

# Hydroxymethylation as a Novel Environmental Biosensor

T. Dao · R. Y. S. Cheng · M. P. Revelo · W. Mitzner ·  
W. Y. Tang

Published online: 14 January 2014  
© Springer International Publishing AG 2014

**Abstract** Beyond the genome, epigenetics has become a promising approach in understanding the interactions between the gene and the environment. Epigenetic regulation includes DNA methylation, histone modifications, and non-coding RNAs. Among these, DNA methylation, which is the addition of a methyl group to the fifth base of cytosine to produce 5-methylcytosine (5-mC), is most commonly studied. Epigenetic regulation has changed given the discovery of 5-hydroxymethylcytosine (5-hmC), considered the “sixth base,” and the nature of TET proteins to catalyze 5-mC oxidation to 5-hmC. 5-Hydroxymethylation has been proposed to be a stable intermediate between methylation and demethylation and has raised questions about the functions of 5-hmC in gene regulation in cells, tissues, and organs in response to environmental exposure. Herein, we have provided an introduction to the chemistry of 5-hydroxymethylation and the techniques for detection of 5-hydroxymethylation. In addition, we have reviewed current reports describing how 5-hmC responds to environmental factors, leading to the development of disease. Finally, we have discussed the potential use of 5-hmC in the study of disease development. All in all, it is our goal to provide innovative and convincing epigenetic studies for understanding the etiology of environmentally related human disease and translate these epigenetic findings into

lifestyle recommendations and clinical practices to prevent and cure disease.

**Keywords** 5-Hydroxymethylcytosine · 5-Methylcytosine · 5-Formylmethylcytosine · 5-Carboxymethylcytosine · DNA methylation · DNA hydroxymethylation · Demethylation · Base excision repair · Oxidation · Glucosylation · Deamination · Ten-eleven translocation proteins · Environmental health · Epigenetics · Environmental biosensor

## Introduction

Each cell of the body has the same genome, yet very different forms and functions. Carefully orchestrated epigenetic mechanisms play a role in the cell/tissue-specific gene expression regulating cell differentiation, maintenance, and proliferation [1]. There are three distinct epigenetic mechanisms, including histone modifications, non-coding RNAs, and DNA methylation, that work singularly or conjointly to epigenetically regulate the chromatin structure [1, 2]. These mechanisms modify downstream gene transcription and function in order to fine-tune biological mechanisms at the cellular level. There is increasing evidence that 5-hydroxymethylcytosine (5-hmC), considered to be the “sixth base” [3], may also play a role in epigenetic regulation. Given the fact that epigenetic changes are sensitive to environmental stimuli, in this review, we focus on the role of DNA hydroxymethylation in gene regulation and how it interacts with environmental factors, leading to disease development.

DNA methylation and hydroxymethylation are proposed as epigenetic modifications of gene regulation. The chemistry of DNA methylation and hydroxymethylation is summarized in Fig. 1. DNA methyltransferases (DNMTs) catalyze the covalent addition of a methyl group from S-adenosyl-methionine (SAM) to the C5 position of cytosine in CpG dinucleotides

---

T. Dao · W. Mitzner · W. Y. Tang (✉)  
Department of Environmental Health Sciences, Johns Hopkins  
Bloomberg School of Public Health, 615 N Wolfe Street,  
Room E7545, Baltimore, MA, USA  
e-mail: wtang@jhsph.edu

R. Y. S. Cheng  
Radiation Biology Branch, National Cancer Institute, National  
Institutes of Health, Bethesda, MA, USA

M. P. Revelo  
Department of Pathology, University of Utah, Salt Lake City, UT,  
USA

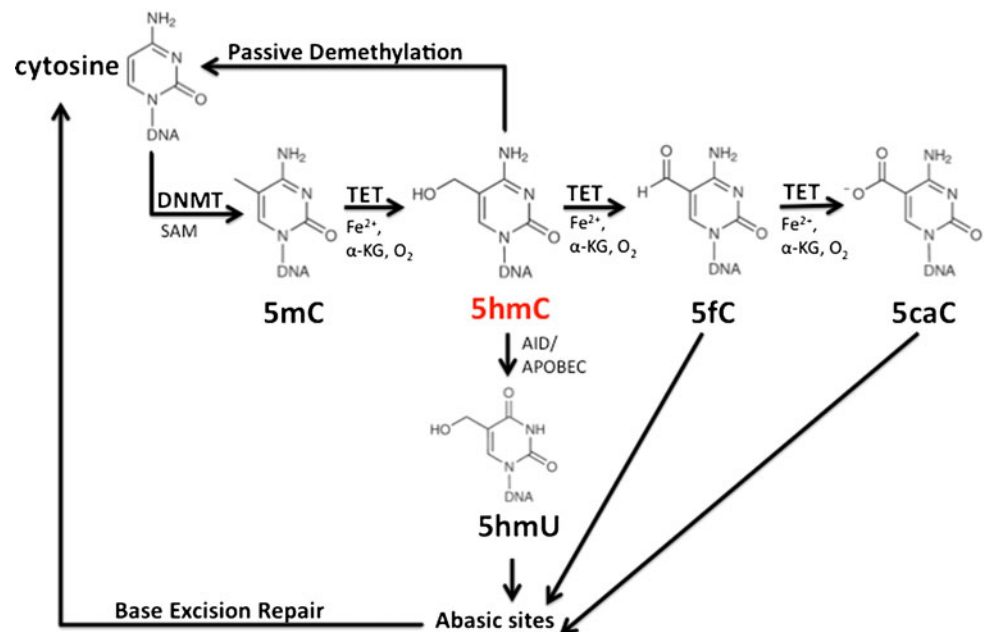
(Fig. 1), which predominantly cluster in densely populated CpG islands (CGI), often located in gene promoters and the first intron and exon [4–7]. Initially, 5-methylcytosine (5-mC) is established by de novo DNMT3A and DNMT3B [8], and during semi-conservative DNA replication, DNMT1 maintains methylation by using the hemi-methylated DNA template to regenerate the symmetrical methylation site on the new DNA strand [5]. 5-mC is found in 1–4 % of all cytosines (or ~80 % of CpG dinucleotides) in the mammalian genome [6, 9, 10]. CGI methylation occurs mainly in the promoter regions, where ~70 % of the CGIs are unmethylated except for those on the inactive X chromosome and some associated with imprinted genes [11]. DNA methylation in critical regulatory regions, such as gene promoters, demonstrates the influential role of 5-mC in gene repression. 5-mC has a high affinity for methyl-CpG-binding proteins (MBPs), and their interaction causes the recruitment of chromatin-modifying enzymes at the promoter, leading to gene repression by sterically hindering the binding of transcription factors and basal transcriptional machinery [5]. Although the mechanisms of DNA methylation are well defined, the opposing role of DNA demethylation has not yet been fully understood. During zygote formation, there is rapid demethylation during zygote cleavage and rapid methylation after implantation, independent of cell division, which indicates that active DNA demethylation pathways are responsible [1, 4]. The discovery of ten-eleven translocation (TET) proteins brought an incredible interest to the 5-hmC-mediated DNA demethylation pathways and the biological role of 5-hmC in development [12]. TET enzymes are responsible for catalyzing the oxidation of 5-mC to 5-hmC and the iterative oxidation of 5-hmC to 5-formylcytosine (5-fC) and 5-carboxycytosine (5-caC) DNA derivatives, believed

