INVITED REVIEW

DNA methylation in insects

L. M. Field*, F. Lyko†, M. Mandrioli‡ and G. Prantera§

*BCH Division, Rothamsted Research, Harpenden, Herts, UK; †Research Group Epigenetics, Deutsches Krebsforschungszentrum, Heidelberg, Germany; ‡Laboratory of Cytogenetics, Department of Animal Biology, University of Modena and Reggio Emilia, Modena, Italy; and §Section of Genetics, Department of Agrobiology & Agrochemistry, University of Tuscia, Viterbo, Italy

Abstract

Cytosine DNA methylation has been demonstrated in numerous eukaryotic organisms and has been shown to play an important role in human disease. The function of DNA methylation has been studied extensively in vertebrates, but establishing its primary role has proved difficult and controversial. Analysing methylation in insects has indicated an apparent functional diversity that seems to argue against a strict functional conservation. To investigate this hypothesis, we here assess the data reported in four different insect species in which DNA methylation has been analysed more thoroughly: the fruit fly Drosophila melanogaster, the cabbage moth Mamestra brassicae, the peach-potato aphid Myzus persicae and the mealybug Planococcus citri.

Keywords: DNA methylation, gene expression, imprinting, transposons, epigenetics.

The function of DNA methylation in vertebrates and plants

It is well known that a variable proportion of cytosine residues in eukaryotic genomes is methylated in the form of 5-methylcytosine. The percentage of methylated cytosines ranges from 0–3% in insects, 5% in mammals and birds, 10% in fish and amphibians to more than 30% in some plants (Adams, 1996). DNA methylation has been associated with numerous functions, depending on the model organism and the experimental context. In general, the presence of DNA methylation, in and around the promoter regions of genes, is associated with gene silencing and loss of methylation accompanies transcription. More rigorous functional analyses have so far only been performed in vertebrates and plant systems that provide an opportunity for the generation of DNA methyltransferase knockout mutants.

In mice, DNA methylation has been shown to be essential for proper embryonic development (Li et al., 1992). Similarly, a loss of DNA methylation caused developmental defects in Xenopus embryos (Stancheva & Meehan, 2000). Analogous results have been reported in plants, in which reduced levels of DNA methylation were shown to be responsible for a large number of developmental abnormalities (Finnegan et al., 1996; Ronemus et al., 1996). Superficial similarities in some of the developmental defects have been interpreted to reflect a conserved function of DNA methylation. However, the molecular consequences of genomic DNA methylation have been found to be surprisingly diverse. It has been generally assumed that DNA methylation has a role in regulating global gene activity even if, to date, only a very limited number of genetic loci have been shown to be misexpressed in DNA methyltransferase mutants. These include a subset of imprinted genes and the IAP retroelement in mice (Li et al., 1993; Walsh et al., 1998) and a small group of developmental genes and the CAC1 transposon in Arabidopsis thaliana (Kakutani et al., 1996; Miura et al., 2001). On the cellular level, loss of DNA methylation has been shown to affect apoptosis in mice (Jackson-Grusby et al., 2001) and Xenopus (Stancheva et al., 2001), X-chromosome inactivation and chromosomal stability in mice (Panning & Jaenisch, 1996; Gaudet et al., 2003) and the overall chromosome organization in Arabidopsis (Soppe et al., 2002). It is possible that the diversity of these effects reflects an as yet unidentified higher functional role of DNA methylation. In this respect, the molecular consequences of DNA methylation would not be uniform but rather depend on the precise context of epigenetic signals.

In vertebrates, 5-methylcytosine is present in CpG dinucleotides and the methylation patterns can arise de novo during
development and be maintained or lost by the action of maintenance methylases and demethylases (Wolffe et al., 1999). There is good evidence that the control of transcription also involves proteins, which bind specifically to methylated DNA, histone modification complexes and local chromatin remodelling. For reviews on the proteins involved in DNA methylation and chromatin formation see Bestor (2000), Ballestar & Wolff (2001) and Wade (2001).

