Development 111, 573-581 (1991)
Printed in Great Britain © The Company of Biologists Limited 1991

Inherited type of allelic methylation variations in a mouse chromosome region where an integrated transgene shows methylation imprinting

HIROYUKI SASAKI[1,*], TSUYOSHI HAMADA[1], TAKAYUKI UEDA[2], RITSUKO SEKI[1], TORU HIGASHINAKAGAWA[2] and YOSHIYUKI SAKAKI[1]

[1]*Research Laboratory for Genetic Information, Kyushu University 18, Fukuoka 812 Japan;* [2]*Department of Developmental Biology, Mitsubishi Kasei Institute of Life Sciences, 11 Minamiooya, Machida, Tokyo 194, Japan*

*Present address: Department of Molecular Embryology, AFRC Institute of Animal Physiology and Genetics Research, Babraham, Cambridge CB2 4AT, UK

Summary

It is still unclear whether or not parent-of-origin-dependent differential methylation observed in some transgenes reflects genomic imprinting of endogenous genes. We have characterized a transgene locus showing such methylation imprinting together with the corresponding-native chromosome region. We show that only part of the transgene is affected by methylation imprinting and the methylation pattern is established before early prophase I during spermatogenesis. Interestingly, the native genomic region, which is mapped to the proximal chromosome 11, shows no evidence of methylation imprinting but displays heritable, strain-specific type of allelic methylation differences. The results demonstrate that transgenes do not necessarily reflect the methylation status of either the surrounding or corresponding chromosome region. In addition, inherited type of allelic methylation variations previously described in human may be widespread in mammals.

Key words: transgenic mouse, DNA methylation, genomic imprinting.

Introduction

There is evidence that certain loci of mammalian autosomes display allelic difference with respect to their DNA methylation levels (reviewed by Monk, 1990). One such example is a heritable variation, which was identified in human pedigrees by discriminating parental alleles using polymorphic tandem repeat markers (Silva and White, 1988). The findings suggest that the programs for tissue-specific DNA methylation can be transmitted through the germ line unchanged. However, some transgenic mice show 'methylation imprinting', in which the methylation level of the transgene is determined by the sex of the parent from which the transgene is transmitted (Reik *et al.* 1987; Sapienza *et al.* 1987; Swain *et al.* 1987; Hadchouel *et al.* 1987; for review, see Surani *et al.* 1988). In most of these cases (with one exception reported by Hadchouel *et al.* 1987), methylation patterns are erased and re-established upon each passage through the germ line and then propagated to somatic tissues.

The parent-of-origin-dependent allelic methylation differences are consistent with a role for DNA methylation in genomic imprinting. Genomic imprinting is a process that modifies maternal and paternal chromosomes differently by an unknown mechanism, resulting in the differential expression of some parental alleles. The importance of imprinting during mouse embryogenesis has been established with ample experimental evidence (for review, see Solter, 1988; Surani *et al.* 1990) and, more recently, the possible importance of imprinting in tumorigenesis and expression of disease phenotypes has been suggested (for review, see Reik, 1989). Since observed methylation imprinting is apparently independent of sequence, size or copy number of the transgenes, the parental influence on methylation can be a consequence of chromosomal position effect. Thus methylation imprinting of transgenes could be a reflection of the methylation behavior of the surrounding or corresponding host chromosome regions and, if DNA methylation were to act as the imprinting mechanism, one might expect to clone 'imprinted' endogenous loci by using the transgenes as a tag.

However, several findings on methylation imprinting of transgenes are not consistent with the possibility discussed above especially because of the discrepancies with the map for chromosome regions containing imprinted genes that was produced by genetic complementation experiments (Searle and Beechey, 1985; Cattanach, 1986). First, if we assume both random integration of transgenes and random distribution of endogenous imprinted loci throughout the genome, the

frequency of transgenes showing methylation imprinting (approximately 30%) is apparently higher than expected. Second, the transgene was undermethylated after paternal transmission in all instances except one (Sapienza *et al.* 1987), but no such extreme bias is expected from the map. Third, the only transgene that has been mapped was present on chromosome 13 (Hadchouel *et al.* 1987) which, based on the genetic experiments, appears not to be subject to imprinting. These observations raise the possibility that transgene methylation imprinting may not be relevant to genomic imprinting of the host chromosomes and, therefore, transgenes cannot be used as probes to identify endogenous imprinted regions.

The first but crucial step to resolve this apparent paradox is to clone and characterize the site of integration of the transgenes, asking whether the flanking or corresponding host sequences also show differential methylation patterns.

We report the analysis of a new transgenic mouse line in which part of the transgene shows methylation imprinting. We describe cloning, chromosome mapping and detailed methylation analysis of the transgene locus. Furthermore, we have tested both the flanking sequences of transgenic mice and the corresponding native region of non-transgenic mice for allelic methylation variation. Our results demonstrate that the methylation level of neither the flanking nor the native chromosome region is changed by the parent of origin. Instead, however, we have found a completely different type of allelic methylation variation in the native chromosomes, which is strain-specific and stably transmitted unchanged through the germ line.

Materials and methods

Mice

The derivation and basic characterization of the transgenic line MPA434 is described in previous reports (Sasaki *et al.* 1986, 1989). Embryos used for injection of DNA were the result of matings between (C57BL/6J×C3H)F_1 females and BALB/cJ males. MOL-MSM, a strain established from *Mus musculus molossinus* mice originally collected at Mishima (Harada *et al.* 1989), was provided from the National Institute of Genetics, Mishima, Japan.

DNA extractions and Southern blotting

Extraction of DNA from mouse tissues and Southern blotting analysis were performed as described (Sasaki *et al.* 1989). To confirm complete digestion, 1 μg of lambda DNA was included in each sample as a marker. Labeling of the probes was done by using the Multiprime DNA labeling system (Amersham, UK).

Genomic libraries and screening

The MPA434 genomic library was constructed from tail DNA of a MPA434 male homozygous for the transgene by using EMBL3 as a vector according to Frischauf *et al.* (1987). The BALB/cJ library (vector: Charon28) was obtained from Japanese Cancer Research Resources Bank and the DBA/2J library (vetor: EMBL3) from Clontech (Palo Alto, USA). These libraries were screened by plaque hybridization as described (Maniatis *et al.* 1982).

