Epigenetics

International Edition: DOI: 10.1002/anie.201504594
German Edition: DOI: 10.1002/ange.201504594

N^6-Methyladenine: A Potential Epigenetic Mark in Eukaryotic Genomes**

*Daniel Summerer**

DNA methylation · epigenetics · N^6-methyladenine · transcriptional regulation

DNA methylation increases the information content of the genome beyond the four canonical nucleobases A, G, C, and T. This epigenetic information can be inherited through cycles of cell division, but it can also be edited on demand by DNA methyltransferases (Dnmt) and demethylases, thereby offering unique possibilities for the dynamic regulation of DNA-associated processes.

DNA methylation is found in all domains of life, but with different occurrences of the individual types of methylated nucleobases and with different associated biological functions. For example, 5-methylcytosine (5mC) is the dominant epigenetic mark in mammals and many other eukaryotes, and it serves as a key regulatory element of gene expression with important roles in genome stability, development, and disease.[1] In bacteria, a dominant epigenetic DNA modification is N^6-methyladenine (6mA, Figure 1 A), which is part of restriction–modification (RM) systems that allow the host to selectively digest foreign DNA with 6mA-sensitive restriction endonucleases. However, 6mA has various additional functions that make it a genuine epigenetic mark in bacteria, with similar importance to 5mC in many eukaryotes. For example, it is involved in the regulation of bacterial replication, mismatch repair, and transcriptional regulation, and it has been shown to be essential in several species.[2] Interestingly, 6mA has previously also been found in the genomes of certain eukaryotic organisms, such as protists, fungi, and plants, in some cases without the presence of a restriction endonuclease counterpart, thus suggesting biological functions beyond mere defense against foreign DNA.[3] However, the levels of 6mA seemed to usually be extremely low and evidence for the involvement of 6mA in epigenetic regulatory processes in eukaryotes has remained elusive.

Three studies have now shown the presence of 6mA—and clues to potential epigenetic functions—in the phylogenetically distinct eukaryotes *Chlamydomonas reinhardtii*,[4] *Drosophila melanogaster*,[5] *and Caenorhabditis elegans*[6] (Figure 1 B).

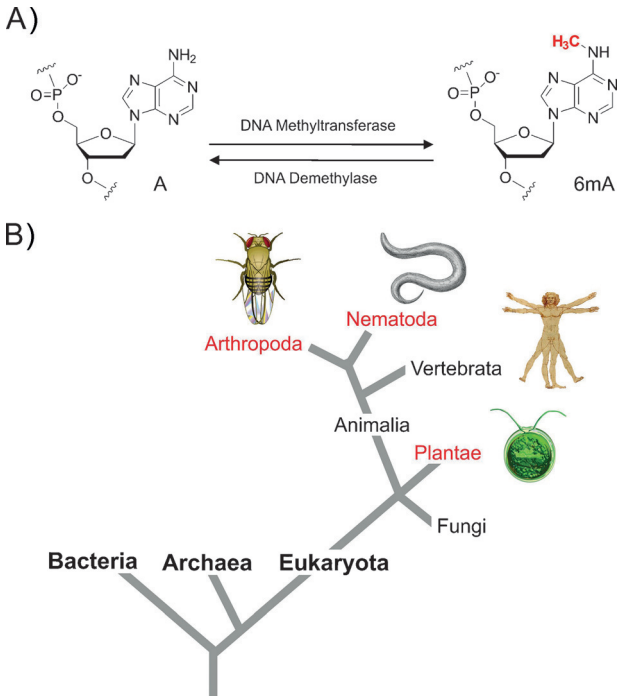

Figure 1. N^6-Methylation of adenine. A) Chemical structures of the adenosine and N^6-methyladenosine nucleotides in DNA, and their interconversion by DNA methyltransferases (Dnmt) and demethylases. B) A simplified phylogenetic tree of life. Organisms in which genomic 6mA has been discovered and/or characterized in the highlighted studies are shown as cartoons, with the respective phylum (or kingdom) marked in red. Humans are shown as a representative of vertebrates that have not been studied yet.

It was previously known that both 6mA and 5mC exist in the genome of the unicellular green alga *C. reinhardtii*, with significant levels (ca. 0.3–0.5 % 6mA/A) of the former.[7] However, Fu et al. used antibody-based methods for affinity enrichment of 6mA-containing DNA fragments in combination with high throughput sequencing approaches[8] to generate the first distribution map of 6mA in the *C. reinhardtii* genome.[4] This revealed the presence of 6mA in more than 14000 genes, corresponding to approximately 84 % of all genes in this organism. 6mA was predominantly found in ApT dinucleotides and showed a preference for certain frequently

[*] Dr. D. Summerer
Department of Chemistry and Konstanz Research School Chemical Biology, University of Konstanz
Universitätsstrasse 10, 78457 Konstanz (Germany)
E-mail: daniel.summerer@uni-konstanz.de