Although there have been many studies on DNA methylation, establishing its primary role that has led to its widespread distribution in higher organisms has proved controversial. Bird (1995) has suggested that during evolution the transition from invertebrates to vertebrates was accompanied by an increase in gene number, made possible by the use of DNA methylation to repress transcriptional background noise. This is supported by the finding that the transition from fractional to global methylation of genomes occurred close to the origin of vertebrates (Tweedie et al., 1997). However, as the number of genes in different organisms is becoming available from genome sequencing the correlation between DNA methylation patterns and the number of genes appears less evident. An alternative view is that the primary function of DNA methylation in vertebrates is to suppress parasitic sequence elements, with the control of gene expression being either secondary or even illusory (Yoder et al., 1997; Walsh & Bestor, 1999). It has also been proposed that DNA methylation acts to ‘memorize’ patterns of gene activity by stabilizing gene silencing brought about by other means (Bird, 2002).

DNA methylation in insects

The presence of 5-methylcytosine has been reported in several insect species belonging to various orders (Table 1). In addition, there are several reports suggesting an absence of methylated DNA from other insect species. However, these reports need to be interpreted with caution because DNA methylation can be restricted to particular developmental stages and to target sequences (see below). It is possible that DNA methylation is conserved in most, if not all, insect species.

The role of DNA methylation in insects is still poorly understood. This is due largely to the lack of suitable model systems for functional analyses. The available data demonstrate varying levels of methylation and do not seem to indicate a conserved function. This apparent functional diversity might be related to the high diversity of insect species and would argue against a strict evolutionary conservation of DNA methylation. In the following sections we discuss results from four different insect species in which DNA methylation has been analysed more thoroughly: the fruit fly Drosophila melanogaster, the cabbage moth Mamestra brassicae, the peach-potato aphid Myzus persicae and the citrus mealybug Planococcus citri. We summarize the DNA methylation data from these species and analyse potential functional associations. This has important implications for understanding the evolution of DNA methylation.

**Drosophila melanogaster: non-CpG methylation mediated by Dnmt2**

DNA methylation in D. melanogaster is characterized by several features that made it highly elusive over a long period of time (Lyko, 2001): (1) the overall methylation level is rather low, with less than 1% of the cytosine residues being methylated; (2) the highest levels of DNA methylation are found in early embryos, which yield only limited amounts of DNA for biochemical or molecular characterization;

<table>
<thead>
<tr>
<th>Table 1. DNA methylation in major insect orders. Not determined indicates the absence of published data on the genomic DNA methylation state</th>
</tr>
</thead>
<tbody>
<tr>
<td>Order</td>
</tr>
<tr>
<td>Coleoptera (beetles)</td>
</tr>
<tr>
<td>Diptera (flies)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Hemiptera (bugs)</td>
</tr>
<tr>
<td>Homoptera (aphids, cicadas, scales)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Hymenoptera (bees, ants, wasps)</td>
</tr>
<tr>
<td>Lepidoptera (butterflies, moths)</td>
</tr>
<tr>
<td>Odonata (dragonflies and damselflies)</td>
</tr>
<tr>
<td>Orthoptera (grasshoppers, crickets)</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

© 2004 The Royal Entomological Society, Insect Molecular Biology, 13, 109–115
(3) most of the 5-methylcytosine is found in the context of non-CpG dinucleotides, which rendered traditional CpG-specific assays ineffective. Together, these characteristics establish a unique profile for DNA methylation in this model organism.

DNA methylation in *D. melanogaster* is catalysed by an enzyme that belongs to the Dnmt2 family of DNA methyltransferases (Kunert *et al.*, 2003). Dnmt2 genes are widely conserved in evolution but their functional characterization has only recently been initiated. Overexpression of the *Drosophila* Dnmt2 homologue caused increased DNA methylation at CpT and CpA dinucleotides (Kunert *et al.*, 2003). This distinguishes the fly from vertebrates and plants, where DNA methylation is concentrated in CpG dinucleotides. The high level of CpG methylation in the latter organisms is presumably due to the dominant CpG-specific activity of maintenance DNA methyltransferases, such as Dnmt1 (Ramsahoye *et al.*, 2000). However, the presence of small amounts of non-CpG methylation has now also been confirmed in vertebrates and plants. An involvement of Dnmt2 proteins in this phenomenon seems likely, but has not yet been demonstrated.

