

Genome-wide antagonism between 5-hydroxymethylcytosine and DNA methylation in the adult mouse brain

Junjie U. GUO^{1,2,3*}, Keith E. SZULWACH^{4*}, Yijing SU^{1,3}, Yujing LI⁴, Bing YAO⁴, Zihui XU⁴, Joo Heon SHIN⁵, Bing XIE⁵, Yuan GAO^{1,5}, Guo-li MING^{1,2,3}, Peng JIN(✉)⁴, Hongjun SONG(✉)^{1,2,3}

¹*Institute for Cell Engineering, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA*

²*The Solomon H. Snyder Department of Neuroscience, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA*

³*Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA.*

⁴*Department of Human Genetics, Emory University School of Medicine, 615 Michael Street, Atlanta, GA 30322, USA*

⁵*Lieber Institute for Brain Development, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA*

Abstract

BACKGROUND: Mounting evidence points to critical roles for DNA modifications, including 5-methylcytosine (5mC) and its oxidized forms, in the development, plasticity and disorders of the mammalian nervous system. The novel DNA base 5-hydroxymethylcytosine (5hmC) is known to be capable of initiating passive or active DNA demethylation, but whether and how extensively 5hmC functions in shaping the post-mitotic neuronal DNA methylome is unclear.

METHODS: Adult mice (8 to 10 weeks old, male, C57BL/6 background) were used for analysis. 5hmC enrichment was performed as previously described with an improved selective chemical labeling method. Five ng of 5hmC-enriched-genomic DNA from 3 independent 5hmC captures or one non-enriched input genomic DNA was end-repaired, adenylated, and ligated to Illumina Genomic DNA Adapters.

RESULTS: 5hmC in the neuronal genome is highly enriched in gene bodies, especially in exons, and correlates with gene expression. Direct genome-wide comparison of 5hmC distribution between embryonic stem cells and neurons reveals extensive differences, reflecting the functional disparity between these two cell types.

CONCLUSIONS: Here we systematically compared the 5hmC distributions between two cell types, ESCs and DGNs, both of which exhibit relatively high abundances of 5hmC. We uncovered both similarities and differences in the 5hmC profiles in these two cell types. In both ESCs and DGNs, 5hmC was enriched near and within genes, especially in coding exons, suggesting a potential role for 5hmC in exon definition and regulation of pre-mRNA splicing. Integrative analysis of 5hmC, overall DNA methylation and gene expression profiles of dentate granule neurons *in vivo* reveals the genome-wide antagonism between these two states of cytosine modifications, supporting a role for 5hmC in shaping the neuronal DNA methylome by promoting active DNA demethylation.

Keywords dentate granule neuron, active DNA demethylation, TET, methylome

Introduction

5-methylcytosine (5mC) is the predominant and most extensively studied epigenetic modification in the mammalian genome (Bird, 2002). Although it is well established that DNA methylation is deposited by a conserved family of DNA methyltransferases (DNMTs) (Goll and Bestor, 2005), the mechanisms underlying the enzymatic removal of this modification remain hotly debated (Ma et al., 2009a; Zhu, 2009; Wu and Zhang, 2010;

Received January 9, 2014; accepted January 11, 2014

Correspondence: ^aHongjun SONG, ^bPeng JIN

E-mail: ^ashongju1@jhmi.edu, ^bpeng.jin@emory.edu

*These authors contributed equally to this work.

Bhutani et al., 2011). Ten-eleven-translocated (TET) family dioxygenases oxidize the 5-methyl group of 5mC, and generate 5hmC (Tahiliani et al., 2009; Ito et al., 2010) as well as the more oxidized 5-formylcytosine and 5-carboxylcytosine (He et al., 2011; Ito et al., 2011). Multiple lines of evidence have supported a role for 5mC-to-5hmC conversion in the process of DNA demethylation. First, 5hmC is known to prevent DNMT1 from recognizing hemi-methylated DNA *in vitro*, which may cause “dilution” of the existing DNA methylation pattern in dividing cells (Valinluck and Sowers, 2007; Frauer et al., 2011). Indeed, such a passive mode of DNA demethylation was recently found to occur in early mouse embryos (Inoue and Zhang, 2011). Second, irreproducible 5hmC-containing DNA can be demethylated in both human cell lines and cultured mouse neurons (Zhang et al., 2010; Guo et al., 2011c), indicating that 5hmC can also function as a precursor of active DNA demethylation. Finally, gain- and loss-of-function manipulations of TET proteins lead to changes in DNA methylation that are consistent with their proposed roles in active DNA demethylation (Kohli and Zhang, 2013; Pastor et al., 2013).

Intriguingly, among all the tissues surveyed thus far, the mammalian brain has the highest abundance of 5hmC (Kriaucionis and Heintz, 2009; Globisch et al., 2010). In the brain, 5hmC accounts for 10%–40% of all methylated cytosines, compared to 1%–2% in other tissues, suggesting the potential functional importance of 5hmC in the nervous system. Previously generated genomic maps of 5hmC in largely post-mitotic mouse hippocampus as well as mouse and human cerebellum tissue indicate that 5hmC associates strongly with genic regions, is depleted at transcription start sites (TSSs), and correlates with gene expression (Song et al., 2011). Our previous studies suggest that 5hmC levels are significantly higher in both the cerebellum and hippocampus of adult mice at six weeks and one year of age than in the early postnatal stage, indicating an age-dependent acquisition of 5hmC and regulation of brain-specific gene expression (Szulwach et al., 2011b). Recent genome-wide studies from us and others have uncovered global changes of 5hmC associated with Rett syndrome as well as a number of neurodegenerative disorders, suggesting potential roles of precise 5hmC distributions in the development and homeostasis of the nervous system (Szulwach et al., 2011a; Szulwach et al., 2011b; Wang et al., 2013; Yao et al., 2013). However, the exact role and scope of 5hmC-mediated active DNA demethylation in shaping the landscape of DNA methylation in neurons *in vivo* are still unclear.

In this study, we systematically characterized the 5hmC distribution in dentate granule neurons (DGNs) *in vivo* and compared it to the 5hmC map of cultured mouse embryonic stem cells (ESCs). We correlated the 5hmC distribution in DGN with both gene expression and transcription factor occupancy. Cross-comparison between 5hmC profiles and overall DNA methylation distributions revealed the global antagonism between these two modification states, supporting a role for 5hmC in shaping the neuronal DNA methylome on a genome-wide scale.

