*4902–4909* *Nucleic Acids Research, 1996, Vol. 24, No. 24* 

# Variegated expression of a globin transgene correlates with chromatin accessibility but not methylation status

**David Garrick, Heidi Sutherland, Graham Robertson and Emma Whitelaw***

Department of Biochemistry, University of Sydney, Sydney, NSW 2006, Australia

Received September 6, 1996; Revised and Accepted November 4, 1996

## ABSTRACT

**There are now many mammalian examples in which single cell assays of transgene activity have revealed variegated patterns of expression. We have previously reported that transgenes in which globin regulatory elements drive the *lacZ* reporter gene exhibit variegated expression patterns in mouse erythrocytes, with transgene activity detectable in only a sub-population of circulating erythroid cells. In order to elucidate the molecular mechanism responsible for variegated expression in this system, we have compared the chromatin structure and methylation status of the transgene locus in expressing and non-expressing populations of erythrocytes. We find that there is a difference in the chromatin conformation of the transgene locus between the two states. Relative to active transgenes, transgene loci which have been silenced exhibit a reduced sensitivity to general digestion by DNase I, as well as a failure to establish a transgene-specific DNase I hypersensitive site, suggesting that silenced transgenes are situated within less accessible chromatin structures. Surprisingly, the restrictive chromatin structure observed at silenced transgene loci did not correlate with increased methylation, with transgenes from both active and inactive loci appearing largely unmethylated following analysis with methylation-sensitive restriction enzymes and by sequencing PCR products derived from bisulphite-converted genomic DNA.**

## INTRODUCTION

A gene is said to exhibit a variegated or mosaic pattern of expression when it is active in some but not all cells of a particular cell type. The best characterized examples of variegation occur in *Drosophila*, where mosaic expression of genes within the eye disk cells results in mottled patterns of eye pigmentation. This variegated expression results from chromosomal rearrangements which juxtapose a normally euchromatic gene with regions of constitutive heterochromatin (1). Cytological observations have revealed a correlation between the transcriptional state of a variegating gene and the degree to which it is compacted within the surrounding chromatin. This has led to the hypothesis that the variegated expression pattern is caused by differences between cells in the extent of heterochromatin spreading (2). Several mutations which were observed to suppress variegation in this system identified genetic loci encoding structural components of heterochromatin (reviewed in 3).

The frequency with which variegated or mosaic patterns of gene expression occur in mammalian systems may have been underestimated. Variegated expression patterns can only be detected when gene activity is assayed in individual cells or in clones derived from individual cells. Since assays of gene expression in mammals are usually performed on lysates obtained from whole cell populations and the observed level of activity is generally assumed to reflect an equivalent rate of expression in all cells, variegated patterns of expression may go unnoticed. However, there are now several studies in which the ability to assay expression in individual cells or clones has revealed variegated expression of either endogenous genes (4) or transgenes (5–13, reviewed in 14) in mammalian systems.

In plant and animal tissues a correlation has frequently been observed between gene inactivity and hypermethylation of cytosine residues (15,16), leading to the suggestion that methylation may be the cause of transcriptional silencing in these instances. The finding that *in vitro* methylated DNA adopts a different chromatin structure to untreated DNA following stable integration into a host genome (17) suggests that the silencing effect of methylation upon gene activity may be mediated by a change in chromatin structure from an active to a repressive state, although what this means at the molecular level is still largely unknown. Methylation has also been implicated as the mechanism of gene silencing responsible for variegated expression patterns in plants and animals, with several studies reporting a correlation between the extent of variegation and the degree of transgene methylation (6,18–20). However, the exact relationship between DNA methylation, chromatin structure and gene expression still remains to be clearly elucidated.

We have previously reported that transgenes in which the *Escherichia coli lacZ* gene is driven by human globin regulatory elements are expressed in a variegated manner in mouse erythroid cells (21,22). Because these transgenes express the cell-autonomous *lacZ* reporter in the erythroid tissue, histochemical staining of circulating erythrocytes with X-gal allows transgene activity to be analysed in individual cells. The proportion of erythroid cells expressing the transgene was consistent within a transgenic line, but varied between different lines bearing the same construct in a manner which was independent of transgene copy number. This copy number independence suggested that the proportion of cells

*To whom correspondence should be addressed. Tel: +61 2 9351 2549; Fax: +61 2 9351 4726; Email: e.whitelaw@biochem.usyd.edu.au

in which the transgene is active is primarily determined by influences specific to the site of integration.

In the work presented here, we have attempted to identify the molecular mechanism by which transgene expression is silenced in some erythroid cells and not others. We have compared chromatin structure and methylation status of the transgene between cells in which the transgene is active and cells in which it has been silenced. We find that there is a clear difference in the chromatin conformation of the transgene between the two states; the transgene is situated within a relatively open chromatin environment in cells in which it is active, while in non-expressing erythrocytes, the chromatin structure of the transgene locus appears to be more restricted. Surprisingly, this difference in the chromatin conformation of the transgene locus between the two states does not reflect a difference in the extent of transgene methylation. These findings suggest that transgene silencing is mediated by locus-specific epigenetic events which do not involve methylation.

## MATERIALS AND METHODS

### Generation of transgenic mice

The production of the DM1 transgenic mouse lines used in this study has been described previously (21). Briefly, the DM1 construct contains a 163 bp promoter (comprised of elements from both the human ζ- and γ-globin promoters) upstream of a 3.3 kb SDK-*lacZ*pA cassette and a 4.2 kb fragment containing the αHS-40 DNase I hypersensitive site (Fig. 1). The DM1 transgene was excised from vector sequences by digestion with *Kpn*I and *Not*I and purified by agarose gel electrophoresis. Fertilized eggs of the outbred PO mouse strain were microinjected with the DM1 construct according to standard techniques (23). Founder mice were identified by Southern blotting of tail DNA and hemizygous transgenic lines were established. Copy number was determined by Southern blotting of tail DNA from $G_1$ mice. Erythroid expression at 12.5 d.p.c. was determined by bleeding individual embryos into PBS, fixing whole blood cells in glutaraldehyde and staining with X-gal as described previously (21). All data presented in this study were obtained from established lines.

