CHAPTER 5

Beyond mCG: DNA Methylation in Noncanonical Sequence Context

E.A. Mukamel¹, R. Lister²,³
¹University of California San Diego, La Jolla, CA, United States; ²The University of Western Australia, Perth, WA, Australia; ³The Harry Perkins Institute of Medical Research, Perth, WA, Australia

INTRODUCTION: BEYOND CG METHYLATION

The genetic code based on the sequence of DNA nucleotides A, C, G, and T is thought to be a universal and invariant feature of life on Earth. Advances in genome sequencing have enabled comprehensive approaches to epigenome profiling that increasingly point to a diverse set of extensions to the genomic code in different cell types, even within the same species. Symbolic diversity is a familiar feature of human languages: French rarely uses the letters k or w, but it has access to a range of diacritical marks such as é that are not used in English. Just as we can easily recognize walked as English and marché as French based on the symbols in each word, recent epigenomic profiling efforts allow a new appreciation of how biological cells and tissues use distinct alphabets of epigenomic marks to regulate their specialized cellular functions. The new findings increasingly show that the mammalian epigenome is a symbolic system that encodes, stores, and transmits information through development and, potentially, across generations. Epigenomic marks, including DNA methylation and covalent modification of histone proteins, enhance the coding capacity of the genome by expanding the number of symbols available for representing gene regulatory information. To understand genomic information processing, we must make sense of the variety of symbolic elements used by particular cell types in each species.

Methylation of cytosine in genomic DNA is an essential epigenetic modification that primarily represses transcription and regulates other genomic processes across most, although not all, plant and animal species. The widespread presence and functional role of methylcytosine at CG dinucleotides has long been recognized (Suzuki & Bird, 2008); however, methylation at CA, CT, and CC positions (collectively called non–CG methylation, or mCH) in mammalian cells has also been established (Ramsahoye et al., 2000). By combining modern whole-genome shotgun DNA sequencing with sodium bisulfite conversion, in a technique called MethylC-seq (Lister & Ecker, 2009), the methylation status of more than 90% of genomic cytosines can now be experimentally determined at single-base resolution. Although MethylC-seq detects both methyl- and hydroxymethyl-cytosine (mC and hmC), techniques
such as Tet-assisted bisulfite sequencing (TAB-seq) profiling (Yu et al., 2012) enable the two modifications to be distinguished at base resolution throughout the genome.

This advance in methylome profiling technology first showed that although IMR90 human fetal lung fibroblast cells contain <0.02% of their methylcytosine in the non–CG context, human embryonic stem (ES) cells harbor nearly a quarter of their methylcytosines at non–CG positions (Lister et al., 2009). Subsequent surveys of a range of cells initially seemed to confirm this pattern, showing abundant non–CG methylation in pluripotent cells (Laurent et al., 2010; Lister et al., 2011), but little or no non–CG methylation across differentiated cell types including primary tissue samples and differentiated cells derived from pluripotent cells (Xie et al., 2013; Ziller et al., 2011). It was surprising, then, that MethylC-seq profiling of brain tissue from mouse (Xie et al., 2012) and human (Lister et al., 2013; Varley et al., 2013; Zeng et al., 2012) revealed a substantial amount of non–CG methylation. By purifying nuclei of neurons expressing the marker NeuN, cell type–specific profiling showed that non–CG methylation accounts for roughly half of all methylcytosine in adult frontal cortex neurons (Lister et al., 2013). This represents the most abundant level of non–CG methylation of any cell type yet observed. TAB-seq profiling in mouse frontal cortex and human ES cells showed that almost all of the non–CG methylation is in the form of mC and not hmC (Lister et al., 2013; Yu et al., 2012).