Materials and methods

Tissue preparation. Adult mice (8 to 10 weeks old, male, C57BL/6 background) were used for analysis in accordance with protocols approved by the Institutional Animal Care and Use Committee. Dentate gyrus tissues were rapidly micro-dissected bilaterally from adult mice. This preparation was highly enriched for mature neurons as shown by immunohistology to contain ~90% NeuN⁺ dentate granule neurons (Ma et al., 2009b). Previous studies of this preparation showed a very similar CpG methylation status at selective loci with FACS purified NeuN⁺ mature neurons (Guo et al., 2011a). Validation experiments were performed using independent biological samples of bilaterally micro-dissected dentate gyri from individual animals.

5hmC DNA capture. 5hmC enrichment was performed as previously described with an improved selective chemical labeling method (Song et al., 2011). 5hmC labeling reactions were performed in a 100 μ L solution containing 50 mM HEPES buffer (pH 7.9), 25 mM MgCl₂, 300 ng/ μ L sonicated genomic DNA (100-500 bp), 250 μ M UDP-6-N₃-Glu, and 2.25 μ M wild-type β -GT. Reactions were incubated for 1 h at 37 °C. DNA substrates were purified via Qiagen DNA purification kit or by phenol-chloroform precipitation and reconstituted in H₂O. Click chemistry was performed with the addition of 150 μ M dibenzocyclooctyne modified biotin into the DNA solution and incubated for 2 h at 37 °C. Samples were purified by Pierce Monomeric Avidin Kit (Thermo) following manufacturer’s recommendations. After elution, biotin-5-N₃-gmC-containing DNA was concentrated by 10K Amicon Ultra-0.5 mL Centrifugal Filters (Millipore) and purified by Qiagen DNA purification kit.

Library construction and high-throughput sequencing. Five ng of 5hmC-enriched-genomic DNA from 3 independent 5hmC captures or one non-enriched input genomic DNA was end-repaired, adenylated, and ligated to Illumina Genomic DNA Adapters (Genomic DNA adapter oligo mix) according to standard Illumina protocols for

ChIP-Seq library construction, maintaining the proper molar ratios of adapter to insert. Adapter-ligated fragments of ~200–350 bp were gel-purified by 2% agarose gel electrophoresis and PCR-amplified for 18 PCR cycles. Libraries were checked for quality and quantified using an Agilent 2100 Bioanalyzer DNA 1000 Chip.

Libraries were sequenced using the Illumina HiScan platform. Cluster generation was performed with Illumina TruSeq cluster kit v2-cBot-HS. Single-read 51-bp sequencing was completed with Illumina TruSeq SBS kit v3-HS. A dedicated PhiX control lane, as well as 1% PhiX spike in all other lanes, was used for automated matrix and phasing calculations. Image analysis and base calling were performed with the standard Illumina pipeline.

Data processing and analysis. FASTQ reads were aligned to NCBIv1/mm9 with Bowtie v0.17.2 retaining non-duplicate, unique matches to the genome, with no more than 3 mismatches in the first 30 bases. Ensembl gene annotations were downloaded from the UCSC Genome Browser (<http://genome.ucsc.edu>). Data analysis and visualization were done using built-in functions of R (<http://www.r-project.org>) and in-house Perl scripts.

qPCR validation of 5hmC-enriched loci. One ng of input or 5mC-enriched DNA, from an independent 5mC capture experiment, was used in triplicate 20 μ L qPCR reactions, each with 1X Power SYBR Green PCR Master Mix (Applied Biosystems), 0.5 μ M forward and reverse primers, and water. Reactions were run on an SDS 7500 Fast Instrument using 7500 Standard cycling conditions. Fold enrichment was calculated as 2^{-dCt} , where $dCt = Ct(5\text{-hmC-enriched}) - Ct(\text{Input})$.

Results

Global properties of the neuronal 5hmC profile

We used a previously developed chemical tagging method to profile the 5hmC distribution in the neuronal genome (Song et al., 2011). The dentate gyrus from an *in vivo* preparation was used because the majority of cells in this region are Prox1⁺NeuN⁺ post-mitotic DGNs (Ma et al., 2009b; Guo et al., 2011a), providing relatively high homogeneity compared to other regions in the mouse brain. A total of ~33 million unique non-duplicate reads were obtained from two biological replicates, which were highly correlated ($r = 0.952$; Fig. 1A). In addition, we also profiled mouse ESCs (~19 million unique non-duplicate reads). To validate the neuronal 5hmC profiles, we used quantitative PCR to measure 5hmC enrichment at specific loci that were identified as 5hmC-marked by sequencing. In total, 16 out of 17 regions showed significant enrichment relative to a negative control region that is not 5hmC-marked (Fig. 1B). Global analysis of input-normalized 5hmC signals showed that all autosomes had comparable 5hmC levels (Fig. 1C), whereas sex chromosomes showed lower 5hmC levels, consistent with previous studies in both ESCs (Pastor et al., 2011; Szulwach et al., 2011a) and the brain (Szulwach et al., 2011b). The mitochondrial genome was depleted in 5hmC, which is consistent with our previous finding that the neuronal mitochondrial genome is virtually non-methylated (Guo et al., 2011a).