to be the intermediates in the DNA demethylation pathways [4, 6, 13] (Fig. 1).

### TET Proteins

TET expression levels vary between cells and organs. TET1 and TET2 are highly expressed in mouse ES cells, while TET3 is seen predominantly in oocytes and one-cell zygotes [4]. TET1 and TET3 are expressed in several adult tissues but most abundant in brain, while TET2 is enriched in hematopoietic cells [3]. There is a limited number of studies about how TET proteins are targeted to specific genes in distinct cell types and developmental stages. Whether TET expression profiles correlate to 5-hmC product formation is also very complicated and controversial. The brain (~0.80 %) and spinal cord (~0.45 %) are the most 5-hmC-enriched cells/organs among different tissues [6, 14]. Studies showed global 5-hmC levels in human tissues are not associated with 5-mC content or TET gene expression [15]. However, additional studies are required to evaluate this finding in other tissue/cell types and in large-scale human population studies. All TET proteins have a catalytic domain that contains both a cysteine-rich region and 2-oxoglutarate-Fe(II) dioxygenase activity [6, 9, 16]. In mammals, TET proteins (TET1, TET2, TET3) catalyze the transfer of a hydroxyl group to 5-mC [9, 13, 16] to form 5-hmC/5-fC/5-caC. TET1 can oxidize both fully and hemi-methylated DNA, and does not require CpG dinucleotides [4]. In the N-terminal, the CXXC domain of the TET proteins (TET1 and TET3 only) is a DNA binding domain, with a high affinity for clustered unmethylated DNA [4, 16]. Subsequent genome-wide mapping of mouse and human embryonic stem cells determined that the CXXC

**Fig. 1** Chemistry of DNA methylation and hydroxymethylation. 5-Methylcytosine (5-mC) is produced from the addition of S-adenosylmethionine (SAM) onto the 5-carbon of cytosine by DNA methyltransferases (DNMT). Ten-eleven translocation (TET) proteins then catalyze the iterative oxidation of 5mC to 5-hydroxymethylcytosine (5-hmC), 5-formylcytosine (5-fC) and 5-carboxycytosine (5-caC), with required cofactors alpha-ketoglutarate (a-KG), iron ( $\text{Fe}^{2+}$ ), and oxygen. 5-hmC, 5-fC, or 5-caC could act as an intermediate in both passive and active DNA demethylation pathways involving DNA repair enzymes like AID or APOBEC (activation-induced cytidine deaminase)



domain of TET1 overlaps with the unmethylated DNA generally located in the CGIs of promoter regions [4], signaling many chromatin-associated proteins. TET1 has been shown to activate gene transcription by reducing the 5-mC mediated recruitment of the polycomb repression complex 2 (PRC2) to chromatin [4]. In contrast, TET proteins have been shown to recruit the Sin3a (directly) and the Ezh2 complex (indirectly), to deacetylate and trimethylate histone H3, and repress target genes [4, 6]. To recap, TET proteins may both repress and activate gene transcription by the recruitment of chromatin remodeling enzymes through the CXXC domain and/or as the intermediate of gene demethylation through the catalytic 5-mC oxidation to 5-hmC/5-fC/5-caC.

### Non-Enzymatic Regulation by 5-hmC

The substitution of 5-mC for 5-hmC can non-catalytically regulate gene transcription by interfering with DNA-protein interactions. Once established, these 5-hmC residues can influence gene regulation. 5-hmC has a large, polar, hydroxymethyl group, which extends into the major groove of DNA, pushing the attached cytosine further away from the duplex and creating a polar cavity that increases solvation dynamics (the interaction of DNA and water molecules) [17]. The solvation effects of 5-hmC causes local geometric changes in intra-base pair fluctuations (shear, stretch, stagger, and buckle), which decreases rigidity around the local helical axis conformation, and increases the ability of the duplex to propeller twist and open [17]. 5-hmC generally functions to activate gene expression by destabilizing the DNA structure to allow transcriptional machinery to access the transcriptional start site [3]. Using synthetic oligonucleotides, the replacement of 5-mC with 5-hmC diminishes the binding affinity of MeCP2 (methyl-CpG binding protein 2), suggesting that 5-hmC residues are functionally equivalent to cytosine residues [5]. The degree of 5-hmC content can reduce methylation binding domain (MBD) interactions to reorganize the chromatin structure and reverse the repressive effects of 5-mC upon transcriptional regulation. 5-hmC modifications generally are transcriptionally activating, but exceptions always arise. In contrast, 5-hmC residues have been demonstrated to interact with MBDs or methyl-CpG-binding proteins (MBPs) to recruit many chromatin remodeling enzymes that maintain the repressive transcriptional regulation. For example, in mouse embryonic stem cells (ESCs), MBD3 (methyl-CpG binding protein 3) directly binds 5-hmC modified DNA (but not 5-mC) and recruits NuRD (nucleosome remodeling and deacetylase) complexes to repress gene transcription [18]. In addition, evidence is emerging that 5-hmC might act as a new landmark that recognizes and recruits specific DNA-binding proteins [6, 16], which can further complicate the role of 5-hmC transcriptional regulation.