To appreciate the potential significance of non–CG methylation, it is important to consider the density of CG and non–CG positions in the human genome (Fig. 5.1). CG

Figure 5.1 The 16 dinucleotides in the human genome are unevenly distributed, with CG dinucleotides (green) greatly depleted compared with non–CG positions (blue). As a result, the average spacing between CG positions is ~100 bp, whereas non–CG positions occur every ~2.1 bp.
Beyond mCG: DNA Methylation in Noncanonical Sequence Context

dinucleotides occur at around 1 in 100 positions in the genome, far less than the 1 in 16 positions expected in a random sequence. CG sequences have been lost during evolution due to the higher rate of mutation of methylcytosine (Saxonov, Berg, & Brutlag, 2006). Around 11% of CG positions are concentrated in CG islands, a small genomic compartment associated with gene promoters and covering ∼1.4% of the genome. By contrast, non–CG positions occur at approximately 1 of every 2.1 bp. Thus, even a low rate of methylation at non–CG sites may have a substantial impact by virtue of the 50-fold increased density of these sites relative to CG positions.

VARIABLE NON–CG METHYLATION ACROSS CELL TYPES

Cellular differentiation requires extensive changes in the use of genomic information without changes to the underlying DNA sequence. Epigenome remodeling accompanies the regulation of gene transcription that defines and maintains the identity of specialized cell types. Although cells throughout the body share the same genetic sequence, apart from limited somatic mutations (Baillie et al., 2011; McConnell et al., 2013; Upton et al., 2015), it is the genome–wide pattern of DNA methylation and chromatin modifications that provide a cell type–specific fingerprint of each cell type’s use of the genetic information. Whereas the cell type–specific transcriptome is a snapshot of a cell’s current biological state, epigenomic information may reflect the past, current, and potential future dynamical regulation (Hon et al., 2013). For example, key neuronal transcription factors such as Npas4, which plays distinct cell type–specific roles in excitatory and inhibitory cortical neurons (Spiegel et al., 2014), are transcribed in response to neuronal depolarization or synaptic input. Cell type–specific epigenomic modifications likely regulate these activity–dependent responses. It is thus critical to understand differences in the epigenomic modifications present in different cell types to elucidate the role they play in cellular specialization.

The presence of dramatically elevated non–CG methylation levels in neurons suggests that this epigenetic mark could be present in one or more of the myriad other classes of human cells. As part of the NIH Roadmap Epigenomics project, MethylC-seq was used to profile the DNA methylomes of 18 human tissues from 36 samples (Schultz et al., 2015). Neurons were found to contain the highest rate of non–CG methylation, far more than any other cell type, with nearly 10% of all non–CG sites methylated. Glia and ES cells were also relatively highly methylated in the non–CG context, although they possess levels far lower than those observed in mature neurons (Lister et al., 2013; Schultz et al., 2015). Notably, the abundance of non–CG methylation in mammalian neurons changes dramatically during brain development (Lister et al., 2013). Negligible non–CG methylation is detected in the fetal mammalian brain, however a rapid accumulation of non–CG methylation occurs during postnatal brain development, specifically between 1 and 4 weeks after birth in mice, and in the first 2 years after birth in humans.
Although Schultz et al. (2015) also demonstrated that non–CG methylation is detectable throughout the genome across most of the tissues profiled, its level is much lower than in neural tissue and pluripotent cells, with the next highest enrichment observed in heart, muscle, and bladder (0.2–0.37%). The pattern of non–CG methylation across these tissues parallels that which occurs in neurons and glia: relatively lower methylation in gene bodies of actively transcribed genes and higher methylation in the bodies of repressed genes. Based on the profile of non–CG methylation, clusters of genes with muscle– or heart–specific non–CG methylation patterns could also be identified (Schultz et al., 2015). In particular, these clusters were enriched for genes with tissue–specific functional annotations.