[**] We acknowledge support by the Zukunftskolleg of the University of Konstanz and the Konstanz Research School Chemical Biology. This work was supported by a grant from the Deutsche Forschungsgemeinschaft (SU-726/5-1 in SPP1784).

occurring sequence motifs that were evenly distributed across the genome. However, despite this even distribution, 6mA was highly enriched at gene promoters, with a bimodal distribution centered at the transcription start sites (TSS; by contrast, 5mC mapping revealed enrichment in gene bodies and not in TSS). High-resolution mapping showed that 6mA occurrence at TSS exhibits a periodic pattern with maxima that are complementary to nucleosome binding sites. This indicates that 6mA at TSS is present predominantly in linker DNA regions between the nucleosomes and may thus potentially be involved in nucleosome positioning. Finally, RNA-Seq analysis correlated the presence of 6mA at TSS with high transcriptional activity, thus suggesting a role in the regulation of gene expression.

The *D. melanogaster* genome contains only minute levels of 5mC[9] and evidence for the existence of 6mA was completely missing. In the second study, Zhang et al. analyzed fully digested genomic DNA from *D. melanogaster* adult tissues and embryos of different developmental stages by ultrahigh-performance liquid chromatography–triple-quadrupole mass spectrometry with multiple reaction monitoring (UHPLC-MRM-MS/MS), a method that enables the extremely sensitive detection of nucleotide modifications.[5] They found low overall genomic levels of approximately 0.001 % 6mA/A in adult tissues and late-stage embryos. However, in very early-stage embryos, levels were 70-fold higher. This confirmed that 6mA can exist in the *D. melanogaster* genome at significant levels and that it is dynamically regulated during embryonic development. To test whether the observed dynamics were a result of enzymatic (active) demethylation, the authors then employed an in vitro assay to quantify 6mA-to-A conversion in enzymatically methylated DNA by embryonic nuclear extracts. Indeed, dose-dependent demethylation activity was observed, with higher activities for extracts from later developmental stages. In general, nucleobase demethylation can occur oxidatively through catalysis by Fe^{II}-α-ketoglutarate-dependent dioxygenases, such as AlkB[10] or the ten-eleven-translocation (TET) dioxygenases.[11] The authors identified the putative dioxygenase *Drosophila* DNA 6mA demethylase (DMAD), which exhibits a similar domain structure to the mammalian TET dioxygenases, in the *D. melanogaster* genome. DMAD was expressed at levels correlating to the levels of 6mA during development, and alteration of its expression led to increased lethality in embryos. Moreover, knockout studies revealed DMAD to be a selective demethylase for 6mA in DNA, without significant activity for mC, hmC, or 6mA in RNA, for example. 6mA was also detectable in different cell types, including germarium cells of the ovary, with decreasing levels in late egg chambers. Finally, 6mA-targeted methylated DNA immunoprecipitation sequencing (MeDIP-Seq)[8] and transcription profiling revealed enrichment of 6mA in transposon bodies, with associated increased transcriptional activity, thus suggesting a role for 6mA in the regulation of gene expression in *D. melanogaster* as well.

In *C. elegans*, both 5mC and enzymes similar to known cytosine DNA methyltransferases (Dnmt) are missing, which led to the view that its genome is devoid of DNA methylation.[12] Nevertheless, a model for epigenetic inheritance in *C. elegans* existed that is based on mutation of the *spr*-5 gene, which encodes a histone 3 lysine 4 dimethyl (H3K4me2, an activating histone mark) demethylase.[13] After successive generations, *spr-5* mutant worms exhibit progressively increasing H3K4me2 levels and an associated increase in infertility (despite possessing virtually identical genomes), thus indicating inheritance of epigenetic information. In the third study, Greer et al. analyzed genomic DNA from wild-type and late-generation *spr-5* mutant worms and were able to detect 6mA but not 5mC or 5hmC.[6] Levels of 6mA were higher in the *spr-5* mutant worms (up to 0.15 % 6mA/A) and increased from generation to generation. 6mA was also detectable in almost all cell types. The authors further performed 6mA-targeted MeDIP-Seq[8] and single-molecule real-time sequencing (SMRT)[14] using worms of mixed developmental stages and found 6mA at approximately 0.7 % of A sites, widely distributed across the genome, with little dependence on genomic features but with a preference for specific sequence motifs. The authors then identified the 6mA demethylase NMAD-1 (one of five AlkB family dealkylating enzymes present in the *C. elegans* genome), which is able to demethylate 6mA in DNA in vitro and led to an accelerated loss of fertility upon knockout in vivo. Moreover, the corresponding Dnmt DAMT-1 was identified and knockout of the corresponding gene led to decreased 6mA levels in *C. elegans* and to decreased progressive infertility in the *spr*-5 mutant (but not in the wild type). Although the specific role of 6mA in this process is not clear yet, these data suggest cross-talk between H3K4me2 and 6mA, both of which act as epigenetic regulators.

