INVITED REVIEW

DNA methylation in insects: on the brink of the epigenomic era

K. M. Glastad*, B. G. Hunt*, S. V. Yi and M. A. D. Goodisman
School of Biology, Georgia Institute of Technology, Atlanta, GA, USA

Abstract

DNA methylation plays an important role in gene regulation in animals. However, the evolution and function of DNA methylation has only recently emerged as the subject of widespread study in insects. In this review we profile the known distribution of DNA methylation systems across insect taxa and synthesize functional inferences from studies of DNA methylation in insects and vertebrates. Unlike vertebrate genomes, which tend to be globally methylated, DNA methylation is primarily targeted to genes in insects. Nevertheless, mounting evidence suggests that a specialized role exists for genic methylation in the regulation of transcription, and possibly mRNA splicing, in both insects and mammals. Investigations in several insect taxa further reveal that DNA methylation is preferentially targeted to ubiquitously expressed genes and may play a key role in the regulation of phenotypic plasticity. We suggest that insects are particularly amenable to advancing our understanding of the biological functions of DNA methylation, because insects are evolutionarily diverse, display several lineage-specific losses of DNA methylation and possess tractable patterns of DNA methylation in moderately sized genomes.

Keywords: comparative genomics, DNA methylation, epigenetics.

Epigenetic information is an important, environmentally responsive mediator of the relationship between genotype and phenotype (Jaenisch & Bird, 2003; Kucharski et al., 2008; Margueron & Reinberg, 2010), which results from mechanisms other than changes in DNA sequence (Berger et al., 2009; Margueron & Reinberg, 2010). Nevertheless, such information is transmissible across mitotic, and sometimes meiotic, cellular divisions (Bonasio et al., 2010a). One of the most important forms of epigenetic information is the methylation of DNA.

DNA methylation is present in all three domains of life (Klose & Bird, 2006; Suzuki & Bird, 2008), suggesting a role in the common ancestor of Metazoa and, possibly, of all multicellular life. The methylation of DNA in animals has been implicated in several important biological processes including developmental progression and regulation (Haines et al., 2001; Futscher et al., 2002; Kucharski et al., 2008), memory formation (Miller & Sweatt, 2007; Lockett et al., 2010) and carcinogenesis (Merlo et al., 1995; Baylin et al., 1998; Jones & Baylin, 2002; Jair et al., 2006). Furthermore, DNA methylation patterns diverge greatly amongst individuals and even monozygotic twins (Fraga et al., 2005; Lister et al., 2009; Javierre et al., 2010). Thus, widespread evidence suggests that DNA methylation may provide critical contributions to developmental and phenotypic variation.

In this review, we explore the broadly conserved DNA methylation system of metazoan taxa, its known function in insects and important gaps in the current knowledge of DNA methylation in insects. Insects provide an integral component of our understanding of the evolutionary diversity of epigenetic systems. In particular, insect taxa encompass multiple states of conservation and loss of DNA methylation. Thus, as the field of comparative epigenomics grows, insects stand to serve as important models of DNA methylation and critical systems for understanding the biological consequences of its loss.

Mediators of the DNA methylome

DNA methylation is a covalent modification that occurs through the addition of a methyl group to DNA, almost exclusively at cytosine bases in animals (but see Vanyushin, 2005). This modification is accomplished by several
key, evolutionarily conserved enzymes known collectively as DNA methyltransferases (DNMTs; Goll & Bestor, 2005; Albalat, 2008). DNMTs are divided into several classes based upon the nature of their activity. Studies in mammalian systems indicate that DNMTs can be separated into 'de novo' and 'maintenance' methyltransferases (Klose & Bird, 2006). De novo methyltransferases are responsible for establishing new methylation patterns within an organism’s genome and are represented by the DNMT3 family of proteins in mammals (Okano et al., 1999; Aapola et al., 2002; Hata et al., 2002; Kato et al., 2007). In contrast, maintenance methyltransferases, represented by the DNMT1 family of proteins, maintain previously established methylation patterns across cell generations by preferentially methylating hemimethylated DNA substrates (Bestor, 2000; Chen et al., 2003). Finally, although DNMT2 was originally believed to be a DNA methyltransferase, it has recently been shown to methylate tRNA and thus differs in function from DNMT1 and DNMT3 (Goll et al., 2006; Jurkowski et al., 2008). The presence of one or more copies of DNMT1 and DNMT3 is generally considered necessary to a functional DNA methylation system (Goll & Bestor, 2005), although emerging data on genome sequences and DNA methylation maps in insects suggest potential exceptions to this pattern (see below).