In situ *hybridization*

Metaphase chromosomes were prepared from primary cultures of lung fibroblast obtained from 18-day mouse embryos. *In situ* hybridization and identification of chromosomes by the replication banding method were performed essentially as described (Buckle and Craig, 1986), except that the post-hybridization wash was done with solutions containing 30–50% formamide at lower temperatures (39–41°C).

Results

Methylation imprinting of the transgene promoter

We previously analyzed the tissue-specific activity and methylation status of a transgene promoter in three transgenic mouse lines (Sasaki *et al.* 1989). These lines were established by injecting a 10.8 kb DNA fragment which contains the mouse metallothionein-I promoter linked to the human transthyretin (also called prealbumin) structural gene, as well as portions of pUC18 plasmid at both ends (Sasaki *et al.* 1986). During the early experiments, we noticed that one line, designated MPA434, showed methylation imprinting at the transgene promoter, and this line was analyzed further. The other two strains showed no evidence of methylation imprinting in that their transgene promoter is always non-methylated irrespective of the parent of origin. As described previously, MPA434 mice contain approximately one copy of the transgene, which is expressed at a very low level to produce the human transthyretin (Sasaki *et al.* 1989).

The methylation status of the promoter in tail DNAs of MPA434 upon transmission through male and female germ lines is summarized in Fig. 1A. The restriction map and design of this experiment is depicted in Fig. 4A,a. If the *Hha*I (a restriction endonuclease that cleaves GCGC but not G^{m}CGC; ^{m}C, 5-methylcytosine) sites close to the transcription start site are non-methylated, a 1.8 kb band appears, but if they are methylated, a 5.6 kb band is seen. As shown in Fig. 1A, when this transgene was derived from the male germ line, these *Hha*I sites were always undermethylated (25/25 cases, of which 10 cases were derived from a homozygous male). By contrast, when derived from the female germ line, the sites were always hypermethylated (13/13). If transgenic females were bred with transgenic males, the offspring showed the undermethylated, hypermethylated and combined (undermethylated plus hypermethylated) patterns at an approximate ratio of 1:1:1 (11/30: 10/30: 9/30). We used BALB/cJ, C57BL/6J, (C57BL/6J×C3H)F_1 and MOL-MSM (see later) mice for mating, but no change in methylation phenotype was observed on these genetic backgrounds. The expression level of the transgene was also examined after maternal and paternal transmissions but we were unable to detect any differences either in somatic tissues or in serum probably because the expression levels were very low.

The methylation pattern of the transgene promoter in other tissues was basically the same as that in the tail, except for the testis (Sasaki *et al.* 1989). The observed undermethylation of maternally derived transgene in

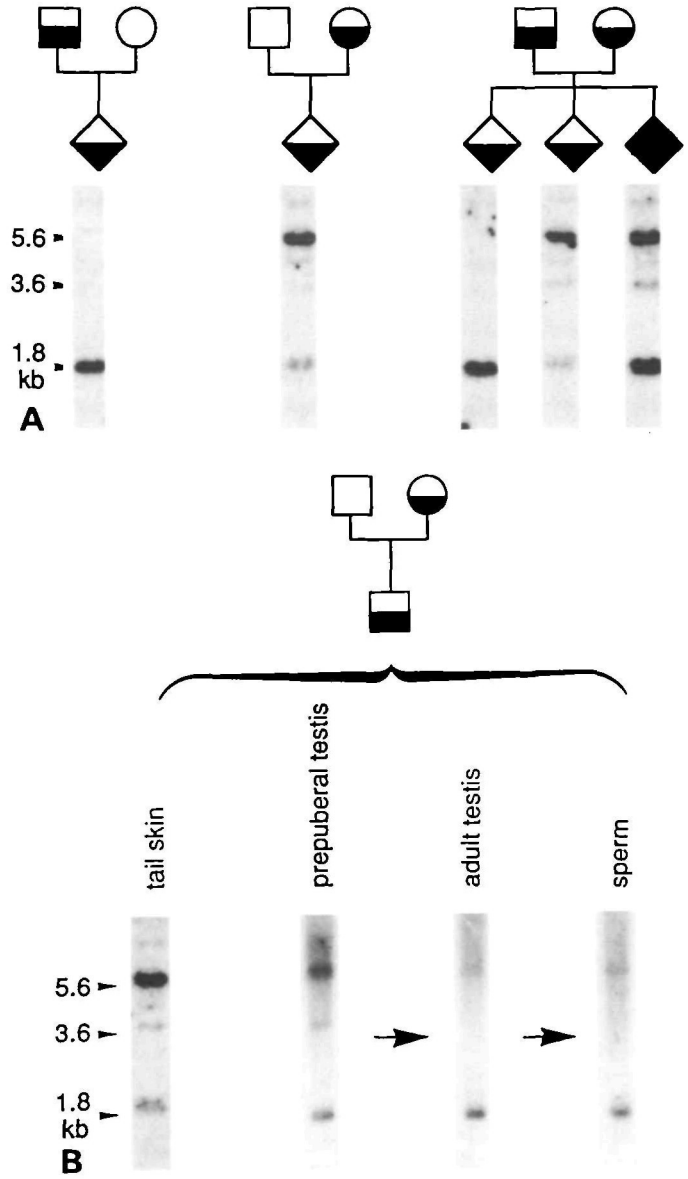

Fig. 1. Methylation imprinting of the transgene promoter in MPA434 strain. The diagram for the interpretation of the restriction fragments is shown in Fig. 4A,a.
(A) Methylation status of the transgene promoter of tail DNA upon paternal and maternal transmissions.
(B) Methylation status of the promoter in male germ cells from mice with maternally derived transgene.

testis suggests that the differential methylation pattern is established during gametogenesis. We have extended the study by analyzing prepuberal testis and sperm from adult epididymis, in both cases from mice with maternally derived transgene (Fig. 1B). The results showed that, already at day 10 after birth, testis DNA is relatively undermethylated at the promoter when compared with the tail DNA. At this point of prepuberal testis development, the germ cells reach the leptotene spermatocyte stage (Bellve *et al.* 1977). The methylation level was further decreased in adult testis and then remained almost unchanged in sperm. In contrast, *Hha*I sites in the body of the transgene, which is methylated in somatic tissues irrespective of the parental origin (see Fig. 4B,b), was also methylated throughout the course of germ cell development (data not shown). These results suggest that the methylation pattern observed in adult somatic tissues is established prior to the zygotene stage of meiosis prophase I during spermatogenesis.