Embryonic stem cells also harbor a substantial amount of non–CG methylation (Lister et al., 2009). This signature is a hallmark of early pluripotency that is shared by induced pluripotent stem cells (Lister et al., 2011; Ziller et al., 2011). Non–CG methylation is rapidly lost upon differentiation of ES cells to a range of lineages, including neural progenitor cells (Xie et al., 2013). However, the functional profile of non–CG methylation in pluripotent cells differs markedly from adult tissues. Whereas the density of non–CG methylation in gene bodies correlates with transcriptional repression in differentiated tissues and cell types, exactly the opposite pattern prevails in pluripotent cells (Lister et al., 2013, 2009; Ziller et al., 2011). Hypermethylation of gene bodies of actively transcribed genes persists in cell lineages derived through in vitro differentiation of ES cells, including mesendoderm, trophoblast, neural progenitor cells, and mesenchymal stem cells (Xie et al., 2013). Although these cell lineages have lower levels of non–CG methylation compared with ES cells, the genomic distribution of this mark is similar across each of these lineages, but highly distinct from adult tissues (Schultz et al., 2015).

The distinct regulation of non–CG methylation in differentiated cell types compared with pluripotent cells is also reflected in their different local sequence contexts. Non–CG methylation occurs at CA and CT positions, with many different flanking sequences, but it is all but undetectable at CC positions. Among all CA and CT positions, methylcytosine is most enriched at CAC sites in neurons and in all differentiated cell types and tissues. In contrast, CAG is highly enriched among methylated non–CG positions in pluripotent cells and lineages derived from them (Lister et al., 2013; Schultz et al., 2015; Varley et al., 2013). This distinct sequence context suggests, at least within these two broad classes of cells, that the methyltransferases responsible for depositing non–CG methylation are either distinct or modulated by different cofactors or posttranslational modifications. Indeed, the de novo methyltransferase DNMT3A is necessary for CA and CT methylation in mouse brain (Gabel et al., 2015; Guo et al., 2014), and, together with DNMT3B, it is responsible for establishing non–CG methylation patterns during early embryogenesis (Okano, Bell, Haber, & Li, 1999) and non–CG methylation in pluripotent cell types (Ziller et al., 2011). Pluripotent cells, but not brain tissue, express the cofactor DNMT3L, which lacks methyltransferase activity but is able to mediate the recruitment of the DNMT3A/B–DNMT3L complex to nucleosomes harboring...
Beyond mCG: DNA Methylation in Noncanonical Sequence Context

unmethylated histone H3 at lysine 4 (H3K4) (Ooi et al., 2007). DNMT3L is required for the establishment of imprinting during early development (Bourc’his, Xu, Lin, Bollman, & Bestor, 2001). Furthermore, DNMT3B, which is expressed at very low levels in the brain compared to DNMT3A, possesses a PWWP domain that mediates its recruitment to H3K36me3, a chromatin modification that is abundant in the body of actively transcribed genes (Baubec et al., 2015). Thus, these differences in the methylation machinery may explain the distinct sequence contexts of non–CG methylation in differentiated compared with pluripotent cells. However, it remains unclear why non–CG methylation shows the opposite association with transcription in these two cell classes.

In addition to non–CG methylation being an effective marker for genes that exhibit particular states of transcriptional activity in pluripotent cells and the brain, large regions of non–CG methylation enrichment or depletion have been identified as effective markers of various functional states in different cell types. In the brain, large genomic regions that are almost completely devoid of non–CG methylation, referred to as mCH deserts, exhibit highly inaccessible chromatin states that seem to resist de novo methylation, and are enriched for olfactory receptor gene and immunoglobulin gene clusters (Lister et al., 2013). In human induced pluripotent stem (iPS) cells, megabase-scale regions of the genome frequently fail to regain non–CG methylation during the reprogramming process and remain hypomethylated compared to ES cells in the non–CG context. These extensive differentially methylated regions (non–CG mega–DMRs) tend to occur at genomic regions which, in the differentiated cells from which the iPS cells were reprogrammed, exist in a partially methylated state in the CG context, referred to as partially methylated domains (PMDs) (Lister et al., 2009, 2011). PMDs are associated with late replicating genomic regions that localize to the nuclear lamina (Berman et al., 2012). They display low transcriptional activity and harbor repressive chromatin modifications such as H3K9me3 (Lister et al., 2009). Non–CG mega–DMRs in iPS cells frequently display high levels of H3K9me3, in contrast to the same regions in ES cells, and serve as highly effective epigenomic markers that allow discrimination of iPS cells from ES cells (Lister et al., 2011). Furthermore, as discussed in further detail below, female genes that escape X chromosome inactivation are marked by hypermethylation in the non–CG context, which allows effective identification of this unique regulatory state (Lister et al., 2013; Schultz et al., 2015). Thus, the pattern of non–CG methylation serves as a highly effective marker for cellular identity and genome regulatory states that can be assessed simply from a genomic DNA sample.