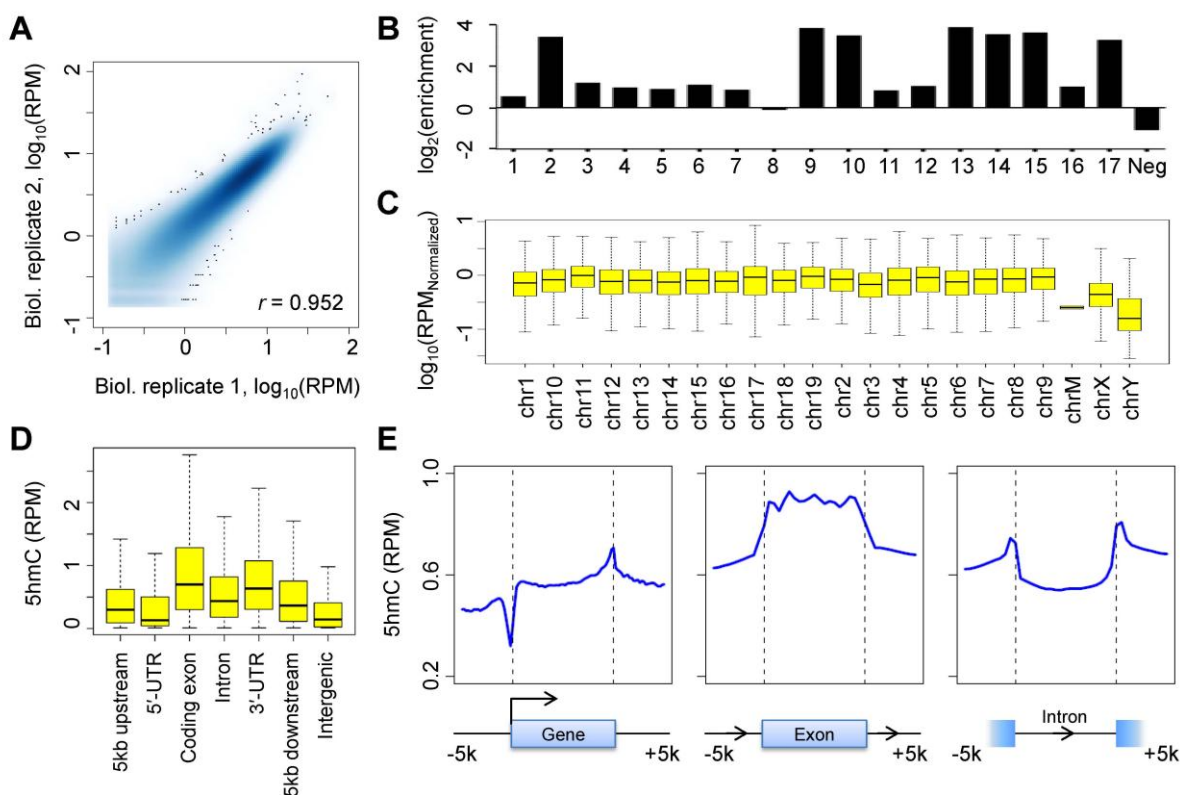


Figure 1 (Guo et al.)

Figure 1 Distribution of 5hmC in the genome of dentate granule neurons (DGNs) *in vivo*. (A) Correlation between 5hmC-seq results from two biologic replicates. Pearson's correlation was calculated using non-overlapping 10-kb windows. (B) qPCR validation of DGN 5hmC-marked regions detected by 5hmC-seq. Regions 1-17 were those marked with 5hmC. Neg was a negative control region with no significant 5hmC. (C) Chromosome-wide 5hmC levels normalized by an input library. Note the lower abundance in the mitochondrial genome and sex chromosomes. (D) 5hmC levels within different genic regions, compared to intergenic regions. E, Meta-gene (left), meta-exon (middle), and meta-intron analyses of 5hmC levels.

We next determined the average 5hmC signal across the genic regions. All regions near or within annotated genes, including TSS upstream regions, 5'-untranslated regions (UTR), coding exons, introns, 3'-UTR, and polyadenylation site (PAS) downstream regions, showed higher 5hmC levels than intergenic regions (Fig. 1D), which is consistent with previous findings that 5hmC is enriched near genes in the brain (Szulwach et al., 2011b; Mellán et al., 2012). Meta-gene analysis showed that gene bodies were enriched in 5hmC (Fig. 1E). Within gene bodies, coding exons are most enriched for 5hmC (Fig. 1D,E). 5hmC levels gradually increased from 5' to 3' ends within gene bodies (Fig. 1E), although to a lesser degree compared to ESCs (Xu et al., 2011).

Relationship between 5hmC and gene expression in neurons

Given the gene-centric property of 5hmC, we determined the relationship between 5hmC levels and associated gene expression levels by correlating 5hmC profiles with previously obtained DGN RNA-seq results (Guo et al., 2013). Meta-gene analysis showed a striking position-specific relationship between 5hmC levels and transcript abundance (Fig. 2A). Near TSSs, 5hmC anti-correlated with gene expression level; 5hmC-marked TSSs were restricted to the genes that had the lowest expression levels. However, in all other structures, including the TSS upstream regions, gene body, and PAS downstream regions, 5hmC levels were positively correlated with gene expression (Fig. 2A). Furthermore, the positive correlation gradually increased from 5' to 3' within the gene body.

Recent studies have revealed an important role for protein-DNA interactions in determining DNA methylation levels at binding sites (Lister et al., 2009; Guo et al., 2011a; Lienert et al., 2011; Stadler et al., 2011b). To test whether protein-DNA interactions also shape the distribution of 5hmC at these binding sites, we averaged 5hmC levels within 10-kb windows across all binding sites for each of several transcription factors for which ChIP-seq profiles were available (Kim et al., 2010). Interestingly, 5hmC was depleted at the binding sites for all the tested neuronal DNA binding factors, including cAMP-responsive element binding protein (CREB), CREB binding protein (CBP), RNA polymerase II (RNAP), and serum response factor (SRF; Fig. 2B). These results suggested that protein-DNA interactions may protect the DNA from methylation and/or hydroxymethylation, or cause DNA demethylation. It is also possible that DNA (hydroxy-) methylation prevents transcription factor binding, which is a general property of many DNA binding factors (Stadler et al., 2011a; Spruijt et al., 2013).