Cloning, structural analysis and chromosome mapping of the transgene locus

In order to analyze the structure of the transgene locus and to determine its chromosomal location, we have cloned the transgene together with the flanking mouse DNA. A genomic library, constructed from tail DNA of a MPA434 mouse, was screened with a mixture of transgene probes A and B (Fig. 2) and, as a result, three overlapping clones were obtained. The cloned region spans 33 kb and its restriction map is shown in Fig. 2. Fine restriction mapping and partial nucleotide sequencing showed that the transgene integrant consists of an almost intact copy of the injected DNA plus a 3′ part (238 bp) of another copy linked to the 5′ end of the intact copy.

To obtain unique sequence probes, various flanking mouse DNA fragments were subcloned and tested by hybridization to the total mouse genomic DNA. As a result, six fragments designated D to I (Fig. 2, for details, see legend) were shown to be unique. Two genomic libraries constructed from normal mice (BALB/cJ and DBA/2J, respectively) DNA were then screened with a mixture of probes D, F, G, H and I to obtain the corresponding mouse chromosome region. A total of six overlapping phage clones covering 32 kb were isolated. By comparing the restriction map of this region with that of the transgene locus, we detected a loss of one cellular *Sal*I site but no other change of the host genome in transgenic mice (Fig. 2). Further analysis by nucleotide sequencing detected only one change in the host DNA, a 34 bp deletion at the site of transgene insertion. The lost *Sal*I site was located just at the 3′ end of the deleted sequence.

In an attempt to characterize this chromosome region, we next determined partial nucleotide sequences of the fragments D to I. However, computer-aided searches of DNA databases failed to identify any known sequences homologous to these. Also, within the cloned region, there was no evidence for CpG islands, which are often found associated with gene promoters (Bird, 1986). In fact, this region contains a relatively small number of sites for methylation-sensitive enzymes whose recognition sequences contain CpG dinucleotides (see Fig. 4A and 5A). Then, the chromosomal location of the region was determined by *in situ* hybridization analysis of metaphase chromosomes using fragment I as a probe. All together, 32 metaphase nuclei were examined. Out of a total of 130 sites of hybridization sites scored, 22 were located on chromosome 11. Fig. 3 shows the distribution of the grains on chromosome 11 and maps the locus to band A5. It is interesting to note that the proximal

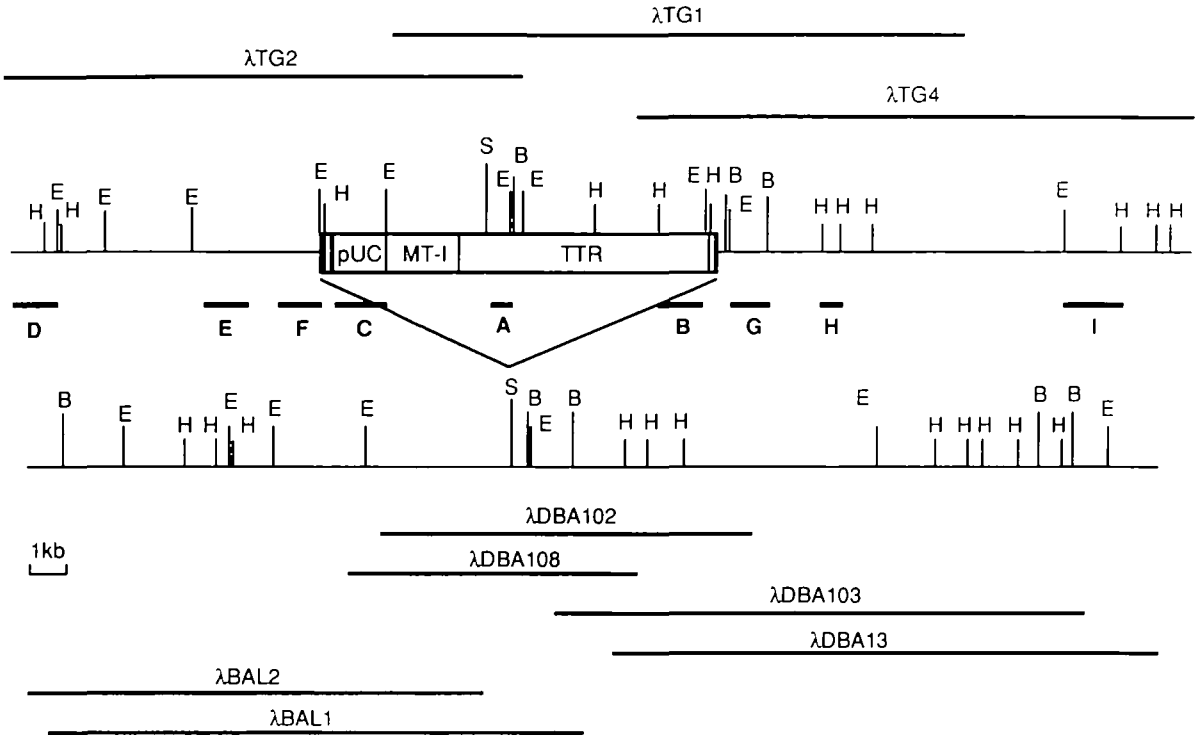

Fig. 2. Structure and restriction map of the transgene locus and the corresponding normal chromosome region. Genomic clone series λTG, λDBA and λBAL were isolated from the libraries constructed from DNA of transgenic mouse, DBA/2J and BALB/cJ, respectively. pUC, pUC18 plasmid sequence; MT-I, mouse metallothionein-I promoter; TTR, human transthyretin gene. Restriction sites shown are: B, *Bam*HI; E, *Eco*RI; H, *Hin*dIII; S, *Sal*I. Transgene probes and isolated unique sequence probes are shown by solid lines and labeled with bold capital letters. These probe fragments are: A, *Hin*cII–*Eco*RI 0.65 kb; B, *Hin*dIII–*Eco*RI 1.05 kb; C, *Eco*RI–*Pvu*I 1.6 kb; D, *Sal*I–*Eco*RI 1.05 kb; E, *Bgl*II–*Xba*I 1.2 kb; F, *Xba*I–*Bgl*II 1.05 kb; G, *Eco*RI–*Bam*HI 1.1 kb; H, *Hin*dIII 0.65 kb; I, *Eco*RI–*Hin*dIII 1.6 kb.

