The EMBO Journal vol.12 no.9 pp.3669–3677, 1993

The ontogeny of allele-specific methylation associated with imprinted genes in the mouse

Michael Brandeis, Tal Kafri, Mira Ariel, J.Richard Chaillet[1], John McCarrey[2], Aharon Razin and Howard Cedar

Department of Cellular Biochemistry, Hebrew University Medical School, Jerusalem, Israel, [1]Howard Hughes Medical Institute, Department of Genetics, Harvard Medical School, Boston, MA 02115 and [2]Department of Genetics, Southwest Foundation for Biomedical Research, San Antonio, TX 78228, USA

Communicated by H.Cedar

We have investigated the DNA methylation patterns in genomically imprinted genes of the mouse. Both Igf2 and H19 are associated with clear-cut regions of allele-specific paternal modification in late embryonic and adult tissues. By using a sensitive PCR assay, it was possible to follow the methylation state of individual *Hpa*II sites in these genes through gametogenesis and embryogenesis. Most of these CpG moieties are not differentially modified in the mature gametes and also become totally demethylated in the early embryo in a manner similar to non-imprinted endogenous genes. Thus, the overall allele-specific methylation pattern at these sites must be established later during embryogenesis after the blastula stage. In contrast, sites in an Igf2r gene intron and one CpG residue in the Igf2 upstream region have allele-specific modification patterns which are established either in the gametes or shortly after fertilization and are preserved throughout pre-implantation embryogenesis. These studies suggest that only a few DNA modifications at selective positions in imprinted genes may be candidates for playing a role in the maintenance of parental identity during development.

Key words: embryogenesis/gametogenesis/imprinting/methylation

Introduction

Genomically imprinted genes represent a class of sequences that are transcribed in an allele-specific manner. Of the four genes identified in the mouse, Igf2r (Barlow *et al.*, 1991) and H19 (Bartolomei *et al.*, 1991) are paternally inactive, while Igf2 (DeChiara *et al.*, 1991) and Snrpn (Leff *et al.*, 1992) have a maternally derived repressed allele. Little is known about the molecular mechanisms involved in this unusual regulatory process, but experiments with transgenic mice suggest that DNA methylation may play a role (Surani *et al.*, 1988). In these studies, several plasmid constructs were shown to acquire different methylation profiles in the adult organism depending on the parental origin of the transgene. In addition, allele-specific differential methylation has been observed for the endogenous Igf2r gene in the mouse (Stoger *et al.*, 1993) and for sequences located within the imprinted Prader Willi/Angelman locus in man (Driscoll *et al.*, 1992; Dittrich *et al.*, 1992).

There are several models which may accommodate a role for this differential methylation in the imprinting process. One possibility is that the allele-specific DNA methylation pattern seen in somatic cells is inherited intact from the gametes and is then stably conserved during embryogenesis. Alternatively, these genes may be initially marked in the gametes at discrete *cis*-acting regulatory elements which then serve as 'imprinting signals' to direct the final differential structure and overall methylation profiles of the individual alleles in cells of the embryo and adult organism. In this latter case, the stable methylation pattern of many and perhaps most sites in these genes would only be established after fertilization. In either event, modified CpG sequences derived from the gametes would have to be preserved during early pre-implantation development and must therefore be specifically recognized and protected from the massive general demethylation which takes place at these stages (Monk *et al.*, 1987; Chaillet *et al.*, 1991; Kafri *et al.*, 1992).

As a step towards understanding how DNA modification might play a role in imprinting, we have analysed the methylation patterns of the endogenous imprinted genes Igf2 and H19 in embryonic and adult tissues and both are clearly associated with regions of allele-specific methylation. By means of a highly sensitive PCR assay (Kafri *et al.*, 1992), it was possible to follow the developmental patterns of DNA modification for the Igf2, H19 and Igf2r genes during embryogenesis and thus determine when differential methylation is actually established and how it is erased in primordial germ line cells. The results indicate that the allele-specific modification patterns seen in the adult, are acquired through a programmed progression of methylation events which begins during gametogenesis and continues through embryogenesis. These data are consistent with a model in which methylation at specific loci may serve as a means for marking and maintaining parental identity in the early embryo.

Results

Differential methylation near the Igf2 gene

By constructing mice containing an Igf2 gene knock-out, it was possible to demonstrate functionally that this locus is maternally repressed in a variety of cell types in post-implantation embryos (De Chiara *et al.*, 1991). Sasaki *et al.* (1992) used normal and chromosome 7 maternally disomic embryos to examine the methylation pattern of this gene. While they found that the structural gene itself and its 5′ CpG islands are stably unmethylated on both alleles, preliminary data suggested that nearby upstream sequences may show some differential methylation (Sasaki *et al.*, 1992). We have identified a region of ~600 bp, located 3 kb upstream to the first promoter of Igf2 (Figure 1), which is methylated in an allele-specific manner. Initial Southern blot analysis of this small region showed that four individual *Hpa*II sites are partially methylated in adult liver tissue

