(75)

A Genetic Investigation of the Establishment of Genomic Imprinting

by

Kerry Lee Tucker

A. B., Biochemistry
Harvard College - 1990

Submitted to the Department of Biology in Partial Fulfillment
of the Requirements for the Degree of

Doctor of Philosophy in Biology

at the Massachusetts Institute of Technology

June 1997

Signature of Author... .....
Department of Biology
May One, 1997

Certified by.. .....
Rudolf Jaenisch
Professor of Biology
Thesis Supervisor

Accepted by ....
Frank Solomon
Chairman, Biology Graduate Committee

A Genetic Investigation of the Establishment of Genomic Imprinting

by

Kerry Lee Tucker

Submitted to the Department of Biology on May One, 1997,
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy In Biology

ABSTRACT

Much evidence indicates that in vertebrates the methylation of DNA at cytosine residues correlates with gene inactivity. A single methyltransferase activity, DNA (cytosine-5)-methyltransferase, has been purified from mammalian cells and its gene (*Dnmt*) cloned. The crucial importance of DNA methylation in development was demonstrated by the targeted mutation of *Dnmt* in murine embryonic stem (ES) cells (Li *et al.*, 1992). Embryos with homozygous disruptions of *Dnmt* die in mid-gestation. In contrast, homozygous mutant ES cells proliferate normally with their DNA highly demethylated, though they die upon differentiation.

The research presented in this thesis focused on expressing the *Dnmt* cDNA in the background of *Dnmt* mutant ES cells. Initial attempts using a cDNA reported by Bestor *et al.* (1988) and a corresponding promoter region reported by Rouleau *et al.* (1992) were unsuccessful in effecting adequate levels of functional MTase expression. A complete cDNA sequence was obtained through the cloning of two new 5'-proximal exons, which extend the open reading frame up to 171 codons upstream of the previously defined start site. This showed the previously-reported promoter sequence to lie in the second intron of the gene, with no evidence that it functions in ES cells. It was found that the extra coding sequence was necessary for functional expression of the gene in ES cells.

Expression of the wild type *Dnmt* cDNA in mutant ES cells caused an increase in methylation of bulk DNA to normal levels, but did not restore the methylation of imprinted genes, whose expression is monoallelic and dependent upon their parental derivation. The expression of these genes was

deregulated because of this hypomethylation, as shown previously in *Dnmt* mutant embryos (Li *et al.*, 1993). Full restoration of monoallelic methylation and expression was imposed on imprinted genes upon germline transmission. These results are consistent with the presence of distinct *de novo* DNA methyltransferase activities during oogenesis and spermatogenesis which specifically recognize imprinted genes but which are absent in the post-implantation embryo and in ES cells.

Because the "rescued" ES cells display biallelic hypomethylation and expression of imprinted genes, their developmental capacities may be impaired, as is seen with cells derived from parthenogenetic or androgenetic embryos, which also display biallelic expression of imprinted genes. Alternatively, the rescued ES cells may have normal developmental capacities, as predicted by a model which explains genomic imprinting as a nonessential, evolutionarily-recent addition to the basic mammalian developmental program. Rescued cells differentiated normally *in vitro*, formed teratomas with a wide variety of differentiated cell types, and contributed substantially to coat color in adult chimeras. Further analysis of tissue-specific distribution of rescued ES cells in chimeras will allow a full assessment of the developmental potential of a mouse lacking an imprinted genome.

Thesis Supervisor: Rudolf Jaenisch

Title: Professor of Biology

Super cuncta est murium fetus, haut sine cunctatione dicendus, quamquam sub auctore Aristotele et Alexandri Magni militibus. generatio eorum lambendo constare, non coitu, dicitur. ex una genitos CXX tradiderunt, apud Persas vero praegnantes in ventre parentis repertas; et salis gustatu fieri praegnantes opinantur. itaque desinit mirum esse unde vis tanta messes populetur murium agrestium; in quibus illud quoque adhuc latet quonam modo illa multitudo repente occidat: nam nec exanimes reperiuntur neque extat qui murem hieme in agro effoderit. plurimi ita ad Troada proveniunt, et iam inde fugaverunt incolas. proventus eorum siccitatibus. tradunt etiam obituris vermiculum in capite gigni. Aegyptiis muribus durus pilus sicut irenaceis; idem bipedes ambulant ceu Alpini quoque.

Plinii, *Naturalis Historia*, Liber X

"The most prolific of all animals whatever is the mouse - one hesitates to state its fertility, even though on the authority of Aristotle and the troops of Alexander the Great. It is stated that impregnation takes place by licking and not by coupling. There is a record of 120 being born from a single mother, and in Persia of mice already pregnant being found in the parent's womb; and it is believed that they are made pregnant by tasting salt. Accordingly it ceases to be surprising how so large an army of field-mice ravages the crops; and in the case of field-mice it is also hitherto unknown exactly how this vast multitude is suddenly destroyed: for they are never found dead, and nobody exists who ever dug up a mouse in a field in winter. Vast numbers thus appear in the Troad, and they have by now banished the inhabitants from that country. They appear during droughts. It is also related that when a mouse is going to die a worm grows in its head. The mice in Egypt have hard hair like hedgehogs, and also they walk on two feet, as also do the Alpine mice."

Translation by H. Rackham

TABLE OF CONTENTS

Introduction

I. DNA Methylation and DNA Methyltransferases

Cytosine methylation in vertebrate genomes

The genomic DNA of vertebrates, higher plants, and many fungi and bacteria contains a covalent modification of cytosine, in which a methyl group is added on to the fifth carbon atom in the pyrimidine ring (for review see Jost and Saluz, 1993). In vertebrates, this methylation occurs mostly at the symmetrically opposed cytosines in CpG dinucleotides (Gruenbaum *et al.*, 1982). This short recognition sequence is underrepresented about five-fold in the mammalian genome, and about 60% of the 5×10^7 CpG dinucleotides in the mammalian genome are methylated (Bestor *et al.*, 1984; Gama-Sosa *et al.*, 1983). Most of the remaining, unmethylated CpG residues are densely clustered together in CpG islands (Bird, 1986), short regions (averaging 1000 bp in length) of GC-rich DNA usually found near the promoters of constitutively-expressed, "housekeeping" genes, which stay hypomethylated throughout development (Bird, 1986).

A strong inverse correlation has been documented between cytosine methylation of genes and their transcriptional activity (Yeivin and Razin, 1993). The mechanism by which cytosine methylation inhibits transcriptional activity is unclear. Some evidence has accumulated that methylation of specific CpG residues prevents the binding of transcription factors (Comb and Goodman, 1990; Iguchi-Ariga and Schaffner, 1989; Prendergast *et al.*, 1991; Watt and Molloy, 1988). However, not all transcription factors are methylation-sensitive to DNA binding (Hoeller *et al.*, 1988), and few transcription factors contain the dinucleotide in their recognition sequence. In addition, the new technique of mapping methylation sites by bisulfite sequencing (Frommer *et al.*, 1992) has shown that, even within one tissue of a single organism, DNA

methylation patterns at any given gene vary considerably (Feil *et al.*, 1994). Another possibility is that repression occurs indirectly through recognition of methylated cytosine residues by nuclear proteins, which can then either directly block access by the transcriptional apparatus or effect a change in local chromatin structure disfavoring such access (Boyes and Bird, 1991). Two proteins have been found which bind specifically to methylated CpG residues (Lewis *et al.*, 1992; Meehan *et al.*, 1989), and some evidence exists that they can repress transcription *in vitro* (Boyes and Bird, 1992; Nan *et al.*, 1997). A similar mechanism can be envisaged to explain activation of genes when they are methylated.

There are various explanations of the biological function of methylation in vertebrates. In prokaryotes methylation-sensitive restriction endonucleases act as a "host defense mechanism" by digesting foreign, unmethylated DNA, while methyltransferase activities in turn protect newly-replicated host DNA by rendering it resistant to the host endonuclease activity (Noyer-Weidner and Trautner, 1993). This has led to the idea that methylation protects the vertebrate genome from invasive parasites such as retroviruses and transposons (Doerfler, 1991), which can damage the functioning of essential genes by integrating within or nearby them. Indeed, methylation of retroviral elements results in their transcriptional and transpositional silencing (Simon *et al.*, 1983; Vardimon *et al.*, 1982), and this unchecked transposition can cause severe mutations in important genes (Jähner and Jaenisch, 1985; Michaud *et al.*, 1994).