### DNase I sensitivity assays

Nuclei were prepared as described by Higgs *et al.* (24) from the pooled blood of five individual 12.5 d.p.c. transgenic embryos. For the DNase I hypersensitivity assay, nuclei were digested with DNase I (Boehringer Mannheim) at $6.25 \times 10^{-2}$ U/ml for 0–50 min at 37°C. At each time point, digestion was stopped by the addition of EDTA to 10 mM and DNA purified by phenol/chloroform extraction following digestion with proteinase K. DNA was then digested with *Bam*HI and resolved on a 1% agarose gel. After Southern blotting, the membrane was probed with the 4.2 kb *Bam*HI fragment of the DM1 transgene containing αHS-40 (Fig. 1). For the general DNase I sensitivity assay, nuclei were digested with DNase I at 0, 0.1, 0.2, 0.5, 1.0 and 2.0 U/ml for 15 min at 37°C. Reactions were stopped and DNA purified as above. After digesting with *Pst*I, Southern analysis was performed using the 3.3 kb SDK-*lacZ*pA cassette as probe. To control for equivalent DNase I digestion, the membrane was rehybridized with a 2.1 kb *Hin*fI–*Sac*I fragment containing the mouse α-globin gene and 5′ flanking sequence. All probes were labelled by random priming using [α-$^{32}$P]dCTP and Klenow DNA polymerase.

### Analysis of methylation by digestion with restriction enzymes

DNA was recovered from circulating erythroid cells of 12.5 d.p.c. transgenic embryos. Aliquots of 2 µg DNA were digested with either *Bam*HI alone or *Bam*HI together with one of the methylation-sensitive restriction enzymes *Sma*I, *Hpa*II or *Hin*PI. To monitor complete digestion, in a separate set of reactions an equivalent amount of wild-type mouse DNA spiked with the DM1 plasmid was digested under the same conditions (not shown). The resulting fragments were analysed by Southern blotting after separation in 1.2% agarose gels. Membranes were probed with either the 4.2 kb *Bam*HI fragment of the DM1 transgene containing αHS-40 or the 3.3 kb SDK-*lacZ*pA cassette. Probes were labelled as described above.

### Bisulphite conversion of genomic DNA and sequencing of sub-cloned PCR products

Genomic DNA was purified from circulating erythroid cells obtained from 12.5 d.p.c. embryos of the transgenic lines 27/5A and 16/7B. Aliquots of 15 µg DNA were digested with *Eco*RI and bisulphite conversion carried out as described previously (25), with removal of free bisulphite achieved by passage through PCR Spinclean columns (Progen Industries) according to the manufacturer's instructions. The PCR primers DFG-1 (5′-GGGAGATTTTTGGATCCGTAGAGATATAAA) and DFG-2 (5′-ATAATAACCTCCCACCTGCAGACATACAA) were used to amplify a 332 bp fragment containing the transgene promoter from the bisulphite-converted coding strand at head–tail transgene junctions. Underlined sequence shows the introduced *Bam*HI and *Pst*I sites used for subsequent sub-cloning. Amplification was performed in 50 µl reactions using 4 µl bisulphite-converted template, 5 µM primers, 200 µM dNTPs, 3 mM $MgCl_2$ and 0.5 U Taq (Boehringer Mannheim) in the buffer provided. After digestion with *Bam*HI and *Pst*I, PCR product was sub-cloned into the polylinker of the plasmid vector KS Bluescript. Dideoxy sequencing was performed on individual clones using the forward Universal primer.

## RESULTS

Transgenic mouse lines were generated bearing the DM1 construct shown in Figure 1. In this construct, a hybrid of the human ζ- and γ-globin promoters drives expression of the *E.coli lacZ* reporter gene in the presence of the major tissue-specific enhancer element of the human α-like globin locus, αHS-40 (24). Use of the cell-autonomous *lacZ* reporter in erythrocytes allows transgene activity to be assayed in individual cells by staining whole blood with the chromogenic substrate 5-bromo-4-chloro-3-indolyl β-D-galactopyranoside (X-gal). We have previously reported that when this cell-by-cell assay was performed on erythrocytes from mice bearing the DM1 construct, in most lines only a sub-population of erythroid cells exhibit β-galactosidase activity (21). Since it was theoretically possible that this heterocellular pattern was the result of somatic mutations within the *lacZ* gene in a sub-population of erythroid cells, PCR was used to amplify the *lacZ* coding region from primitive erythrocytes of lines in which the majority (98%) of cells were not expressing the transgene. After cloning the PCR product into a bacterial expression vector and assaying for β-galactosidase activity, it was found that the *lacZ* gene was fully functional (data not shown). This finding suggests that the inactivity of the transgene locus in a sub-population of cells reflects epigenetic rather than genetic

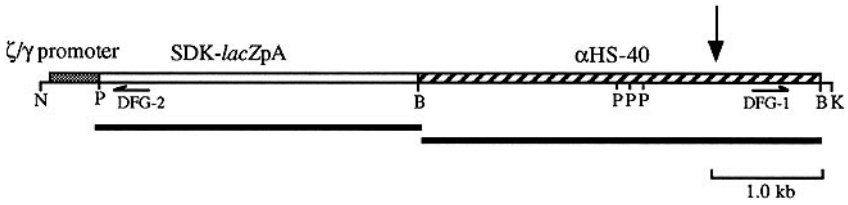


**Figure 1.** The DM1 construct used to generate transgenic mice. The ζ/γ-globin promoter is comprised of nucleotides +6 to –127 of the human ζ-globin promoter but with nucleotides –83 to –112 replaced by bases –134 to –193 of the human γ-globin promoter, thereby exchanging the GATA and CACCC motifs of the two promoters. The SDK-*lac*ZpA cassette contains Shine–Dalgarno and Kozak sequences immediately upstream of the *E.coli lacZ* gene and the SV40 poly(A). αHS-40 is a 4.2 kb *Hin*dIII fragment of the human α-globin locus which contains the αHS-40 DNase I hypersensitive site (vertical arrow) and has been cloned using *Bam*HI linkers. Horizontal black lines represent probes used in this study. Horizontal arrows show the positions of the primers DFG-1 and DFG-2 used to amplify the promoter when transgene monomers are arranged in an array in a head–tail orientation. P, *Pst*I; B, *Bam*HI; N, *Not*I; K, *Kpn*I.

