 5-hydroxymethylcytosine. Nat Methods 2012;9:75-7.

Song J, Rechkoblit O, Bestor TH, Patel DJ. Structure of DNMT1-DNA complex reveals a role for autoinhibition in maintenance DNA methylation. Science 2011b;331:1036-40.

Stroud H, Feng S, Kinney SM, Pradhan S, Jacobsen SE. 5-Hydroxymethylcytosine is associated with enhancers and gene bodies in human embryonic stem cells. Genome Biol 2011;12:R54.

Su DGT, Taylor JSA, Gross ML. A new photoproduct of 5-methylcytosine and adenine characterized by high-performance liquid chromatography and mass spectrometry. Chem Res Toxicol 2010;23:474-9.

Surani MA, Hajkova P. Epigenetic reprogramming of mouse germ cells toward totipotency. Cold Spring Harb Symp Quant Biol 2010;75:211-8.

Suzuki MM, Bird A. DNA methylation landscapes: provocative insights from epigenomics. Nat Rev Genet 2008;9:465-76.

Szulwach KE, Li X, Li Y, Song CX, Han JW, Kim S et al. Integrating 5-hydroxymethylcytosine into the epigenomic landscape of human embryonic stem cells. PLoS Genet 2011a;7:e1002154.

Szulwach KE, Li X, Li Y, Song CX, Wu H, Dai Q et al. 5-hmCmediated epigenetic dynamics during postnatal neurodevelopment and aging. Nat Neurosci 2011b;14:1607-16.

Szwagierczak A, Bultmann S, Schmidt CS, Spada F, Leonhardt H. Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA. Nucleic Acids Res 2010;38:e181.

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.

Tan L, Shi YG. Tet family proteins and 5- hydroxylmethylcytosine in development and disease. Development 2012;139:1895-02.

Teng B, Burant CF, Davidson NO. Molecular cloning of an apolipoprotein B messenger RNA editing protein. Science 1993;260:1816-9.

Tomizawa S, Kobayashi H, Watanabe T, Andrews S, Hata K, Kelsey G et al. Dynamic stage-specific changes in imprinted differentially methylated regions during early mammalian development and prevalence of non-CpG methylation in oocytes. Development 2011; 138:811-20.

Ueda Y, Okano M, Williams C, Chen T, Georgopoulos K, Li E. Roles for Dnmt3b in mammalian development: a mouse model for the ICF syndrome. Development 2006;133:1183-92.

Upadhyay AK, Horton JR, Zhang X, Cheng X. Coordinated methyl-lysine erasure: structural and functional linkage of a Jumonji demethylase domain and a reader domain. Curr Opin Struct Biol 2011;21:750-60.

Valinluck V, Sowers LC. Endogenous cytosine damage products alter the site selectivity of human DNA maintenance methyltransferase DNMT1. Cancer Res 2007;67:946-50.

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.

Wagner JR, Cadet J. Oxidation reactions of cytosine DNA components by hydroxyl radical and one-electron oxidants in aerated aqueous solutions. Acc Chem Res 2010;43:564-71.

Wang H, Guan S, Quimby A, Cohen-Karni D, Pradhan S, Wilson G et al. Comparative characterization of the PvuRts1I family of restriction enzymes and their application in mapping genomic 5-hydroxymethylcytosine. Nucleic Acids Res 2011;39:9294-305.

Wang T, Pan Q, Lin L, Szulwach KE, Song CX, He C et al. Genome-wide DNA hydroxymethylation changes are associated with neurodevelopmental genes in the developing human cerebellum. Hum Mol Genet 2012;21:5500-10.

Wanunu M, Cohen-Karni D, Johnson RR, Fields L, Benner J, Peterman N et al. Discrimination of methylcytosine from hydroxymethylcytosine in DNA molecules. J Am Chem Soc 2010;133:486-92.

Warren RA. Modified bases in bacteriophage DNAs. Annu Rev Microbiol 1980;34:137-58.

Webster KE, O'Bryan MK, Fletcher S, Crewther PE, Aapola U, Craig J et al. Meiotic and epigenetic defects in Dnmt3L-knockout mouse spermatogenesis. Proc Natl Acad Sci U S A 2005;102:4068-73.

Weissmann S, Alpermann T, Grossmann V, Kowarsch A, Nadarajah N, Eder C et al. Landscape of TET2 mutations in acute myeloid leukemia. Leukemia 2012;26:934-42.

Williams K, Christensen J, Pedersen MT, Johansen JV, Cloos PA, Rappsilber J et al. TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. Nature 2011;473:343-48.

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.

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.

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.

Wu H, Zhang Y. Mechanisms and functions of Tet protein-mediated 5-methylcytosine oxidation. Genes Dev 2011a;25:2436-52.

Wu H, Zhang Y. Tet1 and 5-hydroxymethylation- A genome-wide view in mouse embryonic stem cells. Cell Cycle 2011b;10:2428-36.

Wu SC, Zhang Y. Active DNA demethylation: many roads lead to Rome. Nat Rev Mol Cell Biol 2010;11: 607-20.

Wyatt GR, Cohen SS. The bases of the nucleic acids of some bacterial and animal viruses: the occurrence of 5-hydroxymethylcytosine. Biochem J 1953;55:774-82.

Xu C, Bian C, Lam R, Dong A, Min J. The structural basis for selective binding of non-methylated CpG islands by the CFP1 CXXC domain. Nat Commun 2011a;2:227.

Xu YF, Wu FZ, Tan L, Kong LC, Xiong LJ, Deng J et al. Genome-wide regulation of 5-hmc, 5mC, and gene expression by Tet1 hydroxylase in mouse embryonic stem cells. Mol Cell 2011b;42:451-64.

Yildirim O, Li R, Hung JH, Chen PB, Dong X, Ee LS et al. Mbd3/NURD complex regulates expression of 5-hydroxymethylcytosine marked genes in embryonic stem cells. Cell 2011;147:1498-510.

Yu M, Hon GC, Szulwach KE, Song CX, Zhang L, Kim A et al. Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome. Cell 2012; 149:1368-80.

Zemach A, McDaniel IE, Silva P, Zilberman D: Genome-wide evolutionary analysis of eukaryotic DNA methylation. Science 2010;328:916-9.

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.

Zhang L, Lu X, Lu J, Liang H, Dai Q, Xu GL et al. Thymine DNA glycosylase specifically recognizes 5-carboxylcytosinemodified DNA. Nat Chem Biol 2012;8:328-30.

Zhao C, Xue H. A simple method for high-throughput quantification of genome-wide DNA methylation by fluorescence polarization. Epigenetics 2012;7:335-9.