### Enzymatic Regulation by 5-hmC

TET proteins catalyze the generation of 5-hmC residues that act as intermediates for DNA demethylation to affect gene transcription. Proposed mechanisms indicate that 5-hmC is a stable intermediate for both “passive” and “active” DNA demethylation [3] (Fig. 1). “Passive” (replication-dependent) DNA demethylation occurs as a result of the inability of DNMT1 to maintain DNA methylation, resulting in the passive loss of DNA methylation at a specific CpG site [3, 5, 6, 15]. Less defined are the mechanisms of “active” (or replication-independent) DNA demethylation, which suggest that 5-hmC is either spontaneously or enzymatically converted to cytosine through DNA repair pathways [9]. Possible mechanisms of “active” DNA demethylation include: the recognition of 5-hmC by activation-induced cytidine deaminase (AID/APOBEC), which mediates deamination to 5-hydroxymethyluracil (5-hmU), generating an abasic site, which is recognized and subsequently removed by DNA glycosylases (SMUG1, single-strand-selective monofunctional uracil-DNA glycosylase 1) and TDG (thymine DNA glycosylase); and base excision repair (BER) machinery, to restore an unmodified cytosine (Fig. 1) [4, 6]. Regardless of the route of TET-mediated demethylation, 5-hmC is an essential intermediate in the removal of 5-mC residues in the DNA demethylation pathway, which can drastically change global methylation patterns and biological mechanisms. Therefore, further investigations are needed to fully understand the DNA demethylation pathway.

Recent studies have demonstrated that TET enzymatic activity is affected by many factors. Isocitrate dehydrogenase enzymes (IDH1 and IDH2) commonly undergo gain-of-function mutations in human gliomas (~75 %) and acute myeloid leukemia (~20 %) patients that catalyze the production of  $\alpha$ -ketoglutarate ( $\alpha$ -KG) and 2-hydroxyglutarate (2-HG) respectively [19, 20]. 2-HG is a proposed oncometabolite that competitively inhibits  $\alpha$ -KG dioxygenases, such as TET [20]. In TET1-overexpressing HEK293 cells, the coexpression of tumor-derived mutant IDH1 and IDH2 substantially decreased the 5-hmC content compared to the wild type [19]. In vitro, 2-HG was shown to inhibit TET activity, and clinically, glioma samples subjected to immunohistochemical (IHC) staining demonstrated that 5-hmC was significantly decreased in IDH1-mutated gliomas compared to the control, despite consistent 5-mC levels [19]. Oxidative stress has also been proposed to affect TET activity. Under high oxidative stress conditions, when homeostasis cannot be maintained, increased  $\text{NAD}^+$  levels in the mitochondria activate sirtuin  $\text{NAD}^+$  dependent deacetylases (Sirt) [20]. In mammals, increased Sirt3 results in the deacetylation of IDH2 and consequent activation of metabolic activity [20]. Therefore, activated IDH2 increases the conversion of isocitrate to  $\alpha$ -KG and increases TET enzymatic activity,

potentially changing global methylation patterns. Research continues to associate epigenetic mechanisms as responses to environmental sensors that can potentially disrupt cellular and biological processes, resulting in disease development.

### Measurement of DNA Hydroxymethylation

Tools that exist for the detection of methylated DNA, such as bisulfite conversion and methylated-DNA-specific antibodies, have failed to distinguish between 5-mC and 5-hmC DNA, so it is critical to develop techniques that can differentiate between these two modifications. Currently, the most accessible technique developed for hydroxymethylated DNA enrichment and detection is the use of enzymatic and antibody approaches. 5-hmC DNA quantitation, by techniques such as polymerase chain reaction (PCR), high-performance liquid chromatography (HPLC), thin layer chromatography (TLC), or liquid chromatography-mass spectrometry (LC-MS) can detect 5-hmC at single CpG dinucleotides. We summarized several common techniques to reveal (1) global hydroxymethylation, (2) loci-specific hydroxymethylation, and (3) whole genome profiling for hydroxymethylated DNA loci (Table 1). All the techniques described could be used on distinct tissues, cell types or biological specimens.

### Quantitation for Global Hydroxymethylation

The content of 5-hmC across the genome can be measured using enzyme-linked immunosorbent assay (ELISA), with capture and detection antibodies followed by colorimetric quantification. The capture antibody for 5-hmC has no or negligible cross-reactivity to both methylated and unmethylated cytosines. Additionally, there are other protocols to examine the 5-hmC content across the genome, using enzymatic digestion coupled with HPLC, TLC, or LC-MS. Input DNA is first glucosylated by 5-hmC glucosyltransferase (GT), which transfers a glucose moiety from uridine diphosphoglucose (UDPG) onto preexisting 5-hmC's within

DNA [21, 22]. The difference in the chemical structure (mass) between glucosylated 5-hmC and unglucosylated 5-mC or cytosine generates a distinct profile for 5-hmC in the chromatography [23–27]. To locate the 5-hmC distribution in the tissues, immunohistochemistry or immunofluorescence staining using an antibody against 5-hmC is employed to examine the distribution of 5-hmC among different cell types.

### Specific Hydroxymethylated DNA/Loci/Gene DNA Quantitation

Current bisulfite sequencing technology cannot distinguish 5-mC and 5-hmC at specific CpG sites; therefore, real-time polymerase chain reaction (RT-PCR), coupled with a glucosyl-5-hmC sensitive restriction endonuclease (GSRE) digestion [28], is used to measure the percentage of 5-hmC of specific CpG sites. First, the hydroxy group of 5-hmC is protected with a glucosylated moiety by glucosyltransferase prior to PCR. Glucosylated 5-hmC is next cut by a GSRE, like TaqIa, and according to the RT-PCR assay, the cleaved glucosylated 5-hmC has a smaller threshold value (Ct) than the unglucosylated cytosine/5-mC. The percentage 5-hmC at specific CpG sites is quantified using the Ct difference between the glucosylated and unglucosylated samples. It is used for detection of PCR products smaller than 150 base pairs. For longer fragments, we suggest the new TET-1 assisted bisulfite sequencing. It involves the UDPG-mediated protection of 5-hmC and recombinant mouse TET1 (mTET1)-mediated oxidation of 5-mC to 5-caC. After the subsequent bisulfite treatment and PCR amplification, both cytosine and 5-caC (derived from 5-mC) are converted to thymine (T), whereas 5-hmC reads as cytosine (C). The treated genomic DNA is suitable for locus-specific sequencing [9, 29, 30].

