Methylation and expression of amplified esterase genes in the aphid *Myzus persicae* (Sulzer)

Linda M. FIELD[1]

IACR–Rothamsted, Harpenden, Herts. AL5 2JQ, U.K.

Most populations of the aphid *Myzus persicae* have amplified genes (up to 80 copies) encoding the insecticide-detoxifying esterase E4. This paper reports the analysis of methylation of the *E4* gene and its flanking DNA with the use of methylation-sensitive restriction enzymes, CpG profiling and bisulphite sequencing. In combination these show that *E4* has 5-methylcytosine confined to CpG doublets, as previously shown for vertebrate genomes; this is the first such report for an insect gene. The methylation is present within the gene but absent from upstream regions, including the 5′ CpG-rich region around the start of transcription, and from 3′ flanking DNA. Methylated *E4* genes are expressed; loss of the 5-methylcytosine is correlated with a loss of transcription, although this is not accompanied by a global loss of the 5-methylcytosine present in the aphid genome. These results suggest that the methylation of *E4* has a positive role in expression, and call into question the widely held view that methylation in invertebrate genomes is confined to regions that do not contain genes and that methylation is always associated with gene silencing.

Key words: DNA methylation, gene amplification, gene expression, insect.

INTRODUCTION

It is well established that, in addition to the four familiar bases that make up eukaryotic DNA, there is a fifth base, 5-methylcytosine (5mC) that can be present in significant levels. The percentage of cytosine bases with this modification varies between species, ranging from 0–1 % in insects, approx. 5 % in mammals and birds, approx. 10 % in fish and amphibians, to more than 30 % in some plants [1]. In vertebrates 5mC is present in CpG doublets and the methylation patterns can either be inherited or arise *de novo* during development and be maintained or lost by the action of maintenance methylases and demethylases [2].

Many functions for eukaryotic DNA methylation have been proposed, generally involving the presence of 5mC in promoters leading to gene silencing, which in turn can control such diverse cellular activities as genome defence, genetic switching during development, X-chromosome inactivation and imprinting [1,3]. Despite extensive studies, the precise way in which 5mC represses transcription and its relative importance for different aspects of eukaryotic genome functioning have proved difficult questions. The former is now beginning to be resolved by the isolation of proteins that bind specifically to 5′-methylated DNA, recruit histone deacetylase complexes and mediate local chromatin modelling and inactivation of gene expression [4,5]. There is also good evidence from transgene experiments that DNA methylation does repress transcription *in vivo* and is involved in normal development [6]. Despite these findings, the primary role and selective advantage of eukaryotic 5mC, which has led to its widespread distribution, are still highly controversial.

Bird [7] has suggested that, during evolution, the transition from invertebrate to vertebrate animals was accompanied by an increase in gene number, which was made possible by the use of DNA methylation to repress transcriptional background noise. However, this has been questioned by Jablonka and Regev [8] and is in opposition to the view that the primary function of cytosine methylation is the suppression of parasitic sequence elements, with the selective control of gene expression during development as either a secondary or even an illusory effect [9,10]. Bird has countered this with evidence that the colonization of germline genomes has not been significantly limited by methylation [11] and that at least one genome consists of non-methylated transposable elements and methylated genes [12]. These two apparently opposing views can be reconciled by the idea that global silencing mechanisms might have evolved originally to counteract the selfish activity of transposable elements and might then have been subsequently co-opted as host regulatory systems [13,14].

It is well established that vertebrate genomes are mostly methylated, with the short CpG islands at the 5′ ends of genes lacking 5mC [7]. Although it is generally accepted that invertebrate genomes are mostly 5mC-free, with a small fraction of methylated DNA not associated with genes [7], it is becoming evident that some invertebrate genes are methylated [12,15]. There is also good evidence that the transition from fractional to global methylation occurred close to the origin of vertebrates [15]. Exceptional invertebrates seem to be insects, where there is a general assumption that most are like the well-studied *Drosophila*, which has no detectable 5mC, although even here there is evidence for two proteins resembling a cytosine methyltransferase and a mammalian methyl-CpG-binding domain protein respectively [16].

One of the key aspects of the methylation debate is the distribution and role of 5mC in invertebrate genomes; however, there are very few examples in which this has been studied. More than 10 years ago Field et al. [17] reported an association between the presence of 5mC in an aphid gene and changes in its expression [17]. This study used restriction-enzyme analysis with a methylation-sensitive enzyme to study the amplified esterase *E4*

Abbreviations used: 5mC, 5-methylcytosine; RFLP, restriction-fragment-length polymorphism.