Methyl-CpG-binding domain proteins (MBDs) represent another important component of the DNA methylation ‘toolkit’, as MBDs contain a methyl-CpG (cytosine followed by guanine in 5’ to 3’ orientation) recognition motif that allows the selective binding of methylated DNA (Klose & Bird, 2006; Clouaire et al., 2010). Through this selective targeting, MBDs localize chromatin remodelling complexes to the areas of DNA methylation, and can thereby affect epigenetic modifications at multiple levels (Jones et al., 1998; Feng & Zhang, 2001; Jones & Baylin, 2002; Hendrich & Tweedie, 2003; Bogdanovic & Veenstra, 2009). Much like DNMTs, genomes of organisms with functional DNA methylation activity have all been found to contain MBDs, which are highly conserved in all vertebrates (Hendrich & Tweedie, 2003; Clouaire & Stancheva, 2008). However, MBDs are present in many plant and animal taxa that do not display substantial DNA methylation, which suggests that MBDs may have functions other than DNA methylation.

Genomic targets of DNA methylation in animals

DNA methylation is largely confined to CpG dinucleotides in genomes of animals (Bird, 1980; Wang et al., 2006; Feng et al., 2010; Zemach et al., 2010). Although the genomic regions exhibiting CpG methylation vary widely amongst taxa, one of the most broadly conserved patterns of methylation appears to be the targeting of gene bodies (ie, exons and, to a lesser extent, introns). Gene body methylation is observed in plants and animals, but is absent in most fungi (Feng et al., 2010; Zemach et al., 2010). Indeed, there exists a deep phylogenetic signal of gene body methylation across Metazoa, whereas an expanded pattern of global methylation has evolved gradually in deuterostomes (Okamura et al., 2010). For example, DNA methylation in vertebrates occurs throughout the genome (Suzuki & Bird, 2008; Okamura et al., 2010), with between 60–90% of all CpG dinucleotides being subject to methylation in most mammals (Ehrlich et al., 1982; Lister et al., 2009; Li et al., 2010).

Interspersed throughout mammalian genomes are small areas of unmethylated CpGs, termed ‘CpG islands’. CpG islands are approximately 300–3000 base pairs in length and are found in and around approximately 40% of mammalian gene promoters (Fatemi et al., 2005; Saxonov et al., 2006; Elango & Yi, 2008). Importantly, the methylation of promoter regions has been linked to transcriptional repression in vertebrates (Wolffe & Matzke, 1999; Weber et al., 2007). DNA methylation probably inhibits gene expression by interfering with DNA-binding of transcription factors in promoter regions (Watt & Molloy, 1988) or by enhancing the binding of repressive regulatory proteins to methyl-CpG motifs (Boyce & Bird, 1991; Hendrich & Bird, 1998). In vertebrates, DNA methylation also may play a repressive role with respect to the activity of transposable elements (Yoder et al., 1997; O’Neill et al., 1998).

In contrast to the pattern of genome-wide DNA methylation in vertebrates, DNA methylation in invertebrates is relatively sparse (Bird et al., 1979; Suzuki & Bird, 2008). Indeed, the low or absent levels of DNA methylation detected in model invertebrates, such as Drosophila melanogaster (Rae & Steele, 1979; Urei-Shoval et al., 1982) and Caenorhabditis elegans (Simpson et al., 1986), initially suggested diminished functional significance for DNA methylation in invertebrates as a whole. However, recent studies have revealed the persistence of DNA methylation in many invertebrate taxa (Wang et al., 2006; Suzuki et al., 2007; Kronforst et al., 2008; Feng et al., 2010; Nasonia Genome Working Group, 2010; Walsh et al., 2010; Zemach et al., 2010).

DNA methylation is largely confined to genes in invertebrates, whereas intergenic regions remain largely unmethylated (Simmen et al., 1999; Suzuki & Bird, 2008; Feng et al., 2010; Zemach et al., 2010). Moreover, DNA methylation of transposable and repetitive elements has been observed only at moderate levels in basal invertebrates (Feng et al., 2010) and is almost non-existent in insects (Regev et al., 1998; Feng et al., 2010; Schaefer & Lyko, 2010; Zemach et al., 2010). Together, these results suggest that DNA methylation is not preferentially targeted to, and thus plays little role in suppressing the
proliferation of transposable elements in insects and other invertebrates.

