Insect Molecular Biology

Insect Molecular Biology (2011) 20(5), 553–565 doi: 10.1111/j.1365-2583.2011.01092.x

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

First published online 24 June 2011.

Correspondence: Brendan Hunt, School of Biology, Georgia Institute of Technology, Atlanta, GA 30332, USA. Tel.: + 1 404 831 1810; fax: + 1 404 894 0519; e-mail: brendan.hunt@gatech.edu

*These authors contributed equally to this work.

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 (Boyes & 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; Urieli-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,

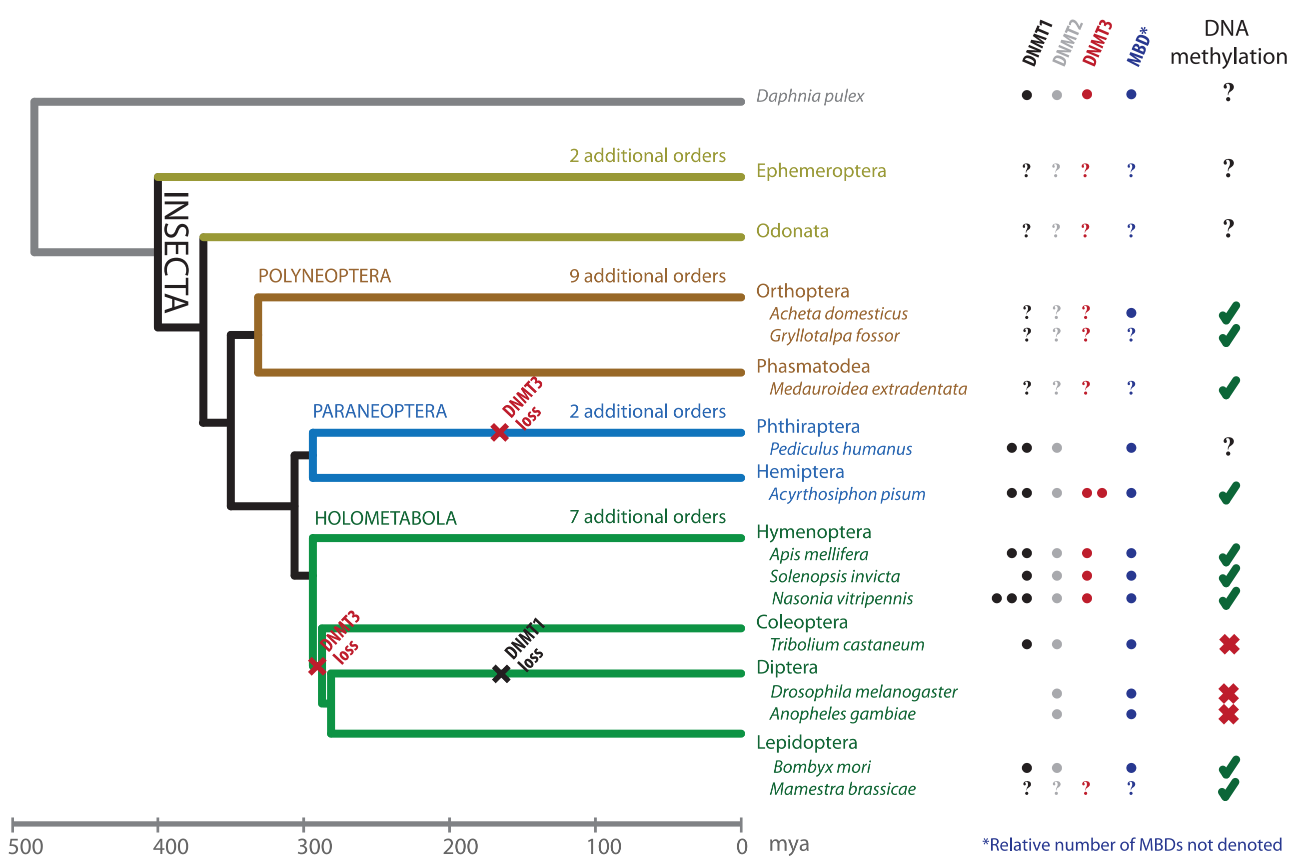

Figure 1. Phylogenetic distribution of DNA methylation in insects. Relationships and approximate divergence times of major insect lineages and an outgroup crustacean, *Daphnia pulex* (according to Gaunt & Miles, 2002; Grimaldi & Engel, 2005; Savard *et al*., 2006). Branches are named for insect orders, 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).

the presence of methylated cytosines has been detected in the orthopteran crickets *Acheta domesticus* (Tweedie *et al*., 1999) and *Gryllotalpa fossor* (Soma & Rao, 1992), as well as in the stick insect *Medauroidea extradentata* (Krauss *et al*., 2009).