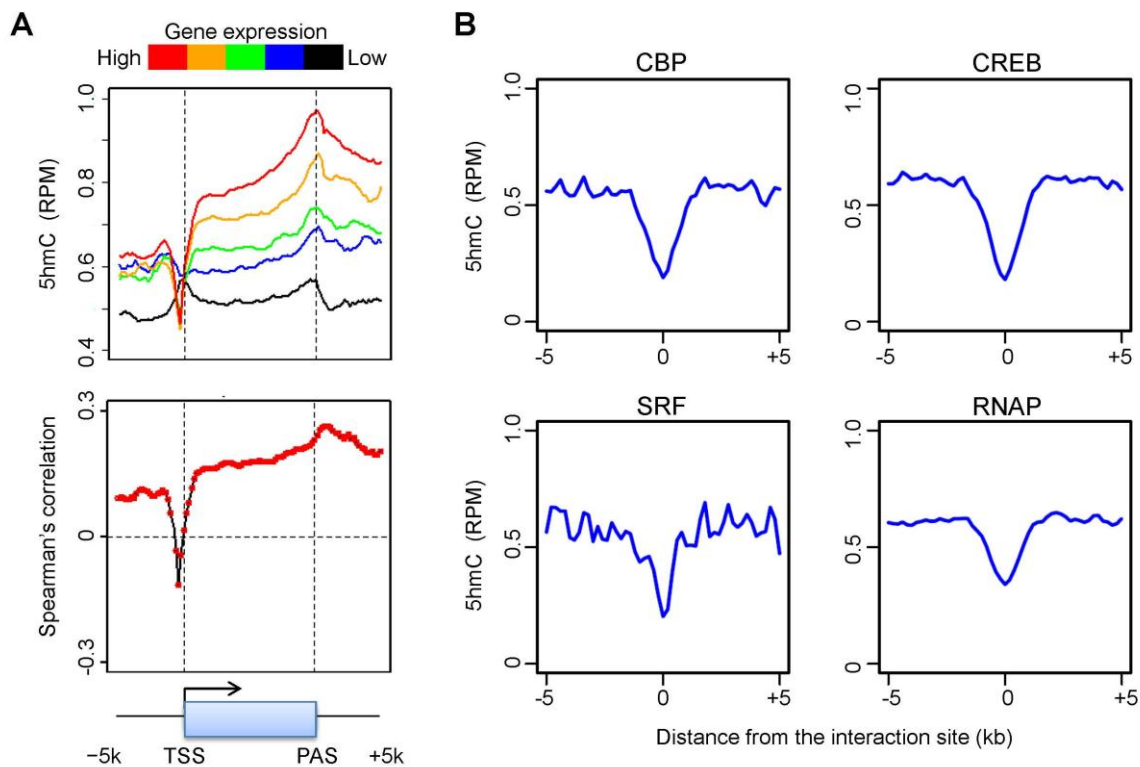


Figure 2 (Guo et al.)

Figure 2 Relationship between 5hmC, gene expression, and protein-DNA interaction in DGNs. (A) Meta-gene analysis of 5hmC levels correlated with gene expression. 5hmC levels were averaged across all genes stratified by their expression rank (top). Spearman's correlations in 500 bp windows were shown (bottom). (B) 5hmC levels near the binding sites of four neuronal DNA-interacting factors previously determined in cultured mouse cortical neurons (Kim et al., 2010).

Extensive differences in 5hmC distributions between ESCs and DGNs

5hmC was first identified in the genomic DNA of ESCs (Tahiliani et al., 2009) and cerebellar neurons (Kriaucionis and Heintz, 2009). Recent studies have shown that these two cell types have higher 5hmC abundance than other somatic cell types (Globisch et al., 2010). We compared the two 5hmC profiles obtained using the same chemical tagging method. Global correlations between ESCs and either DGN sample ($r = 0.69$ and 0.62 , respectively) were lower than between the two DGN samples ($r = 0.83$; Fig. 3A), suggesting extensive differences between the 5hmC distributions in ESC and DGN.

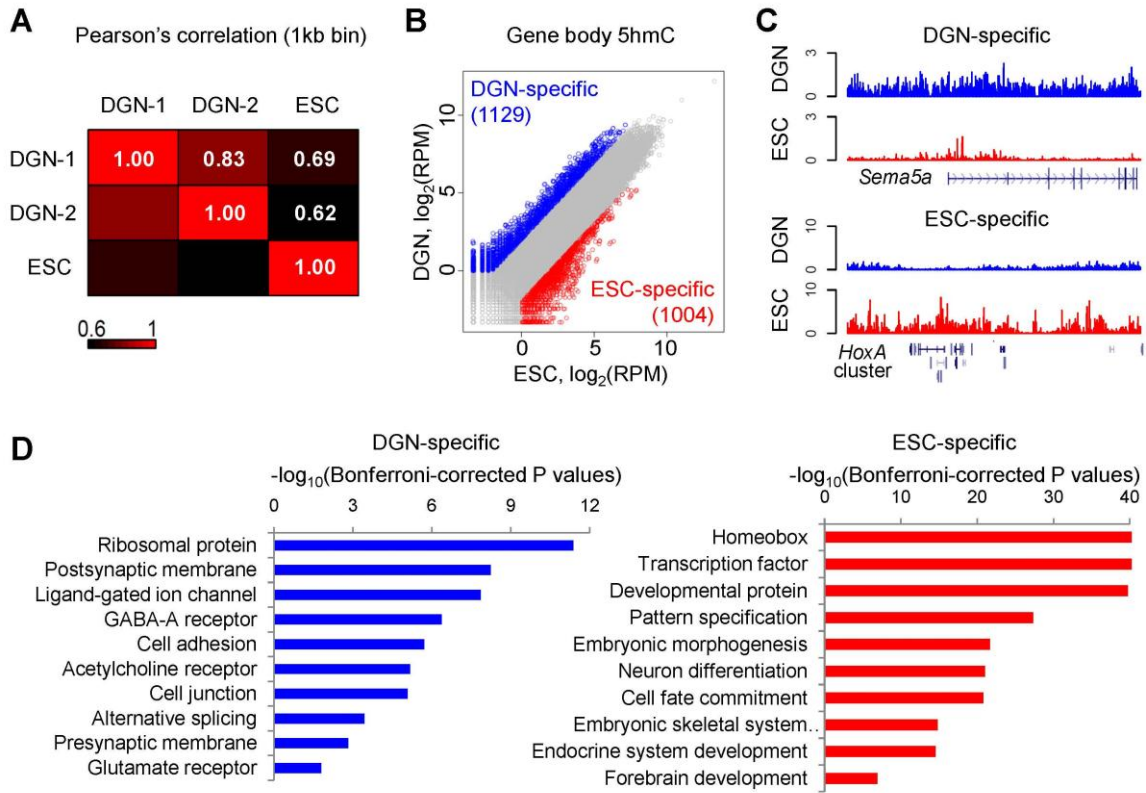


Figure 3 (Guo et al.)

Figure 3 Comparison of 5hmC distributions between DGNs and ESCs. (A) A correlation matrix of 5hmC-seq results from two DGN replicates and on ESC sample. Pearson's correlations were shown. (B) A scatter plot of DGN and ESC gene-body 5hmC levels. DGN- and ESC-specific 5hmC-marked genes were indicated by blue and red circles, respectively. (C) Examples of DGN-specific 5hmC-marked genes (top) and ESC-specific 5hmC-marked genes (bottom). (D) Gene ontology analysis of 1129 DGN-specific 5hmC-marked genes (left) and 1004 ESC-specific 5hmC-marked genes (right).

