Published in final edited form as:

Epigenomics. 2014 December ; 6(6): 563–566. doi:10.2217/epi.14.48.

DNA cytosine methylation and hydroxymethylation at the borders

Melanie Ehrlich and

Tulane Cancer Center, New Orleans, LA 70112, USA and Center for Bioinformatics & Genomics, Tulane Health Sciences Center, New Orleans, LA 70112, USA and Program in Human Genetics, Tulane Medical School, New Orleans, LA 70112, USA

Kenneth C Ehrlich

Center for Bioinformatics & Genomics, Tulane Health Sciences Center, New Orleans, LA 70112, USA

Keywords

5-hydroxymethylcytosine; 5-methylcytosine; chromatin; enhancers; promoters; RNA splicing

“...development-linked changes in DNA methylation and hydroxymethylation within and at the borders of clusters of functionally related genes help to establish multigenic regions for coordinate up- or down-regulation of transcription.”

Previous studies of genomic 5-methylcytosine (5mC) and the recent explosion of research on 5-hydroxymethylcytosine (5hmC) indicate depletion or enrichment of these two modified C residues in mammalian promoters, exons, gene bodies, enhancers and gene clusters [1–4]. Not until 57 years after the discovery of 5hmC in T-even phages [5] was the existence of genomic 5hmC in mammals and its generation by TET-catalyzed oxidation of 5mC proven [6,7]. Genomic 5hmC can act as an intermediate in DNA demethylation by replication-dependent or replication-independent pathways [8] or it can be a relatively stable component of DNA [9,10]. It is implicated in regulating gene expression in mammals as is genomic 5mC [2]. For example, diverse tissues and cell types display a positive association of gene expression with enrichment for 5hmC in the body of genes that are strongly or moderately transcribed versus those that are weakly transcribed [9,11].

Although the involvement of genomic 5hmC and 5mC in regulating transcription is complex [12] and is likely to be highly dependent on DNA, chromatin and cell contexts, much evidence points to these modified bases helping to modulate or stabilize transcription states. This evidence includes cell-type-specific changes in 5hmC and 5mC at cis-acting transcription-regulatory elements, including among functionally related sets of genes

© 2014 Future Medicine Ltd

Author for correspondence: ehrlich@tulane.edu.

Financial & competing interests disclosure

The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

[3,13,14]. Moreover, although the global levels of 5hmC in brain are very much higher than in other tissues, the importance of 5hmC in non-neural tissues is indicated by the finding that compared to brain, skeletal muscle or heart have much higher levels of 5hmC in certain DNA sequences [14,15].

5-Hydroxymethylcytosine and 5mC can influence DNA function in distinct ways although it should be noted that 5mC is usually (but not always) more prevalent than 5hmC at a given modified C site in a population of molecules, and 5hmC can be opposite a 5mC residue [16]. A small bias toward more 5hmC in the sense strand and 5mC in the antisense strand in actively transcribed genes was found in brain [16]. Both 5mC and 5hmC are larger sterically than C, which could affect protein binding [17]. In addition, 5-methylation of C residues in DNA increases hydro-phobicity, which can alter interactions with transcription factors and also locally stabilizes the double helix [18]. In contrast, hydroxymethylation increases the hydrophilicity and destabilizes the double helix [17]. Therefore, conversion of 5mC to 5hmC residues should induce opposite effects on binding of some DNA-binding proteins [19]. Other DNA-binding proteins only distinguish unmodified C from modified C or bind equally well to sequences irrespective of C modification [19]. Modified C residues probably help to establish, maintain or change the chromatin borders of many promoters, enhancers, genes and gene clusters and may influence splicing at exon–intron junctions, as described below.

5hmC at the borders of promoters may stop the spread of repressive 5mC

Analyses of hydroxymethylomes from various human or mouse samples indicate that 5hmC is generally enriched about 0.5–2 kb upstream and downstream of the transcription start site (TSS) but is depleted closer to the TSS at moderately or highly transcribed genes [9,11,20]. Remarkably, poorly transcribed or untranscribed genes tend to have a peak of 5hmC at the TSS itself; therefore, in this sub-region, 5hmC enrichment may be downmodulatory. Repression-associated changes in 5mC levels in the vicinity of the TSS occur, although often with a very different distribution along the extended promoter region than for 5hmC [10,16,21]. Different C modification patterns at mammalian promoters are related to the existence of two distinct classes of promoters based upon CpG density. Most contain the TSS near the middle of a CpG-rich region of up to about 3 kb (CpG islands, CGIs) and are usually unmethylated (70% of total RefSeq genes). Generally, the rest have a low CpG content, like most of the whole genome (about 30% of total) [22]. 5-hmC was found to be enriched at the borders of CGI promoters in a human embryonic kidney cell line [23] and murine hematopoietic stem cells [24].