In the Paraneoptera, early reports suggested that gene-specific 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

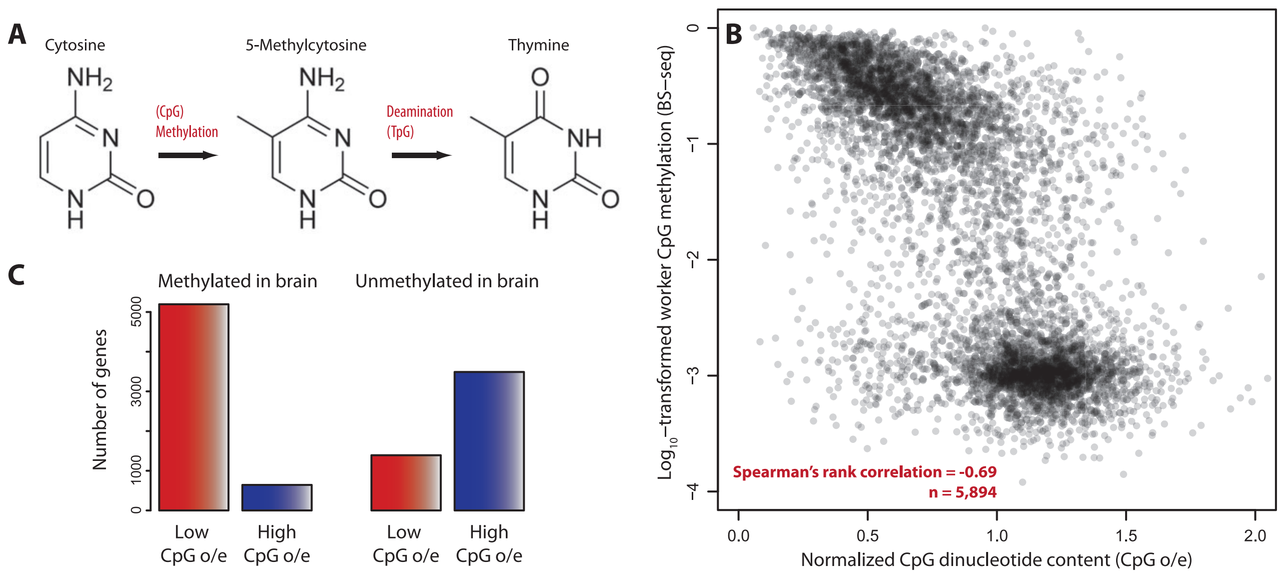

Figure 2. CpG (cytosine followed by guanine in 5′ to 3′ orientation) depletion and empirically measured somatic DNA methylation in the honeybee *Apis mellifera*. (A) DNA methylation results in transitions from CpG to TpG dinucleotides through mutation of methylated cytosines. (B) correlation between normalized CpG content [CpG observed/expected (o/e)] and fractional CpG methylation of genes according to bisulphite-converted sequencing of whole-body worker genomic DNA (Zemach *et al*., 2010) and (C) proportion of genes methylated in honeybee brains belonging to distinct classes of CpG depletion (Lyko *et al*., 2010) demonstrates the strong association between computationally and empirically determined levels of DNA methylation.

methylation level in somatic tissues (Fig. 2B,C). This relationship is particularly striking given that normalized CpG content profiles are inherently shaped by the methylation of germline cells (where mutations are transmitted across generations), and suggests that many genes are methylated in both somatic and germline cells over evolutionary time.