A comparison of gene-body 5hmC levels for individual genes between the two cell types identified 1129 genes that showed ≥ 4 -fold higher 5hmC density in neurons than in ESCs (DGN-specific) and 1004 genes that are ESC-specific (Fig. 3B). For example, *Sema5a*, an autism susceptibility gene that encodes an axon guidance cue, showed a higher 5hmC density in DGNs than in ESCs (Fig. 3C). In contrast, the *HoxA* gene cluster, which encodes an important group of developmental regulators, had much higher 5hmC levels in ESCs than in DGNs (Fig. 3C). Gene ontology analysis further revealed that neuron-specific 5hmC-marked genes were enriched in ribosomal proteins and synaptic proteins, such as neurotransmitter receptors (Fig. 3D), whereas ESC-specific 5hmC-marked genes were enriched in developmental regulatory pathways. Therefore, differentially 5hmC-marked genes were associated with the functional differences between the two cell types.

DNA methylation and 5hmC are antagonistic genome-wide in both DGNs and ESCs

To examine the relationship between overall DNA methylation (i.e. the sum of 5mC and 5hmC) and 5hmC, we overlaid the 5hmC map with the previously determined DNA methylation profile for the same DGNs (Guo et al., 2011a). Strikingly, the global profiles of these two modification states showed highly complementary patterns (Fig. 4A). The high 5hmC density regions showed low levels of DNA methylation and vice versa. Across the genome, 5hmC and overall DNA methylation were significantly anti-correlated (Fig. 4B).

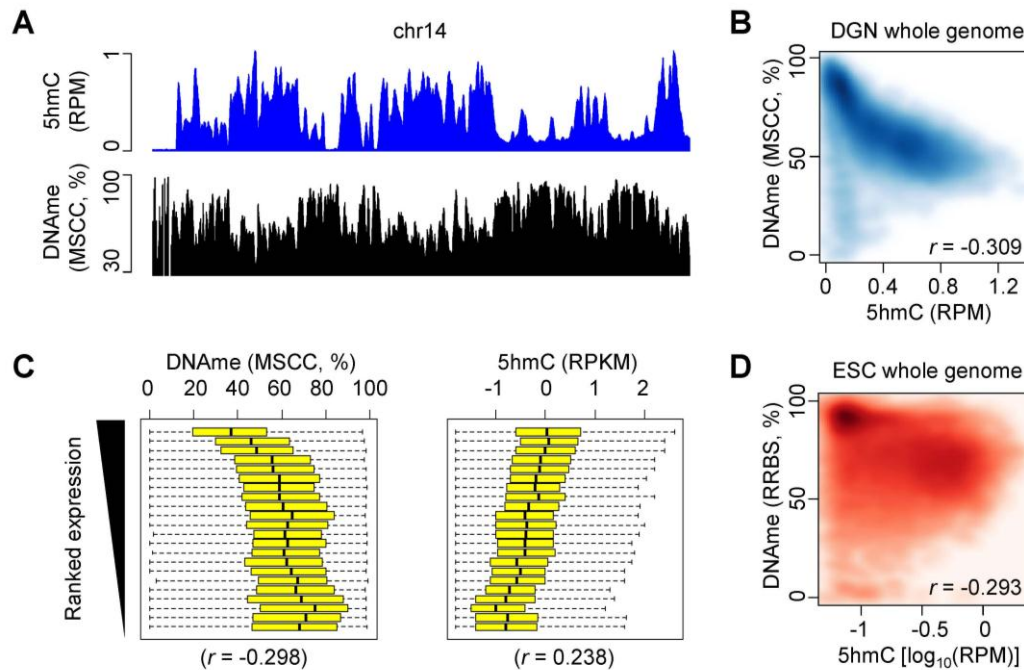


Figure 4 (Guo et al.)

Figure 4 Global antagonism between 5hmC and overall DNA methylation in both DGNs and ESCs. (A) A chromosome-wide view of 5hmC (top) and overall DNA methylation (bottom) levels. Signals were moving-averaged using 10-kb windows. (B) A density plots of genome-wide paired values of DGN overall DNA methylation and 5hmC levels of 10-kb windows. (C) Boxplots of gene-body DNA methylation (left) and 5hmC (right) levels stratified by the expression ranks of all genes. Spearman's correlations were shown at the bottom. (D) A density plots of genome-wide paired values of ESC overall DNA methylation and 5hmC levels of 10-kb windows. 5hmC levels were log10-transformed.

To rule out the possibility that the global antagonism between 5hmC and DNA methylation was merely due to the association of 5hmC and gene-rich regions in the genome, we calculated the averaged gene-body 5hmC and overall DNA methylation levels for each gene and ranked all the genes by their expression levels. As we showed previously, gene-body 5hmC levels were correlated with gene expression (Fig. 4C), whereas DNA methylation was anti-correlated with gene expression, suggesting that the antagonism between 5hmC and overall DNA methylation also exists at the individual gene level.

Finally, we tested whether 5hmC and DNA methylation were also antagonistic genome-wide in ESCs. Using the previously determined ESC methylome data (Meissner et al., 2008), we found a significant anti-correlation between 5hmC and overall DNA methylation in ESCs (Fig. 4D), suggesting that the global antagonism between these two states of cytosine modification is not restricted to post-mitotic neurons.

Discussion

Since the identification of 5hmC in the genomic DNA from ESCs and Purkinje cells, the biological role of this DNA base has been the subject of intensive study. Here we systematically compared the 5hmC distributions between two cell types, ESCs and DGNs, both of which exhibit relatively high abundances of 5hmC. We uncovered both similarities and differences in the 5hmC profiles in these two cell types. In both ESCs and DGNs, 5hmC was enriched near and within genes, especially in coding exons, suggesting a potential role for 5hmC in exon definition and regulation of pre-mRNA splicing, as has been indicated for 5mC (Shukla et al., 2011). Intragenic 5hmC levels were correlated with gene expression in both cell types. As a result, differentially 5hmC-marked genes between ESCs and DGNs reflected the functional differences between the two cell types, supporting a potential role for 5hmC in the regulation of cell type-specific gene expression.

A central question of 5hmC biology is whether it indeed functions as an intermediate product in active DNA demethylation (Wu and Zhang, 2010; Guo et al., 2011b; Branco et al., 2012; Kohli and Zhang, 2013). We have previously shown that fully hydroxymethylated exogenous linear DNA can be demethylated in both human cell lines and mouse neurons in culture (Guo et al., 2011c), suggesting that 5hmC can be converted to non-methylated cytosines by an active mechanism. Here, by correlating the DGN 5hmC profile with our previously reported DGN methylome (Guo et al., 2011a), we showed that 5hmC and DNA methylation were antagonistic genome-wide, further supporting a role for 5hmC in shaping the genomic DNA methylome by promoting active DNA demethylation in post-mitotic neurons. In addition, the global antagonism between 5hmC and DNA methylation was also observed in ESCs, suggesting that maintaining DNA methylation at low levels is a cell-type-independent function of 5hmC.