### Whole Genome DNA Quantitation

Researchers who are interested in quantifying the whole genome profiling hydroxymethylation loci could use the anti-5-

**Table 1** Common methods for 5-hmC detection

Type	Detection level	Detection principle	Pre-modification of DNA <sup>1</sup>	Processing Time <sup>2</sup>	Specialized instrument <sup>3</sup>
Global	Whole sample	ELISA <sup>4</sup>	Required	6–8 h	Not required,
Loci specific	Loci	GSRE, TET1-oxidation, PCR, and sequencing	Required	1–3 days	Not required
Whole genome	Loci	Massive parallel sequencing	Required	Variable	Required <sup>5</sup>

<sup>1</sup> Antibody binding (anti-5-hydroxymethylcytosine) or enzyme modifications (i.e., GT, GSRE, TET1)

<sup>2</sup> Time for completion of ELISA is usually less than one working day and depends on the commercial vendor's manual. For loci 5-hmC detection, it requires 1–3 days for pretreatment of DNA, PCR, and/or sequencing. Whole genome massive parallel sequencing requires library construction, massive sequencing, data analysis, data mining, and gene validation. Time for completion varies between facilities

<sup>3</sup> Instruments other than the thermocycler and plate reader

<sup>4</sup> Techniques using chromatography and mass spectroscopy to detect global hydroxymethylation are available but not commonly used

<sup>5</sup> Massive parallel sequencing on different types of sequencers is available in core facilities or commercial vendors



hmC antibody to immunoprecipitate and enrich the sonicated DNA-protein complexes. The enriched DNA products are subjected to a methylation oligonucleotide array (hMeDIP-chip) or next generation sequencing (hMeDIP-Seq) [31, 32].

### Environmental Exposure Linked to 5-hmC Perturbations

Cytosine modifications, such as 5-mC and 5-hmC, are epigenetic mechanisms that enable genes to respond to external environmental cues. Environmental exposure has been shown to disrupt epigenetic mechanisms and interfere with 5-mC and 5-hmC profiles in many diseases, as summarized in Tables 2 and 3. TET proteins are also apt to environmental perturbations. Therefore, researchers have been investigating 5-hmC content as a novel epigenetic marker to understand the link between epigenetic signatures and disease states.

#### Ascorbate Acid

Vitamin C, or ascorbic acid (AA), stimulates the drastic erasure of 5-mC by promoting TET activity and generating 5-hmC-mediated DNA demethylation [33, 34, 35•]. TET enzymes require both Fe<sup>2+</sup> and  $\alpha$ -KG co-factors [3, 9•], and AA is another co-factor required for full catalytic activity of TET proteins [34]. AA directly interacts with the catalytic domains of TET1 and TET2 to reduce Fe<sup>3+</sup> to Fe<sup>2+</sup> and enhance TET-mediated oxidation [33]. AA not only increases 5-hmC, but also the iterative oxidation products 5-fC and 5-caC in mouse ESCs [33]. The reducing effects of AA are specific, since other antioxidant enzymes (glutathione, selenite, vitamin B1, vitamin E, L-carnitine, lipoic acid) do not induce the enzymatic activity of TET [35•]. Mouse ESCs treated with AA for 12 and 72 h displayed a progressive increase in 5-hmC and decrease in 5-mC at the gene promoters, but it was not influenced by the gene expression of TET or DNMT [35•]. In addition, the global induction of 5-hmC in response to AA treatment was reversible; after 3 days of AA removal, both the 5-hmC and 5-mC profiles returned to the baseline levels. These findings confirmed the dynamic interplay between 5-mC and 5-hmC from exogenous exposures mediated by TET

oxidation. Additionally, these investigations suggest the nutritional value of AA, especially during embryonic development, where drastic global changes in DNA methylation occur at the fetus genome.

#### Phenobarbital

Phenobarbital, a widely used barbiturate medication to control seizures, is classified as a non-genotoxic carcinogen. In the mouse liver, PB perturbs 5-mC and 5-hmC profiles associated with tumor formation [14, 36]. Mice administered PB through drinking water displayed significant differential hydroxymethylation in juvenile (30–33 days of phenobarbital exposure) and mature (91 days of phenobarbital exposure) liver tissue. In mature liver tissue, phenobarbital induced proximal promoter enrichment of 5-hmC, coupled with a decrease in 5-mC, to increase the gene transcription of a collection of candidate genes with a potential role in liver tumorigenesis [36]. Additionally, phenobarbital has provoked the rapid and prolonged gene expression of cytochrome P450 2b10 (*Cyp2b10*), a member of the Cyp family of genes, critical for xenobiotic metabolism in the liver [36]. Therefore, by epigenetically inducing the expression of *Cyp2b10* in the liver, the increased rate of phenobarbital metabolites produced, including carcinogens, could promote liver tumor formation. *Wisp1*, a Wnt signaling pathway gene, was overexpressed and associated with increased 5-hmC levels following phenobarbital exposure [14]. It supports the previous findings in both rodents and humans of *Wisp1* gene expression increasing during the disease progression of hepatocellular carcinoma. These findings exemplify the use of 5-hmC as a sensitive biosensor, which can be used to distinguish control versus phenobarbital-exposed rodents.

#### Diethylstilbestrol

Neonatal exposure to diethylstilbestrol, an endocrine disrupting chemical, resulted in the adult onset of uterine and vaginal cancer in women. Mice postnatally exposed to diethylstilbestrol showed global 5-hmC reduction in uterine tissues of diethylstilbestrol treated mice (8 weeks old) when compared to the control [37]. Furthermore, diethylstilbestrol mediated the reduction of *TET1*, *TET2*, *Dnmt1*, and *Dnmt3a*

**Table 2** Examples of environmental exposures contributed to TET regulation and 5-hydroxymethylation changes

Environmental exposures	Observed changes	References
Ascorbic acid (vitamin C)	Increased TET1 and TET2 activity and increased 5hmC in mouse embryonic stem cells	[33, 34, 35•]
Phenobarbital	Increased 5hmC in mature mouse liver tissue	[36]
Diethylstilbestrol	Decreased 5hmC in mouse adult uterine tissue	[37]
Hydroquinone	Increased TET1 and 5hmC in human embryonic kidney cell culture	[38]
House dust mite	Increased 5hmC in mouse lung tissue	[39]

genes, which remained repressed in adult uterine tissue. The decreased 5-hmC profile observed in adult mice in response to diethylstilbestrol treatment could be due to the impaired TET oxidation of 5-mC and/or declining 5-mC preservation due to DNMT reduction. Results suggest that neonatal exposure to endocrine disruption chemicals may result in aberrant 5-mC/5-hmC-mediated DNA methylation, altering gene expression patterns throughout adulthood and possibly increasing disease risk.