variation. To investigate the epigenetic mechanism responsible for variegated expression of the DM1 transgene, three transgenic lines (27/5A, 22/7A and 16/7B) were selected for analysis, since we had previously found that at 12.5 d.p.c. a very high proportion of erythrocytes express the transgene in 27/5A, while in both 22/7A and 16/7B the transgene is silent in the majority of erythroid cells. Analysis was performed in erythrocytes purified from 12.5 d.p.c. transgenic embryos, since at this stage of development, nucleated primitive cells represent ~95% of circulating erythrocytes (26). Consistent with previous data, when peripheral blood from 12.5 d.p.c. embryos of these three lines was stained with X-gal, line 27/5A expressed β-galactosidase in 95% of erythrocytes, while lines 22/7A and 16/7B expressed the transgene in only 8 and 2% of erythroid cells respectively (Table 1). Erythrocytes isolated from 12.5 d.p.c. transgenic embryos of these lines therefore approximate pure populations of transgene-active (27/5A) and transgene-inactive (22/7A and 16/7B) erythroid cells. The transgene copy numbers of these three lines are also shown in Table 1.

**Table 1.** Copy number and erythroid expression at 12.5 d.p.c. of the DM1 transgenic lines

| Transgenic line | Erythrocytes expressing *lacZ* at 12.5 d.p.c. (%) | Copy number |
|---|---|---|
| 27/5A | 95 ± 1 (*n* = 3) | ~30 |
| 22/7A | 8 ± 3 (*n* = 6) | ~50 |
| 16/7B | 2 ± 1 (*n* = 10) | ~40 |

12.5 d.p.c. transgenic embryos were bled into PBS and after fixing in glutaraldehyde, cells were stained with X-gal as described previously (21). The percentage of erythroid cells expressing *lacZ* was determined by scoring a minimum of 200 cells for detectable blue colour. Data are presented as mean ± SD and the number of individual transgenic embryos assayed for each line (*n*) is shown in parentheses. Approximate copy number was determined using standard Southern blotting of tail DNA by comparing the intensity of the signal with those of known copy number standards.

## Inactivity of the DM1 transgene correlates with a less accessible chromatin structure

DNase I hypersensitivity mapping was performed to assay for the presence of the transgene-specific αHS-40 hypersensitive site in these three lines. Nuclei were isolated from 12.5 d.p.c. erythrocytes of the lines 27/5A, 22/7A and 16/7B and subjected to DNase I digestion for increasing periods of time. DNA was then purified and digested with *Bam*HI to release a 4.2 kb parent fragment which is reduced by DNase I cleavage at the αHS-40 hypersensitive site to produce sub-bands of ~3 and 1 kb (27). The αHS-40 hypersensitive site could be detected in erythrocytes from line 27/5A, where the majority of cells express the transgene, but was not present in erythrocytes from 22/7A and 16/7B, where the transgene has been silenced in most cells (Fig. 2). Since the αHS-40 DNase I hypersensitive site is associated with the binding of a variety of transcription factors within a 350 bp core element (27), the failure of this site to become established in the transgenes of inactive erythroid cells suggests that access of these transcription factors to their cognate recognition sequences within the transgene has been prevented by a restrictive chromatin structure.

As an indicator of chromatin structure over the entire transgene locus, we compared the general DNase I sensitivity of the transgene in expressing and non-expressing erythrocytes. Resistance of a fragment to general digestion by DNase I reflects its location within a domain of closed or tightly confined chromatin (28) which is believed to restrict access of the nuclease. Nuclei isolated from 12.5 d.p.c. erythrocytes of lines 27/5A and 22/7A were digested with increasing concentrations of DNase I. After purification, DNA was digested with *Pst*I to release a fragment of 5.4 kb containing the entire *lacZ* gene and part of αHS-40 and the resistance of this fragment to DNase I digestion was analysed on Southern blots. To ensure that the extent of DNase I digestion was equivalent for the two lines, the membrane was also hybridized with a probe for the endogenous mouse α-globin gene. While the endogenous fragment was digested to the same extent in the two lines, the *lacZ* transgene fragment decreases more rapidly in response to DNase I in erythrocytes from line 27/5A (95% expressing cells) than in those from 22/7A (8% expressing cells) (Fig. 3). Since the SDK-*lacZ* fragment assayed does not include the αHS-40 hypersensitive site, the rate of loss of this fragment is a measure of its general DNase I sensitivity and does not merely reflect the status of the αHS-40 site. Therefore, as well as lacking a detectable hypersensitive site, transgenes which have been inactivated also appear less sensitive to general digestion by DNase I than expressing loci.