[1] e-mail lin.field@bbsrc.ac.uk

genes responsible for insecticide resistance in the aphid *Myzus persicae*. Amplified sequences containing expressed *E4* genes had methylated *Hpa*II sites; however, this was absent from the DNA of aphids that had spontaneously lost both the *E4* protein and the corresponding mRNA. A subsequent similar study showed that this loss of mRNA occurred simultaneously with the loss of 5mC [18], and quantitative PCR has recently identified aphids with approx. 80 copies of the *E4* gene, very low levels of mRNA and little DNA methylation [19]. These results are unusual in that methylation is present in an insect gene and its presence is correlated with expression, whereas a sudden loss of 5mC occurs with loss of expression. This clearly raises questions about the possible role of methylation, which is normally a repressor of transcription, in control of *E4* gene expression. However, because this methylation occurs in a multicopy gene, it might be simply a manifestation of the defence system in which silencing mechanisms can be directed specifically at repeat DNA [20], although clearly in this case the system fails.

The above restriction-fragment-length polymorphism (RFLP) studies of the methylation of aphid *E4* genes did not identify the precise location of the sites being studied and were very limited in their approach. There are now more sophisticated ways of identifying methylation in DNA [21]: this paper reports the use of bisulphite sequencing to study 5mC in and around the aphid *E4* gene and the use of CpG profiling as an index of the extent of CpG methylation [22]. In addition, methylation-sensitive enzymes are used to assess the wider extent of 5mC in the aphid genome [14].

MATERIALS AND METHODS

Aphid material

Two parthenogenic clones of *M. persicae* were reared on Chinese cabbage leaves at 20 °C with a 16 h light/8 h dark regime. The R_3 clone has approx. 80 copies of the *E4* gene that are expressed [19] and the Rev clone has a similar number of silent genes [17,18].

Sequencing of the *M. persicae E4* gene and flanking DNA

The complete *E4* gene and some of its flanking DNA were available on two contiguous genomic DNAs: an approx. 3.5 kb *Eco*RI 5′ fragment obtained by inverse PCR [23] and an approx. 8 kb *Eco*RI region cloned by fragment enrichment [24]. The 3.5 kb fragment had been sequenced previously [23], as had some of the coding region of the 8 kb fragment [24]. In the present study the sequence of the 8 kb plasmid clone was completed by using vector primers and oligonucleotides designed to known sequences. The sequencing reactions were done with an ABI PRISM™ Dye Terminator Ready Reaction Kit, and run on an ABI automated sequencer, type 373A (PE Applied Biosystems, Warrington, Cheshire, U.K.). This gave a complete sequence covering approx. 1.1 kb upstream of the *E4* transcription start site, approx. 6.2 kb containing the *E4* gene and approx. 4.2 kb of 3′ flanking DNA.

Calculation of CpG observed-to-expected ratios

CpG observed-to-expected ratios were calculated in 500 bp windows moving across the gene sequence in 10 bp steps. CpG observed-to-expected ratio is defined as the ratio of the CpG density to that expected from the local base composition, calculated as [number of CpG/(number of C × number of G) × ($w^2/w-1$)], where w is the number of nucleotides in a window.

Table 1 Primers used for PCR of bisulphite-treated DNA

Region	Primary	Nested
1	5′-ATTATTTGATGTTTAAAAT-3′	5′-GTGTTTGAAAGTTTTGGTTGGTTGTAT-3′
	5′-TATAATAAATTCAATATAA-3′	5′-CTCTCAAACCTTAAACAAACATTATAA-3′
2	5′-GTTTTGTTTTATATTAATAT-3′	5′-TTAAATAGTTATTATAAGAG-3′
	5′-TTATAAATAAAAACATTTAC-3′	5′-CATACCTTTAAACATTAATA-3′
3	5′-GATGTTTTATTTAAATATAG-3′	5′-GATAAAATATGATAATGTAT-3′
	5′-TAAATACTAATATCAAATTC-3′	5′-CTCATTACAATAATTAATTA-3′
4	5′-GTTAAATGAGTTTAAATTTT-3′	5′-GAATTTAGTATTTTTAAATG-3′
	5′-TAAACTAAAAAAATAACAAC-3′	5′-AACAAAATAAATCAAAATTC-3′