The evolution of DNA methylation in insects: a patchwork of persistence and loss

The first investigations of DNA methylation in insects were undertaken in *Dr. melanogaster* (Rae & Steele, 1979; Urieli-Shoval *et al*., 1982). These studies indicated that *Dr. melanogaster*'s genome lacked both de novo and maintenance methyltransferases and featured a near-total lack of DNA methylation (but see Tweedie *et al*., 1999 and Marhold *et al*., 2004). Importantly, this result suggests that the functional role of DNA methylation can be readily compensated by other molecular mechanisms in some taxa. Nevertheless, a growing number of investigations has since demonstrated that DNA methylation persists in many insect lineages (Fig. 1). Although the most basal insect lineages have yet to be interrogated with respect to DNA methylation, the genome of the outgroup crustacean *Daphnia pulex* contains both DNMT1 and DNMT3 (Albalat, 2008; Colbourne *et al*., 2011). Furthermore, the presence of methylated cytosine has been observed in its sister taxon *Daphnia magna* (Vandegehuchte *et al*., 2009). Taken together, these results suggest that DNA methylation may have been ancestral to Insecta, and the lineage-specific loss of DNA methylation probably occurred during the evolutionary diversification of insects (Fig. 1). We note that although MBDs remain poorly studied in insects, their presence is phylogenetically widespread, even in insects without substantial DNA methylation (Fig. 1). As mentioned above, this suggests that MBD functions may extend beyond DNA methylation (Hendrich & Tweedie, 2003).

DNA methylation has now been empirically detected in each of the three major groups of Neoptera (winged insects; Grimaldi & Engel, 2005), including Polyneoptera, Paraneoptera and Holometabola (Fig. 1). Although none of the Polyneoptera has been subject to genome sequencing or analysis of methylation-related proteins,

![Figure 1. Phylogenetic distribution of DNA methylation in insects. Relationships and approximate divergence times of major insect lineages and an outgroup crustacean, *Daphnia pulex* (according to Gaunt & Miles, 2002; Grimaldi & Engel, 2005; Savard *et al*., 2006). Branches are named for insect orders, with representative species for which DNA methylation information has been obtained listed below. Dots represent the number of DNA methyltransferases (DNMTs) present in a sequenced genome and the presence of methyl-CpG-binding domain proteins (MBDs; absence indicates no DNMTs of a given family or no MBDs detected, whereas question marks indicate no data). Putative DNMT loss is marked on branches based on currently available data. The detection of DNA methylation is indicated by a check mark and the validation of a near-total lack of DNA methylation is indicated by an ‘X’ (references provided in text).](image-url)

© 2011 The Authors
Insect Molecular Biology © 2011 The Royal Entomological Society, 20, 553–565
the presence of methylated cytosines has been detected in the orthopteran crickets *Acheta domesticus* (Tweedie et al., 1999) and * Gryllotalpa f ossor* (Soma & Rao, 1992), as well as in the stick insect *Medauroidea extradentata* (Krauss et al., 2009).

In the Paraneoptera, early reports suggested that genespecific DNA methylation played a role in mediating insecticide resistance in the hemipteran aphids *Myzus persicae* (Field et al., 1996; Field, 2000) and *Schizaphis graminum* (Ono et al., 1999). Subsequently, the genome sequencing of the pea aphid *Acyrthosiphon pisum* revealed the presence of two isoforms of both DNMT1 and DNMT3 (International Aphid Genomics Consortium, 2010; Walsh et al., 2010), and multiple empirical methods have confirmed the presence of methylated cytosines in this taxon (Walsh et al., 2010). In contrast, the compact genome of the phthirapteran body louse *Pediculus humanus* revealed the apparent loss of DNMT3, suggesting that *Pe. humanus* may not display fully functional DNA methylation (Kirkness et al., 2010; Nasonia Genome Working Group, 2010).

The Holometabola has been the overwhelming focus of DNA methylation study in insects in recent years. For instance, genomic analyses have revealed the evolutionary persistence of DNA methylation across Hymenoptera (Kronforst et al., 2008; Nasonia Genome Working Group, 2010). In fact, de novo and maintenance DNMTs in insects were first fully discovered in the honeybee *Apis mellifera* (Wang et al., 2006). *Ap. mellifera* has since become a model for understanding DNA methylation in insects. In addition, a fully functional methylation toolkit was found in the two ants *Harpegnathos saltator* and *Camponotus floridanus*, with DNA methylation confirmed by the densitometric detection of 5-methylcytosine (Bonasio et al., 2010b). Interestingly, *H. saltator*, which possesses a simpler social system than *C. floridanus*, also exhibits lower levels of DNA methylation than *C. floridanus* (Bonasio et al., 2010b). Furthermore, four other ant genomes (from *Solenopsis invicta*, *Pogonomyrmex barbatus*, *Linepithema humile* and *Atta cephalotes*) were found to possess DNMT1 and DNMT3 (Smith C.D. et al., 2011; Smith C.R. et al., 2011; Suen et al., 2011; Wurm et al., 2011). DNA methylation was confirmed in *Po. barbatus* by methylation-sensitive amplified fragment length polymorphism analysis (Smith C.R. et al., 2011) and methylation in *S. invicta* was confirmed by methylated DNA immunoprecipitation followed by targeted sequencing of bisulphite-converted DNA (Wurm et al., 2011).