### Hydroquinone

Hydroquinone, a copious metabolite of benzene (a carcinogen found in car exhaust, industrial emissions, and cigarette smoke), caused an increase in TET1 activity and global 5-hmC in human embryonic kidney cell cultures [38]. Hydroquinone exposure increased the global 5-hmC content and decreased 5-mC profiles through a reactive oxygen species (ROS)-associated mechanism, since the hydroquinone effects were annulled upon N-acetyl-cysteine (NAC) rescue. In addition, the targeted siRNA-mediated knockdown of TET1 caused a reversal of 5-hmC levels back to baseline. It suggests that numerous environmental exposures that induce reactive oxygen species/oxidative stress may influence 5-hmC patterns and increase the disease risk.

### Environmental Allergens

Asthma is a significant public health concern, and its prevalence has increased dramatically in past 30 years among both children and adults [49]. Since the prevalence of asthma cannot be explained by traditional genetic inheritance, much consideration has been given to lifestyle and environmental exposures being related to increased disease risk. House dust mite, a common environmental allergen, induces airway hyperresponsiveness and inflammation in mouse models through epigenetic alterations in the lung [39]. Exposure to dust mites has resulted in global increases in both 5-mC and 5-

hmC in the lung compared to saline controls. In addition, gene-specific methylation of *Pde4d*, *Pom121l2*, and *Ncx3* has been linked to dysregulated gene expression and cell functions of the lungs. How 5-mC and 5-hmC mediate gene regulation toward environmental allergens requires further investigation.

## 5-hmC in Disease Development

### Cancer

Epigenetic dysregulation is well described in cancer pathogenesis. Using human lung, brain, kidney, liver, skin, and small intestine tissues, the levels of 5-hmC are remarkably depleted in malignant tissues compared to the benign tissues [40, 41, 50]. One explanation is that tumor formation involves rapid cell proliferation, which could lead to the passive loss of 5-hmC [50]. In support of this theory, proliferating cells stained with Ki67 did not express 5-hmC [3, 50]. However, more research is needed to understand the mechanisms linking the loss of 5-hmC in proliferating cells, which leads to increased cancer risks. If this theory is proved, genome-wide 5-hmC and 5-mC profiles in conjunction with Ki67 staining could be valuable biomarkers for cancer diagnosis and management.

Gliomas exhibit epigenetic changes in response to disease progression. Immunohistochemistry (IHC) detection of 5-hmC in formalin-fixed, paraffin-embedded, human glioblastomas revealed that 5-hmC decreased significantly with the progression of the tumor grade, regardless of fetal or adult origin [40]. Upon stratifying the 5-hmC IHC stains into high- and low-content tissues, 5-hmC reduction became a significant prognostic indicator of decreased life expectancy in fetal and adult glioblastomas. This indicates the critical role of 5-hmC in disease state progression and disease management.

In hepatocellular carcinoma, clinical and animal data indicate that 5-hmC may serve as a prognostic marker for disease

**Table 3** Examples of diseases associated with changes in 5-hydroxymethylation or TET regulation

Diseases	Observed changes	References
<b>Cancer</b>		
Gliomas	Decreased 5hmC	[40]
Hepatocellular carcinoma	Decreased 5hmC in both human and rodent HCC	[41]
Melanomas	Decreased TET2 and 5hmC	[42, 43]
Acute myeloid leukemia	Increased TET1 and 5hmC in bone marrow samples	[44]
Myelodysplastic syndrome	Decreased TET2 and 5hmC in bone marrow samples	[45]
<b>Late-onset diseases</b>		
Huntington's disease	Decreased 5hmC in mouse brain	[46]
Alzheimer's disease	Increased 5hmC in mouse brain of females offspring experiencing prenatal stress	[47]
Psychosis	Increased TET1 and 5hmC in brain cortex tissue of psychotic and depressed patients	[48]

outcomes. IHC detection of 5-hmC on hepatocellular carcinoma samples showed a drastic decrease in 5-hmC content in the tumor tissue compared to the benign normal liver tissue [41]. Subsequently, lower levels of 5-hmC were associated with decreased overall survival in hepatocellular carcinoma. Male rats exposed to diethylnitrosamine, a potent carcinogen for hepatocellular carcinoma, exhibited a gradual dose-dependent decrease in 5-hmC IHC detection up through 24 weeks of treatment and showed remarkably reduced 5-hmC in tumor versus non-tumor tissue [41]. These data further confirm the role of 5-hmC depletion as a biomarker of cancer development.

Melanomas are aggressive cancers that are capable of transforming from benign nevi into lethal metastases in a small surface area. Using tissue sections from patients in different stages of melanoma progression, IHC staining exhibited significant decreases in global 5-hmC levels depending on the degree of disease [42, 43], even though the 5-mC profiles among the samples were constant, regardless of the disease state. IHC staining of TET2 was robust in nevi samples and reduced in advanced melanomas [43]. Therefore, a reduction in TET2 and 5-hmC indicates the epigenetic hallmark representing the disease progression of melanoma. These findings also further implicate the dangers of UV exposure in melanoma disease pathogenesis [42].

Myeloid malignancies are cancers originating in blood stem cells in the bone marrow and include acute myeloid leukemia (AML) and myelodysplastic syndrome (MDS) [51]. In bone marrow samples from mixed lineage leukemia (MLL)-rearranged AML patients (TET1 is a fusion partner of the MLL gene), TET1 is overexpressed [44]. Using small hairpin RNA (shRNA) against TET1 in mouse MLL-rearranged leukemic cells, it led to decreased 5-hmC in a dose-dependent manner and induced apoptosis [44]. In vivo, TET1-knockout mice exhibited similar results to TET1-shRNA models, with TET1 reduction correlating to significant decreases in 5-hmC compared to the controls [44]. Results suggested that TET1 plays a role in the pathogenesis of MLL-rearranged leukemia, which is marked by a global increase in 5-hmC and the transcriptional activation of many targets promoting cell proliferation and inhibiting apoptosis, resulting in cell transformation and AML disease pathogenesis. In MDS, TET2 and 5-hmC expression appears to be necessary for the normal differentiation of hematopoietic stem cells [52, 53]. TET2 mutations in bone marrow mononuclear cells were associated with ~16 % of MDS patients, and the level of 5-hmC was significantly lower in MDS patients with mutant TET2 [45]. Upon risk stratifying the 5-hmC patients, high 5-hmC levels had increased MDS survival rates when compared to those with low 5-hmC levels [45]. This suggested that the depletion of 5-hmC can be used as a diagnostic and prognostic tool for MDS patients.