Another model proposes that methylation leads to heterochromatic condensation of "junk" DNA in the large genomes of vertebrates and angiosperms (Bird, 1995). This compartmentalization allows the transcriptional apparatus faster access to those genes necessary for cell

survival, which are characterized by the presence of unmethylated CpG islands. This theory would explain why organisms with smaller genomes, such as *Drosophila* and *Saccharomyces*, do not possess the methylation modification system (Proffitt *et al.*, 1984; Urieli-Shoval *et al.*, 1982). However, the causal relationship between methylation and chromatin structure is unclear. Little evidence indicates that DNA methylation precedes the formation of a condensed, transcriptionally inactive chromatin structure (Buschhausen *et al.*, 1987), and it remains to be seen how closely genome size correlates with the presence of methylation systems. These two theories are not mutually exclusive, nor do they exclude the major role that DNA methylation is thought to play in the control of specific genes during differentiation and development (Yeivin and Razin, 1993).

Changes in methylation patterns during mammalian development

Methylation patterns change dramatically during murine development. During gametogenesis, the unmethylated DNA of the primordial germ cells becomes *de novo* methylated (Monk *et al.*, 1987), so that specific interspersed repetitive (Sanford *et al.*, 1987) and single-copy (Kafri *et al.*, 1992) gene sequences in sperm are hypermethylated relative to the sequences in oocytes. The methylation present in the DNA of the zygote is erased during cleavage so that the genome of the blastocyst is almost completely unmethylated (Chaillet *et al.*, 1991; Kafri *et al.*, 1992; Monk *et al.*, 1987). After implantation, the embryo undergoes a wave of global *de novo* methylation which restores the genomic methylation levels of the gastrulating embryo to levels seen in the adult (Jähner *et al.*, 1982; Monk *et al.*, 1987), though to a lesser degree in extraembryonic tissues than in somatic tissues (Chapman *et al.*, 1984; Monk *et al.*, 1987). The purpose of this erasure and restoration of methylation patterns

is unclear, but it may be to erase the different methylation patterns inherited from the male and female gametes and to allow epigenetic differences in the various gastrulating germ layers to be subsequently established. The only genes which escape this remethylation are imprinted genes (see below), genes active during embryogenesis, such as a-fetoprotein (Vedel *et al.*, 1983), and CpG island-bearing genes (Kafri *et al.*, 1992), which are consistently unmethylated throughout development. Later in development, different tissues undergo their own specific demethylation programs which are implicated in activating genes required in those tissues, though the role methylation has to play in this is still very unclear.

The *Dnmt*-encoded MTase has maintenance methylation activity both *in vitro* and *in vivo*

A single methyltransferase activity, the DNA (cytosine-5)-methyltransferase (MTase, EC 2.1.1.37), has been purified from mammalian cells (Bestor and Ingram, 1983; Bestor and Ingram, 1985) and its gene cloned (Bestor *et al.*, 1988). Two lines of evidence have led to the commonly-accepted view that MTase acts *in vivo* as a *maintenance* methyltransferase, which restores a double-stranded methylation profile to newly-synthesized, hemimethylated DNA (Fig. 1). Firstly, MTase shows a marked preference for hemimethylated DNA *in vitro* (Bestor and Ingram, 1983; Gruenbaum *et al.*, 1982). Secondly, MTase is shown to co-localize to sites of active DNA replication in S-phase cells called replication foci (Leonhardt *et al.*, 1992). The maintenance activity of MTase stands in contrast to *de novo* methyltransferase activity, which occurs on completely unmethylated DNA substrates. Though faithful replication of methylation patterns is probably its

Fig. 1- *Dnmt*: A Maintenance Methyltransferase

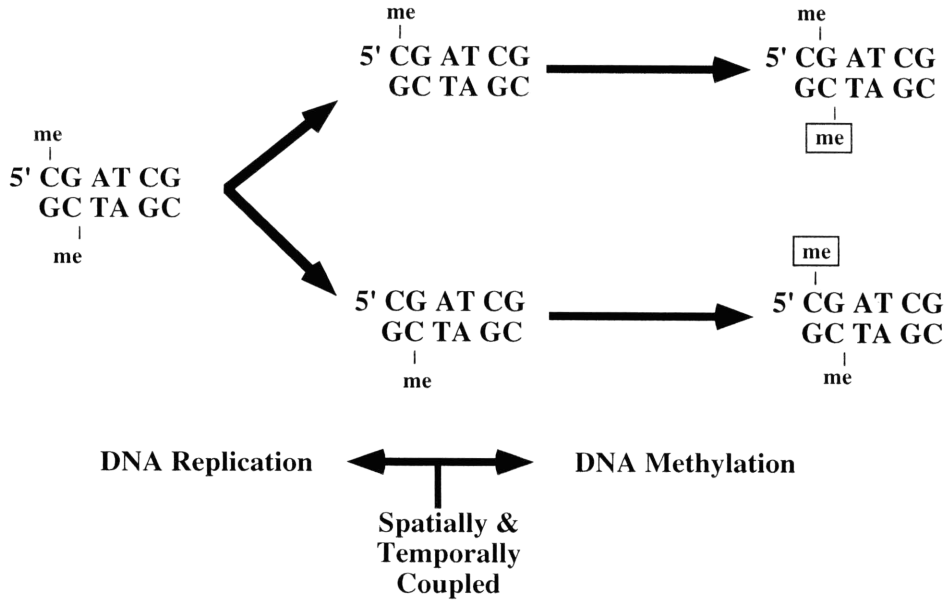
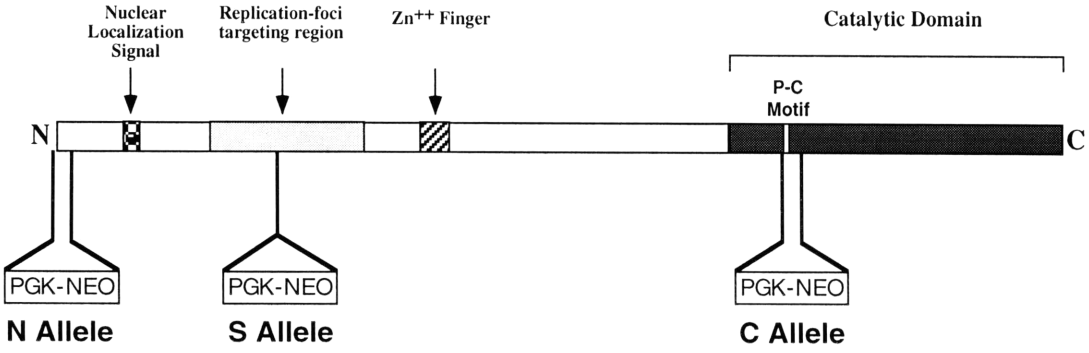


Fig. 2 - Targeted Mutations in *Dnmt*



Homozygous Mutants:

- **Reduced Genomic Methylation**
- **Embryonic Lethal (~ 9.5 d.p.c.)**
- **ES Cells are Viable**

major role in mammals, it is possible that MTase possesses *de novo* activity *in vivo*.

The enzyme consists of a 1000-amino acid amino-terminal domain with three peptide motifs (Fig. 2): an amino-terminal nuclear localization signal, a 250-amino acid domain responsible for targeting MTase to replication foci in the nuclei of dividing cells (Leonhardt *et al.*, 1992), and a zinc-binding zinc-finger motif (Bestor, 1992). This is followed by a 500 amino acid carboxyl-terminal domain, assumed to contain the catalytic center from its homology to cytosine-methylating bacterial type II methyltransferases (Bestor *et al.*, 1988; Kumar *et al.*, 1994). This domain contains all of the ten consensus motifs found in this class of enzymes in their invariant order, including the proline-cysteine amino acid doublet responsible for enzymatic addition to the cytosine ring (Leonhardt and Bestor, 1993). The two domains of MTase are joined together by a run of 13 glycyl - lysyl amino acid pairs.

The crystal structure of the bacterial *Hha*I methyltransferase has recently been solved (Cheng *et al.*, 1993; Klimasauskas *et al.*, 1994), and the results are worth mentioning because the highly-conserved protein sequence motifs shared by MTase and M.*Hha*I are probably reflective of conservation in structure and enzymatic function. The crystal structures show M.*Hha*I to be organized into two domains, a larger one containing most of the conserved catalytic motifs, and a smaller one containing residues important for sequence specificity of the enzyme. The ubiquitous cofactor S-adenosyl-L-methionine serves as the methyl donor to the C5 residue of the cytosine ring, and the methyl group is added after a covalent bond has been formed between a highly-conserved cysteine residue in the enzyme and the C6 residue in the pyrimidine ring. Amazingly, the crystal structure of the DNA-bound enzyme shows that the cytosine residue is completely extruded from the B-form DNA

helix and inserted into the active-site pocket of the enzyme, allowing for the stereoelectronic attack trajectories, normal to the plane of the cytosine ring, which are thought necessary for this catalytic mechanism.