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

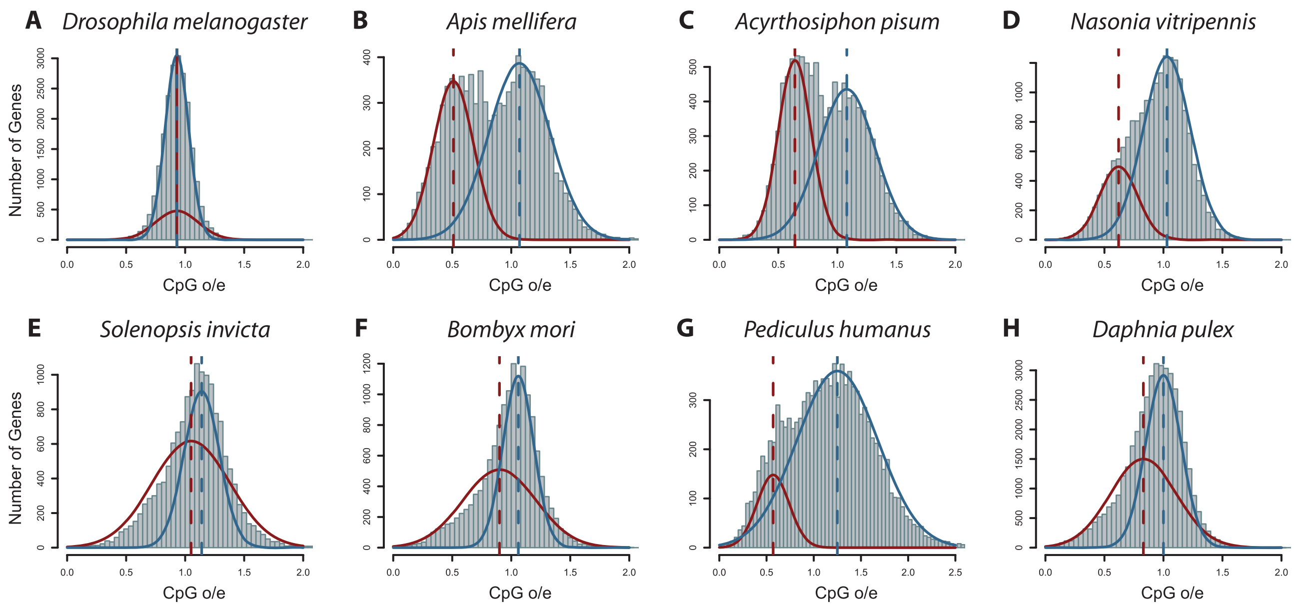

Figure 3. Diversity of evolutionary CpG (cytosine followed by guanine in 5′ to 3′ orientation) depletion in protein-coding sequences amongst insect taxa. Distributions of normalized CpG content [CpG observed/expected (o/e)] in coding sequences with a mixture of two normal distributions fitted to the data using NOCOM (Ott, 1979). Dashed lines indicate the mean of each component. (A) *Drosophila melanogaster* has a genome that is almost entirely devoid of DNA methylation and exhibits a qualitatively unimodal normalized CpG content (CpG o/e) distribution, with the same mean for both of two components fitted to the data (as is typical of insects lacking DNA methylation; Elango *et al*., 2009). The genes of (B) the honeybee *Apis mellifera* and (C) the pea aphid *Acyrthosiphon pisum* are each targeted by DNA methylation and exhibit striking bimodality. In these cases, the component with a lower mean largely represents those genes depleted of CpGs by DNA methylation. In contrast, the genomes of several insects with genic DNA methylation exhibit less striking signatures of CpG depletion, including (D) the jewel wasp *Nasonia vitripennis*, (E) the fire ant *Solenopsis invicta* and (F) the silkworm *Bombyx mori*. The presence of DNA methylation has yet to be confirmed in (G) the body louse *Pediculus humanus* or (H) the crustacean waterflea *Daphnia pulex*, but their profiles of genic CpG depletion are suggestive of its presence.

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 (eg, 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 (Wolffe & 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, 2007; Suzuki *et al*., 2007; Foret *et al*., 2009; Hunt *et al*., 2010; Lyko *et al*., 2010; Maunakea *et al*., 2010).

The idea that genic 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. pisum* 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. pisum* 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 (Illingworth *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 (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. melanogaster*, 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

Aapola, U., Liiv, I. and Peterson, P. (2002) Imprinting regulator DNMT3L is a transcriptional repressor associated with histone deacetylase activity. *Nucleic Acids Res* **30**: 3602–3608.

Albalat, R. (2008) Evolution of DNA-methylation machinery: DNA methyltransferases and methyl-DNA binding proteins in the amphioxus *Branchiostoma floridae*. *Dev Genes Evol* **218**: 691–701.

Ast, G. (2004) How did alternative splicing evolve? *Nat Rev Genet* **5**: 773–782.

Bai, S.M., Ghoshal, K., Datta, J., Majumder, S., Yoon, S.O. and Jacob, S.T. (2005) DNA methyltransferase 3b regulates nerve growth factor-induced differentiation of PC12 cells by recruiting histone deacetylase 2. *Mol Cell Biol* **25**: 751–766.

Barberan-Soler, S. and Zahler, A.M. (2008) Alternative splicing regulation during *C. elegans* development: splicing factors as regulated targets. *PLoS Genet* **4**: e1000001.

Barchuk, A.R., Cristino, A.S., Kucharski, R., Costa, L.F., Simoes, Z.L.P. and Maleszka, R. (2007) Molecular determinants of caste differentiation in the highly eusocial honeybee *Apis mellifera*. *BMC Dev Biol* **7**: 70.

Baylin, S.B., Herman, J.G., Graff, J.R., Vertino, P.M. and Issa, J.P. (1998) Alterations in DNA methylation: a fundamental aspect of neoplasia. *Adv Cancer Res* **72**: 141–196.

Ben-Porath, I. and Cedar, H. (2001) Epigenetic crosstalk. *Mol Cell* **8**: 933–935.

Berger, S.L., Kouzarides, T., Shiekhattar, R. and Shilatifard, A. (2009) An operational definition of epigenetics. *Genes Dev* **23**: 781–783.

Bestor, T.H. (1990) DNA methylation – evolution of a bacterial immune function into a regulator of gene-expression and genome structure in higher eukaryotes. *Philos Trans R Soc B Biol Sci* **326**: 179–187.

Bestor, T.H. (2000) The DNA methyltransferases of mammals. *Hum Mol Genet* **9**: 2395–2402.

Bird, A.P. (1980) DNA methylation and the frequency of CpG in animal DNA. *Nucleic Acids Res* **8**: 1499–1504.

Bird, A.P. (1995) Gene number, noise-reduction and biological complexity. *Trends Genet* **11**: 94–100.

Bird, A.P., Taggart, M.H. and Smith, B.A. (1979) Methylated and unmethylated DNA compartments in the sea urchin genome. *Cell* **17**: 889–901.

Bock, C. and Lengauer, T. (2008) Computational epigenetics. *Bioinformatics* **24**: 1–10.

Bogdanovic, O. and Veenstra, G.J.C. (2009) DNA methylation and methyl-CpG binding proteins: developmental requirements and function. *Chromosoma* **118**: 549–565.

Bonasio, R., Tu, S. and Reinberg, D. (2010a) Molecular signals of epigenetic states. *Science* **330**: 612–616.