The apparent restriction of *D. melanogaster* methylation to non-CpG sequence contexts raises a number of intriguing questions. In most other organisms, DNA methylation patterns are maintained by post-replicative copying from the parental strand to the newly synthesized strand. However, this mechanism requires the presence of symmetric methylation in the context of CpG or CpNpG sequences. There is presently no evidence for such symmetric methylation in the *D. melanogaster* genome, and it is likely that epigenetic information is maintained by other mechanisms. An attractive possibility is provided by methyl-DNA binding proteins (Hendrich & Tweedie, 2003). These proteins have been thoroughly characterized in vertebrate systems and they provide a link between DNA methylation and epigenetic chromatin structures. The *D. melanogaster* genome also encodes a protein (MBD2/3) with extended homologies to vertebrate methyl-DNA binding proteins. Interestingly, MBD2/3 shows a dynamic association with embryonic DNA that coincides with the peak of DNA methylation (Marhold *et al.*, 2002). MBD2/3 might thus initiate the maintenance of epigenetic information by binding to methylated DNA and by simultaneous recruitment of histone-modifying enzymes. This mechanism would result in the establishment of repressive chromatin structures that could be maintained stably over many cell generations (Vermaak *et al.*, 2003).

The role of these processes in *D. melanogaster* development represents an issue that remains to be analyzed. Unfortunately, Dnmt2 mutant strains are not yet available. Depletion of Dnmt2 protein from embryos by RNA interference did not have a major effect on embryonic development (Kunert *et al.*, 2003). In this respect, *Drosophila* might be similar to the filamentous fungus *Neospora crassa*, which has been shown to tolerate a complete loss of genomic DNA methylation (Kouzminova & Selker, 2001). Similarly, Dnmt2 knockout mice did not show any major phenotypes (E. Li, personal communication). A more subtle function remains a viable possibility, but its experimental analysis will require the generation of a mutant allele.

**Mamestra brassicae**: high levels of genome methylation

High-performance liquid chromatography (HPLC) analysis on *M. brassicae* genomic DNA, isolated from larval and adult tissues and from *in vitro* cultured cells, revealed that the 5-methylcytosine content was approximately 10% (Mandrioli & Volpi, 2003). The cabbage moth DNA methylation level therefore represents the highest reported in insects and is similar to that of vertebrates. Methylation analyses with restriction enzymes showed that a portion of the *M. brassicae* genome was methylated at CpG sites and that CpG targets were not clustered (Mandrioli & Volpi, 2003). Additional experiments indicated the presence of methylation not only in the CpG doublets but also at the outer C of the 5′-CCGG-3′ sequence (Mandrioli & Volpi, 2003). This hypothesis has been confirmed using the de-methylating agent 5-aza-cytidine. *Msp*I digestion is, in fact, inhibited by the methylation of the outer C of the 5′-CCGG-3′ sequence that could be erased using a de-methylating agent. After 5-aza-cytidine treatment, the digestion pattern of *Msp*I DNAs was far more extensive than that of untreated genomic DNA, showing that a significant level of methylation was present at the outer C of the 5′-CCGG-3′ sequence (Mandrioli & Volpi, 2003).

In order to begin to analyze the role of methylation in *M. brassicae*, the methylation status of repeated DNAs has been studied. Comparison of the restriction pattern of *Msp*I and *Hpa*II after hybridization with the satellite DNA *MBSAT*1, with the genes coding for 5S and 28S rDNA and with the transposons *hobo*, *mariner*, R1 and TRAS1 did not reveal any differences, indicating the absence of CpG methylation in all of these repetitive sequences (Mandrioli, 2002, 2003a,b; Mandrioli *et al.*, 2002; Mandrioli & Volpi, 2003). The results on transposons are intriguing because mobile DNAs are generally heavily methylated in vertebrate and plant genomes, suggesting that cytosine methylation could represent a conserved defence mechanism against transposition (Yoder *et al.*, 1997; Bestor, 1999). Methylation of cytosine residues can, in fact, silence the expression of transposon-encoded genes, prevent transposon-mediated DNA rearrangements and silence the read-through transcription from transposon promoters into neighbouring host genes (Yoder *et al.*, 1997; Bender, 1998; Bestor, 1999). However, this role of methylation is still controversial and some authors have suggested that the host-defence function of methylation represents a vertebrate-specific
adaptation instead of the primary function (Bird, 2002). Indeed, transposable elements were also found to be non-methylated in other invertebrates (Tweedie et al., 1997; Simmen et al., 1999).