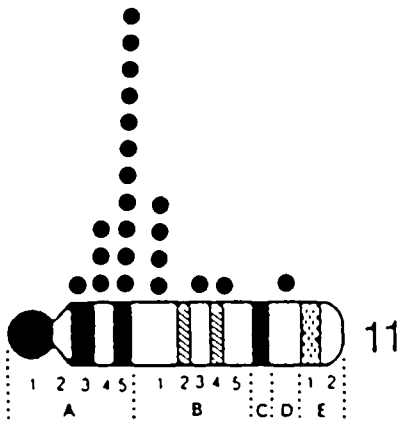

Fig. 3. Chromosomal localization of the locus to chromosome 11 band A5 by *in situ* hybridization.

chromosome 11 including band A5 (Searle *et al.* 1989) has been shown to contain imprinted gene(s) by genetic studies (Cattanach and Kirk, 1985).

Methylation status of the other part of the transgene and the flanking chromosome sequences

By using probes from the transgene and the flanking region, the methylation status of various portions of the transgene and flanking mouse sequences was analyzed following paternal and maternal transmission. Fig. 4A shows the overall structure of the transgene locus and diagrams for the observed restriction bands. Two methylation-sensitive enzymes, *Hha*I and *Hae*II (*Hae*II cuts PuGCGCPy but not PuG^{m}CGCPy), were used and each sample was doubly digested with one of these enzymes in combination with another restriction enzyme as indicated.

As already shown in Fig. 1A, the *Hha*I site cluster within the transgene promoter derived from the mouse metallothionein-I gene was undermethylated when transmitted through the male germ line but methylated when transmitted through the female germ line (Fig. 4B,a). The same parental effect was seen on the two *Hae*II sites within the pUC sequences located upstream of the promoter (b). (One of the expected bands, which should be 0.4 kb in size, could not be detected in this experiment.) In contrast, when three *Hha*I sites within the body of the transgene were examined, a 4.9 kb band resulting from the *Eco*RI digestion and a faint 4.2 kb band were always seen irrespective of the parental origin (c). This suggests that the most upstream *Hha*I site is always partially demethylated but the other two sites are always completely methylated.

In the downstream flanking region, three *Hae*II sites were tested by the transgene probe B. The result showed that these sites were always methylated

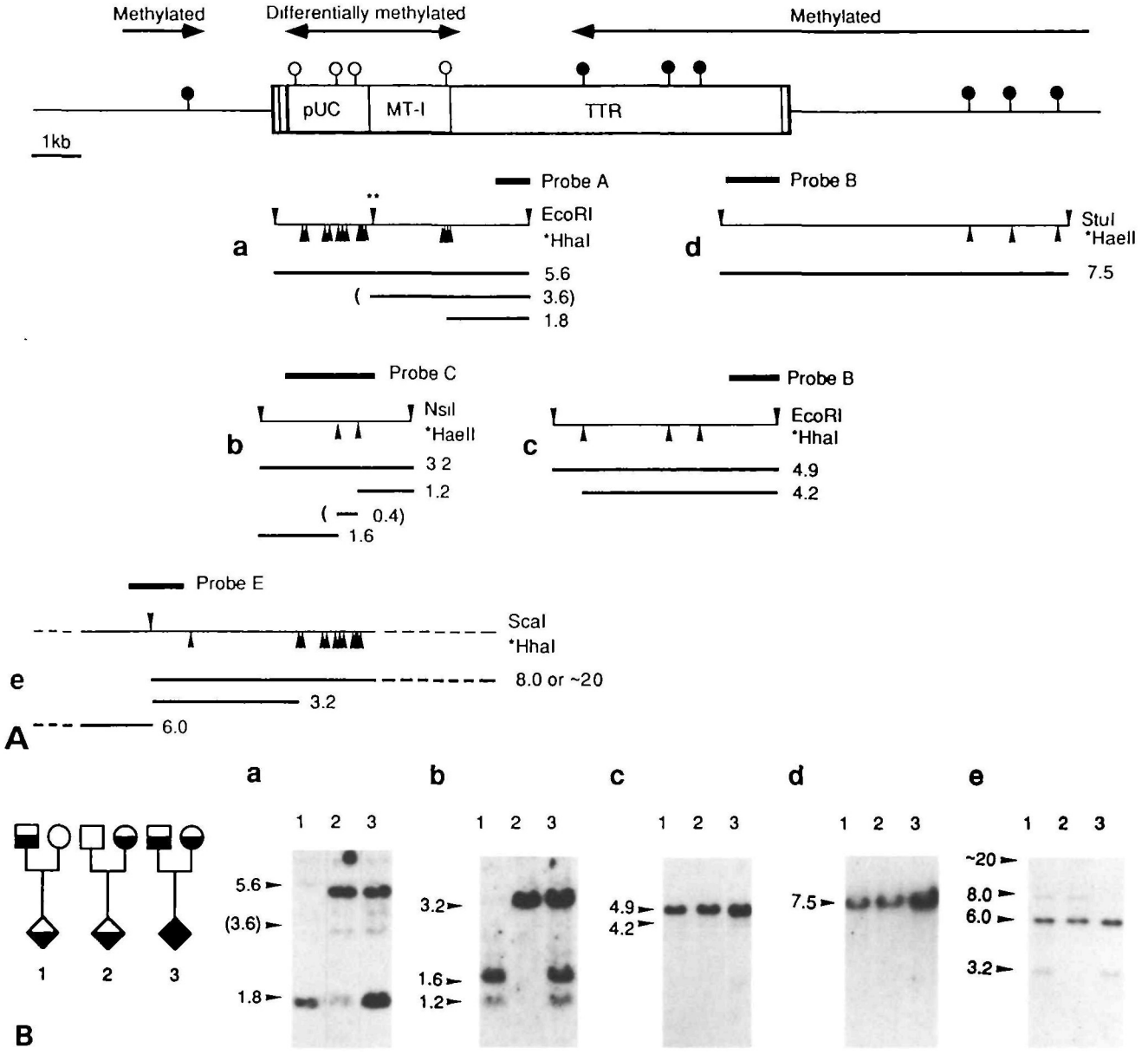

Fig. 4. Detailed methylation analysis of the transgene and the flanking mouse DNA. (A) Diagrams for the interpretation of the restriction fragments observed. Overall structure of the locus and summary of the results obtained are also shown (top). Open circles indicate the sites showing methylation imprinting and filled circles, the sites always methylated irrespective of the parental origin. Samples were digested with the combination of enzymes (methylation-sensitive enzymes are marked with *) shown and hybridized with the probe indicated (diagrams a-e). The *Eco*RI site marked with ** was cut very little when the surrounding pUC and metallothionein sequences were methylated, giving rise to a 5.6 kb band instead of an expected 3.6 kb band. This is probably because the cytosine residue of this *Eco*RI recognition sequence is subject to methylation. In fact, the residue next to the cytosine is guanine on the pUC side and thus creates a CpG dinucleotide, the target sequence of methylase. It is known that *Eco*RI does not cut GAATT^{m}C. (B) Autoradiograms representing the results. Each autoradiogram corresponds to the diagram marked with the same letter in A. Tail DNA samples analyzed were from mice with paternally (lane 1) and maternally (lane 2) derived transgenes and mice having both maternal and paternal copies (lane 3).