© Oxford University Press

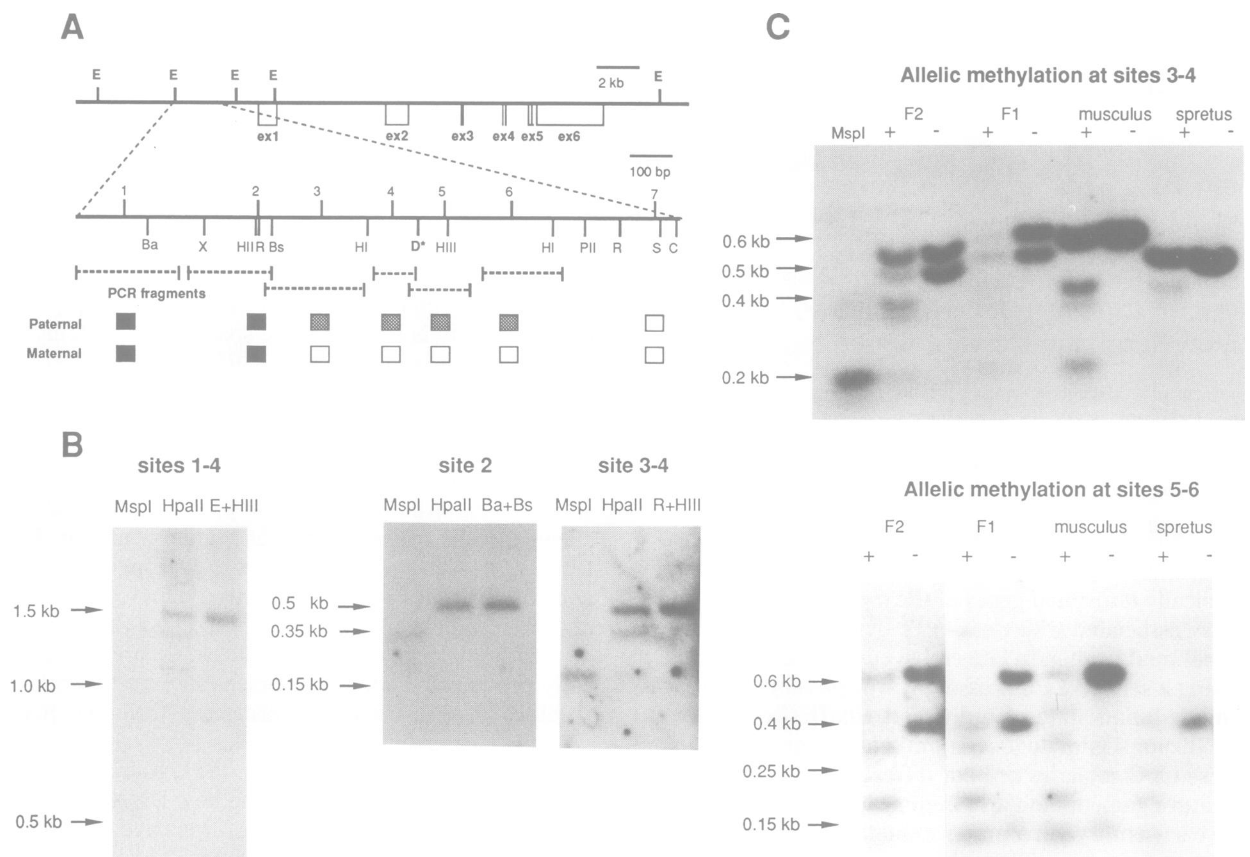

Fig. 1. Allele-specific methylation in the Igf2 upstream region. (**A**) Map of Igf2 showing the overall structure of the gene with its *Eco*RI (E) restriction sites. The upstream region is shown in greater detail along with the seven *Hpa*II sites and the positions of cleavage for *Ban*II (Ba), *Xba*I (X), *Hin*fI (HI), *Hin*dII (HII), *Hin*dIII (HIII), *Rsa*I (R), *Pvu*II (PII), *Sma*I (S), *Cfo*I (C) and *Bst*EII (Bs). D* represents a polymorphic *Dra*I site present only in *M.spretus* DNA. The PCR fragments used for analysis of *Hpa*II sites 1−6 are also indicated. The diagram under the map shows a schematic representation of the methylation state of all *Hpa*II sites on the paternal and maternal alleles in adult liver DNA as determined by the blot analysis shown in this figure. These sites were either uniformly methylated (1 and 2), differentially methylated (4−6) or unmethylated (7). (**B**) Adult liver DNA from *M.musculus* was digested with *Eco*RI, *Hin*dIII and *Hpa*II or *Msp*I, subjected to electrophoresis, blotted and hybridized with an *Eco*RI−*Xba*I probe. As indicated, ~50% of the molecules in this region are either fully methylated (1.5 kb) or partially methylated at site 3 (1.1 kb) or site 4 (1.3 kb). Using *Bam*HI and *Bst*EII digestion it was shown that site 2 is fully methylated, whereas digestion with *Rsa*I and *Hin*dIII indicated that sites 3 and 4 are ~50% methylated. (**C**) Adult liver DNA of *M.spretus*, *M.musculus*, F1 (*M.spretus* × *M.musculus* where *M.spretus* is the father) and F2 (*M.musculus* × F1 where *M.musculus* is the father), were digested with *Hin*dII, *Hin*dIII and *Dra*I with and without *Hpa*II. *Dra*I is a polymorphic site found only in *M.spretus* DNA. Following electrophoresis, blots were hybridized to the *Hin*dII−*Hin*dIII probe. This protocol is aimed at observing the methylation states of sites 3 and 4 which are apparently methylated only on the paternal allele in both F1 and F2 mice. This analysis is slightly complicated by the presence of an additional *Hpa*II site between 3 and 4 in *M.spretus* DNA, and this probably accounts for the weakness of the paternally derived *M.spretus* band in the F1 DNA. A similar strategy for *Hin*fI digestion was used to assay allelic methylation of sites 5 and 6. In addition to the parent-specific *M.spretus* (0.4 kb) and *M.musculus* (0.6 kb) bands, one should also note the appearance of smaller parent-specific products in both F1 and F2.

(Figure 1B) and in 15−18 day embryos (data not shown). This fragment of DNA appears to represent a fairly well-defined locus, since sites immediately 5′ (sites 1 and 2) are fully methylated on both alleles, while the closest 3′ sites are located in a small uniformly unmethylated CpG island (Figure 1A).

In order to determine whether this partial methylation is allele-specific, we analysed DNA from *Mus musculus* × *Mus spretus* interspecific crosses. Sequencing of this region from both strains revealed polymorphic restriction sites for several enzymes including *Hin*fI and *Dra*I in *M.spretus* DNA which allowed us to identify each allele independently. F1 mice all contain a paternal spretus allele and females from this mating were then back-crossed to normal *M.musculus* males in order to obtain F2 animals with a maternally derived spretus allele. As shown in Figure 1C, the pair of sites 3 and 4 is clearly differentially methylated on the paternally derived allele in both F1 and F2 animals and the same is true for sites 5 and 6, located at the 3′ end of the critical fragment. By measuring the intensities of the fragments generated by *Hpa*II digestion, this series of experiments suggest that the paternal gene is partially methylated, while the maternal allele is mostly unmodified at all of the tested sites.

Igf2 gamete methylation

Having shown that the upstream region of Igf2 is differentially methylated, we then set about to determine at what stage this pattern is established during development. If modification of this locus serves as an imprinting signal for identifying the parental alleles, these sites should be methylated in sperm DNA while unmethylated in the oocyte. In order to test this hypothesis, we first analysed vas deferens sperm by blot hybridization. While all of the sites within

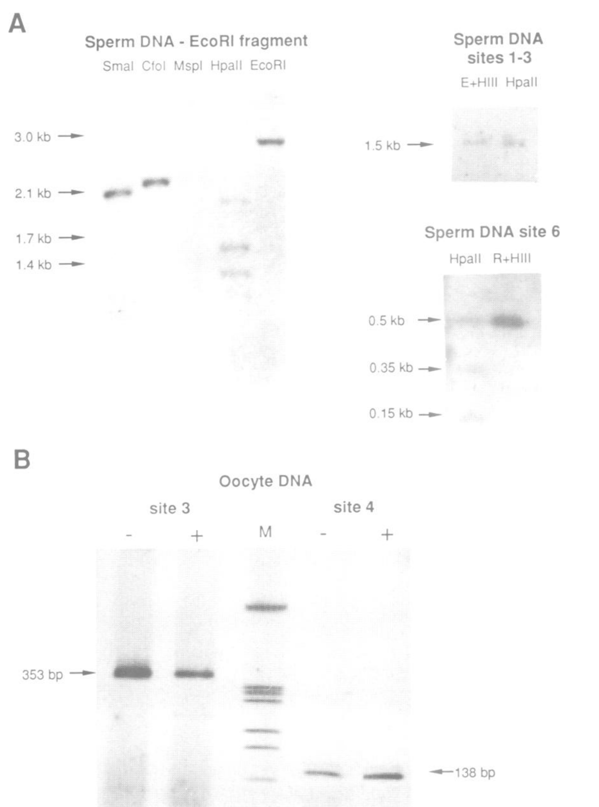

Fig. 2. Igf2 DNA methylation in the gametes. (**A**) Vas deferens DNA was digested with a combination of restriction enzymes in order to assess the methylation pattern of the Igf2 upstream region. The entire domain was assayed by digestion with *Eco*RI using the *Eco*RI–*Xba*I fragment as a probe. The *Sma*I and *Cfo*I sites located in the CpG island near site 7 are completely unmethylated. The *Hpa*II digest shows that sites 4, 5 and 6 are partially modified, while sites 1, 2 and 3 are completely methylated. Sites 1, 2 and 3 were assayed independently by restriction with *Eco*RI and *Hin*dIII and partial methylation at site 6 was confirmed by evaluating this locus using restriction enzymes *Hin*dIII and *Rsa*I. (**B**) PCR methylation analysis of Igf2 upstream region *Hpa*II sites 3 and 4 in oocyte DNA. Amplification products were electrophoresed on acrylamide gels and visualized by silver staining. For each site, PCR was performed either with DNA digested only with *Pvu*II (−) or DNA digested with *Pvu*II and *Hpa*II (+). Similar results were obtained for sites 5 and 6 (data not shown). All assays were performed on the same oocyte DNA preparation and an analysis of site 4 in the Igf2r gene region 2 serves as a control for *Hpa*II digestion (see Figure 5B). ΦX174 *Hae*III digestion products serve as size markers.

the 600 bp fragment were methylated, only site 3 was fully methylated and the remaining loci had a partially modified pattern similar to that seen in somatic cells (Figure 2A). DNA from mature oocytes was assayed for site-specific methylation using PCR analysis. In this method, oocyte DNA is digested with methyl-sensitive restriction enzymes and the degree of cleavage is assessed by measuring the amount of amplified product using primers flanking individual CpG sites. In striking contrast to our expectations, all of the sites in the imprinted locus were found to be in a methylated state in the oocyte, as well (Figures 2B and 7). Thus, the differential methylation pattern found in somatic cells cannot be derived directly from the prefertilization parental gametes and must therefore be established some time during post-fertilization embryonic development.