NON–CG METHYLATION AND BRAIN CELL DIVERSITY

Maintaining a diverse population of specialized cell types is important in most tissues, and it is especially critical for the function of brain circuits. Cognitive processes such as perception, memory, and motor control rely on the balanced interaction, via
coordinated gene expression, protein synthesis, and synaptic signaling, of a myriad of excitatory and inhibitory neuron types. Distinct neuronal cell types express unique DNA-binding transcription factors, have different histone modification profiles, and also differ substantially in their landscape of CG and non–CG DNA methylation. Among these cell type–specific characteristics, the pattern of non–CG methylation is one of the most distinctive molecular-genetic fingerprints of cell type–specific identity. In a comparison of three neuronal cell types (excitatory pyramidal cells, and inhibitory cells expressing parvalbumin or vasoactive intestinal peptide), almost half of all protein-coding genes were found to harbor neuron type–specific non–CG methylation (Mo et al., 2015). The correlation of gene body mCH levels between neuronal cell types ($r = 0.83–0.86$) was lower than the corresponding correlation of transcriptional levels by RNA-seq ($r = 0.95–0.96$), again suggesting that mCH captures essential aspects of neuronal epigenetic diversity. By contrast, both CG and non–CG methylation are precisely conserved in replicate experiments using tissue samples from different individuals, and non–CG methylation is conserved at homologous sequences between human and mouse brain neurons (Lister et al., 2013).

In cell types that harbor non–CG methylation, the mark appears across almost all genomic compartments, including exons, introns, and outside of gene bodies. Active regulatory elements located outside of gene promoters, as defined by a suite of epigenetic and chromatin marks, are depleted of non–CG methylation (Lister et al., 2009; Mo et al., 2015). DMRs, which are classically defined based on statistically significant differences in the level of CG methylation, are also marked by differential non–CG methylation as well as cell type–specific active histone marks (Lister et al., 2013; Mo et al., 2015). These active regulatory regions are also marked by open chromatin as assayed by in vitro transposition of native chromatin by Tn5 transposase (ATAC-seq) (Buenrostro, Giresi, Zaba, Chang, & Greenleaf, 2013). By computationally identifying ATAC-seq footprints lying over transcription factor sequence motifs, putative cell type–specific transcription factor binding sites could be profiled. Cell type–specific footprints corresponding to more than 100 transcription factors were depleted for both CG and non–CG methylation. Only two factors, CTCF and Zfp410, seemed to lack this correlation between transcription factor binding and methylation level (Mo et al., 2015), which suggests a tight relationship between methylation level and regulatory activity. However, this association remains correlational and the causal role of dynamic methylation at these sites remains to be determined (Schübeler, 2015).

In addition to gene expression and transcription factor binding, DNA methylation in both the CG and non–CG contexts can be modulated by the pattern of nucleosome occupancy (Chodavarapu et al., 2010; Lister et al., 2009). In neurons, autocorrelation analysis of non–CG methylation shows a weak modulation with periods of $\sim 10$ bp, corresponding to the DNA helix coil length, and $\sim 180$ bp, corresponding to nucleosome spacing (Lister et al., 2013). This phased modulation of DNA methylation is similar in
Beyond mCG: DNA Methylation in Noncanonical Sequence Context

plants and mammals (Chodavarapu et al., 2010), suggesting it is a universal feature of DNA methyltransferase activity. Integrating MethylC-seq data with nucleosome positions estimated from ATAC-seq chromatin accessibility data showed that both CG and non–CG methylation are depleted at the center of nucleosomes and enriched in the space between nucleosomes (Mo et al., 2015).