In contrast to the Hymenoptera, where DNA methylation appears to be widespread, several other insect taxa exhibit diminished levels of DNA methylation. For example, the coleopteran flour beetle *Tribolium castaneum* has lost DNMT3 and is apparently unable to methylate its DNA (Tribolium Genome Sequencing Consortium, 2008; Zemach et al., 2010). Furthermore, the most dramatic loss of DNA methylation proteins in insects has been observed in the Diptera, where genome sequencing projects have not detected DNMT1 or DNMT3 proteins (Hung et al., 1999; Tweedie et al., 1999; Marhold et al., 2004). As predicted based on the absence of DNMTs, CpG methylation is virtually undetectable in most developmental stages of *Dr. melanogaster* (Zemach et al., 2010). Intriguingly, although DNA methylation has been detected in the lepidopterans *Mamestra brassicae* (Mandrioli & Volpi, 2003) and Bombyx mori (Xiang et al., 2010), the draft *B. mori* genome does not contain a detectable orthologue of DNMT3. *B. mori* was nevertheless the first insect to have its ‘DNA methylome’ profiled by the sequencing of bisulphite-converted DNA on a genomic scale (Xiang et al., 2010), and has become an important model for understanding the genomic targets of DNA methylation in insects (Xiang et al., 2010; Zemach et al., 2010).

**Diverse evolutionary signatures of DNA methylation in insects**

DNA methylation can be identified using molecular genetic and biochemical techniques, as described above. However, CpG methylation also leaves an evolutionary signature in the genome that can be detected by analysing normalized CpG dinucleotide content [CpG observed/expected (o/e); see Yi & Goodisman, 2009]. Normalized CpG content represents the observed frequency of CpG dinucleotides relative to that expected based on the frequency of C and G nucleotides in the genomic region of interest. Normalized CpG content acts as a proxy for DNA methylation because DNA methylation is almost entirely targeted to CpG dinucleotides in animals and methylated cytosines tend to undergo spontaneous deamination to thymine with high frequency (Shen et al., 1994). Consequently, areas of genomic DNA that contain high levels of CpG methylation often exhibit a marked reduction in CpG dinucleotides (Fig. 2; Bird, 1980; Shimizu et al., 1997; Bock & Lengauer, 2008; Yi & Goodisman, 2009). It is notable that a conceptually similar approach to the analysis of CpG o/e, based instead on the measurement of CpG-to-TpG polymorphism, has recently been applied in several ant taxa (Smith C.D. et al., 2011; Smith C.R. et al., 2011).

Many of the functional inferences about DNA methylation in the honeybee *Ap. mellifera* were first achieved using analyses of normalized CpG content (Elango et al., 2009; Foret et al., 2009; Wang & Leung, 2009; Zeng & Yi, 2010). The subsequent empirically derived DNA methylomes of a whole-body worker honeybee (Zemach et al., 2010) and of honeybee brains (Lyko et al., 2010) have provided strong evidence of the negative correlation between normalized CpG content and DNA.
methyltransferases.

Normalized CpG content analysis can readily be conducted on different regions of the genome to provide information on which regions are targets of methylation. For example, the genomes of many vertebrates are globally methylated. As expected, normalized CpG profiles of nearly all genomic regions of vertebrates exhibit a mean value far less than one, indicating the depletion of CpG dinucleotides (Okamura et al., 2010). In contrast, many animals with no detectable levels of DNA methylation exhibit a mean normalized CpG value for genes and other genomic regions of around one, as expected in the absence of DNA methylation (Elango et al., 2009; Yi & Goodisman, 2009). In Ap. mellifera, analyses of normalized CpG content of different genomic regions suggested that genes alone harbour substantial CpG depletion and are thus the dominant targets of DNA methylation (Elango et al., 2009). This result has subsequently been confirmed by empirical analyses in numerous insect taxa with functional DNA methylation (Feng et al., 2010; Lyko et al., 2010; Xiang et al., 2010; Zemach et al., 2010).

