Genome **Biology**

RESEARCH HIGHLIGHT

Birds do it, bees do it, worms and ciliates do it too: DNA methylation from unexpected corners of the tree of life

Soojin Yi[1]*

See related research articles http://genomebiology.com/2012/13/10/R99 and http://genomebiology.com/2012/13/10/R100

Abstract
Studies describing intricate patterns of DNA methylation in nematode and ciliate are controversial due to the uncertainty of genomic evolutionary conservation of DNA methylation enzymes.

The emerging evolutionary framework of DNA methylation

DNA methylation refers to a chemical modification of genomic DNA by the addition of a methyl (-CH_3) group to specific nucleotide bases. The most common form of DNA methylation is cytosine methylation, occurring predominantly in CpGs in animal genomes, and enriched in CpHs and CpHpHs (where H stands for bases other than G) in plant and fungal genomes [1].

DNA methylation is phylogenetically highly variable. Notably, most of the commonly used non-mammalian model organisms (yeast, fruit fly, and worm but not *Arabidopsis*) lack genomic DNA methylation. Genome projects that have emerged in the past few years, however, have repeatedly demonstrated that DNA methylation is far more widespread than one would expect from the lack of DNA methylation in model organisms. These studies converge to establish DNA methylation as an evolutionarily ancient regulatory mechanism and indicate that the loss of DNA methylation is derived and is generally a lineage-restricted evolutionary event.

The distribution of DNA methylation enzymes across the tree of life provides a complementary view. Methylations of DNA templates are achieved by two distinct classes of DNA methylation enzymes, *dnmt1* and *dnmt3*.

*Correspondence: soojinyi@gatech.edu
School of Biology, Georgia Institute of Technology, 310 Ferst Drive, Atlanta, GA 30332, USA

These enzymes are widely distributed in eukaryotic genomes, yet are frequently gained or lost from genomes as a result of gene duplications and losses in specific lineages [2]. Species exhibiting functional DNA methylation generally encode a 'complete' set of both *dnmt1* and *dnmt3* in their genomes. Species lacking DNA methylation, such as the model nematode *Caenorhabditis elegans*, seem to have lost DNA methylation enzymes from their genomes. Furthermore, functional studies have begun to elucidate the regulatory significance of DNA methylation in processes such as alternative splicing, gene expression, and phenotypic plasticity in non-model organisms [3,4].

Two articles in this issue of *Genome Biology* [5,6] further our understanding of the phylogenetic distribution and functional roles of DNA methylation. At the same time, they raise many questions. These studies describe DNA methylation from organisms that were traditionally considered to lack functional DNA methylation. Uncertainty regarding the evolutionary conservation of DNA methylation enzymes in the genomes of the study organisms makes these reports rather controversial.

Nematode DNA methylation and the history of DNA methylation enzymes

Gao *et al.* [5] report evidence of functional DNA methylation in the nematode *Trichinella spiralis*. This species is a parasitic worm that diverged early in the evolution of nematodes. Unlike the free-living *C. elegans*, *T. spiralis* spends most of its life cycle within mammalian hosts, causing trichinellosis, which is a worldwide zoonotic disease. The life cycle of *T. spiralis* is roughly divided into three stages. The first stage is muscle larvae (MLs), which rapidly develop to sexual adults. After sexual adults mate, newborn larvae (NBLs) are produced. These NBLs then localize to various muscular areas via the bloodstream, and form a new generation of MLs.

Gao *et al.* [5] examined the protein repertoire encoded by the *T. spiralis* genome and found that it contains a full set of DNA methylation enzymes. Specifically, they

identified genes that seem to be homologous to *dnmt1* and *dnmt3*. They then probed for the presence of DNA methylation directly by several methods, including liquid chromatography/tandem mass spectrometry, targeted bisulfite PCR, methylated DNA immunoprecipitation (MeDIP) followed by qPCR, and whole genome sequencing of bisulfite-converted genomic DNA.

These analyses reveal a complex picture of DNA methylation. The level of DNA methylation in *T. spiralis* varies dramatically between life stages. The authors estimate that the adult and ML genomes show low levels of DNA methylation, in which approximately 1.5% of all cytosines are methylated, roughly similar to the level of DNA methylation observed in hymenopteran insects. Surprisingly, however, DNA methylation was almost undetectable in the NBL genome. Gao *et al.* [5] further compared differential genomic methylation between life stages with differential gene expression (using RNA sequencing methods). They uncovered a generally negative correlation between gene expression and DNA methylation of upstream regions. Moreover, some of the genes well known to be involved in the parasitism of *T. spiralis* showed differential methylation and differential expression between stages.