Previous studies have shown that DNMT1 has a lower methyltransferase activity toward hemi-hydroxymethylated DNA *in vitro* (Valinluck and Sowers, 2007), raising the possibility that 5hmC may also promote passive DNA demethylation. However, DNMT1-associated factor UHRF1 can recognize hemi-hydroxymethylated DNA (Frauer et al., 2011; Spruijt et al., 2013). Therefore, whether 5hmC has an influence on the maintenance of the symmetric DNA methylation patterns after DNA replication remains to be tested. On the other hand, TET proteins have been shown to play important roles in locus-specific DNA demethylation in several systems, including preimplantation embryos (Gu et al., 2011), primordial germ cells (Dawlaty et al., 2013; Yamaguchi et al., 2013), and various brain regions (Kaas et al., 2013; Rudenko et al., 2013). While it remains unclear to what extent DNA replication and passive dilution of 5hmC play roles in TET-mediated DNA demethylation in embryonic stages, the post-mitotic nature of neurons provides a suitable system for studying the role of 5hmC in active DNA demethylation without any contribution from DNA replication. Future analyses using methods that can distinguish 5mC and its oxidized forms at base resolution (Booth et al., 2012; Yu et al., 2012; Song et al., 2013) will potentially reveal the kinetic details in each individual steps of 5hmC-mediated DNA demethylation, which may yield novel insights into the dynamic regulation of DNA modification states in post-mitotic neurons.

DNA methylation has been shown to play a critical role in synaptic plasticity related to learning and memory in mature neurons, likely owing to the regulation of specific gene expression (Miller and Sweatt, 2007; Feng et al., 2010). Previous studies of DNA methyltransferases have pointed to a mechanism whereby individual DNMTs play distinctive roles during neurodevelopment and are orchestrated to maintain long-term proper neuronal functions (Goto et al., 1994; Feng et al., 2005). Our understanding of the detailed functions and molecular mechanisms of 5hmC and TET proteins during neurodevelopment and in mature neurons is still very limited. A recent study of the dynamic change of TET proteins and 5hmC during neurogenesis revealed increased 5hmC levels during neuronal differentiation (Hahn et al., 2013). It also showed a negative correlation of 5hmC with repressive histone mark H3K27 tri-methylation and its effector Polycomb protein complex. These data indicate a potential role of 5hmC and TET proteins in neurodevelopment, as shown in *Xenopus* (Xu et al., 2011). Several recent publications simultaneously reported key roles of Tet1 in neurogenesis, cognition and memory formation (Kaas et al., 2013; Rudenko et al., 2013; Zhang et al., 2013). Either depletion or overexpression of Tet1 in mouse brain could lead to severe consequences. For example, Tet1 KO mice display impaired neurogenesis, poor learning and memory, abnormal long-term depression and impaired memory extinction, whereas Tet1 overexpression could result in impaired long-term memory formation. Mechanistically, up- or downregulation of Tet1 were found to lead to dysregulation of genes that are involved in critical neuronal activities accompanied by local 5hmC changes. These data together strongly indicate that the level and genomic distribution of Tet1 must be precisely controlled, possibly in a distinct manner in different neural cell types, during neurodevelopment and in mature neurons. This notion was also supported by the observation that the 5hmC distribution and gene expression in different neuronal cell types, such as Purkinje neurons, granule cells and Bergmann glial cells, displayed strong cell-type bias (Mellán et al., 2012). In that study, MeCP2 was also identified as 5hmC binding protein, establishing a dual role for MeCP2 in the orchestration of neuronal plasticity by coordinating different cytosine modifications (Mellán et al., 2012). Our work presented here on the integrative analyses of 5hmC, overall DNA methylation and gene expression profiles in the *in vivo* dentate granule neurons, lay the foundation for the future study of 5hmC in shaping the neuronal DNA methylomes during neurodevelopment and in mature neurons.

Compliance with ethics guidelines

Junjie Guo, Keith E. Szulwach, Yijing Su, Yujing Li, Bing Yao, Zihui Xu, Joo Heon Shin, Bing Xie, Yuan Gao, Guo-li Ming, Peng Jin and Hongjun Song declare that they have no conflict of interest.

All institutional and national guidelines for the care and use of laboratory animals were followed.

Acknowledgments

We thank Cheryl Strauss and Kimberley Christian for critical reading of the manuscript. This study was supported in part by the National Institutes of Health (NS051630 and MH076090 to P.J.; NS047344, MH087874, ES021957 to H.S., HD0679184 and NS048271 to G.L.M.), the Emory Genetics Discovery Fund (P.J.), the Simons Foundation Autism Research Initiative (P.J. and H.S.), Dr. Miriam & Sheldon G. Adelson Medical Research Foundation (G.L.M.), Maryland Stem Cell Research Foundation (G.L.M.), and John Hopkins Brain Science Institute (G.L.M.). J.U.G. was a Damon Runyon fellow supported by the Damon Runyon Cancer Research Foundation. Y.S. was supported by a postdoctoral fellowship from Maryland Stem Cell Research Foundation.