## Developmental Syndromes

In normal human tissues, average 5-hmC values vary depending on tissue type. Brain tissues have the highest 5-hmC content, up to 10- to 20-fold more than peripheral tissues and ESCs [15, 46]. Many neurodegenerative disorders are associated with aberrant methylation once the epigenetic signature has been impaired. Since neuronal cells in the adult brain do not divide mitotically, high levels of 5-hmC are maintained. High 5-hmC content indicates that active demethylation pathways may be upregulated [46, 54]. Under normal conditions in the central nervous system, abundant MeCP2s have a higher affinity for 5-hmC residues compared to 5-mC, facilitating active gene transcription [54]. In classical Rett syndrome, a neurodevelopmental disorder that affects primarily girls and results in the loss of language and motor skills, the loss-of-function MeCP2 mutation disrupts the neuronal maturation through the epigenetic dysregulation of MeCP2 function [55]. Given the abundant levels of 5-hmC in the central nervous system, decoding the role of 5-hmC in neuronal gene regulation and neurodevelopmental disorders is a growing interest among researchers.

## Late-Onset Diseases

Huntington's disease is a fatal genetic disorder characterized by chorea, and cognitive and psychiatric decline. In a mouse model of Huntington's disease, 5-hmC was significantly reduced in the striatum and cortex of Huntington's disease mouse brains compared to the control, and the IHC stain colocalized with NeuN, a marker of mature neurons [46]. Upon deep sequencing and genomic mapping, 436 striatal and 199 cortical genes were differentially hydroxymethylated [46]. Therefore, 5-hmC reduction in Huntington's disease mouse brains further complements previous findings on aging and the decrease in 5-hmC in neurodegenerative disorders.

Alzheimer's disease is the most common form of dementia among the aging population and results from the aberrant production and cleavage of amyloid- $\beta$  (A $\beta$ ) plaques and tau protein tangles. The etiology of Alzheimer's disease is unknown, since disease onset occurs after decades of exposure at different stages of life. One factor investigated in Alzheimer's disease mouse models is the role of stress during prenatal development, which has been shown to negatively affect cognition and hippocampal functioning, predominantly in females [47]. In female offspring experiencing prenatal stress, compared to the males and un-stressed controls, the dorsal and ventral subregions of the hippocampus showed increased 5-hmC. This suggests that in utero prenatal stress can impact epigenetic mechanisms via 5-hydroxymethylation that increase Alzheimer's disease risk.

Increasing evidence is implicating the role of epigenetic modifications in neuropsychiatric disorders. Using post-

mortem brain cortex tissue from psychotic (bipolar disorder, schizophrenia) and depressed patients for examining the TET1 mRNA and protein expression, it was shown that TET1 expression levels were increased in both the psychotic and depressed patients when compared to the controls [48]. In another study, the global 5-hmC levels were significantly increased in the psychotic patients compared to controls [56]. In addition, decreased 5-hydroxymethylation was correlated to the decreased glutamic acid decarboxylase67 (GAD67) mRNA expression in psychotic subjects [48, 56]. Hence, 5-hmC can be used to investigate the pathogenesis of psychosis.

### Conclusion and Future Perspective

Given the nature of 5-hmC in the epigenetic regulation of cell functions, and its possible link to disease development, there is the possibility of employing global 5-hmC or specific hydroxymethylated genes as environmental biosensors. 5-hmC, the “sixth base,” is as stable as an epigenetic marker as 5-mC in the genome. The discovery of TET proteins, which was a breakthrough in the field of understanding gene regulation by DNA hydroxymethylation, allows us to track the development of 5-hmC via TET proteins in response to environmental exposure. Similar to other epigenetic studies, there are several challenges in performing 5-hmC studies. First, the exact mechanism by which TET proteins or 5-hmC regulates gene transcription is unclear. What are the other epigenetic factors that modulate the TET proteins and process the 5-hmC-mediated transcription signal? Second, the distribution of 5-hmC varies among the cell types and cell tissues. Are there any cell type-specific 5-hmC signatures? Most of the 5-hmC studies are focused on prenatal development, especially the stem cell differentiation and lineage. There are limited studies on the biological roles of 5-hmC in adult tissues. If 5-hmC markers could be validated in various biological samples, the 5-hmC signature could be applied in clinical research, such as in the development of non-invasive diagnostic and prognostic approaches, by testing for epigenetic markers of body fluids (i.e., serum and urine). Third, is there any critical window for the establishment of 5-hmC marks under the exposure of environmental factors? Not many studies have demonstrated the establishment of 5-hmC from the prenatal to the aging or diseased stages. No doubt it could be assayed in the rodent model first to understand its role in disease development and manipulate the studies on human tissues.

With the advance of “omic” profiling techniques and bioinformatic database searches, 5-hmC signatures can be discovered at every stage of cell development or in response to any environmental exposure. Proper bioinformatics and statistical analyses on high-throughput epigenomic technologies, such as microarray or deep sequencing, are required because

of the huge amount of data that these tests provide. A set of follow-up studies, *in vitro* or *in vivo*, are also needed to validate and characterize a new set of molecular participants in disease development. We expect to identify markers that could allow us to reset the disruptive hydroxymethylome in disease development. Additionally, techniques for detecting single CpG hydroxymethylation have been established and applied for detecting gene-specific hydroxymethylation. It helps to understand the epigenetic modulation of cell functions via modification at a specific gene promoter, leading to new designs for therapeutic drugs. In order to apply 5-hmC as a promising environmental biosensor, we must track epigenetic changes in different tissues and cell types, and the time course in a number of subjects. Applying 5-hmC studies to human population studies would speed up the understanding of the use of 5-hmC in predicting exposure and disease risk. However, methods of statistical analysis, such as data modeling and multiple testing, must be carefully chosen when we apply genetic and epigenetic analyses in environmental diseases. All in all, the studies of 5-hmC and its interactions with environmental exposure are exciting; however, many questions remain unanswered. It is believed that with the new basic knowledge of 5-hmC biology, advanced sequencing technology, and proper statistical and epidemiological strategies, we could expect the findings from 5-hmC studies to be applied for lifestyle recommendations and/or changes in clinical practices, which can lead to an improvement in disease management and, ultimately, in public health.