These findings indicate that DNA methylation is likely to have been present in the ancestral nematode lineages, and to be potentially involved in regulation of the life cycle in *T. spiralis*. Interestingly, in the *T. spiralis* genome, transposable elements and upstream regions seem to be more heavily methylated than the genomic background. This finding is noteworthy given that methylation of transposable elements and upstream regulatory regions was thought to be mostly absent from invertebrates and was thus viewed as a unique feature of vertebrate genomes [1,7]. The finding from *T. spiralis* demands a re-evaluation of the evolutionary history of genomic DNA methylation: methylation of transposable elements and regulatory elements may have evolved multiple times in different lineages.

Gao *et al.* [5] make an even more provocative claim: they suggest that *dnmt1* homologs are widely distributed in nematode lineages, including in *C. elegans*. They classify the *C. elegans* gene *CE24669* as a potential homolog of *dnmt1*. The exact function of the protein encoded by this gene is currently unknown, although its sequence has been long recognized to include a Cys-rich region closely related to that of Dnmt1 [8]. However, given the lack of DNA methylation in the *C. elegans* genome, it was thought that this protein is likely to function in a methylation-independent process. The unexpected discovery of functional DNA methylation in the genome of *T. spiralis* raises the possibility that the *C. elegans* gene *CE24669* may indeed represent a remnant of ancestral *dnmt* homolog in the evolutionary past (although it is likely to have evolved a new function since the loss of genomic DNA methylation in *C. elegans*). Functional studies of *dnmt1* and *dnmt3* in *T. spiralis* and other nematodes may help us to understand whether *CE24669* and other *dnmt*-like genes in nematodes indicate the presence of functional DNA methylation in early nematode evolution.

Genome gymnastics and DNA methylation in ciliates

Another article in this issue of *Genome Biology* widens the focus on DNA methylation even more dramatically. Bracht *et al.* [6] report extensive cytosine methylation in the genome of the ciliate *Oxytricha trifallax*.

O. trifallax is notable for its complex developmental genomic rearrangements, dubbed 'genome gymnastics'. Each *O. trifallax* contains two kinds of nuclei. These two nuclei house distinctive types of genome that differ tremendously from each other in their sizes, contents and functions. The small, germline micronucleus (MIC) contains a copy of the parental genome that remains transcriptionally silent in vegetative cells. The large, somatic macronucleus (MAC) encodes a fraction (about 5%) of the MIC genome, amplified in numerous fragments called 'nanochromosomes'. These nanochromosomes often encode only a single gene, exist in thousands of amplified copies within a MAC, and are actively transcribed. Thus, MACs provide an active transcriptional facility in the cells, whereas MICs remain as a genomic reservoir. During sexual reproduction (by conjugation), MICs from mating cells fuse and form a diploid MIC, which then gives rise to the new MAC and MIC. This is accompanied by a precisely choreographed genomic elimination process to restore the genomic contents of *O. trifallax* cells. First, the existing, 'old' MAC degrades, allowing the newly formed MACs to take over the transcriptional processes of the daughter cells. Second, the development of a new MAC from the MIC elicits the degradation of about 95% of the parental MIC genome, mostly composed of repetitive elements. The remainder of the parental genome is then extensively rearranged to form numerous nanochromosomes of the MAC.

There is already some evidence that genomic rearrangement processes in ciliates, including *O. trifallax,* are guided by epigenetic mechanisms [9]. Bracht *et al.* [6] now adds DNA methylation to the repertoire of epigenetic processes that are intimately associated with genome rearrangements in this species. Previous studies on cytosine methylation of ciliates had conflicting results. Bracht *et al.* [6] demonstrates that there is, in fact, cytosine methylation in *O. trifallax* cells, although it is extremely transient. Antibody staining and mass spectrometry both detect methylcytosines, most prominently in cells 36-40 hours after conjugation, but not in vegetative cells. This specificity potentially explains why some

previous studies, which typically analyzed vegetative cells, failed to detect cytosine methylation. It also hints at the functional significance of cytosine methylation in the genome elimination process, because the window of time 36-40 hours after conjugation marks active degradation of the 'old' MACs. Indeed, deep sequencing following methyl-DNA immunoprecipitation (MeDIP-seq) reveals specific subsets of parental MACs that house methylcytosines. In addition, some repetitive sequences of MIC origin that are destined for degradation, including transposons and satellite DNA, also appear in the methylated portions. These observations are further validated by sequencing of bisulfite-converted genomic DNA. Synthesizing these findings, Bracht *et al.* [6] propose that the transient cytosine methylation in *O. trifallax* marks specific genomic regions that are destined for degradation.