NON–CG METHYLATION IN X CHROMOSOME INACTIVATION

An intriguing case study of the highly specific, conserved distribution of non–CG methylation across the genome is the X chromosome. In mammals, female cells inactivate one of the two X chromosomes via transcription of the noncoding RNA Xist and subsequent epigenetic silencing of transcription from most of the chromosome. X chromosome inactivation is critical for maintaining a balanced dose of X-linked genes (Deng, Berletch, Nguyen, & Disteche, 2014). In neurons and glia, non–CG methylation is lower throughout the X chromosome in females than males, suggesting a differential distribution of this mark on the active versus inactive X (Lister et al., 2013). This female-specific depletion of non–CG methylation contrasts with the established role of CG methylation, which is enriched at the promoters of transcriptionally silenced genes on the inactive X (Sharp et al., 2011).

Although non–CG methylation is depleted across the X chromosome in female cells, it is locally enriched at genes that escape X inactivation and are consequently expressed from both alleles (Lister et al., 2013). This female non–CG methylation signature of escape from X inactivation appears not only in neurons and glia but also in a range of human tissues with much lower overall levels of mCH (Schultz et al., 2015). Non–CG methylation is thus a useful mark for identifying genes with differential epigenetic regulation on the inactive X chromosome. Indeed, analysis of female versus male human tissues identified 109 genes with significantly higher non–CG methylation in females; only nine of these genes showed the female hypermethylation in all tissues, and the remainder were hypermethylated in a tissue-specific manner in one or more tissues (Schultz et al., 2015). This observation also hints at a functional role for non–CG methylation, which could rein in the expression of escape genes that otherwise might experience a two-fold imbalance in transcription between female and male cells (Johnston et al., 2008).

POSSIBLE FUNCTIONS OF NON–CG METHYLATION

The conservation of cell type–specific non–CG methylation patterns between individuals and even across mammalian species, and the strong correlation between non–CG methylation and particular states of transcriptional activity, points to a functional role for this epigenetic mark. To begin to directly address the causal role of non–CG methylation
in regulating gene expression, Guo et al. (2014) used an in vitro reporter assay to show that CG methylation (mCG) by the bacterial methyltransferase M.Sssl strongly represses transcription; methylation of GC positions by using M.CviPI resulted a similar level of repression. Although GC positions include both CG and non–CG sequences, the majority of such sites are non–CG. This result thus supports a repressive role for non–CG methylation in the regulation of gene expression.

One likely mechanism by which non–CG methylation could act is through the recruitment of methylated DNA binding proteins that reconfigure the local chromatin environment and repress transcription. Methyl–CpG–binding protein 2 (MeCP2) binds methylated CG, CA, and CT sites (Chen et al., 2015; Guo et al., 2014), as well as hmC (Mellén, Ayata, Dewell, Kriaucionis, & Heintz, 2012). Electrophoretic mobility shift assays showed that oligonucleotides containing methylated CA positions compete for MeCP2 binding with equal affinity to oligonucleotides containing methylated CG. Hydroxymethylation at CG positions does not change the affinity for MeCP2 binding, although hmC at CA positions increases the affinity. It seems that hmC is limited to CG positions in vivo (Lister et al., 2013; Yu et al., 2012), suggesting that the major determinants of MeCP2 binding are methylation at CG and CA positions. hmC can function either as a stable base modification or as an intermediate in the TET-mediated active DNA demethylation pathway (Bachman et al., 2014; Guo, Su, Zhong, Ming, & Song, 2011). The finding that hmC does not seem to be present in the non–CG context to any significant level may in part explain why non–CG methylation is so abundant in postmitotic neurons, which at non–CG positions may lack both active (Tet/hmC-mediated) and passive (replication–dependent) DNA demethylation.