One of the more interesting results from analyses of normalized CpG content in invertebrates is the presence of bimodal methylation profiles amongst the genes of several species (Suzuki et al., 2007; Elango et al., 2009; Foret et al., 2009; Wang & Leung, 2009; Walsh et al., 2010). This bimodal profile indicates the presence of two distinct classes of genes with respect to DNA methylation: those with high mean normalized CpG content (and thus low methylation), and those with low mean normalized CpG content (and high levels of methylation). To date, insects with bimodal distributions of normalized CpG content of genes include the honeybee Ap. mellifera (Fig. 3B; Elango et al., 2009; Wang & Leung, 2009) and the pea aphid Ac. pisum (Fig. 3C; Walsh et al., 2010). This pattern stands in stark contrast to the unimodal distribution observed in Drosophila (Fig. 3A) and other taxa lacking DNA methylation (Elango et al., 2009).

Normalized CpG dinucleotide content of protein-coding sequences also provides important clues as to the presence of DNA methylation in systems where DNA methylation has not been directly detected empirically. For example, Pe. humanus, which possesses DNMT1 but may be lacking DNMT3, displays a normalized CpG content distribution that is exceptionally broad, with evidence for CpG depletion in many genes (Fig. 3G). This result suggests that DNA methylation probably occurs in the Pe. humanus genome, which would provide a second example of a functional methylation system in a genome where DNMT3 has not been detected (Xiang et al., 2010). Moreover, the normalized CpG content profile for genes of Da. pulex (Fig. 3H) is similar to that observed for S. invicta (Fig. 3E) and B. mori (Fig. 3F). This finding, together with the identification of a complete suite of methylation
proteins, suggests that DNA methylation may also be present in the *Da. pulex* genome.

Interestingly, it appears that the presence of a bimodal gene methylation profile is sufficient evidence for the presence of a functional methylation system, but is not a universal consequence of DNA methylation. For example, normalized CpG profiles of genes from several insect species known to harbour functional DNA methylation systems lack clearly defined bimodality (e.g., all investigated ant taxa, including *S. invicta*, and the lepidopteran *B. mori*; Fig. 3E,F; Bonasio et al., 2010b; Smith C.D. et al., 2011; Smith C.R. et al., 2011). The evolutionary mechanisms underlying differences in the degree of genic CpG depletion amongst taxa with functional methylation systems (Fig. 3) is presently unclear.

In *Ap. mellifera*, it has been predicted that the process of biased gene conversion, a mechanism by which CG content can be increased during meiosis, may explain the excess of CpG dinucleotides observed genome-wide (Marais, 2003; Elango et al., 2009). Biased gene conversion could also explain how CpG dinucleotides are maintained in methylated genes of some insects. However, genes with high CpG content do not exhibit different recombination rates from genes depleted of CpGs in *Ap. mellifera*, despite the fact that recombination is expected to increase the efficiency of biased gene conversion (Zeng & Yi, 2010). Thus, gene conversion appears unlikely to be solely responsible for preserving CpG content in insect genomes. Differences in the degree of CpG depletion amongst species may instead reflect unrecognized differences in the lineage-specific evolutionary age of methylation targeting, differences in the proportion of methylated copies of DNA or differences in the strength of selective pressures acting to retain CpG dinucleotides. The sequencing of DNA methylomes from both germline and somatic cells of multiple taxa will be necessary to fully understand this enigmatic variation in CpG depletion in insect genomes.

**Insight into the role of DNA methylation in insects**  
As described above, genes rather than entire genomes are targeted by DNA methylation in insects. However, instead of serving as a ubiquitous repressor of transcription, as appears to be the case with the methylation of gene promoter regions in vertebrates (Wolfe & Matzke, 1999), mounting evidence suggests that DNA methylation in gene bodies may play a specialized role in the maintenance of transcript integrity, as well as the regulation of mRNA initiation or splice patterns (Young et al., 2006; Mandrioli, 2011; The Authors).
The idea that genomic DNA methylation may regulate alternative intragenic promoters affecting alternative transcription was first raised by mammalian studies (Cheong et al., 2006; Maunakea et al., 2010). Indeed, the regulation of alternative transcription or splicing may be achieved through DNA methylation’s negative interaction with the elongation efficiency of RNA polymerase (Rountree & Selker, 1997; Zilberman et al., 2007) or the direct interaction of DNA methylation machinery with splicing factors, as in the case of humans (Young et al., 2006). Importantly, alternative splicing and transcription patterns have been shown to vary dramatically through the course of an organism’s development (Barberan-Soler & Zahler, 2008) and probably play a fundamental role in generating phenotypic variation (Ast, 2004).