**Acknowledgment** The authors acknowledge the following support: NIEHS (ES016887-WY Tang, P30ES03819) and NIH (HL10342-W Mitzner).

### Compliance with Ethics Guidelines

**Conflict of Interest** T. Dao, R.Y.S. Cheng, M.P. Revelo, W. Mitzner, and WY Tang declare that they have no conflict of interest.

**Human and Animal Rights and Informed Consent** This article does not contain any studies with human or animal subjects performed by any of the authors.

### References

Papers of particular interest, published recently, have been highlighted as:

- Of importance

1. Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Genet.* 2003;33(3):245–54.
2. Egger G, Liang G, Aparicio A, Jones PA. Epigenetics in human disease and prospects for epigenetic therapy. *Nature.* 2004;429(6990):457–63.



3. Pfeifer GP, Kadam S, Jin SG. 5-hydroxymethylcytosine and its potential roles in development and cancer. *Epigenetics Chromatin*. 2013;6(10):1–9.
4. Wu H, Zhang Y. Mechanisms and functions of Tet protein-mediated 5-methylcytosine oxidation. *Genes Dev*. 2011;25(23):2436–52.
5. 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(14):4100–8.
6. Delatte B, Fuks F. TET proteins: on the frenetic hunt for new cytosine modifications. *Brief Funct Genomics*. 2013;12(3):191–204.
7. Kristensen LS, Treppendahl MB, Grønbaek K. Analysis of epigenetic modifications of DNA in human cells. *Curr Protoc Hum Genet*. 2013;77:20.2.1–20.2.22.
8. Costello JF, Plass C. Methylation matters. *J Med Genet*. 2001;38(5):285–303.
9. 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(5929):930–5. *This study demonstrated the discovery of the TET proteins, which catalyze the oxidation of 5-methylcytosine to 5-hydroxymethylation in mammalian DNA. This breakthrough suggested a novel mechanism of regulation of DNA methylation.*
10. Thomson JP, Moggs JG, Wolf CR, Meehan RR. Epigenetic profiles as defined signatures of xenobiotic exposure. *Mutat Res Genet Toxicol Environ Mutagen*. 1–22, Aug. 2013.
11. Wu H, Zhang Y. Tet1 and 5-hydroxymethylation: a genome-wide view in mouse embryonic stem cells. *Cell Cycle*. 2011;10(15):2428–36.
12. Zhao H, Chen T. Tet family of 5-methylcytosine dioxygenases in mammalian development. *J Human Genet*. 2013;58(7):421–7.
13. Williams K, Christensen J, Helin K. DNA methylation: TET proteins-guardians of CpG islands? *EMBO Rep*. 2011;13(1):28–35.
14. Thomson JP, Hunter JM, Lempiäinen H, Muller A, Terranova R, Moggs JG, et al. Dynamic changes in 5-hydroxymethylation signatures underpin early and late events in drug exposed liver. *Nucleic Acids Res*. 2013;41(11):5639–54.
15. Nestor CE, Ottaviano R, Reddington J, Sproul D, Reinhardt D, Dunican D, et al. Tissue type is a major modifier of the 5-hydroxymethylcytosine content of human genes. *Genome Res*. 2012;22(3):467–77.
16. Tan L, Shi YG. Tet family proteins and 5-hydroxymethylcytosine in development and disease. *Development*. 2012;139(11):1895–902.
17. 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*. 2011;133(3):486–92.
18. Yildirim O, Li R, Hung J-H, Chen PB, Dong X, Ee L-S, et al. Mbd3/NURD complex regulates expression of 5-hydroxymethylcytosine marked genes in embryonic stem cells. *Cell*. 2011;147(7):1498–510.
19. 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(1):17–30.
20. Chia N, Wang L, Lu X, Senut M, Brenner C, Ruden DM. Environmental regulation of 5-hydroxymethylcytosine by oxidative stress. *Epigenetics*. 2011;6(7):853–6.
21. Morera S, Imbert A, Aschke-Sonnenborn U, Ruger W, Freemont PS. T4 phage beta-glucosyltransferase: substrate binding and proposed catalytic mechanism. *J Mol Biol*. 1999;292(3):717–30.
22. Song C-X, 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(1):68–72.
23. Bhattacharyya S, Yu Y, Suzuki M, Campbell N, Mazdo J, Vasanthakumar A, et al. Genome-wide hydroxymethylation tested using the HELP-GT assay shows redistribution in cancer. *Nucleic Acids Res*. 2013;41(16):e157.
24. Chen L, Zhang L, Zhang L, Cai C. Analysis of global deoxyribonucleic acid 5-hydroxymethylcytosine in tissue by liquid chromatography-tandem mass spectrometry. *Se Pu*. 2012;30(5):533–7.
25. Gomez C, Pozo OJ, Diaz R, Sancho JV, Vilaroca E, Salvador JP, et al. Mass spectrometric characterization of urinary toremifene metabolites for doping control analyses. *J Chromatogr A*. 2011;1218(29):4727–37.
26. Mao W, Hu J, Hong T, Xing X, Wang S, Chen X, et al. A convenient method for selective detection of 5-hydroxymethylcytosine and 5-formylcytosine sites in DNA sequences. *Org Biomol Chem*. 2013;11(21):3568–72.
27. Zhang L, Zhang L, Zhou K, Ye X, Zhang J, Xie A, et al. Simultaneous determination of global DNA methylation and hydroxymethylation levels by hydrophilic interaction liquid chromatography-tandem mass spectrometry. *J Biomol Screen*. 2012;17(7):877–84.
28. Szwagierczak A, Bultmann S, Schmidt CS, Spada F, Leonhardt H. Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA. *Nucleic Acids Res*. 2010;38(19):e181.
29. Yu M, Hon GC, Szulwach KE, Song C-X, Jin P, Ren B, et al. Tet-assisted bisulfite sequencing of 5-hydroxymethylcytosine. *Nat Protoc*. 2012;7(12):2159–70.
30. Yu M, Hon GC, Szulwach KE, Song C-X, Zhang L, Kim A, et al. Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome. *Cell*. 2012;149(6):1368–80. *This technique could differentiate 5-mC and 5-hmC at a single-base resolution and could be applied for detection of either locus-specific or genome-wide 5-hydroxymethylation.*
31. Olkhov-Mitsel E, Bapat B. Strategies for discovery and validation of methylated and hydroxymethylated DNA biomarkers. *Cancer Med*. 2012;1(2):237–60.
32. Tan L, Xiong L, Xu W, Wu F, Huang N, Xu Y, et al. Genome-wide comparison of DNA hydroxymethylation in mouse embryonic stem cells and neural progenitor cells by a new comparative hMeDIP-seq method. *Nucleic Acids Res*. 2013;41(7):e84.
33. Yin R, Mao S-Q, Zhao B, Chong Z, Yang Y, Zhao C, et al. Ascorbic acid enhances Tet-mediated 5-methylcytosine oxidation and promotes DNA demethylation in mammals. *J Am Chem Soc*. 2013;135(28):10396–403.
34. Minor EA, Court BL, Young JI, Wang G. Ascorbate induces ten-eleven translocation (tet) methylcytosine dioxygenase-mediated generation of 5-hydroxymethylcytosine. *J Biol Chem*. 2013;288(19):13669–74.
35. Blaschke K, Ebata KT, Karimi MM, Zepeda-Martínez JA, Goyal P, Mahapatra S, et al. Vitamin C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells. *Nature*. 2013;500(7461):222–6. *This study investigates the role of vitamin C in regulating TET activity and DNA methylation profiles in mouse embryonic stem cells.*
36. Thomson JP, Lempiäinen H, Tefferi A, Nestor CE, Müller A, Bolognani F, et al. Non-genotoxic carcinogen exposure induces defined changes in the 5-hydroxymethylome. *Genome Biol*. 2012;13(10):R93.
37. Jefferson WN, Chevalier DM, Phelps JY, Cantor AM, Padilla-Banks E, Newbold RR, et al. Persistently altered epigenetic marks in the mouse uterus after neonatal estrogen exposure. *Mol Endocrinol*. 2013;27(10):1666–77.
38. Coulter JB, O'Driscoll CM, Bressler JP. Hydroquinone increases 5-hydroxymethylcytosine formation through ten eleven translocation 1 (Tet1) 5-methylcytosine dioxygenase. *J Biol Chem*. 2013;288(40):28792–800.
39. Shang Y, Das S, Rabold R, Sham JS, Mitzner W, Tang W-Y. Epigenetic alterations by DNA methylation in house dust mite-