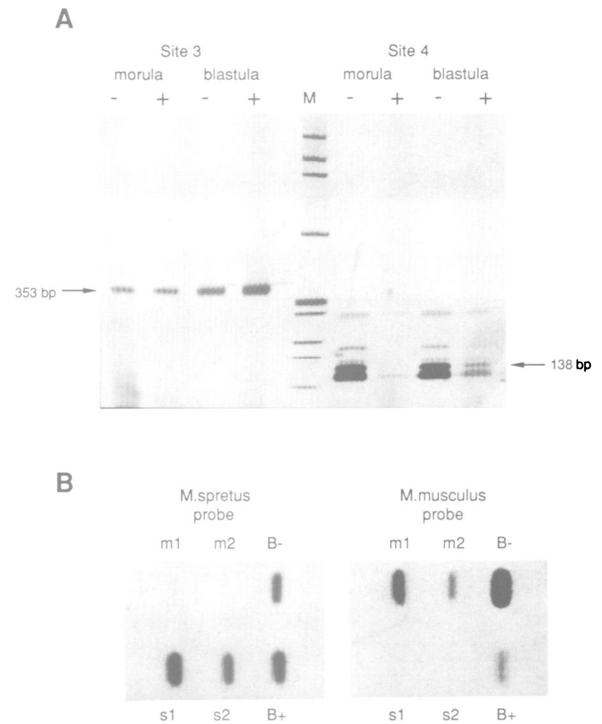

Fig. 3. Igf2 DNA methylation patterns during early embryogenesis. (**A**) PCR methylation analysis of Igf2 upstream region *Hpa*II sites 3 and 4 in the morula and blastula. Amplification products were electrophoresed on acrylamide gels and visualized by silver staining. For each site PCR was performed either on *Pvu*II (−) or *Pvu*II and *Hpa*II digested DNA (+). Sites 5 and 6 were also assayed in this same experiment and found to be unmethylated at both embryonic stages (data not shown). All sites were analysed in the same morula and blastula preparations. Size markers are from ΦX174 DNA digested with *Hae*III (M). (**B**) Site 3 was assayed in a similar manner using blastula DNA from an interspecific cross between *M.spretus* (father) and *M.musculus* (mother). The amplification products (~50 ng) from uncut DNA (B−) and *Hpa*II cut DNA (B+) were placed on two identical filters and subjected to slot blot hybridization using oligonucleotide probes specific for each species (*M.musculus*: GATTAGGCTCTGAAACAG; *M.spretus*: GATTAGACTCTGAAACAG). Also included on the blot are 50 ng (m1) or 15 ng (m2) of *M.musculus* DNA and *M.spretus* DNA (s1 and s2), amplification products.

Methylation changes in the Igf2 locus during development

The DNA methylation pattern of normal endogenous gene sequences during pre-implantation development is characterized by dynamic global changes. Most non-island CpG residues in the mouse genome are methylated in sperm (Yeivin and Razin, 1993) and some of these are also modified in the oocyte (Kafri *et al.*, 1992). However, all methyl moieties present on the original parental chromosomes appear to be removed from the DNA by the 16 cell stage and this gives rise to a predominantly unmethylated genome which remains in this state at least through blastulation (Monk *et al.*, 1987; Chaillet *et al.*, 1991; Shemer *et al.*, 1991; Kafri *et al.*, 1992). The final bimodal methylation profile of the somatic genome is then re-established after implantation by *de novo* methylation at all non-CpG island residues (Frank *et al.*, 1991). In light of this general picture of the methylation reactions occurring during embryogenesis, it was of interest to follow the developmental modification pattern of CpG sites in the Igf2 imprinted locus.