The most remarkable functional difference between the bacterial type II methyltransferases and the mammalian homologues is that the latter discriminate between hemimethylated and unmethylated DNA substrates (Bestor and Ingram, 1983). Treatment of purified murine MTase with protease V8 caused cleavage near the linker separating the catalytic and the amino-terminal domains, resulting in a large increase in the rate of *de novo* methylation without affecting the rate of methylation of hemimethylated DNA (Bestor, 1992). Unfortunately, the proteolyzed fragments were not purified, so it is unclear whether MTase possesses *de novo* activity in the absence of the amino-terminal domain. It is also unclear by what means the amino-terminus of MTase could influence the discrimination of unmethylated versus hemi-methylated substrates, given that steric considerations probably prevent its direct contact with the major groove (Bestor, 1992). Proteolytic truncation of the first 350 amino acids does not affect substrate specificity or *in vitro* activity (Bestor, 1992; Bestor and Ingram, 1985), so presumably the region of the amino terminus lying C-terminal to the replication-foci targeting domain is involved in this substrate discrimination. An elucidation of the atomic structure of MTase bound to its DNA substrate is eagerly awaited.

Based upon homology to murine and bacterial cytosine methyltransferases, genes encoding presumptive methyltransferase activities have been cloned from human (Yen *et al.*, 1992), chicken (Tajima *et al.*, 1995), sea urchin (Aniello *et al.*, 1996), *Xenopus laevis* (Kimura *et al.*, 1996) and *Arabidopsis thaliana* (Finnegan and Dennis, 1993). Of these five genes, the

first four are presumed to encode maintenance methyltransferase activities because of the similarity of their amino-terminal domains to that of the murine MTase, while the amino-terminal domain from *A. thaliana* bears no resemblance to the murine gene and may encode a *de novo* activity. Weak evidence also exists for a DNA methyltransferase gene in *Drosophila* (Vanyushin and Poirier, 1996). Interestingly, a gene (*pmt1*$^+$) bearing significant homology to both the bacterial and mammalian cytosine methyltransferases has been cloned from the fission yeast *Schizosaccharomyces pombe*, despite the apparent lack of DNA methylation in this organism (Wilkinson *et al.*, 1995). The enzyme encoded by this gene possesses no *in vitro* methylation activity. However, deletion of a serine residue found in *pmt1*$^+$ between the canonical proline-cysteine doublet in the catalytic domain allows for production of an enzyme bearing *in vitro* and *in vivo* methyltransferase activity with the same recognition sequence as both the *E. coli dcm* gene product and M.*Eco*RII (Pinarbasi *et al.*, 1996).

Transcriptional and translational control of *Dnmt* expression

The transcriptional control of *Dnmt* is complex. The steady-state level of *Dnmt* mRNA increases as synchronized Balb/c 3T3 cells enter S phase of the cell cycle (Szyf *et al.*, 1991), resulting in a concomitant increase of MTase that closely matches the synthesis of DNA (Szyf *et al.*, 1985). Despite this increase in mRNA levels, nuclear run-on assays indicated that *Dnmt* was transcribed in resting-cell nuclei at the same rate as in S-phase nuclei, indicating that the major determinant of mRNA abundance is post-transcriptional (Szyf *et al.*, 1991). Post-transcriptional down-regulation of *Dnmt* gene activity has also been implicated in the differentiation of both F9 mouse embryonal carcinoma cells (Teubner and Schulz, 1995) and myoblasts

(Liu *et al.*, 1996), where actinomycin-treated nuclei were used to indicate an increase in *Dnmt* mRNA turnover.

A putative promoter has been reported for *Dnmt* (Rouleau *et al.*, 1992). The authors presented data mapping transcriptional start sites of the gene in various cell lines using RNase protection and primer extension assays, and they also employed standard CAT assays in transient transfections to establish a minimal promoter sequence. As described in Chapter One, this putative promoter sequence was found incapable of supporting expression of the *Dnmt* cDNA in embryonic stem (ES) cells, and no evidence was found for transcriptional start sites in the places reported by Rouleau *et al.* Instead, a region 12 kb upstream of this putative promoter has been found to drive appropriate expression of *Dnmt* in ES cells, and is believed to contain the actual promoter (Tucker *et al.*, 1996b).

The 5.2-kb *Dnmt* transcript is found in all tissues of the mouse, and at high levels in ovary, testis, cerebellum, spleen, and heart (Trasler *et al.*, 1992). These relative differences are reflected in protein levels (D. Talbot, unpublished observation). Intriguingly, *Dnmt* was found to be expressed at low levels in postmitotic neurons in the cerebral cortex, the cerebellar granular layer, and hippocampal neuronal layers (Goto *et al.*, 1994), while another report has shown high levels of methyltransferase activity in extracts made from cerebellum (Brooks *et al.*, 1996). In testis, an additional 6.2-kb transcript was found. The gradual appearance of this transcript during spermatocyte maturation correlated with the disappearance of MTase protein, to the point that the enzyme was undetectable in post-replicative pachytene spermatocytes, which expressed only the 6.2-kb message (Jue *et al.*, 1995; Trasler *et al.*, 1992). The significance of this is not at all understood.

In addition to transcriptional control of *Dnmt* gene expression, post-translational control of MTase levels has been observed. In differentiating myoblasts, levels of MTase decreased more rapidly than markers for cellular proliferation (Liu *et al.*, 1996), which is accounted for by an increased turnover rate of MTase, as indicated by pulse-chase experiments.

Subcellular localization of MTase

MTase displays a cell cycle-regulated cellular distribution (Leonhardt *et al.*, 1992; Vogel *et al.*, 1988). As elucidated by Leonhardt *et al.* (1992), MTase colocalizes during S phase with the 1-3 μm toroidal "replication foci" that are known centers of active DNA replication. Dissection of the protein revealed a 250-amino acid domain in the amino terminus of the protein that is both sufficient and necessary for localization to replication foci. In addition, immunofluorescent observations of MTase in developing mouse embryos has revealed a remarkable mobility of the enzyme. Levels of MTase are very high throughout preimplantation development, though they decline by more than a factor of 10 from oocytes to blastocysts (Carlson *et al.*, 1992; Howlett and Reik, 1991; Monk *et al.*, 1991). Despite these high levels, MTase is excluded from the nucleus in one- to four-cell stage embryos, concentrating in small cytoplasmic granules near the nuclear periphery (Carlson *et al.*, 1992). By the compacted 8-cell stage, MTase has moved into the nucleus, but it displays a uniform staining, unlike the replication foci seen in cultured cells. Blastocysts show a reversal of this trend, with only faint staining of the nuclei of both trophoblast and inner cell mass cells. In 5.5 days post coitum (dpc) embryos, MTase has moved back into the nucleus, and stays localized there throughout postimplantation development (Trasler *et al.*, 1996). These observations are difficult to reconcile with observations of declining genomic methylation

levels in the preimplantation embryo (Chaillet *et al.*, 1991; Monk, 1990; Monk *et al.*, 1987), because concentrations of MTase in the nucleus are at their highest during a period when the rate of demethylation of genomic DNA is at a maximum (Monk, 1990).

Dnmt is essential for normal murine development

The crucial importance of DNA methylation in development was demonstrated by the targeted mutation of the single known mammalian DNA methyltransferase gene (*Dnmt*) in ES cells. Three mutant alleles of the *Dnmt* gene have been generated (Fig. 2) (Lei *et al.*, 1996; Li *et al.*, 1993; Li *et al.*, 1992). A partially inactivated allele ($Dnmt^n$) caused lethality of homozygous mutant embryos at the 25 somite stage, whereas embryos homozygous for completely inactivated alleles ($Dnmt^s$ and $Dnmt^c$) died at the 5 somite stage. Embryos displayed morphologically normal major organ rudiments and appeared to be developmentally delayed by one day, but histological examination revealed widespread cell death and a low mitotic index (Lei *et al.*, 1996; Trasler *et al.*, 1996). ES cells homozygous for any of the disrupted *Dnmt* alleles proliferate normally with their genomic DNA highly demethylated but die upon induction of differentiation. Importantly, ES cells homozygous for the $Dnmt^c$-allele gene disruption retain the ability to *de novo* methylate transduced retroviral genomes, as seen in wild-type preimplantation embryos, embryonic carcinoma (EC), and ES cells (Lei *et al.*, 1996; Stewart *et al.*, 1982). Because this gene disruption destroys the catalytic capacity of MTase by removing the proline-cysteine motif, this result strongly suggests the expression in ES cells of an independently-encoded *de novo* DNA methyltransferase (Lei *et al.*, 1996). The identification of this activity is actively being pursued.