Bonasio, R., Zhang, G., Ye, C., Mutti, N.S., Fang, X., Qin, N. *et al.* (2010b) Genomic comparison of the ants *Camponotus floridanus* and *Harpegnathos saltator*. *Science* **329**: 1068–1071.

Boyes, J. and Bird, A. (1991) DNA methylation inhibits transcription indirectly via a Methyl-cpg binding-protein. *Cell* **64**: 1123–1134.

Chen, T.P., Ueda, Y., Dodge, J.E., Wang, Z.J. and Li, E. (2003) Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b. *Mol Cell Biol* **23**: 5594–5605.

Cheong, J., Yamada, Y., Yamashita, R., Irie, T., Kanai, A., Wakaguri, H. *et al.* (2006) Diverse DNA methylation statuses at alternative promoters of human genes in various tissues. *DNA Res* **13**: 155–167.

Clouaire, T. and Stancheva, I. (2008) Methyl-CpG binding proteins: specialized transcriptional repressors or structural components of chromatin? *Cell Mol Life Sci* **65**: 1509–1522.

Clouaire, T., Heras, J., Merusi, C. and Stancheva, I. (2010) Recruitment of MBD1 to target genes requires sequence-specific interaction of the MBD domain with methylated DNA. *Nucleic Acids Res* **38**: 4620–4634.

Colbourne, J.K., Pfrender, M.E., Gilbert, D., Thomas, W.K., Tucker, A., Oakley, T.H. *et al.* (2011) The ecoresponsive genome of *Daphnia pulex*. *Science* **331**: 555–561.

Duret, L. and Mouchiroud, D. (2000) Determinants of substitution rates in mammalian genes: expression pattern affects selection intensity but not mutation rate. *Mol Biol Evol* **17**: 68–74.

Ehrlich, M., Gamasosa, M.A., Huang, L.H., Midgett, R.M., Kuo, K.C., McCune, R.A. *et al.* (1982) Amount and distribution of 5-methylcytosine in human DNA from different types of tissues or cells. *Nucleic Acids Res* **10**: 2709–2721.

Elango, N. and Yi, S.V. (2008) DNA methylation and structural and functional bimodality of vertebrate promoters. *Mol Biol Evol* **25**: 1602–1608.

Elango, N. and Yi, S.V. (2011) Functional relevance of CpG island length for regulation of gene expression. *Genetics* **187**: 1077–1083.

Elango, N., Kim, S.H., Vigoda, E. and Yi, S.V. (2008) Mutations of different molecular origins exhibit contrasting patterns of regional substitution rate variation. *PLoS Comput Biol* **4**: e1000015.

Elango, N., Hunt, B.G., Goodisman, M.A.D. and Yi, S.V. (2009) DNA methylation is widespread and associated with differential gene expression in castes of the honeybee, *Apis mellifera*. *Proc Natl Acad Sci USA* **106**: 11206–11211.

Evans, J.D. and Wheeler, D.E. (2001) Gene expression and the evolution of insect polyphenisms. *BioEssays* **23**: 62–68.

Fatemi, M., Pao, M.M., Jeong, S., Gal-Yam, E.N., Egger, G., Weisenberger, D.J. *et al.* (2005) Footprinting of mammalian promoters: use of a CpG DNA methyltransferase revealing nucleosome positions at a single molecule level. *Nucleic Acids Res* **33**: e176.

Feng, Q. and Zhang, Y. (2001) The MeCP1 complex represses transcription through preferential binding, remodeling, and deacetylating methylated nucleosomes. *Genes Dev* **15**: 827–832.

Feng, S.H., Cokus, S.J., Zhang, X.Y., Chen, P.Y., Bostick, M., Goll, M.G. *et al.* (2010) Conservation and divergence of methylation patterning in plants and animals. *Proc Natl Acad Sci USA* **107**: 8689–8694.

Field, L.M. (2000) Methylation and expression of amplified esterase genes in the aphid *Myzus persicae* (Sulzer). *Biochem J* **349**: 863–868.

Field, L.M., Crick, S.E. and Devonshire, A.L. (1996) Polymerase chain reaction-based identification of insecticide resistance genes and DNA methylation in the aphid *Myzus persicae* (Sulzer). *Insect Mol Biol* **5**: 197–202.

Foret, S., Kucharski, R., Pittelkow, Y., Lockett, G.A. and Maleszka, R. (2009) Epigenetic regulation of the honey bee transcriptome: unravelling the nature of methylated genes. *BMC Genomics* **10**: 472.

Fraga, M.F., Ballestar, E., Paz, M.F., Ropero, S., Setien, F., Ballestart, M.L. *et al.* (2005) Epigenetic differences arise during the lifetime of monozygotic twins. *Proc Natl Acad Sci USA* **102**: 10604–10609.

Fuks, F., Hurd, P.J., Deplus, R. and Kouzarides, T. (2003) The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase. *Nucleic Acids Res* **31**: 2305–2312.

Futscher, B.W., Oshiro, M.M., Wozniak, R.J., Holtan, N., Hanigan, C.L., Duan, H. *et al.* (2002) Role for DNA methylation in the control of cell type-specific maspin expression. *Nat Genet* **31**: 175–179.