It has been suggested that DNA methylation evolved from the restriction-modification system of ancestral bacteria (Bestor, 1990) and was later co-opted to mediate developmental and biological complexity (Bird, 1995; Jablonka & Regev, 1995). In a broad study of invertebrates, Regev et al. (1998) revealed that the amount of cell turnover in an organism is positively associated with levels of DNA methylation, which suggests an increased need for epigenetic information in conjunction with developmental complexity. Furthermore, de novo DNA methylation is hypothesized to play an important role in developmental responsiveness to environmental factors and the regulation of developmental plasticity, as is apparently the case in the honeybee Ap. mellifera (see below; Jaenisch & Bird, 2003; Kucharski et al., 2008; Maleszka, 2008). Thus, through the addition of epigenetic information during the course of organismal development, newly introduced variation in DNA methylation may lead to variation in the regulation of gene transcription that could enhance developmental plasticity and provide an important mechanism for responsiveness to environmental stimuli.

**DNA methylation and phenotypic specialization: the case of the honeybee**

In most social insects, such as the honeybee Ap. mellifera, distinct queen and worker castes result from differential expression of genes during development (Evans & Wheeler, 2001; Barchuk et al., 2007; Smith et al., 2008). Typically, most hymenopteran social insect larvae develop into workers, whereas a select few develop into future queens based on environmental input (Weaver, 1966; Wheeler, 1986). In contrast, following the knockdown of the de novo methyltransferase gene Dnmt3 in Ap. mellifera, a majority of lab-reared larvae developed a queen phenotype (Kucharski et al., 2008). This landmark result suggested a direct link between de novo methylation and the development of specific castes (Kucharski et al., 2008). Indeed, Kucharski et al.’s study stands as one of the most striking links between DNA methylation and developmental plasticity in any taxon (Moczek & Snell-Rood, 2008).

Somewhat surprisingly, however, in the above study only 14 genes were significantly differentially expressed between third instar Dnmt3-silenced and control Ap. mellifera larvae (Kucharski et al., 2008), as compared to 37 genes in a study of wild-type queen and worker larvae of the same stage (Barchuk et al., 2007). Furthermore, only two genes were found in common amongst the top 50 differentially expressed genes in comparisons of Dnmt3-silenced versus control individuals and wild-type queens versus workers (Kucharski et al., 2008). One explanation for these findings is that several developmental pathways (or networks of co-expressed genes) have the potential to act in the differentiation of castes. Alternatively, DNA methylation may affect the production of caste-specific protein isoforms in the honeybee. In other words, rather than modulating expression of different genes per se, DNA methylation may promote caste differences via expression of different versions of genes.

A possible link between alternative splicing and differential DNA methylation in the honeybee has recently been provided by Lyko et al. (2010). In a study that documented genome-wide patterns of DNA methylation at a single-base resolution in adult Ap. mellifera queen and worker brains, these authors found that methylated CpGs were significantly co-localized to alternatively spliced exons (when compared to a randomized distribution). Elaborating upon these findings, the authors examined one differentially methylated gene between queens and workers in detail. In this case, differential methylation between the two castes was targeted to an alternative (and, in the case of workers, highly methylated and omitted) exon containing a stop codon (Lyko et al., 2010). This finding was the first to suggest a link between methylation and the outcome of alternative splicing in insects, which may also be associated with the distinct behavioural repertoires in Ap. mellifera. However, it must be emphasized that the relationship between alternative splicing and caste differences remains strictly hypothetical at present. The number of alternatively spliced genes between queens and workers and their functional consequences will need to be investigated further in order to test this hypothesis. Furthermore, how differential methylation manipulates the activity of mRNA splicing machinery is largely unknown.

**Evolutionary implications of DNA methylation in insects**

As described above, functional inferences from the study of the honeybee have provided substantial insight into the
putative roles of genic DNA methylation in insects and other taxa. These insights have been further enhanced by the recent implementation of comparative genomic analyses of DNA methylation. For example, in an effort to assess whether a common functional role exists for DNA methylation in diverse insects, Hunt et al. (2010) examined the conservation of methylation targets between the highly diverged (~300 Mya) pea aphid Ac. pismum and Ap. mellifera. Interestingly, genes with low levels of methylation were less likely to maintain their methylation status over evolutionary time, whereas heavily methylated genes were more likely to conserve their hypermethylated status, as indicated by analysis of normalized CpG content. Thus, if genes were heavily methylated in the common ancestor of Ac. pismum and Ap. mellifera, they were apparently more likely to stay heavily methylated over evolutionary time.