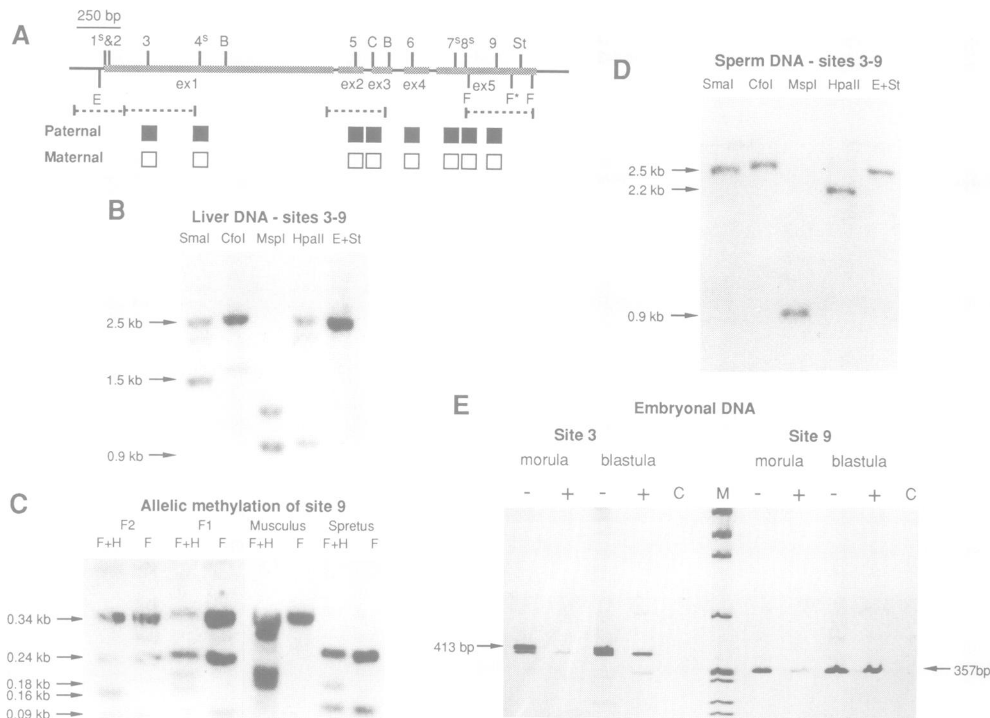

Fig. 4. Allelic methylation of H19 genomic DNA. (**A**) Map of H19 gene showing the five exons, *Hpa*II and *Sma*I (s) sites and the fragments used for PCR methylation analysis of individual CpG residues. The locations of restriction sites *Cfo*I (C), *Stu*I (St), *Eco*RI (E), *Bam*HI (B) and *Fok*I (F) are also shown. The F* polymorphic site present only in *M.spretus* DNA, was used to analyse allele-specific methylation of site 9. The maternal and paternal methylation states of adult liver DNA is shown diagrammatically. This pattern was deduced from the data in this figure showing that 50% of the molecules are fully methylated. Since at least one of these sites is modified in an allele-specific manner, this implies that the same must be true for the other tested loci on the same *Eco*RI–*Stu*I fragment. (**B**) *M.musculus* DNA was digested with *Eco*RI and *Stu*I together with *Hpa*II, *Msp*I, *Cfo*I or *Sma*I and probed with the *Bam*HI fragment. The results indicate that the entire domain is ~50% methylated. (**C**) Adult liver DNA of *M.spretus*, *M.musculus*, F1 and F2 were digested with *Fok*I, with and without *Hpa*II and probed with the *Fok*I fragment. The *M.spretus* DNA yields 0.24 and 0.1 kb bands, which are further cut by *Hpa*II into 0.16 and 0.09 kb fragments. *M.musculus* DNA yields a 0.34 kb fragment which is cut by *Hpa*II into 0.18 and 0.16 kb fragments. The additional hybridization signal observed in this lane is apparently an artefact. The F1 and F2 DNA both had a preferentially methylated paternal allele and the digestion products indeed correspond to that expected from the maternal allele. (**D**) Sperm DNA was cut with *Eco*RI and *Stu*I with or without *Hpa*II, *Msp*I, *Cfo*I and *Sma*I, and probed with the *Bam*HI fragment. While most tested sites were fully methylated, these data indicate that sites 1 and 3 in the 5′ CpG island are unmethylated in sperm. (**E**) PCR methylation analysis of H19 sites 3 and 9 in morula and blastula DNA. Amplification products were electrophoresed on acrylamide gels and visualized by silver staining. For each site PCR was performed with *Pvu*II (−), *Pvu*II and *Hpa*II digested DNA (+) or in the absence of DNA substrate (C). Size markers are from ΦX174 DNA digested with *Hae*III (M). Using these same preparations, we also assayed (data not shown) site 5 in both morula and blastula and it was found to be unmethylated (see Figure 7).

Analysis of the Igf2 upstream region revealed that most of the *Hpa*II sites (4–6) at this locus were unmethylated in the 16–32 cell morula and are thus subject to the genome-wide demodification process which occurs in the early embryo (Figures 3A and 7). In striking contrast, site 3 was methylated both at the morula and blastula stages of development. Since this site is clearly differentially modified in adult tissues, it was of interest to determine whether methylation is allele-specific at the early pre-implantation stages of development. To this end, we prepared 3.5 day interspecific hybrid blastulae using *M.spretus* males and *M.musculus* females. Following PCR, the amplification product was hybridized to labelled oligomers which distinguish between the two alleles. As shown in Figure 3B, the paternal allele was significantly more methylated than the maternal. This suggests that for site 3, this differential methylation pattern is generated quite early in development, most likely through a process involving specific demethylation of the maternal allele. Other sites in the Igf2 upstream region are relatively unmethylated at the blastula stage, and blot analysis indeed demonstrated that allelic modification is only reached at a later time in development, probably prior to 15 days post-coitus (p.c.) (data not shown).

Methylation patterns of the imprinted H19 gene

The mouse H19 gene is first expressed at the blastula stage of development and has been shown to have a post-implantation imprinted pattern of transcription in both embryonic and extra-embryonic tissues with an inactive paternal allele (Bartolomei *et al.*, 1991). Initial Southern blot analysis of the 2.5 kb region containing the coding sequences indicated (Figure 4A) that ~50% of the DNA molecules are methylated in the adult liver (Figure 4B). This is true for single restriction sites such as *Cfo*I and also for more frequent sites, such as *Sma*I and *Hpa*II. The fact that about half of the genomic *Eco*RI–*Stu*I fragment is resistant to *Hpa*II strongly suggests that the two parental alleles may be differentially modified in a regional manner. In order to assay parent-specific methylation, we sequenced a portion of the *M.spretus* H19 gene and identified a *Fok*I restriction site polymorphism in exon 5 (F*) which enabled us to distinguish between the maternal and paternal alleles in interspecific hybrid mice. Blot hybridization analysis of *Hpa*II site 9 in this polymorphic region was carried out on the DNA from both F1 and F2 mice and these studies clearly indicate that the inactive paternal allele is preferentially methylated in adult somatic cells. Additional experiments taking advantage

of a polymorphic *BglII* site showed that two *CfoI* sites upstream of the coding sequence are also modified on the paternal, but not the maternal, allele (data not shown). When taken together, these results suggest that the H19 gene is methylated in an allele-specific manner in the adult liver. With some small exceptions at the 3′ end of this gene sequence, this differential pattern is similar to that observed in early 9–10 day p.c. embryos as well (Ferguson-Smith *et al.*, 1993).

In the light of these observations, we then investigated how the differential modification pattern of H19 may be established during development. If these methyl moieties are derived from the gametes, all of these sites should be 100% methylated in sperm DNA. Southern blot analysis indeed indicated that this was the case for a number of *HpaII*, *CfoI* and *SmaI* loci, but site 3 is actually located within a small CpG island domain that is completely unmethylated in sperm (Figure 4D and Ferguson-Smith *et al.*, 1993). We have not investigated the methylation status of these CpG sites in mature oocytes. However, it is clear that the adult pattern cannot be derived from allele-specific modification in the gametes, since methylation at all of these loci is erased during early stages of embryogenesis in the morula and blastula (Figures 4E and 7).

Methylation patterns of the imprinted Igf2r gene

The Igf2r gene is paternally repressed in the mouse (Barlow *et al.*, 1991). Its coding sequences extend over a 90 kb region on chromosome 17 and within this domain there are two loci which show differential methylation (Stoger *et al.*, 1993). Region 1, which overlaps the gene promoter, is methylated exclusively on the paternal allele in late embryos and adult tissues, but this differential pattern must be established post-fertilization, since both alleles were found to be unmethylated in the gametes, in embryonic stem cells (Stoger *et al.*, 1993) and in the normal morula and blastula (unpublished data). A second differentially methylated domain (region 2) is located in an intron 27 kb downstream from the start of transcription. In contrast to region 1, sites in region 2 are methylated exclusively on the maternal allele and we have previously demonstrated that modification of two specific *HpaII* sites in this region is derived directly from the oocyte and preserved throughout all stages of pre-implantation development (Stoger *et al.*, 1993). These results strongly suggested that methylation in region 2 may serve as an imprinting signal from gamete to early embryo. To evaluate the extent of this signal domain, we studied the methylation status of additional sites in region 2 during early embryogenesis. Although this entire 1 kb locus is allelically modified in somatic cells, it is clear that not all of these methyl moieties are derived directly from the oocyte. For example, *HpaII* sites 1 and 4, positioned at the 5′ and 3′ ends of region 2 (Figure 5A), were found to be unmethylated in mature female germ cells, but became quickly modified prior to the morula stage of embryonic development (Figures 5B and 7). Thus, DNA methylation from only one small segment of region 2 may actually be established prior to fertilization.