Gaunt, M.W. and Miles, M.A. (2002) An insect molecular clock dates the origin of the insects and accords with palaeontological and biogeographic landmarks. *Mol Biol Evol* **19**: 748–761.

Geiman, T.M., Sankpal, U.T., Robertson, A.K., Zhao, Y.X., Zhao, Y.M. and Robertson, K.D. (2004) DNMT3B interacts with hSNF2H chromatin remodeling enzyme, HDACs 1 and 2, and components of the histone methylation system. *Biochem Biophys Res Commun* **318**: 544–555.

Goll, M.G. and Bestor, T.H. (2005) Eukaryotic cytosine methyltransferases. *Annu Rev Biochem* **74**: 481–514.

Goll, M.G., Kirpekar, F., Maggert, K.A., Yoder, J.A., Hsieh, C.L., Zhang, X.Y. *et al*. (2006) Methylation of tRNA(AsP) by the DNA methyltransferase homolog Dnmt2. *Science* **311**: 395–398.

Grimaldi, D. and Engel, M. (2005) *Evolution of the Insects*. Cambridge University Press, Cambridge.

Haig, D. (2000) The kinship theory of genomic imprinting. *Annu Rev Ecol Syst* **31**: 9–32.

Haines, T.R., Rodenhiser, D.I. and Ainsworth, P.J. (2001) Allele-specific non-CpG methylation of the Nf1 gene during early mouse development. *Dev Biol* **240**: 585–598.

Hata, K., Okano, M., Lei, H. and Li, E. (2002) Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. *Development* **129**: 1983–1993.

Hendrich, B. and Bird, A. (1998) Identification and characterization of a family of mammalian methyl-CpG binding proteins. *Mol Cell Biol* **18**: 6538–6547.

Hendrich, B. and Tweedie, S. (2003) The methyl-CpG binding domain and the evolving role of DNA methylation in animals. *Trends Genet* **19**: 269–277.

Hung, M.S., Karthikeyan, N., Huang, B.L., Koo, H.C., Kiger, J. and Shen, C.K.J. (1999) Drosophila proteins related to vertebrate DNA (5-cytosine) methyltransferases. *Proc Natl Acad Sci USA* **96**: 11940–11945.

Hunt, B.G., Brisson, J.A., Yi, S.V. and Goodisman, M.A.D. (2010) Functional conservation of DNA methylation in the pea aphid and the honeybee. *Genome Biol Evol* **2**: 719–728.

Illingworth, R., Kerr, A., DeSousa, D., Jorgensen, H., Ellis, P., Stalker, J. *et al*. (2008) A novel CpG island set identifies tissue-specific methylation at developmental gene loci. *PLoS Biol* **6**: e22.

International Aphid Genomics Consortium (2010) Genome sequence of the pea aphid *Acyrthosiphon pisum*. *PLoS Biol* **8**: e1000313.

Jablonka, E. and Regev, A. (1995) Gene number, methylation and biological complexity. *Trends Genet* **11**: 383–384.

Jaenisch, R. and Bird, A. (2003) Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Genet* **33**: 245–254.

Jair, K.W., Bachman, K.E., Suzuki, H., Ting, A.H., Rhee, I., Yen, R.W.C. *et al*. (2006) De novo CpG island methylation in human cancer cells. *Cancer Res* **66**: 682–692.

Javierre, B.M., Fernandez, A.F., Richter, J., Al-Shahrour, F., Martin-Subero, J.I., Rodriguez-Ubreva, J. *et al*. (2010) Changes in the pattern of DNA methylation associate with twin discordance in systemic lupus erythematosus. *Genome Res* **20**: 170–179.

Jones, P.A. and Baylin, S.B. (2002) The fundamental role of epigenetic events in cancer. *Nat Rev Genet* **3**: 415–428.

Jones, P.L., Veenstra, G.J.C., Wade, P.A., Vermaak, D., Kass, S.U., Landsberger, N. *et al*. (1998) Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nat Genet* **19**: 187–191.

Jurkowski, T.P., Meusburger, M., Phalke, S., Helm, M., Nellen, W., Reuter, G. *et al*. (2008) Human DNMT2 methylates tRNA(Asp) molecules using a DNA methyltransferase-like catalytic mechanism. *RNA* **14**: 1663–1670.

Kangaspeska, S., Stride, B., Metivier, R., Polycarpou-Schwarz, M., Ibberson, D., Carmouche, R.P. *et al*. (2008) Transient cyclical methylation of promoter DNA. *Nature* **452**: 112–115.

Kato, Y., Kaneda, M., Hata, K., Kumaki, K., Hisano, M., Kohara, Y. *et al*. (2007) Role of the Dnmt3 family in de novo methylation of imprinted and repetitive sequences during male germ cell development in the mouse. *Hum Mol Genet* **16**: 2272–2280.

Kirkness, E.F., Haas, B.J., Sun, W., Braig, H.R., Perotti, M.A., Clark, J.M. *et al*. (2010) Genome sequences of the human body louse and its primary endosymbiont provide insights into the permanent parasitic lifestyle. *Proc Natl Acad Sci USA* **107**: 12168–12173.

Klose, R.J. and Bird, A.P. (2006) Genomic DNA methylation: the mark and its mediators. *Trends Biochem Sci* **31**: 89–97.