References

- Bhutani N, Burns D M, Blau H M (2011). DNA demethylation dynamics. *Cell*, 146(6): 866–872
- Bird A (2002). DNA methylation patterns and epigenetic memory. *Genes Dev*, 16(1): 6–21
- Booth M J, Branco M R, Ficz G, Oxley D, Krueger F, Reik W, Balasubramanian S (2012). Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution. *Science*, 336(6083): 934–937
- Branco M R, Ficz G, Reik W (2012). Uncovering the role of 5-hydroxymethylcytosine in the epigenome. *Nat Rev Genet*, 13(1): 7–13
- Dawlaty M M, Breiling A, Le T, Raddatz G, Barrasa M I, Cheng A W, Gao Q, Powell B E, Li Z, Xu M, Faull K F, Lyko F, Jaenisch R (2013). Combined deficiency of Tet1 and Tet2 causes epigenetic abnormalities but is compatible with postnatal development. *Dev Cell*, 24(3): 310–323
- Feng J, Chang H, Li E, Fan G (2005). Dynamic expression of de novo DNA methyltransferases Dnmt3a and Dnmt3b in the central nervous system. *J Neurosci Res*, 79(6): 734–746
- Feng J, Zhou Y, Campbell S L, Le T, Li E, Sweatt J D, Silva A J, Fan G (2010). Dnmt1 and Dnmt3a maintain DNA methylation and regulate synaptic function in adult forebrain neurons. *Nat Neurosci*, 13(4): 423–430
- Frauer C, Hoffmann T, Bultmann S, Casa V, Cardoso M C, Antes I, Leonhardt H (2011). Recognition of 5-hydroxymethylcytosine by the Uhrf1 SRA domain. *PLoS ONE*, 6(6): e21306
- Globisch D, Münzel M, Müller M, Michalakakis S, Wagner M, Koch S, Brückl T, Biel M, Carell T (2010). Tissue distribution of 5-hydroxymethylcytosine and search for active demethylation intermediates. *PLoS ONE*, 5(12): e15367
- Goll M G, Bestor T H (2005). Eukaryotic cytosine methyltransferases. *Annu Rev Biochem*, 74(1): 481–514
- Goto K, Numata M, Komura J I, Ono T, Bestor T H, Kondo H (1994). Expression of DNA methyltransferase gene in mature and immature neurons as well as proliferating cells in mice. *Differentiation*, 56(1-2): 39–44
- Gu T P, Guo F, Yang H, Wu H P, Xu G F, Liu W, Xie Z G, Shi L, He X, Jin S G, Iqbal K, Shi Y G, Deng Z, Szabo P E, Pfeifer G P, Li J, Xu G L (2011). The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. *Nature*, 477(7366): 606–610
- Guo J U, Ma D K, Mo H, Ball M P, Jang M H, Bonaguidi M A, Balazer J A, Eaves H L, Xie B, Ford E, Zhang K, Ming G L, Gao Y, Song H (2011a). Neuronal activity modifies the DNA methylation landscape in the adult brain. *Nat Neurosci*, 14(10): 1345–1351
- Guo J U, Su Y, Shin J H, Shin J, Li H, Xie B, Zhong C, Hu S, Le T, Fan G, Zhu H, Chang Q, Gao Y, Ming G L, Song H (2013). Distribution, recognition and regulation of non-CpG methylation in the adult mammalian brain. *Nat Neurosci*,

- Guo J U, Su Y, Zhong C, Ming G L, Song H (2011b). Emerging roles of TET proteins and 5-hydroxymethylcytosines in active DNA demethylation and beyond. *Cell Cycle*, 10(16): 2662–2668
- Guo J U, Su Y, Zhong C, Ming G L, Song H (2011c). Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. *Cell*, 145(3): 423–434
- Hahn M A, Qiu R, Wu X, Li A X, Zhang H, Wang J, Jui J, Jin S G, Jiang Y, Pfeifer G P, Lu Q (2013). Dynamics of 5-hydroxymethylcytosine and chromatin marks in Mammalian neurogenesis. *Cell Rep*, 3: 291–300
- He Y F, Li B Z, Li Z, Liu P, Wang Y, Tang Q, Ding J, Jia Y, Chen Z, Li L, Sun Y, Li X, Dai Q, Song C X, Zhang K, He C, Xu G L (2011). Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science*, 333(6047): 1303–1307
- Inoue A, Zhang Y (2011). Replication-dependent loss of 5-hydroxymethylcytosine in mouse preimplantation embryos. *Science*, 334(6053): 194
- Ito S, D'Alessio A C, Taranova O V, Hong K, Sowers L C, Zhang Y (2010). Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature*, 466(7310): 1129–1133
- Ito S, Shen L, Dai Q, Wu S C, Collins L B, Swenberg J A, He C, Zhang Y (2011). Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science*, 333(6047): 1300–1303
- Kaas G A, Zhong C, Eason D E, Ross D L, Vachhani R V, Ming G L, King J R, Song H, Sweatt J D (2013). TET1 controls CNS 5-methylcytosine hydroxylation, active DNA demethylation, gene transcription, and memory formation. *Neuron*, 79(6): 1086–1093
- Kim T K, Hemberg M, Gray J M, Costa A M, Bear D M, Wu J, Harmin D A, Laptewicz M, Barbara-Haley K, Kuersten S, Markenscoff-Papadimitriou E, Kuhl D, Bito H, Worley P F, Kreiman G, Greenberg M E (2010). Widespread transcription at neuronal activity-regulated enhancers. *Nature*, 465(7295): 182–187
- Kohli R M, Zhang Y (2013). TET enzymes, TDG and the dynamics of DNA demethylation. *Nature*, 502(7472): 472–479
- Kriaucionis S, Heintz N (2009). The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science*, 324(5929): 929–930
- Lienert F, Wirbelauer C, Som I, Dean A, Mohn F, Schübeler D (2011). Identification of genetic elements that autonomously determine DNA methylation states. *Nat Genet*, 43(11): 1091–1097
- Lister R, Pelizzola M, Dowen R H, Hawkins R D, Hon G, Tonti-Filippini J, Nery J R, Lee L, Ye Z, Ngo Q M, Edsall L, Antosiewicz-Bourget J, Stewart R, Ruotti V, Millar A H, Thomson J A, Ren B, Ecker J R (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature*, 462(7271): 315–322
- Ma D K, Guo J U, Ming G L, Song H (2009a). DNA excision repair proteins and Gadd45 as molecular players for active DNA demethylation. *Cell Cycle*, 8(10): 1526–1531
- Ma D K, Ponnusamy K, Song M R, Ming G L, Song H (2009b). Molecular genetic analysis of FGFR1 signalling reveals distinct roles of MAPK and PLCgamma1 activation for self-renewal of adult neural stem cells. *Mol Brain*, 2(1): 16
- Meissner A, Mikkelsen T S, Gu H, Wernig M, Hanna J, Sivachenko A, Zhang X, Bernstein B E, Nusbaum C, Jaffe D B, Gnirke A, Jaenisch R, Lander E S (2008). Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature*, 454(7205): 766–770
- Mellén M, Ayata P, Dewell S, Kriaucionis S, Heintz N (2012). MeCP2 binds to 5hmC enriched within active genes and accessible chromatin in the nervous system. *Cell*, 151(7): 1417–1430
- Miller C A, Sweatt J D (2007). Covalent modification of DNA regulates memory formation. *Neuron*, 53(6): 857–869
- Pastor W A, Aravind L, Rao A (2013). TETonic shift: biological roles of TET proteins in DNA demethylation and transcription. *Nat Rev Mol Cell Biol*, 14(6): 341–356