The accumulation of non–CG methylation in neurons extends throughout the stages of child and adolescent brain development, spanning the first two decades of life in the human prefrontal cortex (Lister et al., 2013). This period coincides with profound changes to the connectivity of postmitotic neurons, with the mature cortical circuitry emerging through the processes of synaptogenesis and subsequent synaptic pruning. Disruption of brain development during this time can have profound and lasting consequences. It is therefore of considerable interest to determine what role, if any, non–CG methylation plays in directing the proper development of cortical circuitry during early life.

Rett syndrome is a neurodevelopmental disorder that affects girls, with autism–like symptoms emerging around 6–18 months of age at the same time that non–CG methylation begins to accumulate in cortical neurons. Notably, Rett syndrome is caused by mutations in the X-linked gene MeCP2 (Chahrour & Zoghbi, 2007), suggesting a potential causal link between disease onset and disruption of the targeting of MeCP2 to methylated CA sites. Knowledge of the genetic cause of Rett syndrome has enabled development of mouse models of the disease (Ricceri, De Filippis, & Laviola, 2008), which can help define the gene networks that are up– or downregulated by loss of MeCP2 in brain cells.
Notably, genes that are upregulated in MeCP2 knockout animals tend to be long, spanning >100 kb (Gabel et al., 2015). This pattern is observed in a variety of cortical and cerebellar neuron types, including both glutamatergic (excitatory) and GABAergic (inhibitory) cells (Sugino et al., 2014). These long genes, which include many genes with key synaptic function like the calcium/calmodulin–dependent kinase Camk2d, are also enriched for gene body CA methylation (Gabel et al., 2015). Moreover, comparing long genes with short genes showing similar levels of CA methylation, MeCP2 seems to selectively repress expression of long genes (Gabel et al., 2015). These data suggest that MeCP2 may repress genes harboring CA methylation in a dose–dependent manner, exerting the strongest influence on long genes with a high total amount of methylcytosine. Furthermore, the developmental timing of the manifestation of Rett syndrome symptoms closely matches the period when non–CG methylation is accumulating in the developing human brain (Chen et al., 2015). This suggests that disruption of MeCP2–dependent recognition of the accumulating non–CG methylation may be linked to the emergence of Rett syndrome.

OUTLOOK AND FUTURE DIRECTIONS

Advances in DNA sequencing technology have enabled a comprehensive assessment of the DNA methylation landscape across the mammalian genome at single-base resolution. Combined with new technology for sorting cell types from a range of different tissues, DNA methylation profiling has revealed a surprising diversity in the sequence context of genomic methylcytosine, including the presence of abundant non–CG methylation. These rapid advances in characterizing the abundance, distribution, and dynamics of non–CG methylation in the genome of mammalian cells have raised a myriad of questions about the form and functional relevance of this distinct type of DNA methylation, which we outline below:

1. **Cell-type specificity**: Despite a number of studies examining non–CG methylation in the mammalian genome, our understanding of the variation in non–CG methylation within and between cell types is very limited. Although improvements have been achieved in resolving the non–CG methylation patterns of a limited number of cell types, most studies have analyzed highly heterogeneous tissues or cellular populations, and we have effectively no understanding of the consistency or variation of non–CG methylation between individual cells. Even in neurons, with the highest concentration of non–CG methylation so far observed, the majority of methylated non–CG sites are only methylated in ~20% of cells (Lister et al., 2013; Schultz et al., 2015). This distribution suggests substantial variability in the non–CG methylation pattern among individual cells. However, recent advances in base resolution DNA methylome profiling from single mammalian cells suggests that this level of analysis may be possible in the near future (Farlik et al., 2015; Smallwood et al., 2014).
2. **Establishment of non–CG methylation**: Although several lines of evidence indicate that DNMT3A establishes non–CG methylation in mature cell types, little is known about how it is recruited to specific regions of the genome to establish this mark. Furthermore, non–CG methylation patterns in a particular tissue type or cell type show very high positional conservation between individuals; however, the cellular factors or genome features that sculpt these precise patterns are unknown. Gaining a better understanding of the mechanistic basis of how the cell targets particular regions of the genome for establishment of, or protection from, non–CG methylation will shed further light on how the observed tissue– or cell–type patterns of non–CG methylation are controlled, enabling subsequent investigation of the potential effects of their disruption upon genome activity.