Ectopic expression of cytosine methyltransferases in yeast and mammalian cell lines

Several attempts at transgenic expression of cytosine methyltransferases have been reported. Bacterial methyltransferases have been transfected into yeast and mammalian cells under the control of heterologous promoters. The *Bacillus sphaericus* R methyltransferase was expressed in yeast and reported to cause methylation of endogenous genes (Fehér *et al.*, 1983). Expression of the *Bacillus subtilis* SPR phage cytosine methyltransferase in yeast resulted in methylation of CCGG sequences in various genes, which increased when the expression of the examined genes was induced (Singh and Klar, 1992). The bacterial *Hha*I methyltransferase, which methylates the internal C in the sequence GCGC, has been expressed in NIH 3T3 cells (Wu *et al.*, 1996). High levels of expression led to cell death, while lower levels of expression led to increased methylation of genomic *Hha*I recognition sequences and to increased clonability in soft agar, indicating a transformed phenotype.

Two reports of transgenic expression of the mammalian MTase have appeared in the literature (Vertino *et al.*, 1996; Wu *et al.*, 1993). Overexpression of the murine MTase in NIH 3T3 cells was claimed to cause a transformed phenotype (Wu *et al.*, 1993). Definitive evidence of overexpression was not demonstrated, however, and the results are called into question by the fact that the cDNA used in the experiments was incomplete (see Chapter One). The same group reported overexpression of the human MTase in a human fetal lung fibroblast line immortalized with SV40 (Vertino *et al.*, 1996). In this paper, up to 20-fold levels of MTase overexpression were convincingly demonstrated by Western blot and by an increase in *in vitro* enzyme activity in cellular extracts of the transgenic

clones. *De novo* methylation of 5 / 12 CpG island-bearing genes was demonstrated for three clones expressing >9-fold levels of MTase. Again, however, the authors were working with a truncated cDNA, as demonstrated in the paper by a recombinant MTase with an apparent molecular mass approximately 10 kDa shorter than the endogenous MTase, which migrates in SDS-PAGE at 190 kDa. Though this recombinant enzyme showed no increase in *de novo* activity *in vitro,* its apparent toxicity in this and my studies (see Chapter One) demands caution when interpreting results obtained with it.

Ectopic expression of *Dnmt* in embryonic stem cells

The embryonic stem (ES) cell offers enormous advantages as a cell line for experimental manipulation of murine gene expression, because of the high frequency of homologous recombination in gene-targeting experiments and the cell's subsequent pluripotency in the production of mice bearing a desired genetic alteration. Primary cell lines derived from mice bearing such genetic alterations are favorable to individual clones derived from the transfection of standard immortalized cells, as described above, because extended culturing of cells leads to many unpredictable epigenetic changes which can make interpretations problematic (Antequera *et al.*, 1990). In order to establish a model of ectopic *Dnmt* expression in the mouse, I have attempted to express the *Dnmt* gene as a transgene in ES cells. Homozygous mutant *Dnmt* ES cells are viable and their gross methylation status is easily ascertained by Southern blot analysis, making amenable the functional analysis of *Dnmt* expression in this genetic background.

In contrast to previous published reports (Vertino *et al.*, 1996; Wu *et al.*, 1993), I was unable to establish stable transfectants when a *Dnmt* cDNA sequence was placed under the control of two heterologous promoters,

murine *Pgk-1* and human EF-1α. Subsequent analysis of the 5' end of the *Dnmt* mRNA produced in ES cell lines revealed the existence of two new exons, upstream of the published 5' cDNA sequence terminus (Bestor *et al.*, 1988), which extend the open reading frame up to 171 codons upstream of the previously defined start site and show the previously-reported promoter sequence (Rouleau *et al.*, 1992) to lie in the second intron of the gene, with no evidence that it functions in ES cells. It was found that the extra coding sequence was necessary for functional expression of the gene in ES cells.

Expression of the full-length *Dnmt* cDNA in mutant ES cells was achieved through a mini-gene construct employing the endogenous *Dnmt* promoter (Tucker *et al.*, 1996b). Only when this sequence was restored could the *Dnmt* gene product rescue the hypomethylation phenotype of *Dnmt* mutant cells. Restoration of normal methylation levels in *Dnmt* mutant cells restored their developmental potential; in contrast to the parental mutant ES cells, the rescued cells were capable of differentiating *in vitro* and generating teratomas with a wide variety of differentiated cell types. This suggests that genomic remethylation resulted in the restoration of the proper transcriptional program required for differentiation of the ES cells into different tissues. These data also suggest the existence of an amino-terminal region of the protein necessary for normal function *in vivo*.

II. DNA Methylation and Genomic Imprinting

Parthenogenotes and androgenotes are developmentally compromised

For many years it had been known that parthenogenotes, the products of accidental activation of the unfertilized oocyte, were inviable, but the reasons for this were unclear (Graham, 1974). The first suggestion that the

maternal and paternal genomes may be non-equivalent was provided by the observation that in the extraembryonic tissues of the mouse the paternally-derived X-chromosome is preferentially inactivated (Takagi and Sasaki, 1975). The non-equivalence of the maternal and paternal genomes was conclusively demonstrated by the experimental manipulation of fertilized mouse eggs. Single maternally- or paternally-derived pronuclei were surgically removed from one-cell stage embryos and transplanted to recipient embryos, which themselves lacked either a paternally- or maternally-derived pronucleus, respectively (McGrath and Solter, 1983). In this fashion biparental diploid gynogenetic (two maternally-derived pronuclei), androgenetic (two paternally-derived pronuclei), and normal embryos could be tested for their ability to develop *in utero*. Despite the normal developmental potential of parentally "balanced" control embryos, gynogenotes and androgenotes did not develop to term, exhibiting complementary deficiencies in developmental potential (Barton *et al.*, 1984; Mann and Lovell-Badge, 1984; McGrath and Solter, 1984; Surani *et al.*, 1984). Parthenogenotes/Gynogenotes were found to develop poorly, becoming arrested at key steps in peri-implantation development (Sturm *et al.*, 1994; Varmuza *et al.*, 1993). With only about 20% of embryos surviving to the 10-20-somite stage, the most advanced embryos appeared morphologically normal but smaller and with very poor development of extraembryonic membranes and trophoblast (Sturm *et al.*, 1994; Surani and Barton, 1983; Surani *et al.*, 1984). In contrast, the preimplantation development of androgenotes was severely compromised (Barton *et al.*, 1984). Very few embryos progressed to the somite stage, and none made more than 6-8 somites. However, androgenotes exhibited extensive proliferation of the extraembryonic tissue. These data clearly showed that the paternal and maternal genomes are functionally different,

and that the embryo needs one maternally- and one paternally-derived haploid genome for normal development to occur. The complementary phenotypes of parthenogenotes and androgenotes was puzzling, and was noticed later in analysis of chimeras made with embryonic cells derived from parthenogenotes or androgenotes (see Conclusion).

Imprinted chromosomal regions in the mouse and human

Ascertaining the importance of genomic imprinting to mammalian development was hampered by the early-embryonic lethality of androgenetic- and gynogenetic-embryos. These data have been supplemented by examining the effects of regional uniparental disomy (UPD) on development of mice and humans. A series of mice bearing UPD were produced, in which both homologous copies of a specific chromosome (or portion thereof) were inherited exclusively from the mother or the father (Cattanach and Kirk, 1985). In this manner, eight autosomal regions were identified which resulted in developmental and behavioral abnormalities when exhibiting UPD (Beechey and Cattanach, 1996). The "genetic imprinting map" thus constructed also suggested that these genomic regions, and presumably specific genes within them, were responsible for the more severe developmental abnormalities seen in parthenogenotes and androgenotes.