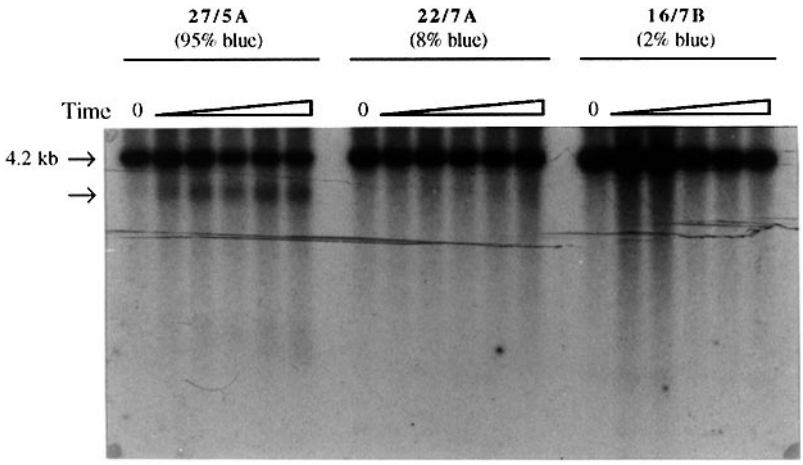


**Figure 2.** Assay for the presence of the αHS-40 DNase I hypersensitive site in 12.5 d.p.c. erythrocyte nuclei of the lines 27/5A, 22/7A and 16/7B. 12.5 d.p.c. erythrocyte nuclei were digested for increasing periods of time with DNase I. After purification, the DNA was digested with *Bam*HI and separated on a 1% agarose gel. Following Southern blotting, the membrane was probed with the 4.2 kb *Bam*HI αHS-40 fragment. The positions of the parent fragment and the ~3 kb daughter fragment generated by specific cleavage at the hypersensitive site are indicated. The daughter fragment is clearly visible in nuclei from line 27/5A, in which 95% of erythrocytes express the transgene.

## No differences in methylation can be detected between active and inactive transgenes

Since the silencing of transgenes and the adoption of an inactive chromatin state has frequently been found to correlate with increased transgene methylation, we compared the methylation status of the DM1 transgene in the expressing and non-expressing erythrocytes. Transgene methylation was initially investigated by digesting with methylation-sensitive restriction enzymes. Chromosomal DNA from 12.5 d.p.c. erythrocytes of lines 27/5A, 22/7A and 16/7B was digested with *Bam*HI to release a 4.2 kb fragment containing αHS-40 and a 3.5 kb fragment containing the promoter/*lacZ*. DNA was then digested with one of the methylation-sensitive restriction enzymes *Sma*I, *Hpa*II or *Hin*PI. The resulting products were analysed by Southern blotting and hybridization with the *Bam*HI parent fragments as probes (Fig. 4). Surprisingly, in all three lines these methylation-sensitive restriction enzymes digested the αHS-40 parent band almost to completion and no difference could be detected between these lines in the pattern of fragments generated (Fig. 4a). Although some copies of the αHS-40 fragment had remained uncut following digestion with *Sma*I, the proportion of digested to undigested fragments in this lane is the same in all three lines. Similarly, a faint 2.0 kb fragment resulting from incomplete *Hpa*II digestion was observed in the DNA of both transgene-expressing (27/5A) and silenced erythrocytes (22/7A and 16/7B). There is therefore no detectable difference in the frequency with which these restriction sites are methylated between active and inactive transgenes.

A similar analysis was performed on the remaining fragment of the transgene, which contains the promoter and *lacZ* gene. CpG dinucleotides, the primary target of mammalian DNA methyltransferase, occur with a much greater frequency in the bacterial *lacZ* reporter gene than is typically observed in eukaryotic DNA and it was therefore possible that an unusually high frequency of methylated cytosines within this region was responsible for transgene silencing. However, it was found that when the 3.5 kb *Bam*HI fragment of the transgene containing the promoter and entire *lacZ* gene was digested with *Hpa*II or *Hin*PI (for which there are 14 and 28 sites respectively) it produced sub-fragments too small for detection by Southern blotting (Fig. 4b). This extensive digestion with *Hpa*II and *Hin*PI indicates that there has been no blanket methylation of this region of prokaryotic DNA in silenced transgene loci. We have performed similar methylation-sensitive restriction analyses on transgenes in a further 15 lines which exhibit variegated expression in mouse erythrocytes. No evidence of methylation was detected in any line containing an active transgene, regardless of the percentage of expressing erythroid cells (data not shown). The various transgene constructs contained in these lines include globin (human α and ζ) and non-globin (thymidine kinase and metallothionine) promoter elements driving expression of *lacZ* both in the presence and absence of αHS-40.

The use of methylation-sensitive enzymes restricts the analysis of methylation status to a limited number of specific sites within the transgene and differences in methylation between active and inactive transgenes at some other sites may go undetected. For this reason, the methylation status of the transgene promoter in active and inactive loci was analysed more thoroughly by sequencing PCR-amplified fragments derived from bisulphite-treated genomic DNA, since it is within this region that methylation at specific sites is most likely to affect expression. During bisulphite treatment, cytosine residues are selectively converted to uracil, while 5-methylcytosines remain unchanged (25). When PCR amplification products from the region of interest are sequenced, only cytosine residues which were methylated in the original genomic template remain, with unmethylated cytosines now appearing as thymine. In this way, the methylation status of every cytosine residue within the amplified region can be determined. Further, by sequencing sub-cloned PCR products, the methylation status of single molecules rather than whole populations can be analysed.

Genomic DNA was obtained from the erythrocytes of 12.5 d.p.c. transgenic embryos of lines 27/5A (95% expressing cells) and 16/7B (2% expressing cells) and subjected to bisulphite conversion. Since transgenes contained within a multi-copy array are usually arranged in a head–tail orientation, PCR was used to amplify across head–tail transgene junctions in order to obtain a product