Methylation of Igf2r in the blastula is allelic

PCR analysis indicated that all of the tested sites in Igf2r gene region 2 are modified in the pre-implantation embryo (Figures 5B and 7). Since this method is designed to assess

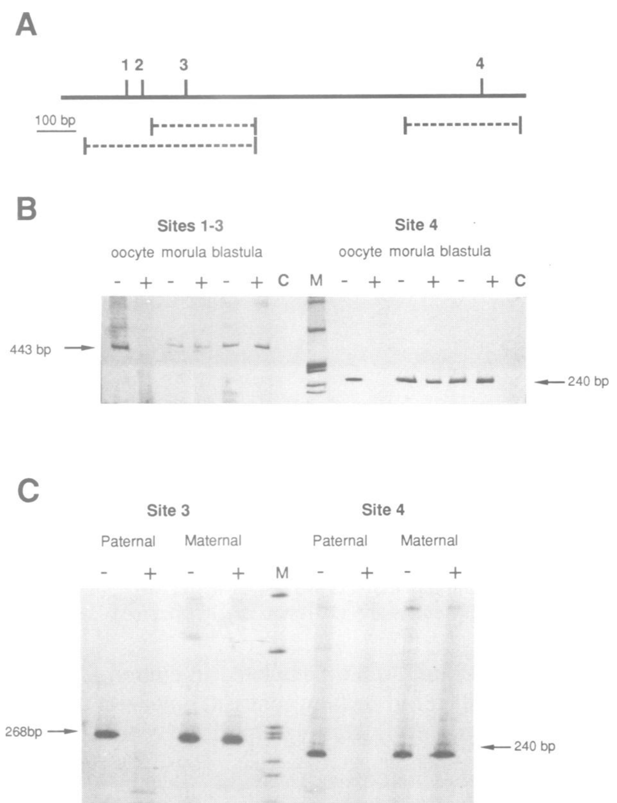

Fig. 5. Methylation patterns of the Igf2r gene in the oocyte and early embryo. (**A**) Map of Igf2r region 2 showing four *HpaII* sites and the PCR fragments used for analysis of their methylation state. (**B**) PCR analysis was performed on sites 1 and 4 of the Igf2r gene region 2 in oocytes, morula and blastula preparations. Oligonucleotide primers for site 1 produce a fragment which includes sites 2 and 3 as well. Since these were previously found to be methylated in all three tissue samples (Stoger *et al.*, 1993), the analysis here is indeed informative for site 1. Unlike sites 2 and 3, sites 1 and 4 are clearly unmethylated in the oocyte. (**C**) Blastulae carrying only a paternal or maternal allele of the Igf2r gene were prepared by crossing C3H T^{hp}-BALB/c hemizygous mice with BALB/c and identifying embryos with single alleles (see Materials and methods). DNA was cut with *PvuII* (−) or *PvuII* and *HpaII* (+) and amplified with primers flanking sites 3 and 4. Both sites are methylated exclusively on the maternal allele.

the overall degree of DNA methylation, it was not possible to determine whether these sites are already allelically modified at this stage. In order to address this question, we evaluated DNA methylation of individual alleles in blastulae of T^{hp} mice which contain uniparental deletions encompassing the entire Igf2r gene domain (Barlow *et al.*, 1991). Embryos containing exclusively either the maternal or paternal Igf2r gene were generated and identified. To this end, male C3H mice containing the T^{hp} deletion were mated with BALB/c females to produce F1 hairpin animals containing the Igf2r deletion on a C3H chromosome 17 and a normal allele on the homologous chromosome from BALB/c. These strain-specific alleles were distinguishable by virtue of a variable repeat length polymorphism located in the MHC domain (Hearne *et al.*, 1991) distal to the Igf2r gene but not imbedded in the deleted region. F1 animals were then backcrossed to BALB/c and individual blastulae assayed by PCR to identify embryos with the C3H polymorphism that are hemizygous for the Igf2r gene. DNA from these informative embryos was then pooled and subjected to PCR methylation analysis (Figure 5C). As expected, site 3 in

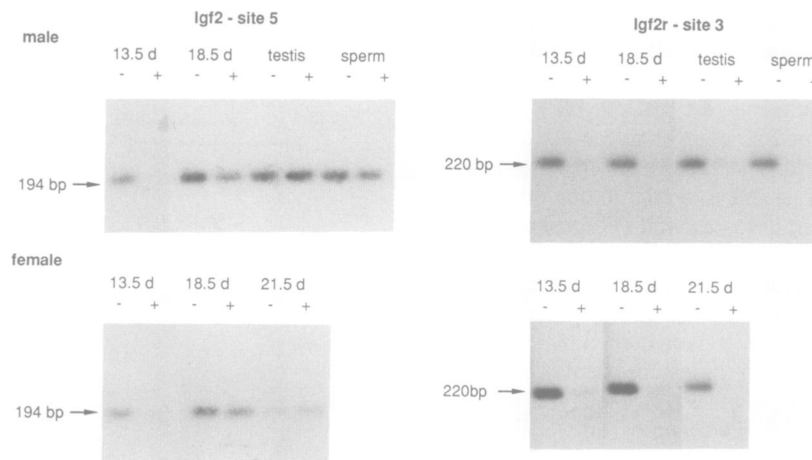

Fig. 6. Methylation patterns of imprinted genes during gametogenesis. DNA from male and female 13.5 and 18.5 day p.c. fetal germ cells, 21.5 day p.c. oogonia, testis and vas deferens sperm were cut with *Pst*I (−) or *Pst*I and *Hpa*II (+) and subjected to specific PCR amplification for site 5 in the Igf2 upstream region and site 3 of the Igf2r region 2. The amplification products were electrophoresed and blot hybridized with probes in these regions. In addition to these sites we also analysed (data not shown) all of the other *Hpa*II loci in the Igf2, Igf2r and H19 genes at all stages of male and female gametogenesis and the results are diagrammed in Figure 7.

region 2 was found to be methylated in embryos carrying the maternal allele. In striking contrast, however, this same site was unmodified in blastulae from the reciprocal cross which contains only the paternal allele. Thus, gamete-specific modification at this locus is faithfully maintained in the early embryo. A similar pattern of paternal-specific methylation was also seen for site 4, indicating that this CpG residue must become *de novo* methylated in an allele-specific manner soon after fertilization. These experiments clearly show that for region 2, differential modification is already well established in the pre-implantation embryo.

Methylation patterns during gametogenesis

While all three genes analysed in this study demonstrated differential and allele-specific methylation in adult and embryonic tissues, in most cases, these patterns were generated subsequent to fertilization. For such CpG sites, the ontogeny of DNA methylation during gametogenesis clearly does not play a role in the imprinting process. In fact, a close examination of these specific sites in the developing gamete demonstrates that most of these imprinted loci undergo normal modification changes characteristic of endogenous non-imprinted genes. In a previous study of 15 *Hpa*II and *Hha*I sites from a variety of different DNA sequences, it was shown that all of these loci are unmethylated in primordial germ cells but become stably remethylated at non-CpG island residues at 15.5−18.5 days p.c. (Kafri *et al.*, 1992). As shown in Figures 6 and 7, a similar developmental methylation pattern is also observed for all tested CpG sites in the allelically imprinted regions of the Igf2 and H19 genes. The *Hpa*II sites in region 1 of the Igf2r gene are constitutively unmethylated throughout gametogenesis in both sexes (data not shown) as is typical of other CpG island sequences. In region 2, all of the tested sites follow this same pattern and are clearly still unmodified in new born ooctyes (Figures 6 and 7). We have already demonstrated, however, that sites 2 and 3 within this region are modified in the mature oocyte (Stoger *et al.*, 1993), suggesting that they become *de novo* methylated during the growth phase of the oocyte prior to meiotic metaphase I in the adult animal. In this regard, it should be noted that the process of gametogenesis in the female does not end with the production of a mature oocyte in the ovary and meiosis is only completed in the female pronucleus of the post-fertilization zygote. Thus, additional gametic changes in methylation may still occur at this stage.