Krauss, V., Eisenhardt, C. and Unger, T. (2009) The genome of the stick insect *Medauroidea extradentata* is strongly methylated within genes and repetitive DNA. *PLoS One* **4**: e7223.

Kronauer, D.J.C. (2008) Genomic imprinting and kinship in the social Hymenoptera: what are the predictions? *J Theor Biol* **254**: 737–740.

Kronforst, M.R., Gilley, D.C., Strassmann, J.E. and Queller, D.C. (2008) DNA methylation is widespread across social Hymenoptera. *Curr Biol* **18**: R287–R288.

Kucharski, R., Maleszka, J., Foret, S. and Maleszka, R. (2008) Nutritional control of reproductive status in honeybees via DNA methylation. *Science* **319**: 1827–1830.

Li, E., Beard, C. and Jaenisch, R. (1993) Role for DNA methylation in genomic imprinting. *Nature* **366**: 362–365.

Li, Y.R., Zhu, J.D., Tian, G., Li, N., Li, Q.B., Ye, M.Z. *et al*. (2010) The DNA methylome of human peripheral blood mononuclear cells. *PLoS Biol* **8**: e1000533.

Lister, R., Pelizzola, M., Dowen, R.H., Hawkins, R.D., Hon, G., Tonti-Filippini, J. *et al*. (2009) Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **462**: 315–322.

Lockett, G.A., Helliwell, P. and Maleszka, R. (2010) Involvement of DNA methylation in memory processing in the honey bee. *Neuroreport* **21**: 812–816.

Lopezrodas, G., Brosch, G., Georgieva, E.I., Sendra, R., Franco, L. and Loidl, P. (1993) Histone deacetylase - a key enzyme for the binding of regulatory proteins to chromatin. *FEBS Lett* **317**: 175–180.

Lyko, F., Foret, S., Kucharski, R., Wolf, S., Falckenhayn, C. and Maleszka, R. (2010) The honey bee epigenomes: differential methylation of brain DNA in queens and workers. *PLoS Biol* **8**: e1000506.

Maleszka, R. (2008) Epigenetic integration of environmental and genomic signals in honey bees. *Epigenetics* **3**: 188–192.

Mandrioli, M. (2007) A new synthesis in epigenetics: towards a unified function of DNA methylation from invertebrates to vertebrates. *Cell Mol Life Sci* **64**: 2522–2524.

Mandrioli, M. and Volpi, N. (2003) The genome of the lepidopteran *Mamestra brassicae* has a vertebrate-like content of methyl-cytosine. *Genetica* **119**: 187–191.