Bisulphite methylation analysis

Bisulphite sequencing [25] was performed with a slight modification of a published protocol (at Technical Tips Online http://www.elsevier.com/locate/tto tip number p01242). Genomic DNA (2 μg), extracted from *M. persicae* and digested with *Hin*dIII (which did not cut within the regions to be analysed), was denatured in 0.3 M NaOH at 43 °C for 30 min in a volume of 110 μl. The bisulphite mutagenesis was done by adding 10 μl of sterile distilled water, 60 μl of 10 mM hydroquinone and 1020 μl of 40.5 % (w/v) sodium bisulphite and incubating overnight in the dark at 55 °C. The modified DNA was purified with Prep-A-Gene® (Bio-Rad) and eluted in a final volume of 100 μl. Hot-start PCR reactions were set up with 8 μl of template DNA in 50 μl reactions containing 62.5 ng of each primer, 8 μl of dNTP species (1.25 mM in water), 2.5 μl of 10 × Taq buffer and 0.75 unit of *Taq* DNA polymerase (MBI Fermentas).

The PCR proceeded for 35 cycles (94 °C for 30 s, 40 °C for 2 min, 72 °C for 2 min). Nested PCRs were done as above with 1 μl of the primary reaction. Four regions were amplified (see the Results section) with the use of primers designed to bind to DNA after bisulphite modification (i.e. complementary to sequences with no CpG and assuming that all other C bases have been converted into T); see Table 1.

PCR products were checked on agarose gels and then ligated into pGEM® T Easy Vector System 1 (Promega Corporation) in accordance with the manufacturer's instructions. The plasmids were used to transform *Escherichia coli* XL1-Blue cells; white colonies were chosen for plasmid purification with Perfect preparative plasmid DNA-purification spin columns (5 Prime → 3 Prime®). Plasmids were sequenced in both strands with M13 forward and reverse primers that bound to the vector.

Restriction digests of genomic DNA

Genomic DNA (2 μg of each) was digested in 25 μl containing 2.5 μl of 10 × buffer (supplied with the enzyme) and 25 units of restriction enzyme (MBI Fermentas for *Hpa*II and New England Biolabs for *Mcr*BC) for 2 h at 37 °C. The digests were subjected to electrophoresis in 0.8 % agarose gels and stained with ethidium bromide.

RESULTS

Methylation of *Msp*I/*Hpa*II and other methylation-sensitive restriction sites in the *M. persicae E4* gene and flanking DNA

Previous studies [17,18] with *Msp*I (which cuts CCGG sites regardless of mCpG) and *Hpa*II (which cuts only if the internal C is unmethylated) had shown that such sites, producing fragments detectable with an *E4* cDNA probe, were methylated

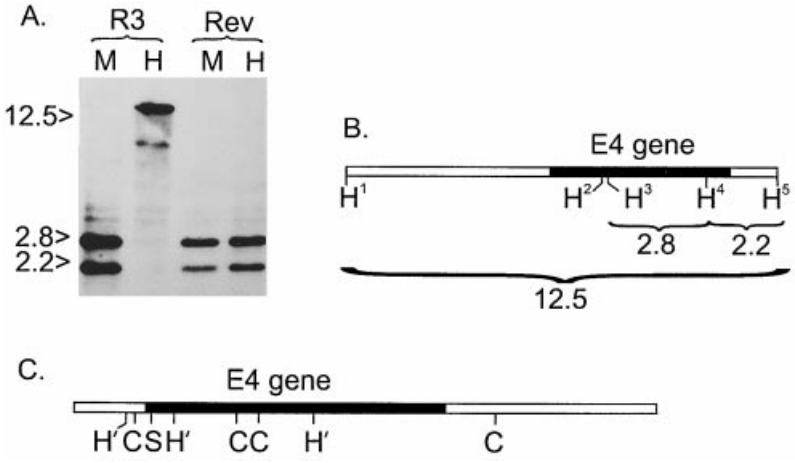

Figure 1 Analysis of the *E4* gene and flanking DNA with the use of methylation-sensitive restriction enzymes

(**A**) Restriction digests of genomic DNA from aphids with either expressing (R_3) or silent (Rev) *E4* genes probed with an *E4* cDNA. Numbers at the left indicate sizes of fragments (in kb). (**B**) Map of *E4* gene and flanking DNA, showing positions of *Msp*I/*Hpa*II restriction sites. M, *Msp*I; H, *Hpa*II. Numbers indicate sizes of fragments (in kb). (**C**) Positions of other methylation-sensitive restriction sites: H′, *Hha*I; C, *Cla*I; S, *Sal*I.

in the DNA of aphids that expressed the *E4* gene (R_3) and unmethylated in aphids with silent *E4* genes (Rev). Subsequent mapping and sequencing studies [18,24,26], and the sequencing reported here, now show the positions of the sites responsible for the observed fragments. These results are put together in Figures 1(A) and 1(B).