Analyses in mice have been complemented by the intensive study of several diseases associated with UPD in the human (reviewed in Lalande, 1996). Developmental abnormalities are caused by the uniparental inheritance of four regions of the human genome. Prader-Willi syndrome, which is characterized by hypotonia and slow growth in infancy, mild mental retardation, and obesity, is caused either by paternally-inherited deletions (Butler and Palmer, 1983) or maternal UPD (Nicholls *et al.*, 1989) of

chromosome 15q11-q13. Paternal UPD (Malcolm *et al.*, 1991) or maternally-inherited deletions (Knoll *et al.*, 1989) of the identical chromosomal region result in Angelman syndrome, which is characterized by craniofacial peculiarities, an ataxic gait and jerky arm movements, severe mental retardation, hyperactivity, seizures, and EEG abnormalities. Maternal UPD of the syntenic genomic region in mouse (central chromosome 7) was shown to result in neonatal lethality (Cattanach *et al.*, 1992). Beckwith-Wiedemann Syndrome (BWS) is associated with multiorgan hyperplasia and a predisposition to embryonal tumors (Junien, 1992). Both paternal UPD and paternal duplication of the chromosomal region 11p15.5 results in BWS (Henry *et al.*, 1991). All BWS-causing paternal duplications and disomies include the linked genes *H19* and *Igf2*, which are discussed at length in this thesis. Finally, many other diseases show parent-of-origin effects, such as Wilms' tumor, rhabdomyosarcoma, neuroblastoma, and Philadelphia-chromosome-positive leukemia (Reeve, 1995). These results confirm the observations made in mice, that several discrete regions of the genome cause phenotypic abnormalities when inherited in a uniparental fashion.

In summary, many lines of evidence have shown that normal development of both mice and humans requires a balanced contribution of homologous chromosomes from both parents. Inbred mouse strains have been used to show that epigenetic differences exist between the haploid genome donated to the embryo by the father and by the mother. Unbalanced contribution of maternally- and paternally-derived genomic material to mice and humans leads to phenotypes ranging from early embryonic lethality, as seen in murine androgenotes, to milder defects such as the distinguishing facial characteristics of children with Angelman syndrome. Many phenotypes were reciprocal (e.g. the ability of parthenogenotes and androgenotes to form

embryonic and extraembryonic structures, respectively), suggesting the absence of specific gene products in parthenogenotes which were present in androgenotes, and vice versa. Differences in expression patterns between the two cell types may reflect allele-specific differences in transcription levels, which are determined solely by the parental origin of each allele of a given "imprinted" gene. The specific location of these imprinted genes was suggested by the maps of imprinted chromosomal regions developed for the mouse and by the association of genomic deletions with the human diseases discussed above. The past five years have seen a flurry of reports describing the identification of these imprinted genes, and evidence has accumulated to support the hypothesis that it is their deregulated expression which causes the developmental defects seen in mice and humans with UPD.

Imprinted genes show allele-specific transcription

Imprinted genes display parent-specific monoallelic expression, although often not at all developmental stages (Szabó and Mann, 1995a; Szabó and Mann, 1995b) or in all tissues of the organism (Latham, 1995). Genes which are predominantly expressed from the maternal allele[1] include *Igf2r* (Barlow *et al.*, 1991), *H19* (Bartolomei *et al.*, 1991), *Mash2* (Guillemot *et al.*, 1995), $p57^{KIP2}$ (Hatada and Mukai, 1995), and *KVLQT1* (Lee *et al.*, 1997). Genes expressed from the paternal allele include *Igf2* (DeChiara *et al.*, 1991), *Snrpn* (Leff *et al.*, 1992), *U2af1-rs1* (Hatada *et al.*, 1993), *Mas* (Villar and Pedersen, 1994), *PAR1/PAR5* (Sutcliffe *et al.*, 1994), *ins1/ins2* (Giddings *et al.*, 1994), *IPW* (Wevrick *et al.*, 1994), *Peg1/Mest* (Kaneko-Ishino *et al.*, 1995), *ZNF127* (Jong *et al.*, 1994), $Cdc25^{Mm}$ (Plass *et al.*, 1996) and *Peg3* (Kuroiwa *et al.*, 1996). The

[1] "Maternal" and "paternal" allele refers to the allele derived from the mother and father, respectively.

existence of many other imprinted genes has been inferred from genetic criteria (Beechey and Cattanach, 1996), and the total number is estimated between 100 and 200 (Hayashizaki *et al.*, 1994; Solter, 1988).

Imprinted gene expression is deregulated
in parthenogenotes and androgenotes

Several analyses have been made of imprinted gene expression in uniparental embryos and their derivatives. If interactions between maternal and paternal alleles of a given imprinted gene are not necessary to establish normal imprinted expression of that gene in wild-type embryos, then the expression of this gene in androgenotes and parthenogenotes should correlate with the respective uniparental derivation of the embryos. For example, the maternally-expressed *H19* should be biallelically expressed in parthenogenotes and mice with maternal disomy of distal chromosome 7 (the chromosomal location of *H19*), while this gene should be transcriptionally inactive in androgenotes and mice with paternal disomy of distal chromosome 7. Several laboratories have reported analyses of post-implantation stage embryos, in which they examined the transcriptional levels of the genes *H19* (Ferguson-Smith *et al.*, 1993), *Igf2* (Ferguson-Smith *et al.*, 1991; Sasaki *et al.*, 1992), and *Snrpn* (Cattanach *et al.*, 1992) in maternally-disomic mice, and of *H19* and *Igf2* in paternally-disomic mice (McLaughlin *et al.*, 1996). In all of these cases the expression levels of the genes correlated with the parental origin of the alleles in question, suggesting that the sole determinant of imprinted gene expression is a parent of origin-specific modification made during gametogenesis, such as DNA methylation. Studies of post-implantation androgenotes showed that the maternally-expressed genes *H19* (Sasaki *et al.*, 1995; Walsh *et al.*, 1994) and *Igf2r* (Walsh *et al.*, 1994)

were not expressed, while in parthenogenotes the paternally-expressed gene *Igf2* was not expressed (Spindle *et al.*, 1996; Walsh *et al.*, 1994).

All of these studies suffer from their inability to distinguish the allelic origin of expression, and thus any attribution of overexpression to biallelic expression is a potentially fallible conclusion. Recently, a quantitative, sensitive assay of allelic expression levels in androgenotes and parthenogenotes revealed a dosage-dependent correspondence between parental origin of an allele of *Snrpn* and its expression level, regardless of ploidy of the embryo (Szabó and Mann, 1996). This careful analysis lends strong support to the hypothesis that it is solely the deregulated, biallelic expression of imprinted genes which causes the developmental defects of uniparentally-derived embryos and mice and humans with UPD of specific chromosomes.

Imprinted genes show allele-specific methylation

DNA methylation has long been proposed to constitute the molecular mark which distinguishes the two alleles of imprinted genes (Mann and Lovell-Badge, 1984). Evidence for DNA methylation constituting the imprinting mark includes differential methylation of imprinted transgenes (Chaillet *et al.*, 1991; Reik *et al.*, 1987; Sapienza *et al.*, 1987; Swain *et al.*, 1987) and the imprinted endogenous genes *H19* (Bartolomei *et al.*, 1993; Tremblay *et al.*, 1995), *Igf2r* (Stöger *et al.*, 1993), *Xist* (Ariel *et al.*, 1995; Zuccotti and Monk, 1995), *Igf2* (Sasaki *et al.*, 1992), $p57^{KIP2}$ (Hatada and Mukai, 1995), *Snrpn* (Glenn *et al.*, 1996), and *U2af1-rs1* (Hatada *et al.*, 1995). In fact, the imprinted genes *U2af1-rs1* (Hatada *et al.*, 1993) and $Cdc25^{Mm}$ (Plass *et al.*, 1996) have been cloned based upon the methylation differences between the two alleles in adult mouse DNA. Rigorous analyses now need to be performed upon

imprinted genes in parthenogenotes and androgenotes to see if biallelic methylation of these genes is responsible for directing their biallelic expression, as experiments described below would predict.

Establishment and maintenance of genomic imprints

The acquisition and developmental propagation of the methylation mark has been described in detail for the three imprinted genes *H19*, *Igf2*, and *Igf2r*, and for *Xist*, a partially-imprinted gene encoded at the center for X chromosome inactivation. *H19*, *Igf2r*, and *Xist* all undergo sex-specific *de novo* methylation during gametogenesis which is believed to mark the alleles so that imprinted expression can be established after fertilization. The area of an imprinted gene which receives this "primary imprint" is subsequently referred to as an "imprinting box" (Barlow, 1995). Importantly, the allele-specific methylation differences in these imprinting boxes are resistant to the wave of global demethylation during cleavage and the subsequent wave of global *de novo* methylation before gastrulation (Bartolomei *et al.*, 1993; Brandeis *et al.*, 1993; Stöger *et al.*, 1993; Tremblay *et al.*, 1995). This assures that the methylation mark imposed on the two alleles of imprinted genes during gametogenesis remains unaltered during development and, therefore, distinguishes the two alleles in each cell of the organism.