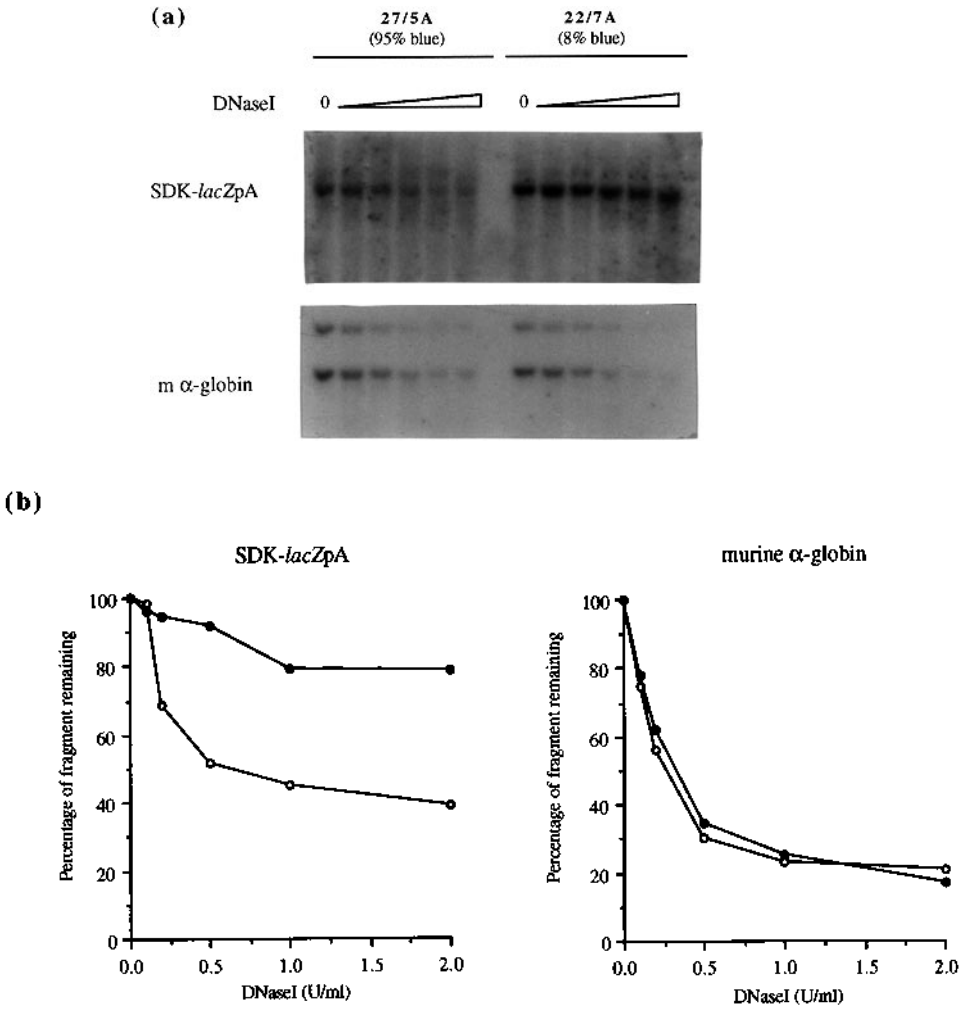


**Figure 3.** General DNase I sensitivity of the transgene in 12.5 d.p.c. erythrocyte nuclei of the lines 27/5A and 22/7A. (**a**) 12.5 d.p.c. erythrocyte nuclei of the lines 27/5A and 22/7A were digested with increasing concentrations of DNase I. After purification, the DNA was digested with*Pst*I, separated on a 0.8% agarose gel and Southern blotted. The membrane was hybridized with the SDK-*lacZ*pA cassette to detect a 5.4 kb fragment containing the entire *lacZ* gene and part of the αHS-40 fragment, but not the DNase I hypersensitive site. The SDK-*lacZ*pA fragment from erythrocytes of line 27/5A is more sensitive to DNase I than that from the 22/7A line. To control for equivalent DNase I digestion, the membrane was stripped and re-hybridized with a murine α-globin probe to detect the endogenous gene. (**b**) Phosphorimager analysis of the membranes shown in (a). For the murine α-globin probed membrane, only quantification of the lower band is shown. The vertical axis shows the proportion of fragment remaining as a percentage of the initial fragment (0 U/ml) at each DNase I concentration for the lines 27/5A (open circles) and 22/7A (closed circles).

which included the entire promoter region. For both 27/5A and 16/7B, diagnostic digests and Southern blotting confirmed that the majority of the transgenes within the array were arranged in a head–tail orientation (data not shown). PCR was performed using bisulphite-converted DNA as a template to generate a product of 332 bp which included the transgene promoter. PCR products were then sub-cloned into the plasmid vector KS Bluescript and individual clones sequenced to determine the methylation status of cytosine residues within the promoter (Fig. 5). There appears to be no significant difference in the overall frequency of cytosine methylation between the promoters of active and inactive transgenes and no specific site showed a consistent difference in methylation between the two states. None of the methylated cytosines detected reside within the recognized transcription factor binding sites of the promoter. All methylated cytosines detected in this study were found to reside in CpTpG trinucleotides and although CpG dinucleotides are the most common target for mammalian DNA methyltransferase, the CpG dinucleotide within the sequenced region was only methylated in one clone. Since the two cytosine residues which showed the highest frequency of methylation (–16 and –106) are situated within the recognition site of the bacterial *dcm* methylase, detection of these sites may not indicate *de novo* methylase activity within the mouse, but may represent maintenance in some transgenes of the methylation pattern imposed during plasmid propagation in *E.coli* prior to injection (29). The finding that promoters from inactive transgenes are no more extensively methylated than those from active loci is in agreement with results obtained using methylation-sensitive enzymes.

## DISCUSSION

Transgenes in which globin regulatory elements drive expression of *lacZ* are expressed in a heterocellular or variegated pattern in