irrespective of the parental origin (only one band of 7.5 kb resulting from *Stu*I digestion was observed in all situations) (d). To test the upstream flanking region, we used probe E from the flanking mouse DNA since the transgene probes could not detect the flanking sequences when the pUC sequences were not methylated. This is because the pUC region contains many methylation-sensitive enzyme sites and is cut into small pieces if non-methylated. When the samples were digested with *Sca*I only, probe E detected a 20 kb band derived from the chromosome with the transgene insertion and an 8.0 kb band derived from the normal counterpart, together with a common 6.0 kb band. A double digestion with *Sca*I and *Hha*I showed that the

*Hha*I site located 2.3 kb upstream of the insertion junction is invariably methylated after maternal and paternal transmissions (e).

The results of these experiments are summarized at the top of Fig. 4A. Only part of the transgene is affected by methylation imprinting and a clustering of the differentially methylated sites within the transgene is evident. Furthermore, our results demonstrate a lack of methylation imprinting in the flanking mouse DNA.

Methylation status of the corresponding native chromosome region in non-transgenic mice

Although we did not observe methylation imprinting in the flanking region, it may be interesting to ask whether the corresponding native chromosome region shows such methylation variation in non-transgenic mice. To obtain a marker for discriminating between paternally and maternally derived chromosomes, we looked for restriction fragment length polymorphisms between different mouse strains. We first tested BALB/cJ and C57BL/6J DNAs with four unique sequence probes (D, E, H and I) and 24 restriction enzymes but failed to find any polymorphisms. We therefore examined DNA from several wild-derived strains of mice established at the National Institute of Genetics, Mishima (samples kindly provided by Dr T. Shiroishi and Dr K. Moriwaki). Among the 11 restriction enzymes examined with probe E, *Eco*RI identified a polymorphism between BALB/c and MOL-MSM, a Japanese strain of wild mice referred to as *Mus musculus molossinus* (Harada *et al.* 1989). Upon *Eco*RI digestion, DNA from BALB/cJ gave a 4.5 kb band while DNA from MOL-MSM gave a 6.3 kb band (Fig. 5B,a). This polymorphism can be attributed to the absence of an *Eco*RI site in MOL-MSM as illustrated in Fig. 5A. Then a BALB/cJ male was bred with a MOL-MSM female and DNA from tail of both parents and an F_1 progeny was analyzed for methylation.

By using the *Eco*RI polymorphism, we could test a total of three methylation-sensitive enzyme sites around the insertion points. First, the *Hpa*II (which cuts CCGG but not C^{m}CGG) site which is only 70 bp upstream from the insertion junction was analyzed. When samples were doubly digested with *Eco*RI plus an isoschizomer of *Hpa*II, *Msp*I (which cleaves both CCGG and C^{m}CGG), BALB/cJ DNA gave a 3.7 kb band, MOL-MSM DNA a 4.9 kb band, and F_1 DNA a combination of these (Fig. 5B,b). However, a double digestion with *Eco*RI and *Hpa*II generated the same patterns as the *Eco*RI digestion only (c), suggesting that this site is always methylated regardless of the parental origin. Second, a *Hha*I site located 2.3 kb upstream of the transgene insertion point (the same *Hha*I site analyzed in Fig. 4B,e) was tested. Although a faint 3.3 kb band indicative of partial demethylation of this site was seen in both MOL-MSM and F_1 samples, this change was not dependent on the parental origin since we did not observe corresponding partial digestion (which should produce a 1.5 kb band) in BALB/cJ samples (d). Third, the *Sal*I site, which was lost in transgenic mice by the deletion, was examined (*Sal*I cuts GTCGAC but not GT^{m}CGAC). Upon double digestion with *Eco*RI and *Sal*I, the samples gave bands diagnostic for the cleavage at the *Sal*I site (3.9 and 5.6 kb) (e). Although a faint band of undigested 4.5 kb *Eco*RI fragment was observed in BALB/cJ and F_1 samples, again, the methylation variation was not dependent on the parental origin since neither of the two homologous chromosomes of MOL-MSM showed corresponding undigested band which should be 6.2 kb in size. Thus, we did not observe methylation imprinting at any of the three sites tested.