Furthermore, methylated genes in divergent taxa exhibited greater overlap in their patterns of functional enrichment than unmethylated genes (Hunt et al., 2010). These results suggest that there is some degree of functional conservation of DNA methylation status over vast evolutionary time. Genes with prominent methylation signatures also appear to be more highly conserved at the sequence level than their unmethylated counterparts (Suzuki et al., 2007; Hunt et al., 2010; Lyko et al., 2010), a result that is particularly striking given the mutational effect of DNA methylation (Elango et al., 2008), and one that is typical of ubiquitously expressed genes (Duret & Mouchiroud, 2000; Pal et al., 2006).

One of the most important evolutionary insights with respect to DNA methylation in insects has been the observation that ubiquitously expressed genes are preferentially targeted by DNA methylation in numerous insect taxa (Elango et al., 2009; Foret et al., 2009; Hunt et al., 2010; Xiang et al., 2010). In contrast, genes that show less evidence of DNA methylation according to normalized CpG content are more likely to be differentially expressed across tissues or alternate phenotypes (Elango et al., 2009; Foret et al., 2009; Hunt et al., 2010). Interestingly, differentially methylated genes themselves are less depleted of CpGs than genes that are similarly methylated in all contexts (Lyko et al., 2010). This may indicate that genes that undergo differential methylation in somatic tissues are less prone to germline DNA methylation than other methylated genes. Alternatively, differentially methylated genes may be under stronger selective pressure to preserve CpG dinucleotides.

The preferential targeting of ubiquitously expressed genes by DNA methylation, together with the implication of DNA methylation in the regulation of alternative transcription (Maunakea et al., 2010), suggests that the regulation or repression of alternative transcription patterns may be particularly important in ubiquitously expressed genes. This hypothesized connection could result either from an enhanced negative fitness effect for spurious transcription initiation and termination in ubiquitously expressed genes, or from a regulatory need to differentiate the tissue- and condition-specific roles of ubiquitously expressed genes. Interestingly, within mammals, CpG island promoter length is associated with tissue expression breadth (Elango & Yi, 2008, 2011; Sharif et al., 2010). Thus, a conserved (or convergent) connection between DNA methylation variation and tissue expression breadth may exist between gene body and promoter methylation (Ililworth et al., 2008; Maunakea et al., 2010).

Prospects for insect epigenomics

Considerable progress has been made in the last several years in understanding the nature and functional significance of DNA methylation in insects. We now have an increased understanding of the scope of DNA methylation in insects and the patterns of methylation within insect genomes. However, the field of true epigenomics remains in its nascent stages of exploration, and considerable further research is required to fully understand the role of DNA methylation in insects.

For example, with the increasing accessibility of DNA methylome sequencing, the degree of polymorphism in methylation status between tissues and individuals can begin to be characterized in models of insect DNA methylation, such as Ap. mellifera or B. mori. The rate of change in methylation profiles amongst taxa, which is poorly understood at present, will also be revealed by DNA methylome data from diverse insect taxa. These advances will lay the groundwork for a more comprehensive understanding of the potential link between the generation of phenotypic novelty and DNA methylation.

One of the more pressing questions regarding DNA methylation in insects is its exact role in the regulation of transcription. The coupling of transcriptome data with single-base resolution maps of DNA methylation from diverse tissue types and species will help to characterize more fully the relationship between gene regulation, including the regulation of alternative splicing, and DNA methylation. Furthermore, the demonstrated utility of RNA interference (Kucharski et al., 2008) and a topical inhibitor of DNMT3 (Lockett et al., 2010) to experimentally perturb de novo DNA methylation in insects suggests that experiments can be undertaken to assess whether DNA methylation itself actively alters patterns of alternative transcription, RNA splicing or condition-specific expression levels of genes.

Another open question lies in the persistence of DNA methylation in B. mori, despite the apparent lack of DNMT3. In fact, a similar proportion of CpG dinucleotides are targeted by methylation in the genomes of B. mori (0.7%) and Ap. mellifera (0.5%; Zemach et al., 2010). How
is this methylation maintained, whereas the loss of DNMT3 in T. castaneum is associated with the loss of DNA methylation (Fig. 1)? If DNMT3 is truly absent in the genome of B. mori and not an artefact of stochastic variation in sequencing coverage, what are the molecular mechanisms perpetuating DNA methylation? Furthermore, what is the mechanism responsible for the distinct patterns of CpG depletion present in, for example, B. mori and Ap. mellifera (Fig. 3)?