H19: *H19* is expressed from the maternal allele, and it acquires paternal-specific methylation in a region upstream of its promoter during spermatogenesis (Tremblay *et al.*, 1995). The cognate region in the imprinted human homologue of *H19* also acquires paternal-specific methylation patterns, despite a lack of any sequence homology between these two regions (Jinno *et al.*, 1996). The sequences in both species display a tandem array of

repeated sequences, though these sequences are not homologous to one another. The two regions remain completely unmethylated in oocytes, so that the zygote already displays a parent-specific difference in methylation status at the two alleles. During post-implantation development, the paternal allele of *H19* acquires extensive methylation throughout its coding sequence, and this is termed the secondary imprint (Bartolomei *et al.*, 1993; Brandeis *et al.*, 1993; Ferguson-Smith *et al.*, 1993).

The gene function of *H19* remains mysterious. The gene is highly conserved between mouse and human, and it lacks an open reading frame (Brannan *et al.*, 1990). Gene knockout studies have shown that *H19* controls in *cis* the expression of two paternally-expressed imprinted genes that are located upstream, *Igf2* and *Ins-2* (Leighton *et al.*, 1995a). In mice lacking the *H19* gene, a maternally-transmitted deletion results in the aberrant transcription of *Ins-2* and the aberrant methylation and transcription of *Igf2* (both from the maternal allele). Another disruption experiment has shown that the transcription of all three genes is dependent on the presence of two enhancer elements located downstream of the three genes (Leighton *et al.*, 1995b). Transgenic experiments examining *H19* imprinting at ectopic loci have identified the imprinting box and the first exon as being crucial for establishment of transgene imprinting (Elson and Bartolomei, 1997; Pfeifer *et al.*, 1996). These experiments indicate that imprinting of *H19* is established and controlled by the imprinting box. When the imprinting box is methylated, as on the paternal allele, *H19* expression is shut off and the neighboring imprinted genes *Ins-2* and *Igf2* are able to utilize the two downstream enhancers for transcription. When *H19* is unmethylated, it "competes" successfully for these enhancers, leaving *Igf2* and *Ins-2* inactive.

Igf2: As mentioned above, the expression of *Igf2* is controlled by *H19* expression. *Igf2* does not possess a region which becomes differentially methylated in the gametes (Brandeis *et al.*, 1993). However, a 1.5-kb region upstream of the first *Igf2* promoter has been shown to display a secondary imprint acquired during embryogenesis (Brandeis *et al.*, 1993; Feil *et al.*, 1994; Sasaki *et al.*, 1992), in which the paternal allele is on average more methylated.

Igf2r: The maternally-expressed gene *Igf2r* (Stöger *et al.*, 1993) acquires a maternal-specific methylation mark (the primary imprint) during oogenesis. As with *H19*, *Igf2r* acquires an additional secondary imprint during pre-implantation development, which is on the paternal allele and in the promoter region (Brandeis *et al.*, 1993; Stöger *et al.*, 1993). Curiously, the human homologue of *Igf2r* is not imprinted, despite the fact that it displays a conservation of the methylation patterns seen in the mouse imprinting box (Smrzka *et al.*, 1995).

Xist: This gene shows imprinted expression only in preimplantation embryos and extraembryonic tissues (Kay *et al.*, 1993). Like *Igf2r*, *Xist* acquires a maternal-specific primary imprint during oogenesis(Ariel *et al.*, 1995; Zuccotti and Monk, 1995). In contrast to *H19* and *Igf2r*, the methylation imprint of the maternal *Xist* allele is erased after implantation in cells of the embryonic lineage, and either the paternal or the maternal allele becomes methylated at the time of X inactivation (Norris *et al.*, 1994).

Other phenomena associated with genomic imprinting

Several other observations of imprinted genes deserve mentioning, for they may explain either the mechanism of monoallelic expression or the evolutionary origin of genomic imprinting in mammals.

a) Fluorescent *in situ* hybridization analysis has suggested asynchronous replication times of the two alleles of many imprinted genes (Kitsberg *et al.*, 1993; Knoll *et al.*, 1994), and this asynchrony may depend upon imprinted transcriptional activity within a given cell type (Kawame *et al.*, 1995).

b) The observation of repeated sequences at many imprinted genes, which resemble the long terminal repeats of retroviral-like sequences, has led to the hypothesis that the molecular means of imprinting may have arisen from the use of DNA methylation as a host defense response (Neumann *et al.*, 1995). This argument was supported by the fact that insertions of a defective retroviral intracisternal particle at the agouti locus leads to partial genomic imprinting of that gene's expression (Duhl *et al.*, 1994; Michaud *et al.*, 1994).

c) Imprinted genes are also claimed to possess fewer and smaller introns than normal (Hurst *et al.*, 1996). The two previous assertions (b and c) were challenged on the basis that exceptions to the rule exist and that the number of genes available to analyze is too small (Haig, 1996).

d) Imprinted genes are found to cluster together on various chromosomes. The imprinted genes *H19*, *Igf2*, *Ins2*, *Mash2* and $p57^{KIP2}$ all lie on distal chromosome 7 in the mouse (Giddings *et al.*, 1994; Guillemot *et al.*, 1995; Hatada and Mukai, 1995; Zemel *et al.*, 1992) and *H19* seems to control the imprinted gene expression of at least *Igf2* and *Ins2* (Leighton *et al.*, 1995a, b). A 400-kb segment of human chromosome 15q, deletions in which cause both Prader-Willi and Angelman syndrome, contains the five imprinted genes *SNRPN*, *ZNF127*, *PAR1*, *PAR5*, and *IPW* (Lalande, 1996).

e) Finally, imprinted genes bear some other curious features, such as sex-specific differences in recombination rates (Pàldi *et al.*, 1995) and a late-S-

phase physical association between the imprinted domains of homologous chromosomes (LaSalle and Lalande, 1996). Any of these traits may reflect the mechanism by which monoallelic expression is established or maintained, but their relevance to the establishment of imprinting patterns is entirely unclear, and will not be discussed further.

III. Genomic Imprinting and DNA Methyltransferase

DNA methylation controls both the initiation and maintenance of imprinted gene expression

Strong evidence for methylation being crucial for the maintenance of monoallelic expression of imprinted genes was provided by the analysis of *Dnmt* mutant embryos (Beard *et al.*, 1995; Li *et al.*, 1993). Hypomethylation of the DNA in mutant embryos or ES cells resulted in loss of monoallelic expression, either in the activation of the inactive *H19* and *Xist* alleles or in the inactivation of the active *Igf2r* and *Igf2* alleles. These experiments did not, however, provide insights into the role of methylation in the establishment of genomic imprinting during gametogenesis.

In the second chapter I address the process by which the allele-specific methylation of imprinted genes is acquired (Tucker *et al.*, 1996a). I expressed the wild-type *Dnmt* cDNA in homozygous *Dnmt* mutant ES cells to assess whether this would restore the low methylation levels of the genomic DNA to normal. While normal methylation levels of non-imprinted DNA were reestablished, imprinted genes were neither remethylated nor appropriately expressed. However, passage of the remethylated ES cells through the germ line restored normal methylation and expression to all imprinted genes

analyzed. These results support the primary role that methylation plays in the control of imprinted gene expression. The data are consistent with the presence of distinct *de novo* DNA methyltransferase activities during oogenesis and spermatogenesis which specifically recognize imprinted genes but which are absent in the post-implantation embryo and in ES cells.

Chapter One

Much evidence indicates that in vertebrates the methylation of DNA at cytosine residues affects gene transcription (Yeivin and Razin, 1993). Methylation is believed to be important in a variety of biological processes, including embryonic development, genomic imprinting, and cancer (Barlow, 1995; Jones *et al.*, 1992). A single gene (*Dnmt*) encoding a DNA (cytosine-5)-methyltransferase (MTase) has been cloned from mammalian cells (Bestor *et al.*, 1988). MTase shows a marked preference for hemimethylated DNA (Gruenbaum *et al.*, 1982), suggesting that its function is to maintain the methylation status of newly replicated DNA. The enzyme consists of an amino-terminal regulatory domain followed by a 500-amino acid domain assumed to contain the catalytic center, due to its homology to cytosine-methylating bacterial type II restriction methyltransferases (Leonhardt and Bestor, 1993).

The crucial importance of methylation in development, genomic imprinting, and cancer was demonstrated by the targeted mutation of the *Dnmt* gene in embryonic stem (ES) cells (Li *et al.*, 1992). Mice homozygous for the mutation die at midgestation; homozygous mutant ES cells proliferate normally with their genomic DNA highly demethylated but die upon differentiation (Lei *et al.*, 1996). Imprinted genes display parent-specific monoallelic expression, and DNA methylation has been proposed to constitute the molecular mark that distinguishes the two alleles (Barlow, 1995). Strong support for this notion was provided by the loss of monoallelic expression of these genes in homozygous *Dnmt* mutant embryos or ES cells (Li *et al.*, 1993). Furthermore, a large body of evidence links hypo- and hypermethylation of genomic DNA to cancer progression (Laird and Jaenisch, 1994). A direct correlation between DNA methylation and intestinal

neoplasia was demonstrated in mice expressing different levels of MTase (Laird *et al.*, 1995).