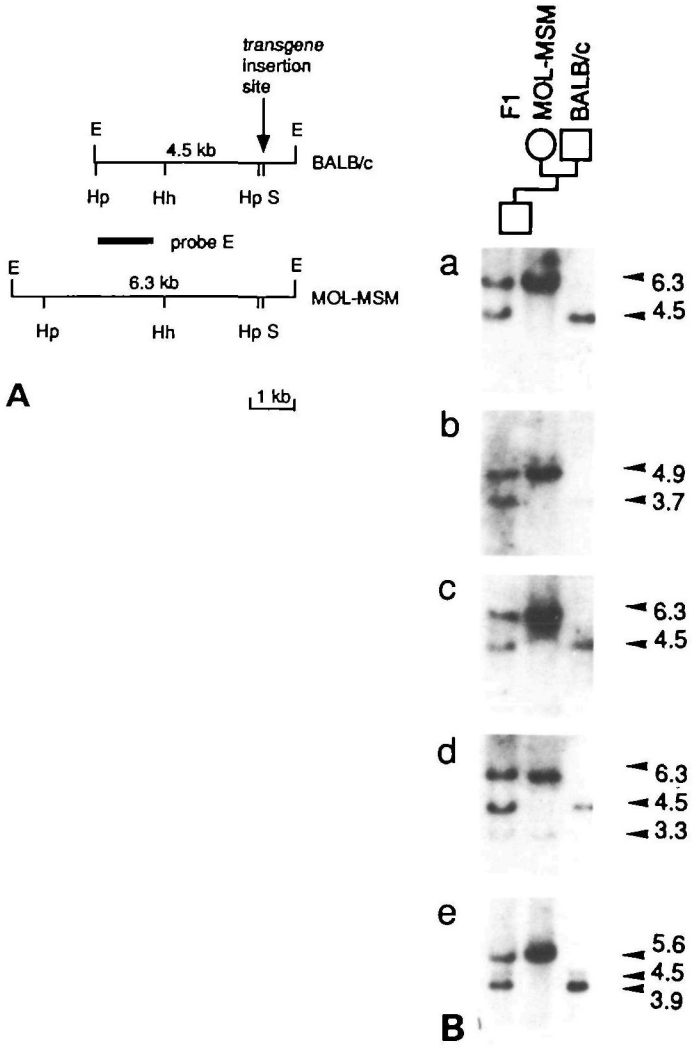

Fig. 5. Methylation analysis of the native chromosome region corresponding to the transgene locus in non-transgenic mice. (A) Map of the BALB/c and MOL-MSM chromosomes showing the *Eco*RI polymorphism and methylation-sensitive enzyme sites tested. E, *Eco*RI; Hh, *Hha*I; Hp, *Hpa*II; S, *Sal*I. (B) Autoradiograms showing the results. Tail DNA samples were digested with: a, *Eco*RI only; b, *Eco*RI+*Msp*I; c, *Eco*RI+*Hpa*II; d, *Eco*RI+*Hha*I; e, *Eco*RI+*Sal*I.

However, the data presented above provide another example of allelic methylation variation. Interestingly, it seems that the allelic methylation patterns are strain-specific. For instance, partial demethylation of the

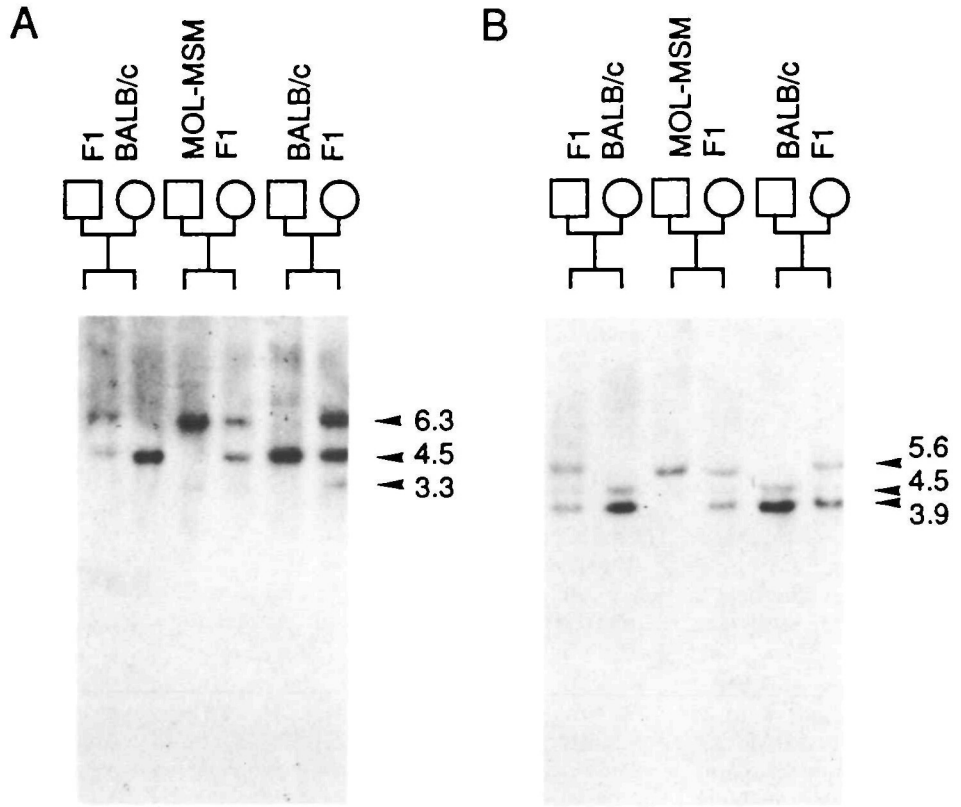

Fig. 6. Inheritance of the allelic methylation differences to F2 progenies of backcrossed F_1. Tail DNA samples were doubly digested with either *Eco*RI+*Hha*I (A) or *Eco*RI+*Sal*I (B).

*Hha*I site was observed for the chromosomes of MOL-MSM origin but not for the chromosomes derived from BALB/cJ. Similarly, partial methylation of the *Sal*I site was likely to be specific to the chromosomes of BALB/cJ origin.

To establish whether these methylation variations are strain-specific and whether they are stably transmitted through the germ line, we backcrossed F_1 mice with either of the parental strains and analyzed the F2 progenies. As shown in Fig. 6A, the *Hha*I site was always partially demethylated on the chromosomes of MOL-MSM origin but never on the BALB/cJ-derived chromosomes. Similarly, the partial methylation of the *Sal*I site invariably co-segregated with the BALB/cJ chromosome (Fig. 6B). These results demonstrate the presence of an inherited type of allelic methylation variation in the mouse, in a chromosome region where an inserted transgene shows methylation imprinting.

Discussion

Transgene methylation imprinting in the MPA434 line

We have identified and characterized a new transgene locus showing methylation imprinting in one of three lines carrying a fusion gene construct. Characteristics of the methylation imprinting observed are basically very similar to previous reports (Reik *et al.* 1987; Sapienza *et al.* 1987; Swain *et al.* 1987): the transgene promoter was more methylated after maternal transmission; the methylation pattern was reversible upon each germ-line passage through the opposite sex; there was little variation in methylation pattern in adult somatic tissues (Sasaki *et al.* 1989). We note, however, that, although maternally derived transgenes were clearly hypermethylated, a small fraction of DNA was still cut at the promoter, giving rise to a faint 1.8 kb band (Fig. 1A). Since only one copy of the transgene is present in each cell, this implies cellular mosaicism with respect to the differential methylation patterns (McGowan *et al.* 1989). Our study on the prepuberal testis suggests that the methylation pattern observed in the somatic tissues is established before the zygotene stage of meiosis prophase I during spermatogenesis. Since this methylation variation was not observed in the other two transgenic lines, it is likely that the phenomenon results from a chromosomal position effect.