The 2.8 and 2.2 kb fragments in *Msp*I digests of the DNA from both R_3 and Rev aphids arise from digestion at H^4, H^5 and either or both of H^2 and H^3. Because the same fragments are present in the *Hpa*II digest of Rev aphids, sites H^4, H^5 and H^2 or H^3 (or both) must be unmethylated in their DNA. In contrast, the *Hpa*II digest of R_3 aphid DNA has a 12.5 kb band, resulting from cutting at H^1 and H^5, with H^2, H^3 and H^4 being uncut and therefore methylated. This suggests that 5mC is present within the expressed *E4* genes, but absent from the silent genes, and the flanking DNA of both is unmethylated.

This is supported by similar RFLPs with methylation-sensitive enzymes *Hha*I, *Cla*I and *Sal*I with sites as shown in Figure 1(C). Probing of Southern blots of R_3 aphid DNA (results not shown) shows that the cluster of sites at the 3′ end of the *E4* gene is unmethylated, the two *Cla*I sites and the *Hha*I site within the gene are methylated and the 3′ *Cla*I site is unmethylated.

Distribution of CpG doublets in the *M. persicae E4* gene and flanking DNA

Because 5mC, present in 5mCpG doublets, is prone to mutation to T, the genomes of species with 5mCpG will, over evolutionary time, become depleted in CpG doublets and the extent of CpG depletion will indicate the extent of CpG methylation [22]. The observed-to-expected CG ratios for points along the *E4* gene and flanking DNA are shown in Figure 2. An observed-to-expected ratio greater than 1 indicates CG enrichment; less than 1 shows CG depletion. Thus a region around the start of transcription is enriched, the main body of the gene is depleted (except for small peaks associated with some exons) and there is generally enrichment 3′ of the gene. This is consistent with the restriction enzyme results, again suggesting that methylation is confined to the gene. The peak around the transcription start sites is typical of a CpG island that is often present at the 5′ ends of genes [7,22,27].

Bisulphite analysis of regions associated with the *M. persicae E4* gene

Treatment of genomic DNA with bisulphite causes the sulphonation at C-6 of susceptible C residues and spontaneous hydrolysis of the amino group at C-4 to produce sulphonated uracil with 5mC residues remaining unconverted [21]. Thus bisulphite treatment followed by PCR with primers designed to match the DNA sequence, assuming that all Cs have been converted into Ts, produces fragments in which all unmethylated C residues have become T but 5mC residues remain as C. The PCR products can then be cloned and sequenced, allowing the methylation patterns of single DNA molecules to be analysed.

In the present study, regions 1–4, associated with the *E4* gene and shown in Figure 2, were amplified with the primers described in the Materials and methods section. In all cases the sequencing showed that all C residues not next to a G had been converted into T, indicating that the bisulphite treatment had been successful. It also showed that 5mC in this aphid gene is confined to CpG doublets as found in vertebrate genes, and is the first example of this being demonstrated in an insect gene. For region 1, a 390 bp stretch containing 29 potentially methylated CpG doublets around the *E4* transcription start site, 10 sequences were obtained from the DNA of aphids expressing the *E4* gene and 13 from aphids with silent *E4* genes (results not shown). In all cases, all the 29 C residues in CpG doublets had been converted into T and must therefore have been unmethylated. A similar result was obtained for the 481 bp region (region 4) in the 3′ flanking DNA, with each of the 30 C residues present in CpG doublets being unmethylated in each of 10 replicates from expressed and silent genes. The consistent results for all strands suggest that all copies of *E4* lack 5mC. The lack of methylation in these regions is also consistent with the enriched CpG content described above.

The results of bisulphite analysis of regions 2 and 3 (covering exons 3, 6 and 7; see Figure 2) are given in Figure 3; here the patterns are more variable. For the R_3 aphids, with expressed *E4*