- Pastor W A, Pape U J, Huang Y, Henderson H R, Lister R, Ko M, McLoughlin E M, Brudno Y, Mahapatra S, Kapranov P, Tahiliani M, Daley G Q, Liu X S, Ecker J R, Milos P M, Agarwal S, Rao A (2011). Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. *Nature*, 473(7347): 394–397
- Rudenko A, Dawlaty M M, Seo J, Cheng A W, Meng J, Le T, Faull K F, Jaenisch R, Tsai L H (2013). Tet1 is critical for neuronal activity-regulated gene expression and memory extinction. *Neuron*, 79(6): 1109–1122
- Shukla S, Kavak E, Gregory M, Imashimizu M, Shutinoski B, Kashlev M, Oberdoerffer P, Sandberg R, Oberdoerffer S (2011). CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. *Nature*, 479(7371): 74–79
- Song C X, Szulwach K E, Dai Q, Fu Y, Mao S Q, Lin L, Street C, Li Y, Poidevin M, Wu H, Gao J, Liu P, Li L, Xu G L, Jin P, He C (2013). Genome-wide profiling of 5-formylcytosine reveals its roles in epigenetic priming. *Cell*, 153(3): 678–691
- Song C X, Szulwach K E, Fu Y, Dai Q, Yi C, Li X, Li Y, Chen C H, Zhang W, Jian X, Wang J, Zhang L, Looney T J, Zhang B, Godley L A, Hicks L M, Lahn B T, Jin P, He C (2011). Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine. *Nat Biotechnol*, 29(1): 68–72
- Spruijt C G, Gnerlich F, Smits A H, Pfaffeneder T, Jansen P W, Bauer C, Münzel M, Wagner M, Müller M, Khan F, Eberl H C, Mensinga A, Brinkman A B, Lephikova K, Müller U, Walter J, Boelens R, van Ingen H, Leonhardt H, Carell T, Vermeulen M (2013). Dynamic readers for 5-(hydroxy)methylcytosine and its oxidized derivatives. *Cell*, 152(5): 1146–1159
- Stadler M B, Murr R, Burger L, Ivanek R, Lienert F, Schöler A, van Nimwegen E, Wirbelauer C, Oakeley E J, Gaidatzis D, Tiwari V K, Schübeler D (2011a). DNA-binding factors shape the mouse methylome at distal regulatory regions. *Nature*, 480(7378): 490–495
- Stadler M B, Murr R, Burger L, Ivanek R, Lienert F, Schöler A, van Nimwegen E, Wirbelauer C, Oakeley E J, Gaidatzis D, Tiwari V K, Schübeler D (2011b). DNA-binding factors shape the mouse methylome at distal regulatory regions. *Nature*, 480(7378): 490–495
- Szulwach K E, Li X, Li Y, Song C X, Han J W, Kim S, Namburi S, Hermetz K, Kim J J, Rudd M K, Yoon Y S, Ren B, He C, Jin P (2011a). Integrating 5-hydroxymethylcytosine into the epigenomic landscape of human embryonic stem cells. *PLoS Genet*, 7(6): e1002154
- Szulwach K E, Li X, Li Y, Song C X, Wu H, Dai Q, Irier H, Upadhyay A K, Gearing M, Levey A I, Vasanthakumar A, Godley L A, Chang Q, Cheng X, He C, Jin P (2011b). 5-hmC-mediated epigenetic dynamics during postnatal neurodevelopment and aging. *Nat Neurosci*, 14(12): 1607–1616
- Tahiliani M, Koh K P, Shen Y, Pastor W A, Bandukwala H, Brudno Y, Agarwal S, Iyer L M, Liu D R, Aravind L, Rao A (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science*, 324(5929): 930–935
- Valinluck V, Sowers L C (2007). Endogenous cytosine damage products alter the site selectivity of human DNA maintenance methyltransferase DNMT1. *Cancer Res*, 67(3): 946–950
- Wang F, Yang Y, Lin X, Wang J Q, Wu Y S, Xie W, Wang D, Zhu S, Liao Y Q, Sun Q, Yang Y G, Luo H R, Guo C, Han C, Tang T S (2013). Genome-wide loss of 5-hmC is a novel epigenetic feature of Huntington's disease. *Hum Mol Genet*, 22(18): 3641–3653
- Wu S C, Zhang Y (2010). Active DNA demethylation: many roads lead to Rome. *Nat Rev Mol Cell Biol*, 11(9): 607–620
- Xu Y, Wu F, Tan L, Kong L, Xiong L, Deng J, Barbera A J, Zheng L, Zhang H, Huang S, Min J, Nicholson T, Chen T, Xu G, Shi Y, Zhang K, Shi Y G (2011). Genome-wide regulation of 5hmC, 5mC, and gene expression by Tet1 hydroxylase in mouse embryonic stem cells. *Mol Cell*, 42(4): 451–464
- Yamaguchi S, Shen L, Liu Y, Sandler D, Zhang Y (2013). Role of Tet1 in erasure of genomic imprinting. *Nature*, 504(7480): 460–464

- Yao B, Lin L, Street R C, Zalewski Z A, Galloway J N, Wu H, Nelson D L, Jin P (2013). Genome-wide alteration of 5-hydroxymethylcytosine in a mouse model of fragile X-associated tremor/ataxia syndrome. *Hum Mol Genet*, (Oct): 20 (Epub ahead of print)
- Yu M, Hon G C, Szulwach K E, Song C X, Zhang L, Kim A, Li X, Dai Q, Shen Y, Park B, Min J H, Jin P, Ren B, He C (2012). Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome. *Cell*, 149(6): 1368–1380
- Zhang H, Zhang X, Clark E, Mulcahey M, Huang S, Shi Y G (2010). TET1 is a DNA-binding protein that modulates DNA methylation and gene transcription via hydroxylation of 5-methylcytosine. *Cell Res*, 20(12): 1390–1393
- Zhang R R, Cui Q Y, Murai K, Lim Y C, Smith Z D, Jin S, Ye P, Rosa L, Lee Y K, Wu H P, Liu W, Xu Z M, Yang L, Ding Y Q, Tang F, Meissner A, Ding C, Shi Y, Xu G L (2013). Tet1 regulates adult hippocampal neurogenesis and cognition. *Cell Stem Cell*, 13(2): 237–245
- Zhu J K (2009). Active DNA demethylation mediated by DNA glycosylases. *Annu Rev Genet*, 43(1): 143–166