3. **Information content and function**: Although the presence of non–CG methylation has now been characterized genome–wide in a range of tissues and cell types (Lister et al., 2013, 2009, 2011; Mo et al., 2015; Xie et al., 2012), these studies have not provided direct evidence for a role of non–CG methylation in mediating transcriptional regulation. Recent studies have provided some evidence that disruption of the reading or writing of non–CG methylation in the brain results in altered transcription (Gabel et al., 2015; Guo et al., 2014); however, we have little or no understanding beyond this of the potential function of the modification. Does non–CG methylation only have an effect upon transcription when it is deposited throughout broad regions, such as gene bodies, or can methylation of single non–CG positions affect gene expression? Does deposition of non–CG methylation induce changes in local histone modifications or chromatin state, or is non–CG methylation establishment dependent upon other local histone modifications and chromatin conformation? In–depth investigations are required to explore these questions regarding the potential functionality of non–CG methylation. Investigation of the role of non–CG methylation by inhibition of DNMT3A/B is potentially challenging due to the simultaneous effects upon both CG and non–CG methylation. Alternatively, emerging tools to specifically target changes to DNA methylation at specific locations in the genome may in the future allow direct evaluation of the causal role of non–CG methylation in a range of genomic contexts (Maeder et al., 2013).

4. **Non–CG readers and writers**: An alternative route to investigating the function of non–CG methylation is disruption of the cellular factors that read and interpret the modification, as exemplified by recent studies of MeCP2 (Chen et al., 2015; Gabel et al., 2015). However, beyond MeCP2, the cellular factors that may bind to, read, and interpret the non–CG methylation are unknown. Identification of such factors through unbiased screens would provide new targets for the functional investigation of non–CG methylation and the consequences of disrupting its use in the cell.
5. **Non–CG methylation dynamics**: Although profiling through brain development has revealed the rapid accumulation of non–CG methylation during postnatal maturation of the neural circuitry, little is known about the timing of non–CG methylation accumulation through development in distinct neural cell types and during adult neurogenesis, or potential changes in non–CG methylation in response to neuronal activity and stimulation. Given evidence for the requirement for DNMT3A2 in the adult mammalian hippocampus for cognitive function, and the loss of cognitive capacity upon knockdown or natural age–related cognitive decline of DNMT3A2 abundance in the hippocampus (Oliveira, Hemstedt, & Bading, 2012), an intriguing possibility is that DNMT3A2 is required during neuronal differentiation or maturation during adult neurogenesis. Further investigation of the temporal dynamics of non–CG methylation during neural differentiation and development and in response to neuronal activity and learning is essential to further reveal its functional relevance.

6. **Disruption in neurological disorders**: Recent insights into the relationship between non–CG methylation and MeCP2, the effects of their disruption on neural transcription, and the potentially central role that non–CG methylation may play in Rett syndrome, underline the importance of gaining a comprehensive understanding of this modification and the effects of its disruption in human health and disease states. DNA methylation changes at CG positions have been linked with neurodegenerative disorders such as Alzheimer disease (De Jager et al., 2014) as well as with neuropsychiatric disorders such as schizophrenia with a stronger developmental role (Pidsley et al., 2014). Given the dramatic changes in non–CG methylation that take place during human brain development, future comprehensive studies of brain non–CG methylation, and its readers and writers, in neurological disorders of developmental origin need to be undertaken to explore the potential role of disruption of this unique epigenomic mark in human neurological disease.

**ACKNOWLEDGMENTS**

This work was supported by NIH/NINDS R00NS080911 (EAM), and an ARC Future Fellowship and Sylvia and Charles Viertel Foundation Senior Medical Research Fellowship (RL).

**REFERENCES**


Beyond mCG: DNA Methylation in Noncanonical Sequence Context