“The unexpected association of hypermethylated DMRs with the borders of certain active-promoter regions could be due to DNA methylation stopping the spread of polycomb silencing into the promoter...”

The association of 5hmC with the borders of promoters of actively transcribed genes could be due to the mostly unmethylated CGIs that overlap the majority of promoter regions [12]. Such CGI promoters need to be protected against DNA methylation spreading [25] from adjacent regions that have much higher 5mC contents [23]. We propose that, for some

genes, transcription facilitates spreading of methylation into a CGI promoter and thus could necessitate increased levels of protective 5hmC bordering active promoters.

5hmC & 5mC may shape the ends of promoters & promoter-enhancer regions

In addition to a protective function, 5hmC surrounding an active promoter might have a direct transcription-regulatory function. Active-promoter chromatin (histone H3K4 trimethylated and H3K27 acetylated) is often bordered by active- or weak-enhancer chromatin (H3K4 monomethylated with or without H3K27 acetylation) [26]. Because elevated 5hmC and low 5mC levels are associated with poised or active enhancers [9,13,16,27], we hypothesize that 5hmC at the borders of active promoters favors forming or maintaining enhancer chromatin there. Although a higher level of 5mC than 5hmC is usually observed at a given 5hmC-containing CpG in a steady-state population of cells [1,16], exceptional CpG sites, especially in non-cycling cell populations, can be found with average 5hmC contents higher than those of 5mC [14,16].

Many studies have shown that extensive promoter hypermethylation *in vivo* is associated with gene repression (e.g., [28]). In contrast, at the borders of some active-promoter chromatin regions [26], we found myoblast-associated differentially methylated regions (DMRs) that were hypermethylated in myoblasts and not present in nonexpressing cell types [15,28]. Although the methylation was monitored by reduced representation bisulfite sequencing, which does not distinguish between 5mC and 5hmC, the low levels of genomic 5hmC in myoblasts [4,14–15] make it likely that these DMRs contain mostly 5mC. The unexpected association of hypermethylated DMRs with the borders of certain active-promoter regions could be due to DNA methylation stopping the spread of polycomb silencing into the promoter [29]; counteracting the activity of a silencer element; or inhibiting repressive noncoding RNA transcription from the promoter flanks. In addition, hypermethylated DMRs that frame an active-promoter region may limit its length and thereby partly downmodulate expression (e.g., *TBX15* is less highly expressed in myoblasts, which have these border DMRs, than in osteoblasts, which do not and which have a longer region of active-promoter chromatin) [28]. Some promoter-adjacent DMRs might affect the choice of TSS and/or alternative splice sites near the 5' end of the nascent RNA by influencing chromatin structure at promoters.

5mC & 5hmC at the borders of exons may affect splicing, & at the 3' exon, transcription termination

Differential DNA methylation of exons may help regulate RNA splicing partly by modulating the binding of CCCTC-binding factor (CTCF), a chromatin-looping protein. Thereby, it may control the rate of transcription elongation [30]. A small peak of enrichment of hydroxymethylation at the 5' splice sites of brain DNA was described [16]. In embryonic stem cells (ESC), peaks of 5hmC enrichment were seen at both the 5' and 3' boundaries of exons, especially in actively transcribed genes [12]. *Tet2* knockdown decreased gene-body hydroxymethylation and resulted in aberrant frequencies of exclusion or inclusion of exons.

The enrichment of CpG in exons, which probably mostly reflects codon restraints on DNA sequence [22], may underlie the higher levels of 5mC and 5hmC in exons versus introns. This difference in CpG composition between exons and introns has apparently been exploited to help the splicing machinery recognize exon–intron boundaries and, with CpG modification, to affect the choice of alternative splice sites. Because last exons (including the 3′ untranslated region) are often enriched in 5hmC and 5mC, we hypothesize that these modified bases at the 3′ terminal exon of genes sometimes demarcate gene ends [9,27]. This might facilitate transcription termination, especially at alternative last exons.