Discussion

Overall methylation patterns of imprinted genes are established progressively after fertilization

In order to understand the relationship between parental imprinting and DNA methylation, we have evaluated the modification patterns of endogenous imprinted genes in the mouse. Igf2 and H19, as well as the Igf2r gene, all contain regions that are differentially methylated in an allele-specific manner in the adult. For most CpG sites in these gene domains, however, the modification pattern is not directly inherited from the gametes, but is rather added to the DNA some time after fertilization. Several lines of evidence support this conclusion. First, the methylation state in the gametes does not always correspond to that of the parental alleles in the offspring. Thus, while all of the tested sites in the Igf2 upstream locus are methylated preferentially on the paternal allele in adult and embryonic liver DNA, these same CpG moieties are found modified both in sperm and in mature oocytes. A similar effect is observed in the Igf2r domain where CpG islands in region 1 (Stoger *et al.*, 1993) and part of region 2 (Figures 5B and 7) are allelically methylated in the adult, yet completely unmodified in both gametes and the same appears to be true for site 3 in the H19 gene domain (see also Ferguson-Smith *et al.*, 1993). For the Igf2 upstream region, the paternal allele appears to be methylated in the male gamete as expected, but surprisingly, some of the sites (5 and 6) are only partially modified in sperm (Figure 2A). Since each embryo can thus inherit gene molecules with different methylation patterns, these particular modifications cannot possibly form the basis for the transfer of imprint information from parent to offspring. Finally and most significantly, our studies show that differentially methylated regions in the imprinted genes Igf2 and H19 are generally subject to the same massive wave

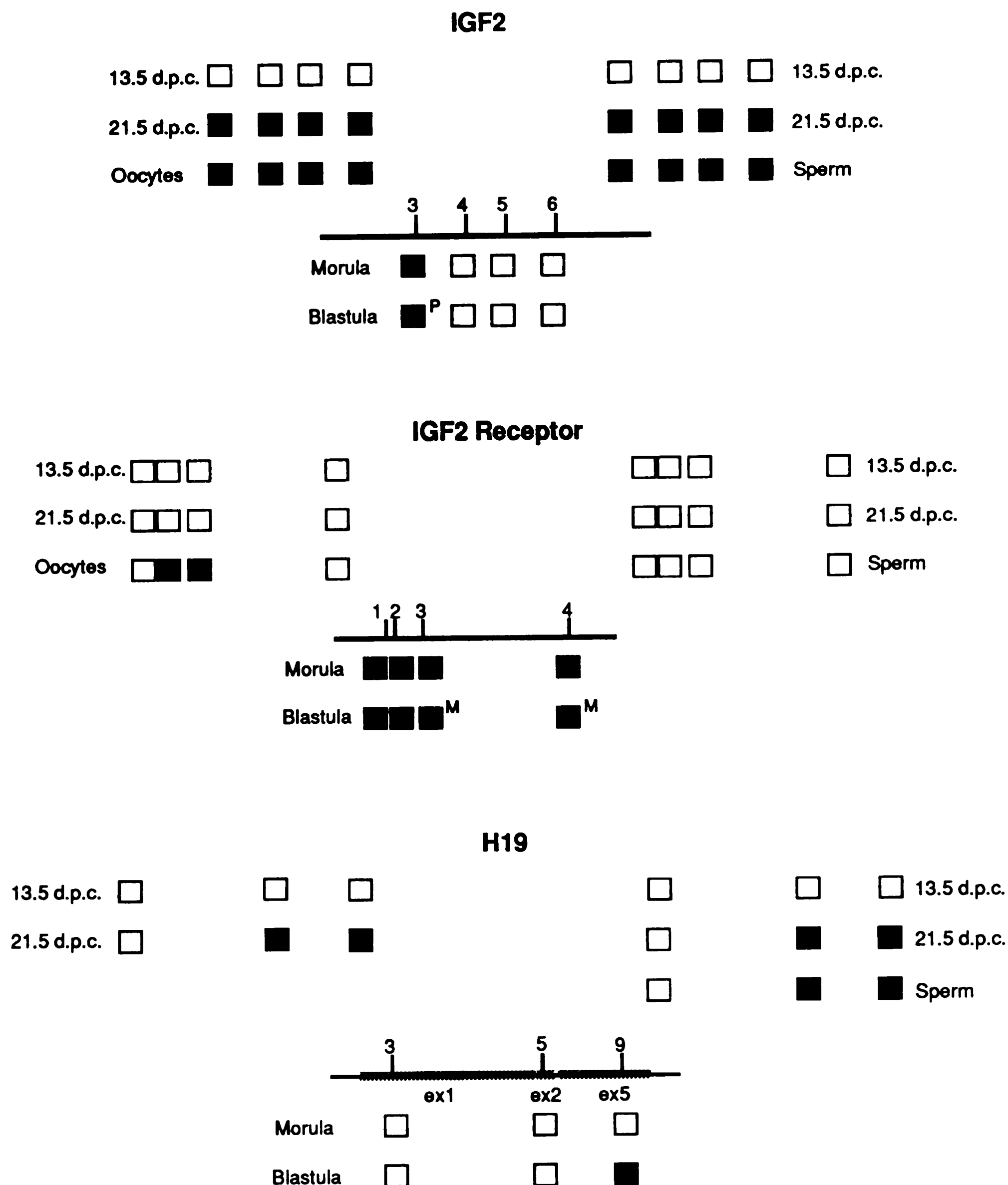

Fig. 7. DNA methylation patterns of imprinted genes—summary. DNA methylation patterns in the Igf2 upstream region, Igf2 receptor region 2 and the H19 gene during female (left side) and male (right side) gametogenesis and early embryogenesis are shown schematically. A black box indicates that the DNA was found to be 50−100% methylated by PCR analysis, while an open box indicates 0−10% methylation. For sperm samples, the results were confirmed by Southern blot hybridization analysis (see Figures 2 and 4; Stoger *et al.*, 1993). Every site was assayed at the indicated points in development although only some of the data were presented in the paper. In addition to the data shown here, all sites were analysed in male and female germ cells at 18.5 d.p.c. and were always the same as that found at 21.5 d.p.c. and the same is true for DNA from spermatogonia, pachytene spermatocytes, round spermatids and testis sperm in the male. Marked sites in the blastula were shown to be allelically methylated. In the case of site 3 in the Igf2 gene, the paternal allele is methylated (Figure 3), while sites 3 and 4 in the Igf2r gene region 2 were proven to be maternally methylated (Figure 5B).

of demethylation which takes place on other endogenous DNA sequences in the morula (Kafri *et al.*, 1992). As a result, almost all of the gametic methyl moieties at these loci are erased in the early embryo.

The final inclusive differential methylation pattern for each of the three imprinted genes is established in a stepwise manner during embryogenesis. Allele-specific modification at some sites is generated soon after fertilization, while for others, this apparently takes place by *de novo* methylation at the blastula stage or during post-implantation development. Igf2r region 1, which is completely unmodified in the blastula, only reaches its maximal allelic methylation pattern after birth (Stoger *et al.*, 1993). Although the mechanism of this process is not known, it is possible that this represents a chain reaction with allelic methylation at specific sites at one stage serving as a parent identity signal for further modification at a subsequent stage.

Methylation sites preserved during embryogenesis could be imprinting signals

In order for overall differential methylation patterns to form in the late embryo, the individual alleles are presumably marked in the gametes and this indicator of parental identity must then be preserved through early pre-implantation development. In the Igf2r gene domain, the intronic region 2 appears to represent a locus that could provide this type

of imprinting signal, since all of the tested sites in this region maintain their methylated state in the early embryo. While the paternal allele is already unmethylated in sperm and remains unmodified during embryonic development, the maternal allele becomes progressively methylated first at sites 2 and 3 in the mature oocyte (Stoger *et al.*, 1993) and then at additional CpG residues (1 and 4) after fertilization. This second stage of modification may take place in the zygote as part of the female pronucleus maturation process required to complete meiosis and in this way would represent a genuine gametically formed marker of parental identity.

An imprinting signal of this nature may also be associated with the Igf2 domain. Site 3 in the upstream region is fully methylated (Figure 2A) in sperm DNA and remains modified and resistant to demethylation on the paternal allele during pre-implantation stages of development (Figure 3B). In contrast, this same CpG residue is strikingly undermethylated on the maternal allele in the early embryo. Since this locus appears modified in the mature oocyte (Figure 2B), it follows that specific demethylation of the maternal allele takes place soon after fertilization.