Marais, G. (2003) Biased gene conversion: implications for genome and sex evolution. *Trends Genet* **19**: 330–338.
Margueron, R. and Reinberg, D. (2010) Chromatin structure and the inheritance of epigenetic information. *Nat Rev Genet* **11**: 285–296.
Marhold, J., Rothe, N., Pauli, A., Mund, C., Kuehle, K., Brueckner, B. *et al*. (2004) Conservation of DNA methylation in dipteran insects. *Insect Mol Biol* **13**: 117–123.
Maunakea, A.K., Nagarajan, R.P., Bilenky, M., Ballinger, T.J., D'Souza, C., Fouse, S.D. *et al*. (2010) Conserved role of intragenic DNA methylation in regulating alternative promoters. *Nature* **466**: 253–257.
Mayer, W., Niveleau, A., Walter, J., Fundele, R. and Haaf, T. (2000) Embryogenesis – demethylation of the zygotic paternal genome. *Nature* **403**: 501–502.
Merlo, A., Herman, J.G., Mao, L., Lee, D.J., Gabrielson, E., Burger, P.C. *et al*. (1995) 5′ CpG island methylation is associated with transcriptional silencing of the tumor-suppressor p16/CDKN2/MTS1 in human cancers. *Nat Med* **1**: 686–692.
Metivier, R., Gallais, R., Tiffoche, C., Le Peron, C., Jurkowska, R.Z., Carmouche, R.P. *et al*. (2008) Cyclical DNA methylation of a transcriptionally active promoter. *Nature* **452**: 45–50.
Miller, C.A. and Sweatt, J.D. (2007) Covalent modification of DNA regulates memory formation. *Neuron* **53**: 857–869.
Moczek, A.P. and Snell-Rood, E.C. (2008) The basis of bee-ing different: the role of gene silencing in plasticity. *Evol Dev* **10**: 511–513.
Monk, M., Boubelik, M. and Lehnert, S. (1987) Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ-cell lineages during mouse embryo development. *Development* **99**: 371–382.
Nasonia Genome Working Group (2010) Functional and evolutionary insights from the genomes of three parasitoid *Nasonia* species. *Science* **327**: 343–348.
Okamura, K., Matsumoto, K. and Nakai, K. (2010) Gradual transition from mosaic to global DNA methylation patterns during deuterostome evolution. *BMC Bioinformatics* **11**: S2.
Okano, M., Bell, D.W., Haber, D.A. and Li, E. (1999) DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**: 247–257.
Okitsu, C.Y. and Hsieh, C.L. (2007) DNA methylation dictates histone H3K4 methylation. *Mol Cell Biol* **27**: 2746–2757.
O'Neill, R.J.W., O'Neill, M.J. and Graves, J.A.M. (1998) Undermethylation associated with retroelement activation and chromosome remodelling in an interspecific mammalian hybrid. *Nature* **393**: 68–72.
Ono, M., Swanson, J.J., Field, L.M., Devonshire, A.L. and Siegfried, B.D. (1999) Amplification and methylation of an esterase gene associated with insecticide-resistance in greenbugs, *Schizaphis graminum* (Rondani) (Homoptera : Aphididae). *Insect Biochem Mol Biol* **29**: 1065–1073.
Ooi, S.K.T. and Bestor, T.H. (2008) The colorful history of active DNA demethylation. *Cell* **133**: 1145–1148.
Ott, J. (1979) Detection of rare major genes in lipid-levels. *Hum Genet* **51**: 79–91.
Pal, C., Papp, B. and Lercher, M.J. (2006) An integrated view of protein evolution. *Nat Rev Genet* **7**: 337–348.
Queller, D.C. (2003) Theory of genomic imprinting conflict in social insects. *BMC Evol Biol* **3**: 23.
Rae, P.M.M. and Steele, R.E. (1979) Absence of cytosine methylation at C-C-G-G and G-C-G-C sites in the rDNA coding regions and intervening sequences of *Drosophila* and the rDNA of other higher insects. *Nucleic Acids Res* **6**: 2987–2995.
Regev, A., Lamb, M. and Jablonka, E. (1998) The role of DNA methylation in invertebrates: developmental regulation or genome defense? *Mol Biol Evol* **15**: 880–891.
Reik, W. and Walter, J. (2001) Genomic imprinting: parental influence on the genome. *Nat Rev Genet* **2**: 21–32.
Rountree, M.R. and Selker, E.U. (1997) DNA methylation inhibits elongation but not initiation of transcription in *Neurospora crassa*. *Genes Dev* **11**: 2383–2395.
Savard, J., Tautz, D., Richards, S., Weinstock, G.M., Gibbs, R.A., Werren, J.H. *et al*. (2006) Phylogenomic analysis reveals bees and wasps (Hymenoptera) at the base of the radiation of Holometabolous insects. *Genome Res* **16**: 1334–1338.
Saxonov, S., Berg, P. and Brutlag, D.L. (2006) A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proc Natl Acad Sci USA* **103**: 1412–1417.
Schaefer, M. and Lyko, F. (2010) Lack of evidence for DNA methylation of Invader4 retroelements in *Drosophila* and implications for Dnmt2-mediated epigenetic regulation. *Nat Genet* **42**: 920–921.
Sharif, J., Endo, T.A., Toyoda, T. and Koseki, H. (2010) Divergence of CpG island promoters: a consequence or cause of evolution? *Dev Growth Differ* **52**: 545–554.
Shen, J.C., Rideout, W.M. and Jones, P.A. (1994) The rate of hydrolytic deamination of 5-methylcytosine in double-stranded DNA. *Nucleic Acids Res* **22**: 972–976.
Shimizu, T.S., Takahashi, K. and Tomita, M. (1997) CpG distribution patterns in methylated and non-methylated species. *Gene* **205**: 103–107.
Simmen, M.W., Leitgeb, S., Charlton, J., Jones, S.J.M., Harris, B.R., Clark, V.H. *et al*. (1999) Nonmethylated transposable elements and methylated genes in a chordate genome. *Science* **283**: 1164–1167.
Simpson, V.J., Johnson, T.E. and Hammen, R.F. (1986) *Caenorhabditis elegans* DNA does not contain 5-methylcytosine at any time during development or aging. *Nucleic Acids Res* **14**: 6711–6719.
Smith, C.D., Zimin, A., Holt, C., Abouheif, E., Benton, R., Cash, E. *et al*. (2011) Draft genome of the globally widespread and invasive Argentine ant (*Linepithema humile*). *Proc Natl Acad Sci USA* **108**: 5673–5678.
Smith, C.R., Toth, A.L., Suarez, A.V. and Robinson, G.E. (2008) Genetic and genomic analyses of the division of labour in insect societies. *Nat Rev Genet* **9**: 735–748.
Smith, C.R., Smith, C.D., Robertson, H.M., Helmkampf, M., Zimin, A., Yandell, M. *et al*. (2011) Draft genome of the red harvester ant *Pogonomyrmex barbatus*. *Proc Natl Acad Sci USA* **108**: 5667–5672.
Soma, S. and Rao, S.R.V. (1992) 5-Methylcytosine content in *Gryllotalpa fossor* (Orthoptera). *Genome* **35**: 163–166.
Strassmann, J.E. and Queller, D.C. (2007) Insect societies as divided organisms: the complexities of purpose and cross-purpose. *Proc Natl Acad Sci USA* **104**: 8619–8626.
Suen, G., Teiling, C., Li, L., Holt, C., Abouheif, E., Bornberg-Bauer, E. *et al*. (2011) The genome sequence of the leaf-cutter ant *Atta cephalotes* reveals insights into its obligate symbiotic lifestyle. *PLoS Genet* **7**: e1002007.

Suzuki, M.M. and Bird, A. (2008) DNA methylation landscapes: provocative insights from epigenomics. *Nat Rev Genet* **9**: 465–476.