The simple structure of the transgene locus (approximately one copy insertion without any host rearrangement except one small deletion) made the efficient cloning and detailed analysis of the locus feasible. One feature revealed is the clear clustering of the differentially methylated sites. This 'island' of methylation imprinting is about 3.5 kb in size and composed of the metallothionein promoter and plasmid sequences. From the description of the previous cases, it is difficult to judge whether such a clustering of the differentially methylated sites is a general feature or not. One interesting possibility suggested by the island is that the differential methylation pattern of CpG sites within the island may be determined as a block rather than individually. Such a control seems possible if we consider configurational changes of chromatin as a mechanism of methylation imprinting (see later).

Another interesting finding is the localization of the locus to the proximal chromosome 11, a region likely to contain imprinted gene(s); maternal duplication/paternal deficiency for this region, as well as maternal disomy 11, resulted in mice that were small in size while paternal disomy 11 resulted in large mice compared to the controls (Cattanach and Kirk, 1985). Although this does not necessarily mean that the flanking host sequences are imprinted, it is tempting to speculate that a signal for endogenous imprinting may be present in the 5′ flanking region, since the island showing methylation imprinting is located at the 5′ end of the transgene (see below).

Transgene methylation imprinting and endogenous genomic imprinting

The main purpose of this study was to ask whether the host sequences around the transgene insertion site show methylation imprinting or not. If the endogenous sequences were to show methylation imprinting, it would strongly suggest both that DNA methylation is actually involved in endogenous imprinting and that this region is subject to genomic imprinting. We examined four sites flanking the transgene and three sites (one of which is the same as one of the above four) on the corresponding native chromosome. However, none of these sites showed methylation imprinting. Instead, we found a completely different type of allelic methylation variation, which was heritable and strain specific, at two of the three sites tested in non-transgenic mice.

One general conclusion drawn from the study is that the methylation behavior of transgenes does not necessarily reflect the methylation status of the host sequences around the insertion site. Therefore, even if differential methylation is the molecular mechanism of endogenous imprinting, transgenes may not serve as markers for imprinted regions. Besides the discrepancies with the genetic maps for endogenous imprinted regions and the possibility of involvement of other chromosome modifications in the imprinting process (Monk, 1988; Surani *et al.* 1990), this raises another possibility that transgenes can be probes for endogenous imprinting.

Regarding the mechanisms of transgene methylation imprinting, the absence of methylation imprinting in the host sequences does not necessarily indicate the absence of the signal for that phenomenon. For instance, the sequence coding for the signal might not exhibit methylation imprinting itself; or, the signal for methylation imprinting might be transmitted from a distant site in a form other than DNA methylation. If, however, such signals were present on chromosomes, they should be widespread because of the observed high frequency of transgene methylation imprinting. Although several explanations may be applicable, the difficulties in detecting endogenous methylation imprinting in both human (Silva and White, 1988) and mouse (Surani *et al.* 1990; this study) are inconsistent with this view. Rather, it seems that there is some special reason for transgenes being prone to methylation imprinting. One possible explanation is that transgenes behave as insertional mutations which interrupt the waves of chromosome modifications that normally occur during gametogenesis and embryonic development (Surani *et al.* 1990). The foreign or displaced DNA may tend to lock the methylation status of the locus, which cannot now be changed. It may be of interest that our transgene is expressed in the testis, most probably in germ cells (Sasaki *et al.* 1989). Transcription beginning at the promoter may interrupt the normal changes in chromatin structure in this region, and the undermethylated state of the island in male germ cells may then persist in adult somatic tissues (although this does not apply to the plasmid sequences). Another possibility is that the plasmid sequences may tend to interfere with such chromosome modifications since all of the transgenes showing methylation imprinting reported thus far contain some kind of prokaryotic sequences: chloramphenicol acetyl transferase gene (Reik *et al.* 1987), pBR322 (Sapienza *et al.* 1987; Swain *et al.* 1987; Hadchouel *et al.* 1987) or pUC18 (this study). Such an effect of prokaryotic sequences could be associated with the fact that they are rich in CpG dinucleotide. It is tempting to speculate that these sequences can behave as CpG islands when introduced into the mouse germ line, just like those found at the 5′ end of endogenous genes such as the mouse metallothionein-I promoter (Bird, 1986). Since CpG islands may be involved in distinctive chromatin structures (Tazi and Bird, 1990), CpG-rich sequences associated with transgenes may be similarly involved in the process and could be targets for differential methylation.

Our unexpected finding made during this study is the identification of an inherited type of allelic methylation variation in this chromosome region. The heritable nature suggests that a signal controlling the methylation variation of this type is apparently associated with the affected sites (Silva and White, 1988). Since only the inherited type of variation is detectable in non-transgenic mice, perhaps the signal for methylation imprinting, if it exists, may be controlled at a different level. In this case, it is particularly interesting to ask whether or not an inherited methylation pattern at certain sites can be altered to show a parent-of-origin type of methylation (or a combined pattern of the two types).

Heritable, strain-specific methylation variations in the mouse

The heritable, strain-specific type of allelic methylation variation is by itself of considerable interest. Strain-dependent changes in methylation patterns were described previously (Sapienza *et al.* 1989; McGowan *et al.* 1989; Allen *et al.* 1990) but these variations were not due to the origin of the alleles but solely due to the genetic background. Inherited types of allelic methylation differences were first demonstrated in human pedigrees by Silva and White (1988). Using polymorphic markers based on variable numbers of tandem repeats, they found that homologous chromosomes

have distinct and heritable 'blueprints' for the tissue-specific control of methylation. Although only one tissue (tail) was tested in our study, the results demonstrate the presence of an inherited type of allelic methylation variation in the mouse genome and in a region not associated with tandem repeat sequences. Thus, this type of variation could be ubiquitous in mammalian genomes.

At present, it is unknown whether the strain-specific methylation variations arise from differences in the adjacent DNA sequences. The frequency of sites showing such methylation variation between different *Mus musculus* subspecies seems quite high since two out of three sites tested were variable. The frequency of variation between certain strains could reflect their relatedness or genetic distance. Another and probably more important question is whether such variation can occur within gene promoters or gene exons. Methylation of CpG dinucleotide can influence gene expression (Cedar, 1984) and, accordingly, there is the interesting possibility that allelic methylation variations of this type may affect phenotypes in heritable fashion. Also, methylated CpG sites are mutation hotspots (Coulondre *et al.* 1978; Bird, 1980) and, therefore, if we assume such variations in germ line, the frequency of mutation or polymorphism at certain loci may vary in different populations depending on the frequency of the methylated alleles.