The work presented in this chapter (see also Appendix A) used the complementation of the *Dnmt* mutation as an assay to define the sequences encoding a functional MTase. Previous reports of high expression of functional MTase in transient transfections (Czank *et al.*, 1991; Glickman and Reich, 1994) contrasted with the inability to achieve high expression levels in stable transfection experiments. In fact, only one report described high levels of MTase expression in stably transfected clones which, however, displayed a transformed phenotype (Wu *et al.*, 1993). I found that the published cDNA was incapable of expressing wild type levels of MTase and complementing the *Dnmt* mutation. A construct containing a previously reported (Rouleau *et al.*, 1992) *Dnmt* promoter fused to the *Dnmt* cDNA also failed in this assay. In contrast, wild type levels of *Dnmt* expression were obtained after targeting the published cDNA to its cognate locus, where expression of the full-length cDNA from its cognate promoter fully restored MTase function to ES cells (Tucker *et al.*, 1996). This result suggested that sequences upstream of the previously identified initiation codon were crucial for stable and functional expression of the enzyme. In this chapter I report a previously unidentified amino-terminal open reading frame which extends the coding capacity of the construct by up to 171 amino acids. Expression of the full-length *Dnmt* cDNA in mutant ES cells restored both normal methylation levels and the capacity to form teratomas, indicating a crucial role of the amino-terminal domain in normal enzymatic activity.

To establish stable wild type and *Dnmt* mutant cell lines expressing recombinant MTase, four different expression constructs carrying the *Dnmt* cDNA were created. The ubiquitously-expressed *Pgk-1* promoter was used in

pMT10 and the human EF-1α promoter in pMT40 to direct expression of the *Dnmt* cDNA. A synthetic intron was inserted into pMT10 to increase gene expression levels, creating pMT20. pMT30 was derived from pMT10 by replacing the *Pgk-1* promoter and the first 180 bp of the *Dnmt* cDNA with a 4.7-kb *Dnmt* genomic fragment containing a previously reported *Dnmt* promoter (Rouleau *et al.*, 1992) and the excised 180 bp of cDNA, restored in its genomic context.

The four constructs were separately co-transfected into wild type *Dnmt* J1 ES cells (Li *et al.*, 1992) and the homozygous mutant $Dnmt^{s/s}$ *(Lei et al., 1996)* and $Dnmt^{n/n}$ cell lines (Li *et al.*, 1992). Only low levels of the *Dnmt* cDNA were expressed in wild type or mutant ES cells stably transfected with any of these constructs. To test the hypothesis that the low mRNA expression was caused by the protein product of the *Dnmt* expression constructs being detrimental to cells, a construct was designed such that translation of the *Dnmt* cDNA would end prematurely, producing truncated and possibly inactive peptides. In contrast to a typical clone carrying the wild type cDNA, high levels of RNA were produced in 4 / 8 clones carrying the mutant cDNA construct. These results are consistent with a detrimental effect of the protein produced from the original expression constructs resulting in selection against cells expressing high levels of the truncated *Dnmt* cDNA.

Both the low expression levels in cells carrying the truncated cDNA in stably transfected ES cells and the restoration of high levels of expression from constructs containing an in-frame stop codon in the ORF, were consistent with a truncated polypeptide being produced from the cDNA, translation of which may be detrimental to ES cells. By 5' RACE and screening of a *Dnmt*-specific cDNA library, I isolated additional upstream cDNA sequence comprising two exons separated by an 11.5-kb intron.

Examination of the genomic sequence surrounding this new primary exon revealed an uninterrupted open reading frame that begins upstream of and continues throughout the two new exons, extending the published coding sequence by 171 codons. Three new in-frame ATG codons were found in this ORF.

To test the possibility that the newly-discovered cDNA sequence would allow normal expression and MTase functioning *in vivo*, a mini-gene expression construct was made and stably transfected into *Dnmt* mutant ES cells. pMT50 was made by insertion of a 5.0-kb region of the *Dnmt* locus that included the first exon into pMT30. pMT30 contained the previously reported *Dnmt* promoter (Rouleau *et al.*, 1992) but had been shown not to function in *Dnmt* mutant ES cells. pMT50 resulted in the *Dnmt* cDNA fused 3' to a 9.7-kb piece of genomic *Dnmt* sequence that ended in the fourth exon and included 3.6 kb of genomic sequence upstream of the first exon, the two new exons reported above, three introns, and the beginning of the fourth exon.

Dnmt$^{s/s}$ cells were electroporated with the linearized construct. 7/12 clones examined showed random integration into the genome and normal levels of genomic methylation, as compared to wild type cells. All clones carrying an intact transgene displayed equivalent methylation profiles at repetitive gene sequences and *Xist*. The imprinted genes *H19* and *Igf2r*, however, were not remethylated, as reported previously for a smaller, similar *Dnmt* cDNA-containing plasmid that was targeted to the endogenous *Dnmt* locus (Tucker *et al.*, 1996). Western blot analysis revealed a range of MTase levels from 10-100% that of a wild type allele, with the minigene-encoded enzyme co-migrating with wild type MTase.

Three of the rescued lines were used to induce teratomas by subcutaneous injection into syngeneic male host animals. While the parental mutant cells failed to form palpable teratomas within 3-4 weeks (n = 0 / 12), the rescued ES cells induced teratoma formation at a similar rate (n = 9 / 10) as the wild type J1 ES cells (n = 6 / 6). Histological analysis of these teratomas revealed a wide variety of differentiated cell types, consisting mostly of mature neural tissue and including discrete patches of epithelial tissue, cartilage, trophectoderm, osteoid, and a relative deficiency of striated muscle (data not shown). From these results we conclude that the additional sequences uncovered above are necessary for functional activity of MTase *in vivo*.

The first exon of *Dnmt* was found to lie within a CpG island, as defined by Bird (Bird, 1986). Our results are not in agreement with the localization of the *Dnmt* promoter localized upstream of exon 3, as reported by Rouleau *et al.* (1992). Expression from this promoter would produce a truncated MTase which is detrimental to stably transfected cells, and *Dnmt* expression constructs utilizing this promoter failed to effect complementation in *Dnmt* mutant cells. Instead we suggest that a promoter contained within the pMT50 construct represents the functionally relevant transcriptional start site.

Chapter Two

Imprinted genes display parent-specific monoallelic expression (Barlow, 1995; Surani, 1994), although often not at all developmental stages (Szabo and Mann, 1995). Genes which are predominantly expressed from the maternal allele include *Igf2r* (Barlow *et al.*, 1991) and *H19* (Bartolomei *et al.*, 1991), while *Igf2* (DeChiara *et al.*, 1991) is expressed paternally. DNA methylation has been proposed to constitute the molecular mark which distinguishes the two alleles of imprinted genes. Evidence for DNA methylation constituting the imprinting mark includes differential methylation of imprinted transgenes (Chaillet *et al.*, 1991; Reik *et al.*, 1987; Sapienza *et al.*, 1987; Swain *et al.*, 1987) and imprinted endogenous genes. The acquisition and developmental propagation of the methylation mark has been described in detail for the two imprinted genes *H19* and *Igf2r*. The *H19* gene acquires paternal-specific methylation in a region upstream of its promoter during spermatogenesis (Tremblay *et al.*, 1995), while *Igf2r* (Stöger *et al.*, 1993) acquires a maternal-specific methylation mark during oogenesis. Importantly, the allele-specific methylation differences in these "imprinting boxes" are resistant to the wave of global demethylation during cleavage and the subsequent wave of global *de novo* methylation before gastrulation (Bartolomei *et al.*, 1993; Brandeis *et al.*, 1993; Stöger *et al.*, 1993; Tremblay *et al.*, 1995). This assures that the methylation mark imposed on the two alleles of imprinted genes during gametogenesis remains unaltered during development and, therefore, distinguishes the two alleles in each cell of the organism.

Strong evidence for methylation being crucial for the maintenance of monoallelic expression of imprinted genes was provided by the analysis of *Dnmt* mutant embryos (Li *et al.*, 1993; Beard *et al.*, 1995). Hypomethylation of the DNA in mutant embryos or ES cells resulted in loss of monoallelic

expression, either in the activation of the inactive *H19* allele or in the inactivation of the active *Igf2r* and *Igf2* alleles. These experiments did not, however, provide insights into the role of methylation in the establishment of genomic imprinting during gametogenesis.