The details of *O. trifallax* DNA methylation that Bracht *et al.* [6] describe are extraordinary in many aspects. First, they describe a mode of DNA methylation that stands in diametric contrast to the canonical cytosine methylation typically localized to CpG, CpH, and CpHpH motifs. Analyses of bisulfite-converted sequences indicate a 'pan-cytosine' methylation in which the majority of cytosines in specific contigs are methylated, irrespective of their genomic sequence context. Second, they show that in addition to methylcytosines, hydroxymethylcytosines are also present in *O. trifallax* genomes. Hydroxymethylcytosines are close chemical relatives of methylcytosines, generated by oxidation of methylcytosines by a dedicated family of enzymes. In mammalian model systems, methylcytosines and hydroxymethylcytosines tend to concentrate in different genomic regions, seemingly marking them for distinctive developmental pathways [10]. Bracht *et al.* [6] find that genomic locations of hydroxymethylcytosines and methylcytosines largely overlap in *O. trifallax* genomes, although hydroxymethylcytosines are present long after methylcytosines disappear (hydroxymethylcytosines were observed up to 64 hours after conjugation). Future studies will be needed to elucidate the evolutionary distribution and significance of DNA hydroxymethylation.

The most provocative aspect of DNA methylation in *O. trifallax* is the lack of canonical DNA methyltransferase homologs in its genome. The authors found no trace of *dnmt* homologs in either macronuclear or micronuclear *O. trifallax* genomes. However, DNA methyltransferase inhibitors (azacitidine and decitabine) induce partial demethylation, indicating similar molecular components of DNA methylation pathways between *O. trifallax* and those of other organisms. Moreover, homologs of TET families, which perform hydroxymethylation in other species, are present in the *O. trifallax* genome. Whether cytosine methylation in the *O. trifallax* genome is mediated by a new, undefined DNA methylation enzyme or by canonical *dnmt* homologs residing in the small portion of the genome that has yet to be sequenced needs to be resolved. The answer to this question will have tremendous implications for our understanding of the evolution of DNA methylation.

The presence of DNA methylation in nematodes and ciliates dramatically expands the roster of organisms harboring functional DNA methylation system. Moreover, these studies [5,6] indicate that crucial regulatory pathways, such as genome elimination and a parasitic life cycle, are mediated by genomic DNA methylation. Thus, DNA methylation is functionally implicated in a wide set of lineage-specific regulatory pathways. Many details, including the evolutionary conservation of DNA methylation machinery in these organisms, need to be resolved in future studies. Moreover, the functional significance of DNA methylation in diverse organisms deepens a fundamental mystery regarding the evolutionary patterns of DNA methylation: if it is so important, why is DNA methylation lost so frequently in many lineages with seemingly little consequence?

Competing interests
The author declares that she has no competing interests.

Published: 17 October 2012

References
1. Suzuki MM, Bird A: **DNA methylation landscapes: provocative insights from epigenomics.** *Nat Rev Genet* 2008, **9:**465-476.
2. Ponger L, Li W-H: **Evolutionary diversification of DNA methytransferases in eukaryotic genomes.** *Mol Biol Evol* 2005, **22:**1119-1128.
3. Herb BR, Wolschin F, Hansen KD, Aryee MJ, Langmead B, Irizarry R, Amdam GV, Feinberg AP: **Reversible switching between epigenetic states in honeybee behavioral subcastes.** *Nat Neurosci* 2012, **15:**1371-1373.
4. Park J, Peng Z, Zeng J, Elango N, Park T, Wheeler D, Werren JH, Yi SV: **Comparative analyses of DNA methylation and sequence evolution using *Nasonia* genomes.** *Mol Biol Evol* 2011, **28:**3345-3354.
5. Gao F, Liu XS, Wu X-P, Wang X-L, Gong D, Lu H, Song Y, Wang J, Du J, Liu S, *et al.*: **Differential methylation in discrete developmental stages of the parasitic nematode *Trichinella spiralis*.** *Genome Biol* 2012, **13:**R100.
6. Bracht JR, Perlman DH, Landweber LF: **Cytosine methylation and hydroxymethylation mark DNA for elimination in *Oxytricha trifallax*.** *Genome Biol* 2012, **13:**R99.
7. Elango N, Yi SV: **DNA methylation and structural and functional bimodality of vertebrate promoters.** *Mol Biol Evol* 2008, **25:**1602-1608.
8. Bestor TH: **The DNA methyltransferases of mammals.** *Hum Mol Genet* 2000, **9:**2395-2402.
9. Chalker DL, Yao M-C: **DNA Elimination in ciliates: transposon domestication and genome surveillance.** *Annu Rev Genet* 2011, **45:**227-246.
10. Branco MR, Ficz G, Reik W: **Uncovering the role of 5-hydroxymethylcytosine in the epigenome.** *Nat Rev Genet* 2012, **13:**7-13.