Whatever the case, it is likely that heritable allelic methylation differences are common in mammalian genomes. The mouse offers a useful experimental system to study the genetic basis and biological significance of this interesting phenomenon as well as those of other types of methylation variations, such as methylation imprinting.

We are grateful to Drs Toshihiko Shiroishi and Kazuo Moriwaki for MOL-MSM mice and DNA samples from various wild-derived mouse strains. We would like to thank Ms Michiko Hanaoka for assistance in *in situ* hybridization experiments. Drs M. Azim Surani, Nicholas D. Allen, Wolf Reik and Peter A. Jones kindly read the manuscript and suggested improvements.

References

Allen, N. D., Norris, M. L. and Surani, M. A. (1990). Epigenetic control of transgene expression and imprinting by genotype-specific modifiers. *Cell* **61**, 853–861.

Bellve, A. R., Cavicchia, J. C., Millette, C. F., O'Brien, D. A., Bhatnagar, Y. M. and Dym, M. (1977). Spermatogenic cells of the prepuberal mouse: Isolation and Morphological characterization. *J. Cell Biol.* **74**, 68–85.

Bird, A. P. (1980). DNA methylation and the frequency of CpG in animal DNA. *Nucl. Acids Res.* **8**, 1499–1503.

Bird, A. P. (1986). CpG islands and the function of DNA methylation. *Nature* **321**, 209–213.

Buckle, V. J. and Craig, I. W. (1986). *In situ* hybridization. In *Human Genetic Diseases: A Practical Approach* (ed. K. E. Davies), pp. 85–100. IRL Press, Oxford.

Cattanach, B. M. (1986). Parental origin effect in mice. *J. Embryol. exp. Morph.* **97 Supplement**, 137–150.

Cattanach, B. M. and Kirk, M. (1985). Differential activity of maternally and paternally derived chromosome regions in mice. *Nature* **315**, 496–498.

Cedar, H. (1984). DNA methylation and gene expression. In *DNA Methylation: Biochemistry and Biological Significance* (ed. A. Razin, H. Cedar and A. D. Riggs), pp. 147–164. Springer-Verlag, New York.

Coulondre, C., Miller, J. H., Farabough, P. J. and Gilbert, W. (1978). Molecular basis of base substitution hotspots in *Escherichia coli*. *Nature* **274**, 775–780.

Frischauf, A. M., Murray, N. and Lehrach, H. (1987). λ phage vectors-EMBL series. *Methods Enzymol.* **153**, 103–115.

Hadchouel, M., Farza, H., Simon, D., Tiollais, D. and Pourcel, C. (1987). Maternal inhibition of hepatitis B surface antigen gene expression in transgenic mice correlates with *de novo* methylation. *Nature* **329**, 454–456.

Harada, Y., Bonhomme, F., Natsuume-Sakai, S., Tomita, T. and Moriwaki, K. (1989). Serological survey of complement factor H in common laboratory and wild mice: A new third allotype. *Immunogenetics* **29**, 148–154.

Maniatis, T., Fritsch, E. F. and Sambrook, J. (1982). *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

McGowan, R., Campbell, R., Peterson, A. and Sapienza, C. (1989). Cellular mosaicism in the methylation and expression of hemizygous loci in the mouse. *Genes Dev.* **3**, 1669–1676.

Monk, M. (1988). Genomic imprinting. *Genes Dev.* **2**, 921–925.

Monk, M. (1990). Variation in epigenetic inheritance. *Trends Genet.* **6**, 110–114.

Reik, W. (1989). Genomic imprinting and genetic disorders in man. *Trends Genet.* **5**, 331–336.

Reik, W., Collick, A., Norris, M. L., Barton, S. C. and Surani, M. A. H. (1987). Genomic imprinting determines methylation of parental alleles in transgenic mice. *Nature* **328**, 248–251.

Sapienza, C., Paquette, J., Tran, T. H. and Peterson, A. (1989). Epigenetic and genetic factors affect transgene methylation imprinting. *Development* **107**, 165–168.

Sapienza, C., Peterson, A. C., Rossant, J. and Balling, R. (1987). Degree of methylation of transgenes is depedndent on gamete of origin. *Nature* **328**, 251–254.

Sasaki, H., Nakazato, M., Saraiva, M. J. M., Matsuo, H. and Sakaki, Y. (1989). Activity of a metallothionein-transthyretin fusion gene in transgenic mice: Possible effect of plasmid sequences on tissue-specific expression. *Molec. biol. Med.* **6**, 345–353.

Sasaki, H., Tone, S., Nakazato, M., Yoshioka, K., Matsuo, H., Kato, Y. and Sakaki, Y. (1986). Generation of transgenic mice producing a human transthyretin variant: A possible mouse model for familial amyloidotic polyneuropathy. *Biochem. biophys. Res. Commun.* **139**, 794–799.

Searle, A. G. and Beechey, C. V. (1985). Noncomplementation phenomena and their bearing on nondisjunctional effects. In *Aneuploidy* (ed. V. L. Dellarco, P. E. Voytek and A. Hollaender), pp. 363–376. Plenum press, New York.

Searle, A. G., Peters, J., Lyon, M. F., Hall, J. G., Evans, E. P., Edwards, J. H. and Buckle, V. J. (1989). Chromosome maps of man and mouse. IV. *Ann. Hum. Genet.* **53**, 89–140.

Silva, A. J. and White, R. (1988). Inheritance of allelic blueprints for methylation patterns. *Cell* **54**, 145–152.

Solter, D. (1988). Differential imprinting and expression of maternal and paternal genomes. *Annu. Rev. Genet.* **22**, 127–146.

Surani, M. A., Allen, N. D., Barton, S. C., Fundele, R., Howlett, S. K., Norris, M. L. and Reik, W. (1990). Developmental consequences of imprinting of parental chromosomes by DNA methylation. *Phil. Trans. R. Soc. Lond.* B **326**, 313–327.

Surani, M. A., Reik, W. and Allen, N. D. (1988). Transgenes as probes for genomic imprinting. *Trends Genet.* **4**, 59–62.

Swain, J. L., Stewart, T. A. and Leder, P. (1987). Parental legacy determines methylation and expression of an autosomal transgene: A molecular mechanism for parental imprinting. *Cell* **50**, 719–727.

Tazi, J. and Bird, A. (1990). Alternative chromatin structure at CpG islands. *Cell* **60**, 909–920.

(Accepted 22 November 1990)