Thus, both the Igf2r and Igf2 gene domains contain highly localized sequence elements which maintain allele-specific methylation during early embryogenesis and could serve as an imprinting signal. Such loci are probably a general characteristic of imprinted gene domains, since a similar marker region has been identified for the TgRSVIgmycA transgene, and in this case as well, some specific CpG sites become differentially methylated only after fertilization (Chaillet *et al.*, 1991). Although there is as yet no evidence that these sites serve as regulatory elements, it is possible that they provide simple markers of parental identity which have a region-wide influence on gene structure and expression. The H19 gene, located only 90 kb downstream from Igf2 (Zemel *et al.*, 1992) may be either part of the Igf2 domain or have its own imprinting signal which has not yet been detected.

We have not presented direct proof that DNA modification serves as an imprinting signal and other types of factors may be involved. In light of the general demethylation which takes place in the early embryo, however, the very fact that imprinted genes have CpG sites which remain modified during development, in and of itself suggests that these loci are specifically recognized and therefore, could play some role in imprinting. This maintenance process probably involves region-specific methylated DNA-binding proteins present in the early embryo which could initially interact uniquely with the modified DNA from one of the gametes and then serve to create an altered conformation which might prevent demethylation on the same allele in the morula and blastula. In this way, protein factors can serve as markers for parental identity, while DNA methylation would provide a mechanism for propagating this allele specificity following DNA replication. In this model, it is the cooperation between a *cis*-acting DNA marker and *trans*-acting proteins which orchestrate the transfer of allele-specific signals from gamete to embryo.

The imprinting signals are erased in the germ line and re-established during gametogenesis

Since maternal DNA methylation in Igf2r region 2 is derived from the oocyte and stably inherited to cells in the embryo, there must be some mechanism for erasing this signal during spermatogenesis and resetting it in the ovary. An analysis of this region in cells of the embryonic gonads indeed showed that all of the tested *Hpa*II sites are completely unmethylated in both male and female primordial germ cells at 12.5 and 13.5 days p.c. Since CpG residues for a variety of different endogenous (Driscoll and Migeon, 1990; Grant *et al.*, 1992; Kafri *et al.*, 1992) and exogenous (Chaillet *et al.*, 1991; Ueda *et al.*, 1992) gene sequences have all been found to be unmethylated at this stage, it is likely that the demodification of sites in the Igf2r region 2 and in other imprinted loci is actually part of a general erasure which occurs during early germ line development. CpG sites in most gene sequences become remodified between 15.5 and 18.5 days p.c. in both sexes, but CpG island sequences stand out as an exception to this rule and they usually remain unmethylated throughout gametogenesis (Kafri *et al.*, 1992). The Igf2r region 2 island also follows this pattern, but then undergoes a late methylation step during oocyte maturation either in the adult ovary (sites 2 and 3) or following fertilization (sites 1 and 4). This mechanism is clearly oogenesis-specific, since these same sequences remain unmethylated throughout spermatogenesis. In terms of DNA modification, this represents the major molecular event in the imprinting process. It is interesting to note that both the TgRSVIgmycA and metallothionein-I/human transthyretin transgenes contain CpG island-like pBR sequences which also undergo modification at the same exact stage of oogenesis (Chaillet *et al.*, 1991; Ueda *et al.*, 1992), suggesting that these sequences and perhaps the CpG island-like plasmid DNA in other imprinted transgenes may be mistakenly recognized by the normal endogenous mechanism for establishing allele-specific methylation.

DNA methylation and gene expression

Studies on three different imprinted genes are consistent in showing that differential methylation is established in a progressive manner during embryogenesis. Although not always directly derived from the gametes, and despite the fact that allelic methylation is often mosaic in nature, this DNA modification could certainly play a role in the control of imprinted gene expression in embryo and adult tissues. Furthermore, its function may not necessarily be as a primary determinant, but could act through a mechanism similar to that for genes on the X chromosome, which first undergo inactivation in the blastula and then utilize subsequent DNA methylation for maintaining their allele-specific expression pattern (Lock *et al.*, 1987; Grant *et al.*, 1992). Both H19 (Ferguson-Smith *et al.*, 1993) and Igf2r region 1 (Stoger *et al.*, 1993), for example, are methylated in their upstream promoter regions and this modification could thus mediate inhibition of transcription from the paternal allele. For the Igf2 upstream region and Igf2r region 2, the DNA methylation pattern is unusual in that modification is associated with the active allele. While it has not yet been shown that these sites are actually gene-specific regulatory elements, it is possible that modification at these positions acts by inhibiting the binding of repressor factors involved in allele-specific regulation.

Materials and methods

Preparation of DNA

DNA was prepared from adult or 18 day embryonic livers of purebred and interspecific hybrid mice as previously described by Ausubel *et al.* (1990).

F2 animals containing *M.spretus* and *M.musculus* Igf2 alleles were identified using the polymorphic *DraI* restriction enzyme (see Figure 1). Sperm DNA from ductus deferens was prepared as previously described by Ariel *et al.* (1991). Oocytes were collected from oviducts of super ovulated F1C57 black × BALB/c females and cleaned carefully from contaminating granulosa cells by treatment with hyalouronidase. They were further incubated in acid tyrode to remove the zona pelucida and individually washed in PBS several times (Hogan *et al.*, 1986). DNA was usually prepared from pools of 200 oocytes (Kafri *et al.*, 1992). Blastulae and morulae were prepared from mated superovulated females which were sacrificed 52 h after injection. The 4-cell morulae were released by shredding of the oviduct into DMEM, 10% FCS, penicillin/streptomycin and grown for 24–48 h at 37°C in 5% CO_2 to obtain 16–32 cell morulae or blastulae. DNA was prepared from pooled embryos after washing them individually several times in PBS. DNA of fetal germ cells from male and female gonads was prepared as previously described by Kafri *et al.* (1992) and McCarrey *et al.* (1987).

Blastulae for analysis of allelic methylation of the Igf2r region 2 were prepared by taking advantage of a difference between C3H and BALB/c alleles at a minisatellite VNTR adjacent to the Igf2r locus which can be identified by PCR analysis (Hearne *et al.*, 1991). BALB/c females were mated with C3H Hairpin (T^{hp}) hemizygous males. The resulting offspring had one BALB/c and either a C3H normal or C3H T^{hp} allele. The latter animals, identified by their distorted tail, were then mated to BALB/c mice in order to obtain 3.5 day blastulae. Samples of DNA, prepared from individual embryos, were then amplified by PCR using the primers TNFa5′ and TNFa3′ to distinguish between the BALB/c homozygotes and hemizygotes carrying a C3H T^{hp} allele. This procedure was used to obtain pooled embryos containing only the maternal or paternal Igf2r allele.

Methylation analysis of DNA

Methylation patterns in liver and sperm DNA were determined by Southern blotting using methyl-sensitive restriction enzymes. PCR methylation analysis of individual *HpaII* sites was performed as described by Kafri *et al.* (1992). Briefly, DNA from each sample was cut with either *PvuII* or *PvuII* plus *HpaII* and amplified with 100 ng of specific flanking primers in a 50–100 μl PCR reaction with 2 mM dNTPs and 2 units of Taq polymerase (Promega or USB). PCR cycles consisted of 1 min at 95°C, 2 min at 55–60°C and 2 min at 72°C, except for the first two cycles where denaturation was carried out for 4 min.