- induced airway hyperresponsiveness. *Am J Respir Cell Mol Biol*. 2013;49(2):279–87.
40. Orr BA, Haffner MC, Nelson WG, Yegnasubramanian S, Eberhart CG. Decreased 5-hydroxymethylcytosine is associated with neural progenitor phenotype in normal brain and shorter survival in malignant glioma. *PLoS ONE*. 2012;7(7):e41036.
  41. Liu C, Liu L, Chen X, Shen J, Shan J, Xu Y, et al. Decrease of 5-hydroxymethylcytosine is associated with progression of hepatocellular carcinoma through downregulation of TET1. *PLoS ONE*. 2013;8(5):e62828.
  42. Lian CG, Xu Y, Ceol C, Wu F, Larson A, Dresser K, et al. Loss of 5-hydroxymethylcytosine is an epigenetic hallmark of melanoma. *Cell*. 2012;150(6):1135–46.
  43. Gambichler T, Sand M, Skrygan M. Loss of 5-hydroxymethylcytosine and ten-eleven translocation 2 protein expression in malignant melanoma. *Melanoma Res*. 2013;23(3):218–20.
  44. Huang H, Jiang X, Li Z, Li Y, Song CX, He C, et al. TET1 plays an essential oncogenic role in MLL-rearranged leukemia. *Proc Natl Acad Sci U S A*. 2013;110(29):11994–9.
  45. Liu X, Guangsen Z, Yi Y, Xiao L, Pei M, Liu S, et al. Decreased 5-hydroxymethylcytosine levels are associated with TET2 mutation and unfavorable overall survival in myelodysplastic syndromes. *Leuk Lymphoma*. 2013;54(11):2466–73.
  46. Wang F, Yang Y, Lin X, Wang JQ, Wu YS, Xie W, et al. Genome-wide loss of 5-hmC is a novel epigenetic feature of Huntington's disease. *Hum Mol Genet*. 2013;22(18):3641–53.
  47. Sierksma ASR, Prickaerts J, Chouliaras L, Rostamian S, Delbroek L, Rutten BPF, et al. Behavioral and neurobiological effects of prenatal stress exposure in male and female APP<sup>sw</sup>/PS1<sup>dE9</sup> mice. *Neurobiol Aging*. 2013;34(1):319–37.
  48. Dong E, Gavin DP, Chen Y, Davis J. Upregulation of TET1 and downregulation of APOBEC3A and APOBEC3C in the parietal cortex of psychotic patients. *Transl Psychiatry*. 2012;2:e159.
  49. Selgrade MK, Blain RB, Fedak KM, Cawley MA. Potential risk of asthma associated with in utero exposure to xenobiotics. *Birth Defect Res C Embryo Today*. 2013;99(1):1–13.
  50. 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*. 2011;71(24):7360–5.
  51. Tefferi A, Thiele J, Vardiman JW. The 2008 World Health Organization classification system for myeloproliferative neoplasms: order out of chaos. *Cancer*. 2009;115(17):3842–7.
  52. Branco MR, Ficz G, Reik W. Uncovering the role of 5-hydroxymethylcytosine in the epigenome. *Nat Rev Genet*. 2011;13(1):7–13.
  53. 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(7325):839–43.
  54. Mellén M, Ayata P, Dewell S, Kriaucionis S, Heintz N. MeCP2 binds to 5hmC enriched within active genes and accessible chromatin in the nervous system. *Cell*. 2012;151(7):1417–30.
  55. Della Sala G, Pizzorusso T. Synaptic plasticity and signaling in Rett syndrome. *Dev Neurobiol*. Aug. 2013.
  56. Grayson DR, Guidotti A. The dynamics of DNA methylation in schizophrenia and related psychiatric disorders. *Neuropsychopharmacology*. 2013;38(1):138–66.