Myzus persicae: methylation of amplified esterase genes

The peach-potato aphid, *My. persicae*, has developed resistance to a number of insecticides by overproducing insecticide-detoxifying esterases. This overproduction results from amplification of esterase genes with up to eighty copies present in the most resistant insects (Field et al., 1999). There are two slightly different forms of the gene, E4 and FE4, and early experiments with methylation-sensitive restriction enzymes revealed that the amplified genes have 5-methylcytosine, present in CpG doublets, in and around the genes, with the single copy gene being unmethylated (Field et al., 1989). Surprisingly, loss of DNA methylation occurred simultaneously with loss of E4 gene expression in clonal lines of *M. persicae* (Hick et al., 1996). One possible interpretation of these data was that the methylation occurred in response to the amplification event, in a similar way to methylation being associated with the presence of parasitic DNA in vertebrates and transgenes in plants. It is also interesting that in another aphid species, which has developed insecticide resistance by amplifying an esterase gene, the amplified sequences are again methylated (Ono et al., 1999). However, in this case there is no evidence for loss of resistance or loss of methylation. In addition, there is no such methylation in the amplified esterase genes found in the mosquito *Culex pipiens quinquefasciatus* (M. Raymond, personal communication) and no reports that other amplified insect genes are methylated. Thus, evidence to date would support the view that amplified genes are only methylated in aphid species.

The use of the more sophisticated techniques of CpG profiling and bisulphite sequencing has allowed the precise location of the 5-methylcytosine, showing that it is present within the E4 genes but absent from the upstream regions including the CpG-rich region around the start of transcription (Field, 2000). This means that the methylation would be unlikely to interfere with transcription. This work also demonstrated that the number of CpG doublets is depleted within the E4 gene. Because mutation of 5-methylcytosine over evolutionary time leads to a depletion of CpGs (Shimizu et al., 1997), this suggests that the E4 genes have been methylated for a long period, raising the question of what selection pressures could have maintained it in the aphid genes. It is possible that the methylation of E4 is maintained because it plays a positive role in its transcription. The idea that methylation within genes can be positively correlated with transcription has been suggested by Simmen et al. (1999), who suggested that methylation within coding regions can prevent the production of incorrectly initiated transcripts by silencing expression from spurious promoters. It is interesting that the bisulphite sequencing of regions in and around the E4 gene shows that, as in vertebrates, the 5-methylcytosine is confined to CpG doublets, the first such report for an insect gene (Field, 2000).

Although there is a great deal of information about the methylation of E4 genes in *My. persicae* it seems that the methylation in the rest of the genome is at a low level, being much less than in human DNA but slightly more than in *Drosophila* (Field, 2000). This study also showed that the aphid genome lacks the highly methylated fraction typical of invertebrates such as sea urchin, and seen as a non-digested region of DNA in HpaII digests. There is also evidence that the loss of methylation of the E4 genes is not accompanied by a global loss of methylation within the aphid genome. Unfortunately, the wider role of 5-methylcytosine in aphids is not known and whether any other genes are methylated has not been established. However, the aphid clearly has the cellular enzymes necessary to methylate and demethylate DNA although it has not proved possible to clone regions of DNA encoding methyl-binding domains, homologous to those found in *Drosophila* and many other species (L. M. Field, unpublished data).

Planococcus citri: DNA methylation and genomic imprinting

The term genomic imprinting indicates the epigenetic phenomenon by which homologous alleles or chromosomes behave differently depending on the sex of the parent from which they are inherited (Crouse, 1960; Moore & Haig, 1991). The most striking examples of genomic (or chromosome) imprinting were first described in insects. It was in the early 1960s that the phenomenon was first observed in Sciaridae flies and the term coined by Helen Crouse (1960). At about the same time, a genome-wide imprinting phenomenon was described in the lecanoid coccids, or mealybugs (Brown & Nur, 1964; Nur, 1990). In male mealybugs, a whole haploid chromosome set becomes precociously heterochromatic, and hence inactive, during embryogenesis (Hughes-Schrader, 1948). In females, both haploid chromosome sets remain euchromatic and active. The haploid complement undergoing facultative heterochromatization is invariably that of paternal origin (Brown & Nelson-Rees, 1961). This implies that in male mealybugs at the onset of heterochromatization, after the seventh cleavage division (Bongiorni et al., 2001; Bongiorni & Prantera, 2003), the two haploid chromosome sets are still distinguishable, and hence differentially imprinted.