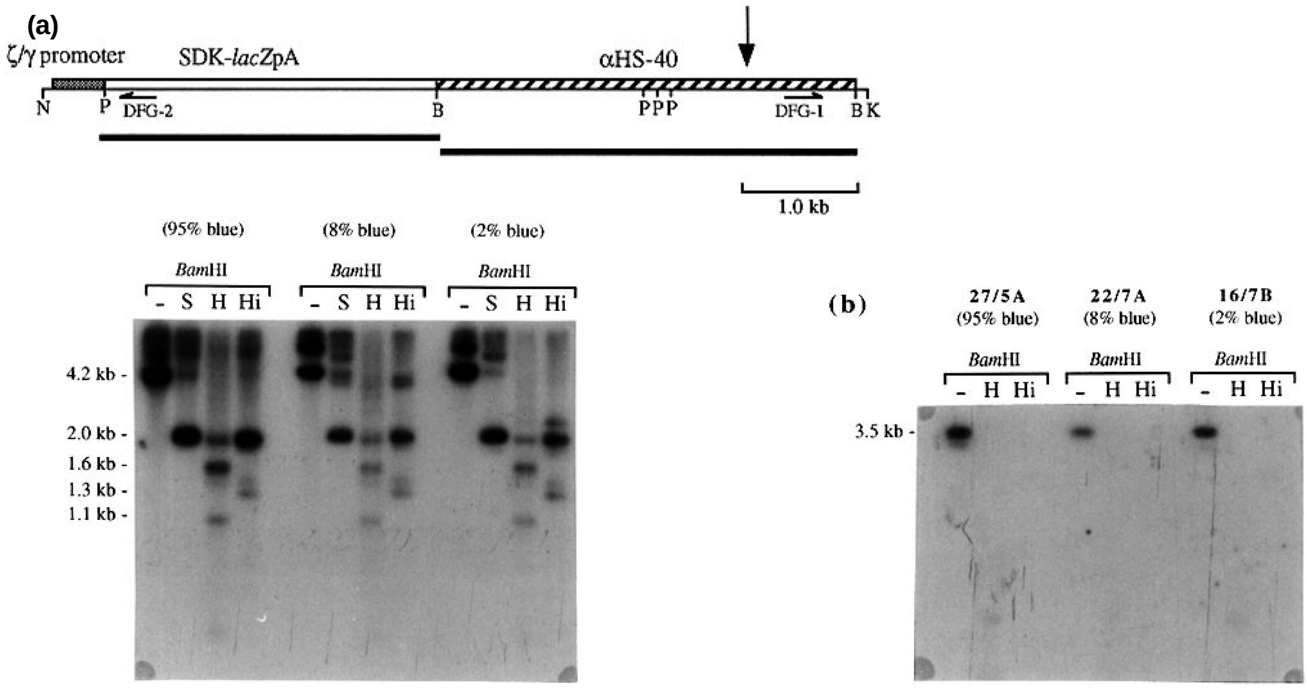


**Figure 4.** Susceptibility of the DM1 transgene to digestion by methylation-sensitive restriction enzymes. (**a**) Digestion of the αHS-40 fragment. The top panel shows the positions of recognition sites for the methylation-sensitive enzymes *Sma*I (S), *Hpa*II (H) and *Hin*PI (Hi) within the 4.2 kb *Bam*HI (B) αHS-40 fragment of the transgene (top line). Horizontal black lines indicate the products expected from complete digestion with the enzymes shown at left which are sufficiently large to be detected following separation on a 1.2% agarose gel and Southern blotting. The bottom panel shows DNA from 12.5 d.p.c. erythrocytes of lines 27/5A, 22/7A and 16/7B digested with the restriction enzymes shown in each lane and analysed by Southern blotting after separation on a 1.2% agarose gel. The membrane was hybridized with the entire 4.2 kb *Bam*HI parent fragment. The sizes of the observed fragments are indicated. (**b**) Digestion of the promoter/*lacZ* fragment. The same DNA as in (a) was digested with the restriction enzymes shown in each lane and analysed by Southern blotting as above. The membrane was hybridized with the SDK*lacZ*pA cassette. *Sma*I digestion was not included since there are no *Sma*I sites in this fragment of the transgene. For both fragments, the transgene from all three lines appears to be mainly unmethylated at the sites tested.

mouse erythrocytes. In the work presented here, a clear difference was observed between the chromatin structure of active and inactive transgenes. Silenced loci are characterized by an absence of the αHS-40 hypersensitive site and a reduced general sensitivity to DNase I relative to that of active loci. The finding that variegation of this transgene is associated with a closed chromatin configuration is consistent with other studies which have compared the chromatin structure of variegating transgenes between expressing and non-expressing populations. For human CD2 transgenes which exhibit variegated expression in mouse thymocytes, it was found that a transgene-specific DNase I hypersensitive site was present in expressing thymocytes but could not be detected in those cells in which the transgene had been silenced (11,13). Similarly, it was found that the variegated silencing of a *hsp26/white* transgene in *Drosophila* was accompanied by packaging within a more regular nucleosome array and decreased sensitivity to micrococcal nuclease (30).

Despite the observed difference in chromatin structure, we were unable to detect any difference in the extent of transgene methylation between active and silenced transgene loci. Transgenes from both active and inactive erythrocytes were equally susceptible to digestion by a variety of methylation-sensitive restriction enzymes. Further, analysis of methylation by sequencing PCR products obtained from bisulphite-converted genomic templates revealed no consistent difference in either the overall density of 5-methylcytosine within the promoter or in the methylation of any specific cytosine residue between active and inactive transgenes. While we have not analysed the methylation of every cytosine within the transgene and therefore cannot definitively rule out differential methylation at an as yet undetected specific site, these results strongly argue against methylation as the cause of variegated transgene silencing in this system.

While there are many reported examples in which the inactivation of a gene has been found to correlate with hypermethylation of its regulatory elements (15,16), there is still considerable controversy as to whether gene methylation is the primary inactivating event or whether it is a secondary modification which takes place on genes which have already been silenced by a different mechanism. Where methylation has been proposed as the cause of gene inactivation, the mechanism by which this is achieved remains unclear. The existence of transcription factors whose binding affinity is reduced by the methylation of cytosine residues within their cognate recognition sites (31–36) suggests that methylation can interfere with gene expression directly. However, other transcription factors show no difference in their ability to bind methylated and unmethylated target sequences (37) and the number of known methylation-sensitive transcription factors cannot account for all the reported observations of gene silencing associated with methylation. The observation that *in vitro* methylated DNA adopts a chromatin structure that is different from that of untreated DNA upon genomic integration (17) suggests that the repressive influence of methylation on gene expression can also be indirect and involve a change in chromatin structure. Consistent with these findings, it was reported that methylation alone was not sufficient for the transcriptional silencing of a HSV TK transgene but that repression was also