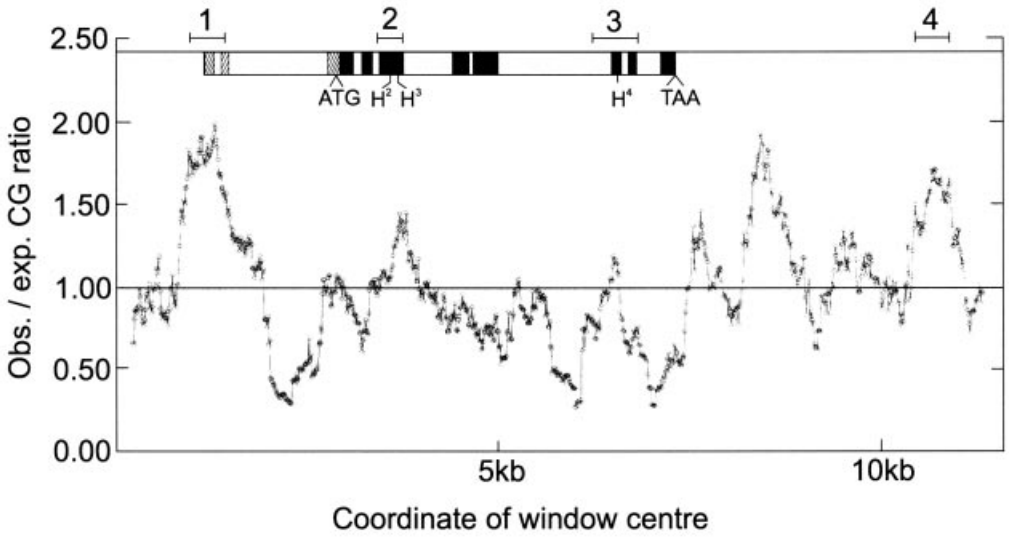

Figure 2 Distribution of CpG dinucleotides in and around the *E4* gene

CpG observed-to-expected ratios are plotted against the central co-ordinate of each window used in the calculation (see the Materials and methods section). In the diagram of the *E4* gene at the top, open rectangles represent introns, hatched rectangles represent untranslated regions and filled rectangles represent translated regions. H indicates *Hpa*II restriction sites (with the same numbering as in Figure 1) and 1–4 indicate regions of bisulphite sequencing (see the Results section and Figure 3).

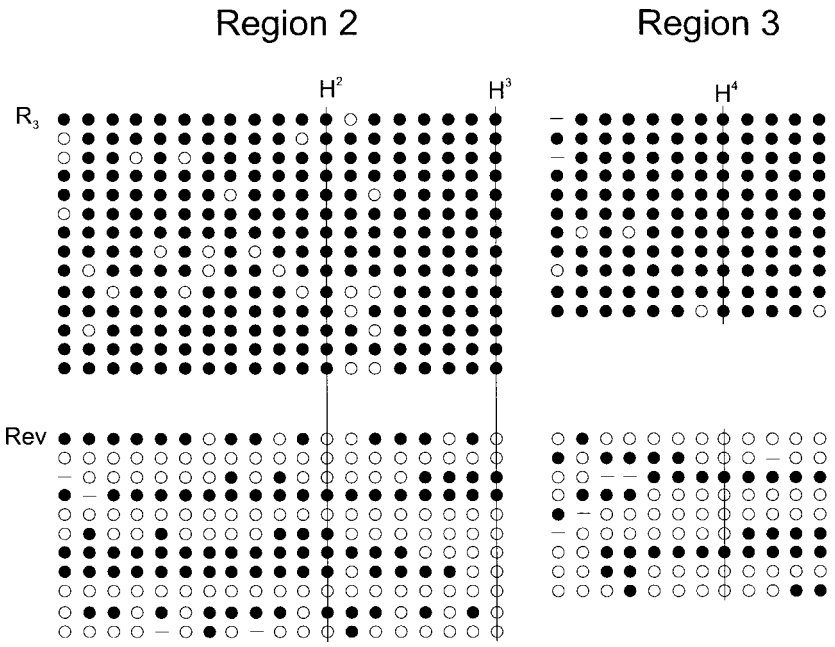

Figure 3 Bisulphite methylation analysis of two regions of the *E4* gene

C (○) and 5mC (●) in the CpG doublets of two regions (2 and 3 of Figure 2) of the *M. persicae E4* gene. Each row represents a single strand of DNA from aphids with either expressing (R_3) or silent (Rev) *E4* genes. H^2, H^3 and H^4 are CpG doublets in *Msp*I/*Hpa*II sites, corresponding to those shown in Figures 1 and 2.

genes, most of the CpG doublets on most of the strands for both regions were methylated, explaining the lack of *Hpa*II digestion at sites H^2, H^3 and H^4. For the DNA from revertant (Rev) aphids with silent *E4* genes, there was much less methylation, although it was not completely absent. For region 2, 10 of the 11 strands analysed would have been digested with *Hpa*II at either H^2 or H^3, and 7 of the 9 strands in region 3 would have been digested at H^4, thus producing the restriction fragments reported above. Overall the presence of 5mC in regions 2 and 3 within the expressed *E4* gene was consistent with the observed low levels of CpG doublets. For the revertants in which the genes had been silenced spontaneously, there was a loss of 5mC, as suggested by the *Msp*I/*Hpa*II analysis, but this was by no means complete: some strands were fully methylated, some were unmethylated and