5hmC & 5mC at the borders & within clusters of genes may help coordinate expression changes

Several large subclusters of *HOX* genes that are selectively active in myoblasts are embedded in an almost continuous domain of interspersed active-promoter and enhancer chromatin segments and also are surrounded by myoblast-hypermethylated DMRs at the borders of the promoter/enhancer (P/E) domain [4]. In comparing diverse cell types, both the hypermethylation and the P/E domains were positively associated with expression. Several sites within the border DMRs were determined by an enzymatic assay to have high 5mC levels and no 5hmC in myoblasts. In murine hematopoietic stem cells, long low-5mC DNA regions, including *HOX* gene clusters, were bordered by regions of high 5mC content that contained 5hmC as well [24]. In a study of a human embryonic carcinoma cell line, increased levels of 5hmC within half of the *HOXA* gene cluster were implicated in coordinate upregulation of expression of genes in this domain upon retinoic acid induction [3]. It is likely that development-linked changes in DNA methylation and hydroxymethylation within and at the borders of clusters of functionally related genes help to establish multigenic regions for coordinate up- or down-regulation of transcription.

Outlook

Many of the biological roles of genomic 5mC and 5hmC probably involve setting or maintaining chromatin boundaries that fine-tune gene expression by various mechanisms. Critical to understanding the functions of DNA hydroxymethylation and methylation is to profile the relative and absolute levels of 5mC and 5hmC residues at single-base resolution in many more cell and tissue types with regard to histone modifications, long-range as well as short-range chromatin interactions, expression and differentiation-, cell physiology-, disease- and aging-related epigenetic changes.

Acknowledgements

The authors thank their collaborators M Lacey, S Pradhan, J Terragni, G Zhang and S Chandra for essential insights into differential DNA methylation and hydroxymethylation.

This work was supported in part by a grant from the National Institutes of Health (NS04885).

Biography

Melanie Ehrlich

Kenneth C Ehrlich

References

1. Yu M, Hon GC, Szulwach KE, et al. Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome. *Cell*. 2012; 149(6):1368–1380. [PubMed: 22608086]
2. Lister R, Mukamel EA, Nery JR, et al. Global epigenomic reconfiguration during mammalian brain development. *Science*. 2013; 341(6146):1237905. [PubMed: 23828890]
3. Bocker MT, Tuorto F, Raddatz G, et al. Hydroxylation of 5-methylcytosine by TET2 maintains the active state of the mammalian HOXA cluster. *Nat. Commun*. 2012; 3:818. [PubMed: 22569366]
4. Tsumagari K, Baribault C, Terragni J, et al. DNA methylation and differentiation: HOX genes in muscle cells. *Epigen. Chromatin*. 2013; 6(1):25.
5. Wyatt GR, Cohen SS. A new pyrimidine base from bacteriophage nucleic acids. *Nature*. 1952; 170(4338):1072–1073. [PubMed: 13013321]
6. Kriaucionis S, Heintz N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science*. 2009; 324(5929):929–930. [PubMed: 19372393]
7. Tahiliani M, Koh KP, Shen Y, et al. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science*. 2009; 324(5929):930–935. [PubMed: 19372391]
8. Hackett JA, Sengupta R, Zylitz JJ, et al. Germline DNA demethylation dynamics and imprint erasure through 5-hydroxymethylcytosine. *Science*. 2013; 339(6118):448–452. [PubMed: 23223451]
9. Neri F, Incarnato D, Krepelova A, et al. Genome-wide analysis identifies a functional association of Tet1 and Polycomb repressive complex 2 in mouse embryonic stem cells. *Genome Biol*. 2013; 14(8):R91. [PubMed: 23987249]
10. Hahn MA, Qiu R, Wu X, et al. Dynamics of 5-hydroxymethylcytosine and chromatin marks in mammalian neurogenesis. *Cell Rep*. 2013; 3(2):291–300. [PubMed: 23403289]
11. Song CX, Szulwach KE, Fu Y, et al. Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine. *Nat. Biotechnol*. 2011; 29(1):68–72. [PubMed: 21151123]