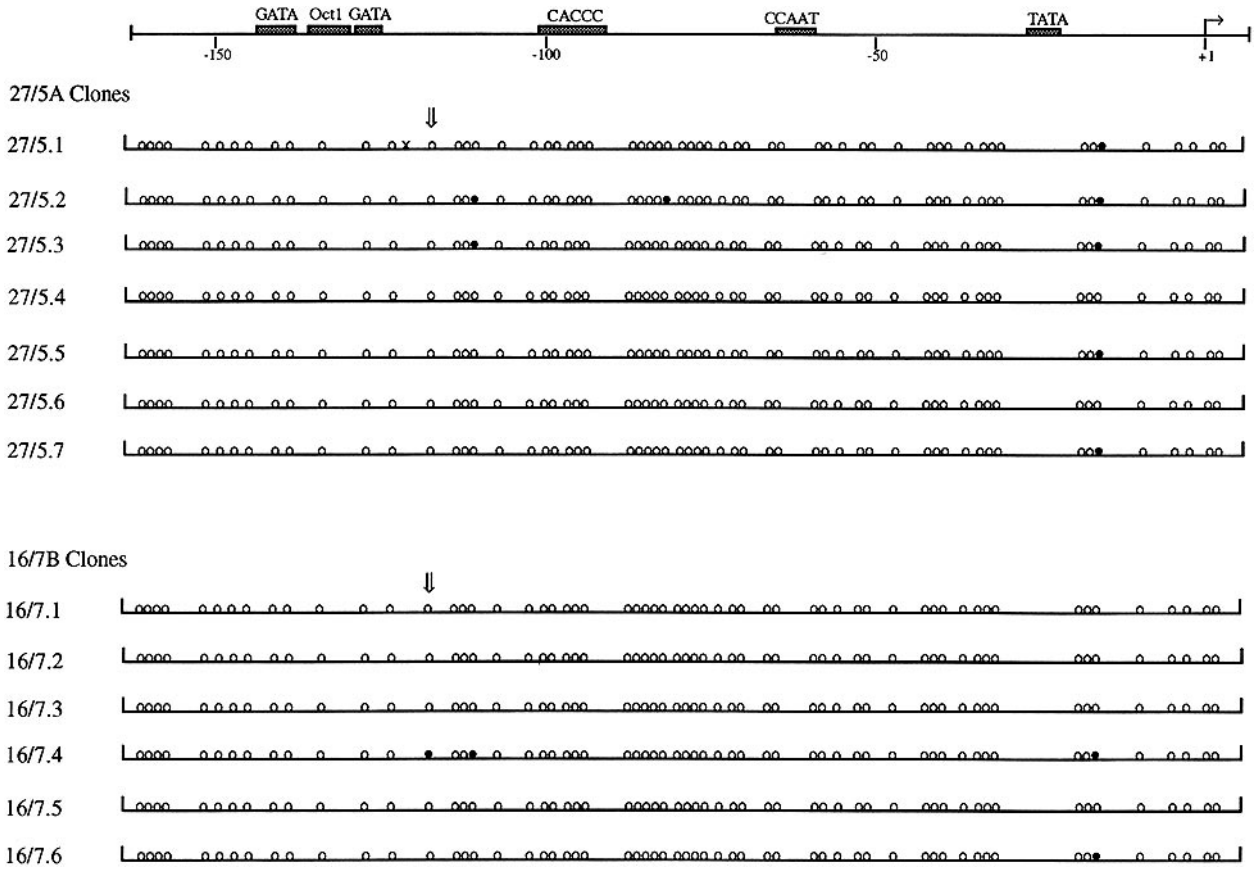


**Figure 5.** Methylation analysis of transgene promoters from lines 27/5A and 16/7B. Genomic DNA from 12.5 d.p.c. erythrocytes of lines 27/5A and 16/7B was bisulphite treated and used to generate a PCR product of 332 bp containing the entire transgene promoter across junctions between adjacent transgenes which are arranged in a head–tail orientation. PCR products were sub-cloned into the vector KS Bluescript and individual clones sequenced. The top line represents the 163 bp promoter contained in the transgene, showing the recognized transcription factor binding motifs and the transcription start site. The status of each cytosine residue in the promoter following bisulphite treatment is shown for each clone. Open circles represent cytosines which were converted to uracil by bisulphite treatment, indicating that they were unmethylated in the original template, and closed circles represent cytosines which remained unconverted following bisulphite treatment. The vertical arrow indicates the position of the CpG dinucleotide within the sequenced region.× represents a departure from the expected sequence (G→T) which was observed in one clone. One clone of the 27/5A PCR product and one clone of the 16/7B product exhibited localized regions in which all cytosines had failed to convert to uracil (data not shown). These cloned PCR products are probably derived from templates in which incomplete denaturation has resulted in local regions resistant to bisulphite conversion (25). The recovery of one of these clones from both 27/5A (95% of cells express the transgene) and 16/7B (2% express) is consistent with this being an artefact of the experimental system.

dependent on the formation of chromatin within the recipient cell (38,39).

A number of studies have investigated the role of methylation in establishing variegated patterns of expression. The proportion of neural tube cells expressing a mouse hsp68/*lacZ* transgene was found to be inversely proportional to the extent of transgene methylation (6) and the variegated expression of a maize A1 cDNA transgene in the flower of *Petunia hybrida* correlated both with high copy number and transgene methylation (18). Similarly, two variegating epi-alleles of the maize *P-rr* gene were found to be hypermethylated relative to the non-variegating parent allele (19). Kennedy *et al.* (40) found that methylation did not correlate with variegated silencing of a human α1 antitrypsin/*lacZ* transgene in mouse hepatocytes, although analysis was performed with only a single methylation-sensitive restriction enzyme. Our finding that the inactivity of the variegating DM1 transgene correlates with a less accessible chromatin structure but not increased methylation contrasts with these earlier studies and is consistent with the proposal that the activity of a gene is determined primarily by its location within a permissive or restrictive chromatin domain and is not dependent on methylation status *per se*. For some genes it may be that methylation of one or more cytosine residues is responsible for initiating the repressive chromatin structure. In other cases, different factors, including the proximity of the locus to spreading heterochromatin or the presence within the locus of multiple repeated copies of a monomeric unit (41), may trigger the formation of a repressive chromatin structure in a methylation-independent manner. The primary importance of chromatin structure in determining the activity of a variegating locus is consistent with the observation of variegated expression patterns in *Drosophila*, where methylation of genomic DNA does not occur.