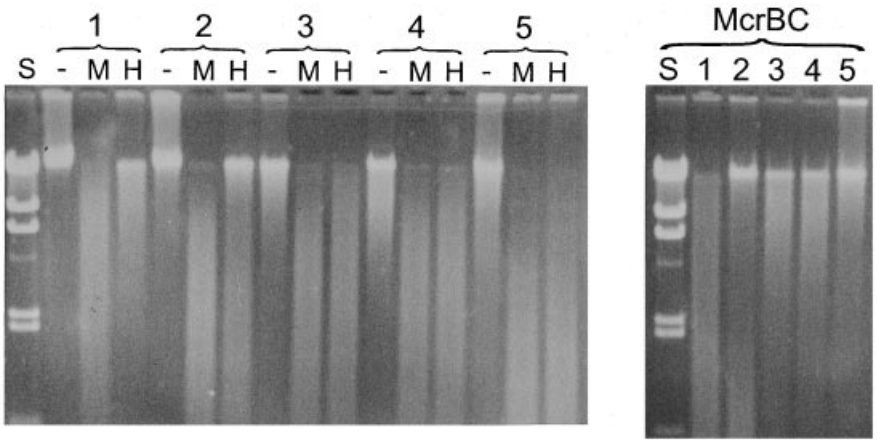

Figure 4 Restriction digests of genomic DNA species stained with ethidium bromide

Genomic DNA species from human (1), sea urchin (2), *M. persicae* with silent *E4* genes (Rev) (3), *M. persicae* with expressed *E4* genes (R_3) (4) and *Drosophila melanogaster* (5) were either undigested (—) or digested with *Msp*I (M) or *Hpa*II (H) (left panel) or *Mcr*BC (right panel). Tracks labelled S contain λ/*Hin*dIII markers with standard bands of 23.1, 9.4, 6.6, 4.3, 2.3, 2.0 and 0.6 kb.

most were a mosaic, possibly reflecting accessibility of the DNA to a methyltransferase.

Genome-wide methylation in *M. persicae*

Ethidium bromide staining of genomic DNA digested with methylation-sensitive restriction enzymes can indicate the extent of genome methylation [15]. For example, vertebrate DNA, which is widely methylated, is largely resistant to *Hpa*II digestion, whereas invertebrate DNA is cleaved to a heterogeneous mixture of fragments that enter the gel and a second apparently undigested fraction [15].

Here, two methylation-sensitive enzymes were used to establish whether loss of methylation within the *E4* gene in revertant aphids was associated with global loss of 5mC. Figure 4 (left panel) shows *Msp*I and *Hpa*II digestions of *M. persicae* DNA from aphids with expressed (R_3) and silent (Rev) *E4* genes, alongside standard DNA species from a vertebrate (human) and two invertebrates (sea urchin and *Drosophila*). The human DNA was largely resistant to *Hpa*II digestion and the sea urchin had the typical two fractions reported [15]. The aphid DNA was clearly much less methylated than human, but did not have an obvious *Hpa*II-insensitive fraction typical of invertebrates and as shown for sea urchin. Both aphid samples were less digested with *Hpa*II than the unmethylated *Drosophila* sample and both had similar digestion patterns, showing that the loss of 5mC within the *E4* gene was not associated with a general loss elsewhere in the genome.

This finding is supported by the results of digestion with *Mcr*BC, which cuts only methylated DNA [at $Pu^mC(N_{40-2000})Pu^mC$], and again resulted in the two aphid DNA samples being digested to the same degree (Figure 4, right panel). However, a comparison of *Msp*I and *Hpa*II digests showed only slight differences for both aphid samples, suggesting that there were only very low levels of genome methylation, making it difficult to draw any firm conclusions.

DISCUSSION

A very wide range of *M. persicae* samples, of wide geographic origin, show resistance to insecticides conferred by high levels of esterase enzyme and resulting from the presence of amplified genes ([22,28]; L. M. Field, unpublished work). In all cases where *Msp*I/*Hpa*II digests have been used to assess the methylation of the amplified sequences, the results have been positive [17–19], with the methylation patterns being maintained through both parthenogenetic and sexual reproduction. Thus the presence of amplified, expressed methylated *E4* genes seems to be the norm, although the copy number can vary considerably [19].