12. Huang Y, Chavez L, Chang X, et al. Distinct roles of the methylcytosine oxidases Tet1 and Tet2 in mouse embryonic stem cells. *Proc. Natl Acad. Sci. USA*. 2014; 111(4):1361–1366. [PubMed: 24474761]
13. Serandour AA, Avner S, Oger F, et al. Dynamic hydroxymethylation of deoxyribonucleic acid marks differentiation-associated enhancers. *Nucleic Acids Res*. 2012; 40(17):8255–8265. [PubMed: 22730288]
14. Terragni J, Zhang G, Sun Z, et al. Notch signaling genes: myogenic DNA hypomethylation and 5-hydroxymethylcytosine. *Epigenetics*. 2014; 9(6):842–850. [PubMed: 24670287]
15. Tsumagari K, Baribault C, Terragni J, et al. Early de novo DNA methylation and prolonged demethylation in the muscle lineage. *Epigenetics*. 2013; 8(3):317–332. [PubMed: 23417056]
16. Wen L, Li X, Yan L, et al. Whole-genome analysis of 5-hydroxymethylcytosine and 5-methylcytosine at base resolution in the human brain. *Genome Biol*. 2014; 15(3):R49. [PubMed: 24594098]
17. Wanunu M, Cohen-Karni D, Johnson RR, et al. Discrimination of methylcytosine from hydroxymethylcytosine in DNA molecules. *J. Am. Chem. Soc*. 2011; 133(3):486–492. [PubMed: 21155562]
18. Ehrlich M, Ehrlich K, Mayo JA. Unusual properties of the DNA from *Xanthomonas* phage XP-12 in which 5-methylcytosine completely replaces cytosine. *Biochim. Biophys. Acta*. 1975; 395(2): 109–119. [PubMed: 1138935]
19. Spruijt CG, Gnerlich F, Smits AH, et al. Dynamic readers for 5-(hydroxy)methylcytosine and its oxidized derivatives. *Cell*. 2013; 152(5):1146–1159. [PubMed: 23434322]
20. Szulwach KE, Li X, Li Y, et al. Integrating 5-hydroxymethylcytosine into the epigenomic landscape of human embryonic stem cells. *PLoS Genet*. 2011; 7(6):e1002154. [PubMed: 21731508]
21. Kim M, Park YK, Kang TW, et al. Dynamic changes in DNA methylation and hydroxymethylation when hES cells undergo differentiation toward a neuronal lineage. *Hum. Mol. Genet*. 2013; 23(9): 657–667. [PubMed: 24087792]
22. Saxonov S, Berg P, Brutlag DL. A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proc. Natl Acad. Sci. USA*. 2006; 103(5): 1412–1417. [PubMed: 16432200]
23. Jin C, Lu Y, Jelinek J, et al. TET1 is a maintenance DNA demethylase that prevents methylation spreading in differentiated cells. *Nucleic Acids Res*. 2014; 42(11):6956–6971. [PubMed: 24875481]
24. Jeong M, Sun D, Luo M, et al. Large conserved domains of low DNA methylation maintained by Dnmt3a. *Nat. Genet*. 2014; 46(1):17–23. [PubMed: 24270360]
25. Orend G, Kuhlmann I, Doerfler W. Spreading of DNA methylation across integrated foreign (adenovirus type 12) genomes in mammalian cells. *J. Virol*. 1991; 65(8):4301–4308. [PubMed: 1712860]
26. Ernst J, Kheradpour P, Mikkelsen TS, et al. Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature*. 2011; 473(7345):43–49. [PubMed: 21441907]
27. Sun Z, Terragni J, Borgaro JG, et al. High-resolution enzymatic mapping of genomic 5-hydroxymethylcytosine in mouse embryonic stem cells. *Cell Rep*. 2013; 3(2):567–576. [PubMed: 23352666]
28. Chandra S, Baribault C, Lacey M, Ehrlich M. Myogenic differential methylation: diverse associations with chromatin structure. *Biology (Basel)*. 2014; 3(2):426–451. [PubMed: 24949935]
29. Wu H, Coskun V, Tao J, et al. Dnmt3a-dependent nonpromoter DNA methylation facilitates transcription of neurogenic genes. *Science*. 2010; 329(5990):444–448. [PubMed: 20651149]
30. Shukla S, Kavak E, Gregory M, et al. CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. *Nature*. 2011; 479(7371):74–79. [PubMed: 21964334]