Suzuki, M.M., Kerr, A.R.W., De Sousa, D. and Bird, A. (2007) CpG methylation is targeted to transcription units in an invertebrate genome. *Genome Res* **17**: 625–631.

Tamaru, H. and Selker, E.U. (2001) A histone H3 methyltransferase controls DNA methylation in *Neurospora crassa*. *Nature* **414**: 277–283.

Tribolium Genome Sequencing Consortium (2008) The genome of the model beetle and pest *Tribolium castaneum*. *Nature* **452**: 949–955.

Tweedie, S., Ng, H.H., Barlow, A.L., Turner, B.M., Hendrich, B. and Bird, A. (1999) Vestiges of a DNA methylation system in Drosophila melanogaster? *Nat Genet* **23**: 389–390.

Urieli-Shoval, S., Gruenbaum, Y., Sedat, J. and Razin, A. (1982) The absence of detectable methylated bases in *Drosophila melanogaster* DNA. *FEBS Lett* **146**: 148–152.

Vandegehuchte, M.B., Lemiere, F. and Janssen, C.R. (2009) Quantitative DNA-methylation in *Daphnia magna* and effects of multigeneration Zn exposure. *Comp Biochem Physiol C Toxicol Pharmacol* **150**: 343–348.

Vanyushin, F. (2005) Methylation of adenine residues in DNA of eukaryotes. *Mol Biol* **39**: 557–566.

Vire, E., Brenner, C., Deplus, R., Blanchon, L., Fraga, M., Didelot, C. *et al*. (2006) The Polycomb group protein EZH2 directly controls DNA methylation. *Nature* **439**: 871–874.

Walsh, T.K., Brisson, J.A., Robertson, H.M., Gordon, K., Jaubert-Possamai, S., Tagu, D. *et al*. (2010) A functional DNA methylation system in the pea aphid, *Acyrthosiphon pisum*. *Insect Mol Biol* **19**: 215–228.

Wang, Y. and Leung, F. (2009) In silico prediction of two classes of honeybee genes with CpG deficiency or CpG enrichment and sorting according to gene ontology classes. *J Mol Evol* **68**: 700–705.

Wang, Y., Jorda, M., Jones, P.L., Maleszka, R., Ling, X., Robertson, H.M. *et al*. (2006) Functional CpG methylation system in a social insect. *Science* **314**: 645–647.

Watt, F. and Molloy, P.L. (1988) Cytosine methylation prevents binding to DNA of a hela-cell transcription factor required for optimal expression of the adenovirus major late promoter. *Genes Dev* **2**: 1136–1143.

Weaver, N. (1966) Physiology of caste determination. *Annu Rev Entomol* **11**: 79–102.

Weber, M., Hellmann, I., Stadler, M.B., Ramos, L., Paabo, S., Rebhan, M. *et al*. (2007) Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat Genet* **39**: 457–466.

Wheeler, D.E. (1986) Developmental and physiological determinants of caste in social Hymenoptera: evolutionary implications. *Am Nat* **128**: 13–34.

Wilkins, J.F. and Haig, D. (2003) What good is genomic imprinting: the function of parent-specific gene expression. *Nat Rev Genet* **4**: 359–368.

Wolffe, A.P. and Matzke, M.A. (1999) Epigenetics: regulation through repression. *Science* **286**: 481–486.

Wu, S.C. and Zhang, Y. (2010) Active DNA demethylation: many roads lead to Rome. *Nat Rev Mol Cell Biol* **11**: 607–620.

Wurm, Y., Wang, J., Riba-Grognuz, O., Corona, M., Nygaard, S., Hunt, B.G. *et al*. (2011) The genome of the fire ant *Solenopsis invicta*. *Proc Natl Acad Sci USA* **108**: 5679–5684.

Xiang, H., Zhu, J.D., Chen, Q., Dai, F.Y., Li, X., Li, M.W. *et al*. (2010) Single base-resolution methylome of the silkworm reveals a sparse epigenomic map. *Nat Biotechnol* **28**: 516–520.

Yi, S.V. and Goodisman, M.A.D. (2009) Computational approaches for understanding the evolution of DNA methylation in animals. *Epigenetics* **4**: 551–556.

Yoder, J.A., Walsh, C.P. and Bestor, T.H. (1997) Cytosine methylation and the ecology of intragenomic parasites. *Trends Genet* **13**: 335–340.

Young, J.I., Hong, E.P., Castle, J.C., Crespo-Barreto, J., Bowman, A.B., Rose, M.F. *et al*. (2006) Regulation of RNA splicing by the methylation-dependent transcriptional repressor methyl-CpG binding protein 2. *Proc Natl Acad Sci USA* **103**: 1656–1656.

Zemach, A., McDaniel, I.E., Silva, P. and Zilberman, D. (2010) Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science* **328**: 916–919.

Zeng, J. and Yi, S.V. (2010) DNA methylation and genome evolution in honeybee: gene length, expression, functional enrichment covary with the evolutionary signature of DNA methylation. *Genome Biol Evol* **2**: 770–780.

Zilberman, D., Gehring, M., Tran, R.K., Ballinger, T. and Henikoff, S. (2007) Genome-wide analysis of *Arabidopsis thaliana* DNA methylation uncovers an interdependence between methylation and transcription. *Nat Genet* **39**: 61–69.