The bisulphite sequencing data and CpG profiles reported here confirm the preliminary RFLP evidence that the expressed *E4* gene is methylated. However, for the highly amplified clone used here, this is now shown to be confined to the body of the gene; 5′ and 3′ flanking DNA are free of methylation. The apparent anomaly of expression of *E4* in the presence of methylation is therefore explained by the present finding that this methylation does not include the promoter region and that the CpG-rich region around the transcription start site is free of methylation. There is therefore no reason to suppose that the observed methylation would interfere with the binding of transcription factors to the *E4* promoter. The methylation pattern is very similar to the situation reported for genes of the invertebrate *Ciona intestinalis* (sea-squirt) [12] and is another example against the view that methylation in invertebrate genomes is not associated with genes.

Because mutation of 5mC will cause CpG depletion only over the course of evolution [22], the depletion of CpG doublets within the *E4* gene suggests that its methylation must have been present for a long period. There are two possible explanations for this: first, the original single-copy gene could have been methylated before a recent amplification event, or secondly, the amplification is ancient and the methylation occurred on the amplified genes. The former cannot be ruled out but is perhaps questioned by the finding that the single-copy gene, which is present in a limited number of field aphids, is not methylated [17]. If the latter is true, the methylation might have occurred in response to the presence of multiple-copy genes [20]; however, it has not affected the promoter and the genes have remained active. What selection pressures might have maintained the presence of methylation of *E4*? One possibility is that it has a role in the suppression of homologous recombination, which might lead to lethal chromosome abnormalities, as reported for the fungus *Ascobolus* [3]. However, it should be noted that the *E4*

amplification is actually associated with a chromosome translocation [29]. It is also possible that the 5mC found within the *E4* gene actually has a positive role in its transcription. The possibility that methylation within genes decreases transcriptional interference has been suggested by Simmen et al. [12] and Jones [27]. This idea is based on the fact that active genes can produce incorrectly initiated transcripts that can interfere with authentic transcription, and methylation within coding regions can silence the expression of spurious promoters, allowing the production of the correct transcript. The idea that the presence of 5mC within the *E4* gene is important for its correct transcription is supported by the results reported here for the bisulphite-treated DNA from aphids with silent *E4* genes (Revs). These show that the spontaneous loss of transcription of *E4* seen in these aphids is correlated with a loss of 5mC within the *E4* gene (regions 2 and 3).

Overall, the results presented here suggest that the *E4* gene amplification in *M. persicae* occurred in ancient evolutionary history. This was followed by methylation of the amplified genes, but not their co-amplified flanking DNA, and allowed the observed CpG profiles to evolve. This would imply that amplified expressed *E4* genes have been present throughout evolutionary history, possibly being selected for by their ability to protect against insecticidal compounds made by some plants. This is in contrast with the view that the amplification is a recent event selected for during the use of agriculturally applied insecticides. However, it is possible that levels of *E4* gene amplification were initially low and have increased since the widespread use of insecticides, with the high copy numbers (up to 80 [19]) now found in very resistant aphids being a relatively recent phenomenon.

Despite the finding of 5mC associated with the *E4* genes in *M. persicae*, the restriction digests with the methylation-sensitive enzymes *Hpa*II and *Mcr*BC suggest that methylation is rare in the aphid genome. It remains to be seen whether it has any role in controlling the transcription of other genes or in general genome defence.

I thank: Professor Adrian Bird (University of Edinburgh, U.K.) for helpful discussions and critical reading of the manuscript; Dr Martin Simmen (University of Edinburgh, Edinburgh, Scotland, U.K.) for the CpG analysis of the sequence results; and Dr Donald Mcleod (University of Edinburgh, Edinburgh, Scotland, U.K.) for technical advice on bisulphite sequencing. IACR–Rothamsted receives grant-aided support from the BBSRC.