Germ cell DNA samples of 1 ng were amplified for 25 cycles, electrophoresed on agarose gels and blotted onto Hybond (Amersham) nylon membranes. Hybridization was performed with specific probes prepared from plasmids or PCR products. Oocyte, morula and blastula DNA samples of 100 pg were amplified for 40 cycles and visualized directly on silver-stained polyacrylamide gels (Sajantilla *et al.*, 1992). Under these conditions, the amount of visualized product is linear over a 50-fold range of DNA concentrations and the degree of modification can be accurately measured. This method does not, however, distinguish methylation levels between 50 and 100% (Kafri *et al.*, 1992). For most DNA samples we carried out both positive and negative controls for *HpaII* digestion by assaying known methylated or unmethylated sites. In addition, most PCR analyses were accompanied by a minus DNA control. These data are not always shown in the figures.

DNA sequencing

The *M.musculus* Igf2 upstream region sequences were cloned from a mouse genomic phage library into plasmids and subsequently sequenced. The H19 sequence was obtained from the Genbank database. Polymorphic species-specific restriction sites in the Igf2 upstream region and H19 gene were identified from PCR amplification products of *M.spretus* genomic DNA using dideoxy double-stranded sequencing with either the Sequenase™ (USB) or Circumvent™ kit (New England Biolabs).

PCR primers

Igf2 upstream *HpaII* sites 3–6 were analysed by PCR with the following primers: site 3, CCTTGAGCCACACTTTGACT (5′ primer) and CCAGAGATGAGCAAGGTTCT (3′ primer); site 4, CCCAGAACCTTGCTCATCTCTG (5′ primer) and GTCCTTGTGAGGCCAGATTAGG (3′ primer); site 5, GCCTAATCTGGCCTCACAAGGACTA (5′ primer) and ATCCACCAGCCTCTATCCCTGGCTT (3′ primer); site 6, ACAGGCCAGATGTTCTTGGAATGGG (5′ primer) and ACCACCCGTAAGAGGATGGTATCAG (3′ primer). Analysis of the H19 *HpaII* sites 3, 5 and 9 was performed with the following primers: site 3, ACCTCATCTGGAGTCTGG (5′ primer) and CCTTTGCTAACTATCCTG (3′ primer); site 5, TGATCGGTGACTCGAAGAGC (5′ primer) and AGACGGCTTCTACGACAAGG (3′ primer); site 9, ACTTCATCATCTCCCTCCTG (5′ primer) and GCATGTTGAACACTTTATGA (3′ primer). Igf2r region 2 *HpaII* sites 1–4 were analysed as follows. Site 3 was detected using the primer pair AACCCTCGGAACCCTGCCCTT (5′ primer) and TAGCACAACTCCAATTGTGCTGCG (3′ primer). Sites 1, 2 and 3 were detected in tandem using the same 3′ primer together with the 5′ primer CCTCGAATGCTCCCCTTGTGCAAG. Site 4 was studied with the primer pair TCAGAACACTGGTGAGCAGTGGG (5′ primer) and GAGGGTAGGATTCCGTTGCAAGG (3′ primer). PCR amplification for identifying BALB/c and C3H chromosome 17 alleles was performed with primers TNFa5′-GTTTCAGTTCTCAGGGTCCTA and TNFa3′ CAGGATTCTGTGGCAATCTGG.

Acknowledgements

The first two authors participated in this study equally. We thank A.Efstratiadis and D.Barlow for the early discussions which led up to this study. This research was supported by grants from the NIH, the Israel Cancer Research Fund, the US–Israel Binational Foundation, the Council for Tobacco Research (USA) and the Foulkes Foundation Fellowship (to T.K.).

References

Ariel,M., McCarrey,J. and Cedar,H. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 2317–2321.

Ausubel,F.M., Brent,R., Kingston,R.E., Moore,D.D., Seidman,J.G., Smith,J.A. and Struhl,K. (1990) *Current Protocols in Molecular Biology*. J.Wiley and Sons, New York.

Barlow,D.P., Stoger,R., Herrmann,B.G., Saito,K. and Schweifer,N. (1991) *Nature*, **349**, 84–87.

Bartolomei,M.S., Zemel,S. and Tilghman,S.M. (1991) *Nature*, **351**, 153–155.

Chaillet,J.R., Vogt,T.F., Beier,D.R. and Leder,P. (1991) *Cell*, **66**, 77–83.

DeChiara,T.M., Robertson,E.J. and Efstratiadis,A. (1991) *Cell*, **64**, 849–859.

Dittrich,B., Robinson,W.P., Knoblauch,H., Buiting,K., Schmidt,K., Gillessen-Kaesbach,G. and Horsthemke,B. (1992) *Hum. Genet.*, **90**, 313–315.

Driscoll,D.J. and Migeon,B.R. (1990) *Somatic Cell. Genet.*, **16**, 267–275.

Driscoll,D.J., Waters,M.F., Williams,C.A., Zori,R.T., Glenn,C.C., Avidano,K.M. and Nicholls,R.D. (1992) *Genomics*, **13**, 917–924.

Ferguson-Smith,A.C., Sasaki,H., Cattanach,B.M. and Surani,M.A. (1993) *Nature*, **362**, 751–755.

Frank,D., Keshet,I., Shani,M., Levine,A., Razin,A. and Cedar,H. (1991) *Nature*, **351**, 239–241.

Grant,M., Zuccotti,M. and Monk,M. (1992) *Nature Genet.*, **2**, 161–166.

Hearne,C.M., McAleer,M.A., Love,J.M., Aitman,T.J., Cornall,R.J., Ghosh,S., Knight,A.M., Prins,J.B. and Todd,J.A. (1991) *Mammalian Genome*, **1**, 273–282.

Hogan,B., Constantini,F. and Lacey,E. (1986) *Manipulating The Mouse Embryo*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Kafri,T., Ariel,M., Brandeis,M., Shemer,R., Urven,L., McCarrey,J., Cedar,H. and Razin,A. (1992) *Genes Dev.*, **6**, 705–714.

Leff,S.E., Brannan,C.I., Reed,M.L., Ozcelik,T., Franke,U., Copeland,N.G. and Jenkins,N.A. (1992) *Nature Genet.*, **2**, 259–264.

Lock,L.F., Takagi,N. and Martin,G.R. (1987) *Cell*, **48**, 39–46.

McCarrey,J.R., Hsu,K.C., Eddy,E.M., Klevecz,R.R. and Bolen,J.L. (1987) *J. Exp. Zool.*, **242**, 107–111.

Monk,M., Boubelik,M. and Lehnert,S. (1987) *Development*, **99**, 371–382.

Sajantila,A., Budowle,B., Strom,M., Johnson,V., Lukka,M., Peltonen,L. and Ehnholm,C. (1992) *Am. J. Human Genet.*, **50**, 816–825.

Sasaki,H., Jones,P.A., Chaillet,J.R., Ferguson-Smith,A.C., Barton,S.C., Reik,W. and Surani,M.A. (1992) *Genes Dev.*, **6**, 1843–1856.

Shemer,R., Kafri,T., O'Connell,A., Eisenberg,S., Breslow,J.L. and Razin,A. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 11300–11304.

Stoger,R., Kubicka, Liu,C.-G., Kafri,T., Razin,A., Cedar,H. and Barlow,D.P. (1993) *Cell*, **73**, 61–71.

Surani,M.A., Reik,W. and Allen,N.D. (1988) *Trends Genet.*, **4**, 59–62.

Ueda,T., Yamazaki,K., Suzuki,R., Fujimoto,H., Sasaki,H., Sakaki,Y. and Higashinakagawa,T. (1992) *Development*, **116**, 831–839.

Yeivin,A. and Razin,A. (1993) In Jost,J.P. and Saluz,H.P. (eds), *DNA Methylation: Molecular Biology and Biological Significance*. Birkhauser Verlag, Basel, pp. 523–568.

Zemel,S., Bartolomei,M.S. and Tilghman,S.M. (1992) *Nature Genet.*, **2**, 61–65.

Received on April 21, 1993; revised on June 7, 1993