Taken together, the three studies show that 6mA exists at significant levels in the genomes of several phylogenetically distinct eukaryotes and that it can exhibit genomic distributions and correlations with local transcriptional activities that are suggestive of roles in the regulation of gene expression. 6mA can further vary over different developmental stages, and new Dnmt and/or demethylases for its introduction and removal have been identified that provide a basis for its dynamic regulation. Since similar enzymes have been conserved during evolution, these studies raise the central question of how widespread 6mA is in eukaryotes, i.e., whether it might be present in vertebrate or even mammalian genomes, including humans. For example, as Greer et al. point out, methyl-transferase-like 4 (METTL4) in humans is similar to DAMT-1 in *C. elegans*, and Zhang et al. showed that DMAD exhibits similarity to human TET dioxygenases and is able to demethylate 6mA.[15] Extended studies on a larger range of higher eukaryotes will reveal the true significance of these exciting findings. The studies finally highlight the importance of the development of advanced, highly sensitive analytical methods such as UHPLC-MRM-MS/MS for studying nucleobase modifications. These will continue to provide new and unexpected insights into the occurrence of epigenetic nucleobase modifications in eukaryotes and their functional significance in the regulation of key biological processes.

How to cite: *Angew. Chem. Int. Ed.* **2015**, *54*, 10714–10716
Angew. Chem. **2015**, *127*, 10862–10864

[1] J. A. Law, S. E. Jacobsen, *Nat. Rev. Genet.* **2010**, *11*, 204 – 220.
[2] D. Wion, J. Casadesus, *Nat. Rev. Microbiol.* **2006**, *4*, 183 – 192.
[3] D. Ratel, J. L. Ravanat, F. Berger, D. Wion, *Bioessays* **2006**, *28*, 309 – 315.
[4] Y. Fu, G. Z. Luo, K. Chen, X. Deng, M. Yu, D. Han, Z. Hao, J. Liu, X. Lu, L. C. Dore, X. Weng, Q. Ji, L. Mets, C. He, *Cell* **2015**, *161*, 879 – 892.
[5] G. Zhang, H. Huang, D. Liu, Y. Cheng, X. Liu, W. Zhang, R. Yin, D. Zhang, P. Zhang, J. Liu, C. Li, B. Liu, Y. Luo, Y. Zhu, N. Zhang, S. He, C. He, H. Wang, D. Chen, *Cell* **2015**, *161*, 893 – 906.
[6] E. L. Greer, M. A. Blanco, L. Gu, E. Sendinc, J. Liu, D. Aristizabal-Corrales, C. H. Hsu, L. Aravind, C. He, Y. Shi, *Cell* **2015**, *161*, 868 – 878.
[7] S. Hattman, C. Kenny, L. Berger, K. Pratt, *J. Bacteriol.* **1978**, *135*, 1156 – 1157.
[8] a) M. J. Booth, E. A. Raiber, S. Balasubramanian, *Chem. Rev.* **2015**, 115, 2240 – 2254; b) N. Plongthongkum, D. H. Diep, K. Zhang, *Nat. Rev. Genet.* **2014**, *15*, 647 – 661.
[9] F. Lyko, R. Maleszka, *Trends Genet.* **2011**, *27*, 127 – 131.
[10] a) S. C. Trewick, T. F. Henshaw, R. P. Hausinger, T. Lindahl, B. Sedgwick, *Nature* **2002**, *419*, 174 – 178; b) P. O. Falnes, R. F. Johansen, E. Seeberg, *Nature* **2002**, *419*, 178 – 182.
[11] M. Tahiliani, K. P. Koh, Y. Shen, W. A. Pastor, H. Bandukwala, Y. Brudno, S. Agarwal, L. M. Iyer, D. R. Liu, L. Aravind, A. Rao, *Science* **2009**, *324*, 930 – 935.
[12] D. Wenzel, F. Palladino, M. Jedrusik-Bode, *Genesis* **2011**, *49*, 647 – 661.
[13] D. J. Katz, T. M. Edwards, V. Reinke, W. G. Kelly, *Cell* **2009**, *137*, 308 – 320.
[14] B. A. Flusberg, D. R. Webster, J. H. Lee, K. J. Travers, E. C. Olivares, T. A. Clark, J. Korlach, S. W. Turner, *Nat. Methods* **2010**, *7*, 461 – 465.
[15] L. M. Iyer, S. Abhiman, L. Aravind, *Prog. Mol. Biol. Transl. Sci.* **2011**, *101*, 25 – 104.

Received: May 21, 2015
Published online: July 17, 2015