Several key aspects of DNA methylation in mammals remain entirely unexplored in invertebrates and insects. For example, global DNA demethylation occurs during early development in mammals, which allows the ‘reprogramming’ of the genome essential for proper development in mammals, which allows the ‘reprogramming’ of the genome essential for proper development (Monk et al., 1987; Mayer et al., 2000). Demethylation has also been shown to play an important role in transcriptional cycling of mammalian gene promoters (Kangaspeska et al., 2008; Metivier et al., 2008). Furthermore, DNA demethylation can occur on the time scale of hours (Kangaspeska et al., 2008; Metivier et al., 2008; Ooi & Bestor, 2008), suggesting that this process may play an involved role in transcription (Wu & Zhang, 2010). Whether DNA demethylation is similarly critical during insect (and, more generally, invertebrate) development is unknown. Likewise, the presence of methylation cycling in insects and other invertebrates has yet to be demonstrated.

The layering and exchange of distinct types of epigenetic information is another exciting and unexplored direction for future study in insects. DNA methylation involves the interaction of a large suite of proteins in fungi and vertebrates (Vire et al., 2006), such as those linked to histone modification systems (Ben-Porath & Cedar, 2001; Tamaru & Selker, 2001; Okitsu & Hsieh, 2007). Moreover, in mammals, DNMTs and MBPs are known to participate in the recruitment of histone modification proteins (Lopezrodas et al., 1993; Jones et al., 1998; Feng & Zhang, 2001; Fuks et al., 2003; Geiman et al., 2004; Bai et al., 2005) and other proteins responsible for the remodelling of chromatin (Geiman et al., 2004; Margueron & Reinberg, 2010). However, the interaction amongst different epigenetic systems in invertebrates has not been explored in detail. It is notable, however, that Dr. melangaster, one of the best-studied model systems of epigenetic protein modifications, lacks DNA methylation. Comparative studies of the interaction between DNA and protein modifications of insect species may thus elucidate evolutionary progression towards the interaction of DNA methylation and other epigenetic modifications.

Another important finding in mammals is that DNA methylation may act as a mechanism for genomic imprinting, which results in the differential expression of parental alleles (Li et al., 1993; Reik & Walter, 2001; Hata et al., 2002). Imprinting is a potential source of conflict between parental genomes (Wilkins & Haig, 2003), and is hypothesized to play a particularly important role in the biology of highly social organisms such as the eusocial Hymenoptera (Haig, 2000; Queller, 2003; Kronauer, 2008). For example, conflict between the relative investment in female queen and worker offspring may arise between males and queens in eusocial hymenopterans where queens mate multiply (because each male would benefit from producing a greater proportion of reproductive offspring). Indeed, imprinting is predicted to occur in many circumstances in the eusocial Hymenoptera (see Queller 2003 for an extensive discussion). The demonstration of a link between DNA methylation and imprinting in social insects would provide new insight into the evolution of conflict and cross-purpose in social colonies (Strassmann & Queller, 2007).

Conclusion

Insects are excellent model systems for studying the evolution of DNA methylation. By investigating evolutionary patterns of DNA methylation in insects, we stand to gain valuable insight into the conservation and function of this widespread epigenetic mark. Furthermore, comparative epigenomic studies of insect taxa have tremendous potential to illuminate the contributions of DNA methylation to developmental regulation. Social insects in particular are exceptionally promising models in this regard because of the presence of outstanding phenotypic plasticity and ample potential for genomic imprinting. In addition, insects are highly amenable to large scale genomic and epigenomic studies owing to their moderately sized genomes and experimental tractability. Undoubtedly, our understanding of DNA methylation will grow with the continued exploration of insect genomic data and the continued sequencing of insect DNA methylomes.

Acknowledgements

This work was supported by the U. S. National Science Foundation (grant numbers MCB-0950896 and DEB-0640690 to S. Y. and M. G.). We thank two anonymous reviewers for their comments.

References


Bai, S.M., Ghoshal, K., Datta, J., Majumder, S., Yoon, S.O. and Jacob, S.T. (2005) DNA methyltransferase 3b regulates nerve...
Gaunt, M.W. and Miles, M.A. (2002) An insect molecular clock dates the origin of the insects and accords with...
DNA methylation in insects

563