## ACKNOWLEDGEMENTS

This work was supported by a grant from the National Health and Medical Research Council of Australia to E.W. D.G. and H.S. are both supported by Australian Postgraduate Awards. We thank

Hugh Morgan for his contribution to the work, David Martin for the DM1 clone, Susan Clark for advice on bisulphite treatment of genomic DNA and Merlin Crossley for helpful comments on the manuscript.

## REFERENCES


1 Schultz,J. (1936) *Proc. Natl. Acad. Sci. USA*, **22**, 27–33.
2 Tartof,K.D., Hobbs,C. and Jones,M. (1984) *Cell*, **37**, 869–878.
3 Eissenberg,J.C. (1989) *BioEssays*, **11**, 14–17.
4 Rubin,D.C., Ong,D.E. and Gordon,J.I. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 1278–1282.
5 Sweetser,D.A., Hauft,S.M., Hoppe,P.C. and Birkenmeier,E.H. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 9611–9615.
6 McGowan,R., Campbell,R., Peterson,A. and Sapienza,C. (1989) *Genes Dev.*, **3**, 1669–1676.
7 Bradl,M., Larue,L. and Mintz,B. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 6447–6451.
8 Mintz,B. and Bradl,M. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 9643–9647.
9 Thorey,I.S., Meneses,J.J., Neznanov,N., Kulesh,D.A., Pedersen,R.A. and Oshima,R.G. (1993) *Dev. Biol.*, **160**, 519–534.
10 Pravtcheva,D.D., Wise,T.L., Ensor,N.J. and Ruddle,F.H. (1994) *J. Exp. Zool.*, **268**, 452–468.
11 Elliott,J.I., Festenstein,R., Tolaini,M. and Kioussis,D. (1995) *EMBO J.*, **14**, 575–584.
12 Perry,W.L., Hustad,C.M., Swing,D.A., Jenkins,N.A. and Copeland,N.G. (1995) *Genetics*, **140**, 267–274.
13 Festenstein,R., Tolaini,M., Corbella,P., Mamalaki,C., Parrington,J., Fox,M., Miliou,A., Jones,M. and Kioussis,D. (1996) *Science*, **271**, 1123–1125.
14 Martin,D.I.K. and Whitelaw,E. (1996) *BioEssays*, in press.
15 Cedar,H. (1988) *Cell*, **53**, 3–4.
16 Razin,A. and Cedar,H. (1991) *Microbiol. Rev.*, **55**, 451–458.
17 Keshet,I., Lieman-Hurwitz,J. and Cedar,H. (1986) *Cell*, **44**, 535–543.
18 Linn,F., Heidmann,I., Saedler,H. and Meyer,P. (1990) *Mol. Gen. Genet.*, **222**, 329–336.
19 Das,O.P. and Messing,J. (1994) *Genetics*, **136**, 1121–1141.
20 Martin,C.C. and McGowan,R. (1995) *Genet. Res. Camb.*, **65**, 21–28.
21 Robertson,G., Garrick,D., Wu,W., Kearns,M., Martin,D. and Whitelaw,E. (1995) *Proc. Natl. Acad. Sci. USA*, **92**, 5371–5375.
22 Robertson,G., Garrick,D., Wilson,M., Martin,D.I.K. and Whitelaw,E. (1996) *Nucleic Acids Res.*, **24**, 1465–1471.
23 Hogan,B., Beddington,R., Costantini,F. and Lacy,E. (1994) *Manipulating the Mouse Embryo: A Laboratory Manual*, 2nd Edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
24 Higgs,D.R., Wood,W.G., Jarman,A.P., Sharpe,J., Lida,J., Pretorius,I.-M. and Ayyub,H. (1990) *Genes Dev.*, **4**, 1588–1601.
25 Clark,S.J., Harrison,J., Paul,C.L. and Frommer,M. (1994) *Nucleic Acids Res.*, **22**, 2990–2997.
26 Brotherton,T.W., Chui,D.H.K., Gauldie,J. and Patterson,M. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 2853–2857.
27 Jarman,A.P., Wood,W.G., Sharpe,J.A., Gourdon,G., Ayyub,H. and Higgs,D.R. (1991) *Mol. Cell. Biol.*, **11**, 4679–4689.
28 Smith,R.D., Yu,J., Annunziato,A. and Seale,R.L. (1984) *Biochemistry*, **23**, 2970–2976.
29 Clark,S.J., Harrison,J. and Frommer,M. (1995) *Nature Genet.*, **10**, 20–27.
30 Wallrath,L.L. and Elgin,S.C.R. (1995) *Genes Dev.*, **9**, 1263–1277.
31 Becker,P.B., Ruppert,S. and Schütz,G. (1987) *Cell*, **51**, 435–443.
32 Kovesdi,I., Reichel,R. and Nevins,J.R. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 2180–2184.
33 Watt,F. and Molloy,P.L. (1988) *Genes Dev.*, **2**, 1136–1143.
34 Hermann,R., Hoeveler,A. and Doerfler,W. (1989) *J. Mol. Biol.*, **210**, 411–415.
35 Iguchi-Ariga,S.M.M. and Schaffner,W. (1989) *Genes Dev.*, **3**, 612–619.
36 Comb,M. and Goodman,H.M. (1990) *Nucleic Acids Res.*, **18**, 3975–3982.
37 Hsller,M., Westin,G., Jiricny,J. and Schaffner,W. (1988) *Genes Dev.*, **2**, 1127–1135.
38 Buschhausen,G., Graessmann,M. and Graessmann,A. (1985) *Nucleic Acids Res.*, **13**, 5503–5513.
39 Buschhausen,G., Wittig,B., Graessmann,M. and Graessmann,A. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 1177–1181.
40 Kennedy,S., Rettinger,S., Flye,M.W. and Ponder,K.P. (1995) *Hepatology*, **22**, 160–168.
41 Dorer,D.R. and Henikoff,S. (1994) *Cell*, **77**, 1–20.