REFERENCES

1 Adams, R. L. P. (1996) DNA methylation. In Principles of Medical Biology, vol. 5, pp. 33–66, JAI Press Inc., New York
2 Wolff, A. P., Jones, P. L. and Wade, P. A. (1999) DNA demethylation. Proc. Natl. Acad. Sci. U.S.A. **96**, 5894–5896
3 Colot, V. and Rossignol, J.-L. (1999) Eukaryotic DNA methylation as an evolutionary device. BioEssays **21**, 402–411
4 Nan, X., Campoy, F. J. and Bird, A. (1997) MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. Cell **88**, 471–481
5 Li, E. (1999) The mojo of methylation. Nat. Genet. **23**, 5–6
6 Siegfried, Z., Eden, S., Mendelsohn, M., Feng, X., Tsuberi, B.-Z. and Cedar, H. (1999) DNA methylation represses transcription *in vivo*. Nat. Genet. **22**, 203–206
7 Bird, A. P. (1995) Gene number, noise reduction and biological complexity. Trends Genet. **11**, 94–100
8 Jablonka, E. and Regev, A. (1995) Gene number, methylation and biological complexity. Trends Genet. **11**, 383–384
9 Yoder, J. A., Walsh, C. P. and Bestor, T. H. (1997) Cytosine methylation and the ecology of intragenomic parasites. Trends Genet. **13**, 335–340
10 Walsh, C. P. and Bestor, T. H. (1999) Cytosine methylation and mammalian development. Genes Dev. **13**, 26–34
11 Bird, A. (1997) Does DNA methylation control transposition of selfish elements in the germline? Trends Genet. **13**, 469–470
12 Simmen, M. W., Leitgeb, S., Charlton, J., Jones, S. J. M., Harris, B. R., Clark, V. H. and Bird, A. (1999) Nonmethylated transposable elements and methylated genes in a chordate genome. Science **283**, 1164–1167
13 Martienssen, R. (1998) Transposons, DNA methylation and gene control. Trends Genet. **14**, 263–264
14 McDonald, J. F. (1998) Transposable elements, gene silencing and macroevolution. Trends Ecol. Evol. **13**, 94–95
15 Tweedie, S., Charlton, J., Clark, V. and Bird, A. (1997) Methylation of genomes and genes at the invertebrate–vertebrate boundary. Mol. Cell. Biol. **17**, 1469–1475
16 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
17 Field, L. M., Devonshire, A. L., Ffrench-Constant, R. H. and Forde, B. G. (1989) Changes in DNA methylation are associated with loss of insecticide resistance in the peach-potato aphid *Myzus persicae* (Sulz.). FEBS Lett. **243**, 323–327
18 Hick, C. A., Field, L. M. and Devonshire, A. L. (1996) Changes in the methylation of amplified esterase DNA during loss and reselection of insecticide resistance in peach-potato aphids. Insect Biochem. Mol. Biol. **26**, 41–47
19 Field, L. M., Blackman, R. L., Tyler-Smith, C. and Devonshire, A. L. (1999) Relationship between amount of esterase and gene copy number in insecticide-resistant *Myzus persicae* (Sulzer). Biochem. J. **339**, 737–742
20 Selker, E. U. (1999) Gene silencing: repeats that count. Cell **97**, 157–160
21 Rein, T., DePamphilis, M. L. and Zorbas, H. (1998) Identfying 5-methylcytosine and related modifications in DNA genomes. Nucleic Acids Res. **26**, 2255–2264
22 Shimizu, T. S., Takahashi, K. and Tomita, M. (1997) CpG distribution patterns in methylated and non-methylated species. Gene **205**, 103–107
23 Field, L. M. and Devonshire, A. L. (1998) Evidence that the *E4* and *FE4* esterase genes responsible for insecticide resistance in the aphid *Myzus persicae* (Sulzer) are part of a gene family. Biochem. J. **330**, 169–173
24 Field, L. M., Williamson, M. S., Moores, G. D. and Devonshire, A. L. (1993) Cloning and analysis of the esterase genes conferring insecticide resistance in the peach-potato aphid, *Myzus persicae* (Sulzer). Biochem. J. **294**, 569–574
25 Frommer, M., McDonald, L. E., Millar, D. S., Collis, C. M., Watt, F., Grigg, G. W., Molloy, P. L. and Paul, C. L. (1992) A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. Proc. Natl. Acad. Sci. U.S.A. **89**, 1827–1831
26 Field, L. M., Javed, N., Stribley, M. F. and Devonshire, A. L. (1994) The peach-potato aphid *Myzus persicae* and the tobacco aphid *Myzus nicotianae* have the same esterase-based mechanisms of insecticide resistance. Insect Mol. Biol. **3**, 143–148
27 Jones, P. A. (1999) The DNA methylation paradox. Trends Genet. **15**, 34–37
28 Field, L. M., Anderson, A. P., Denholm, I., Foster, S. P., Harling, Z. K., Javed, N., Martinez-Torres, D., Moores, G. D., Williamson, M. S. and Devonshire, A. L. (1997) Use of biochemical and DNA diagnostics for characterising multiple mechanisms of insecticide resistance in the peach-potato aphid, *Myzus persicae* (Sulzer). Pestic. Sci. **51**, 283–289
29 Blackman, R. L., Takada, H. and Kawakami, K. (1978) Chromosomal rearrangement involved in the insecticide resistance of *Myzus persicae*. Nature (London) **271**, 450–452

Received 7 April 2000/9 May 2000; accepted 2 June 2000