In mammals, the role of CpG methylation as the epigenetic mark responsible for genomic imprinting has been clearly established (Razin & Cedar, 1994; Feil & Khosla, 1999). In mealybugs, the possible role of DNA methylation
as an imprinting signal was first investigated by Scarbrough et al. (1984). HPLC was used to estimate the total amount of cytosine methylation in male and female DNA, showing that in one of the two species examined, there was a significant difference in DNA methylation between the two sexes, with the DNA of the males more methylated than that of the females. Contrasting results were obtained by Buglia et al. (1999), who failed to demonstrate the presence of cytosine methylation in mealybugs and hence argued against a correlation between DNA methylation and chromosome imprinting. However, these results contrast with several findings showing that coccids not only contain methylcytosine (Scarborough et al., 1984; Bongiorni et al., 1999) but also a CpG methyltransferase (D. Bizzaro, personal communication). Interestingly, a CpA methyltransferase has been also described in coccids (Devajyothi & Brahmachari, 1992).

A methylation difference between the two sexes could not be considered per se as proof of a correlation between DNA methylation and imprinting, unless it could be related directly to the different functional fate of the two haploid chromosome sets in males. This specific issue was addressed by Bongiorni et al. (1999), who investigated the presence of CpG methylation at both the DNA and the chromosome level in the mealybug P. citri and obtained molecular and cytological evidence confirming the existence of DNA methylation. Furthermore, they showed that paternally derived chromosomes were hypomethylated with respect to maternally derived chromosomes in both male and female embryos. This result suggests that DNA methylation also acts as an epigenetic mark in coccids, with DNA hypomethylation being the mark for chromosome inactivation, rather than that for transcriptional activity as occurs in vertebrates (Bongiorni et al., 1999; Bongiorni & Prantera, 2003). The positive correlation between DNA methylation and transcriptional activity found in mealybugs is also consistent with the results in M. persicae discussed above (Field et al., 1989; Hick et al., 1996). As a whole, these results support the notion that in invertebrates DNA methylation is not a mechanism to promote and/or stabilize gene silencing.

Conclusions

Several features of DNA methylation in insects render these species unique and insightful for studying the function and the evolution of this phenomenon. The well-known role of DNA methylation in vertebrates as an epigenetic mechanism to silence inactive genes permanently (Bird, 2002) appears not to have been conserved in the insects studied. Moreover, the data available in these insect species contradict a role of DNA methylation as a genome defence mechanism against mobile elements (Yoder et al., 1997), which supports such reports in other invertebrates (Simmen et al., 1999). Lastly, the presence of cytosine methylation outside the canonical symmetrical CpG doublets indicates a role of DNA methylation that does not require a faithful mitotic transmission of methylcytosine patterns. Consistent with earlier analyses (Colot & Rossignol, 1999), the data reported in insects thus argue against a conserved function retained by DNA methylation (Fig. 1) and confirm a discontinuity in its functional role from invertebrates to vertebrates. Conceivably, the functional manifestations of DNA methylation are much more diverse than originally thought: DNA methylation can be established, propagated and erased. Moreover, methylated DNA can be the preferential target of specific proteins, or, conversely, can inhibit the binding of specific proteins. In this respect, DNA methylation appears to be a conserved, versatile epigenetic mark that can be recruited for different functions in different taxa.

Acknowledgements

We thank E. Li, M. Raymond and D. Bizzaro for allowing us to cite unpublished observations. This work was supported by grants from the Deutsche Forschungsgemeinschaft (to F.L.), University of Modena and Reggio Emilia (M.M.) and from Italian MIUR (to G.P.). Rothamsted Research receives grant-aided support from the BBSRC of the UK.

References