In this chapter I address the process by which the allele-specific methylation of imprinted genes is acquired. I have expressed the wild type *Dnmt* cDNA in homozygous *Dnmt* mutant ES cells to assess whether this would restore the low methylation levels of the genomic DNA to normal. To assure correct quantitative and developmental expression of the construct, the cDNA was inserted by homologous recombination into the cognate *Dnmt* locus upstream of the mutation. Homologous insertion of this hybrid vector was expected to leave all upstream control elements of the gene intact, leading to restoration of normally-regulated MTase expression. By Southern blot analysis, 73 / 182 puromycin-resistant clones had undergone homologous recombination events in which a single copy of the plasmid integrated at a single allele. Northern blot analysis showed that the inserted cDNA was transcribed at a level and size comparable to that of the endogenous allele. All 73 homologous recombinants examined showed near-complete genomic remethylation, as assayed using two repetitive sequence probes: a Moloney Murine Leukemia Virus (MoMuLV) cDNA and a minor satellite centromeric repeat.

I tested whether restoration of genomic DNA methylation levels in mutant ES cells carrying the *Dnmt* cDNA would rescue the lethal mutant phenotype, as exhibited by a failure to differentiate normally *in vitro*, to form teratomas, and to donate to late-gestation and adult chimeras made with wild-type hosts. Rescued ES cells were capable of differentiating normally in all three assays, in contrast to the parental *Dnmt* mutant cells.

The restoration of wild-type methylation levels and consequent transcriptional inactivation was observed for the single-copy gene *Xist*. The partially methylated pattern seen in normal male ES cells (Beard *et al.*, 1995) was restored to the same degree in the rescued ES cells. The remethylation of *Xist*, coupled with the functional complementation assays, suggests that the methylation of all or most other non-imprinted genes is fully restored.

The imprinted genes *H19* and *Igf2r* were not remethylated after introduction of the *Dnmt* cDNA into the mutant ES cells. The regions examined for methylation show a gamete-specific methylation difference and are believed to constitute the primary imprinting mark (Barlow, 1995). The methylation of these regions and expression status of these genes in rescued ES cells was identical to that seen previously in *Dnmt* mutant ES cells and embryos (Li *et al.*, 1993). Furthermore, hypomethylation persisted in ES cell-derived fibroblasts isolated from chimeras, and methylation was restored to normal only after germline transmission. Normal expression patterns of all imprinted genes was observed only after passage through the germ line.

Our results show conclusively that passage through the male germ line is necessary to obtain the allele-specific methylation mark on at least one imprinted gene, *H19*. This suggests the existence of a male-specific *de novo* methyltransferase activity which may specifically recognize all paternally-modified imprinted genes. Presumably a female-specific *de novo* methyltransferase activity exists, which would be active during oogenesis and would be expected to recognize maternally-imprinted genes such as *Igf2r*. We have no evidence whether the gamete-specific activities are encoded by separate genes, are due to modification of the maintenance MTase or the separately-encoded *de novo* methyltransferase activity (Lei *et al.*, 1996), or are reflective of a change in chromatin structure affecting substrate accessibility.

Conclusion

In order to establish a model of ectopic DNA cytosine methyltransferase (MTase) expression, I have attempted to express the *Dnmt* gene as a transgene in embryonic stem (ES) cells homozygous for a *Dnmt* gene disruption. Chapter One reports the cloning of the 5' end of the *Dnmt* mRNA and subsequent expression of the full-length *Dnmt* cDNA in *Dnmt* mutant ES cells with a construct employing the endogenous *Dnmt* promoter to drive transcription. Only when the full-length cDNA was employed could the *Dnmt* gene product rescue the hypomethylation phenotype of *Dnmt* mutant cells, causing a rescue of the inability of the mutant cells to differentiate. Chapter Two details the functional consequence of MTase expression in the background of the completely hypomethylated *Dnmt* mutant genome. While normal methylation levels of non-imprinted DNA were reestablished, imprinted genes were neither remethylated nor appropriately expressed. However, passage of the "rescued" ES cells through the germ line restored normal methylation and expression to all imprinted genes analyzed. The data are consistent with the presence of distinct *de novo* DNA methyltransferase activities during oogenesis and spermatogenesis which specifically recognize imprinted genes but which are absent in the post-implantation embryo and in ES cells.

As the main findings of these two projects are discussed at length in their respective chapters, the remainder of this conclusion will be spent in discussing the experimental utility of rescued ES cell lines for investigating the evolution of genomic imprinting in mammals. A very attractive model to explain the origin of this phenomenon proposes it to be a nonessential, evolutionarily-recent addition to the basic mammalian developmental program. Rescued ES cells lacking methylation at imprinted loci are hypothesized to reflect this "pre-imprinted" genome, and may consequently

have a normal developmental potential. Rescued cells can best be compared with that of cells derived from androgenetic and parthenogenetic embryos. Like the rescued ES cells, these uniparentally-derived embryonic cells show deregulated, biallelic expression of imprinted genes. Chimeras made between wild-type embryos and uniparental embryonic cells are generally inviable, showing highly skewed and complementary distribution of the uniparental cells to different tissues of the chimera. If genomic imprinting is nonessential for development, however, the tissue-specific distribution of rescued ES cells in chimeras should be as widespread as that of wild-type ES cells. Published research which is relevant to this question shall be discussed first, followed by an outline of current theories explaining the evolution of genomic imprinting and the experiments (currently underway) which will address these theories.

Tissue-specific distribution of parthenogenetically- and androgenetically-derived cells in chimeras made with wild-type hosts

The ability of cells derived from parthenogenetic (PG) and androgenetic (AG) embryos to contribute to differentiated tissues in the embryonic and adult mouse has been analyzed by the production of chimeras using wild-type embryos as hosts (reviewed in Fundele and Surani, 1994). Because of their ease of production (Cuthbertson, 1983), aggregation chimeras made with PG cells were the most extensively analyzed, though later analyses with PG ES cells confirmed these results (Allen *et al.*, 1994). PG cells were found in both the inner cell mass and trophectoderm of the blastocyst (Clarke *et al.*, 1988; Thomson and Solter, 1989). PG cells were sequentially eliminated, first from the extraembryonic trophoblast, then the yolk sac endoderm, and finally the yolk sac mesoderm, to the extent that they were completely absent from the

extraembryonic tissues at mid-gestation (Clarke *et al.*, 1988; Nagy *et al.*, 1987; Surani *et al.*, 1988; Thomson and Solter, 1988). From midgestation, the number of PG cells in the fetus decreased steadily (Bender *et al.*, 1995; Fundele *et al.*, 1990; Nagy *et al.*, 1987). In the adult chimera, PG cells contributed most extensively to tissues of ectodermal origin (e.g. brain, epidermis, dorsal root ganglia) (Fundele *et al.*, 1989; Fundele *et al.*, 1990; Nagy *et al.*, 1989) and to the ovary, including oocytes capable of supporting normal development upon fertilization (Stevens, 1978). Relative deficiencies were seen in all other organs, especially the liver and pancreas (Fundele *et al.*, 1989; Fundele *et al.*, 1990; Nagy *et al.*, 1989).

Most strikingly, PG cells were completely excluded from skeletal muscle, though not from cardiac or smooth muscle (Fundele *et al.*, 1989; Fundele *et al.*, 1990; Nagy *et al.*, 1989). More extensive analysis of this phenotype revealed that by 13 dpc, any PG cells found in chimeric muscle blastemas were excluded from the myogenic lineage (Bender *et al.*, 1995). As PG cells contribute well to the myotome of 11 dpc chimeras (Fundele *et al.*, 1994), and as occasional PG cell-derived myotubes were observed (Bender *et al.*, 1995), it is unclear whether the defect lies in an impaired ability of PG cells to form myoblasts or for those myoblasts to fuse.

Reflecting the differences in developmental potential of parthenogenotes and androgenotes, AG cells showed a distribution in chimeras made with wild-type hosts which was strikingly complementary to PG cells. Initial studies indicated that AG cells were allocated mostly to the extraembryonic tissues of pre-implantation embryos (Surani *et al.*, 1988; Thomson and Solter, 1988). Interestingly, PG cell <-> AG cell chimeras displayed this complementary lineage segregation without rescuing embryonic lethality (Surani *et al.*, 1987). Donation of AG cells to the embryo

proper was not observed until ES cells derived from androgenotes were used to make chimeras (Mann *et al.*, 1990). AG cells could populate all cell types examined, sometimes to a high degree of chimerism (Barton *et al.*, 1991; Fundele *et al.*, 1995a). A disproportionate representation of AG cells was found in mesodermally-derived tissues, especially in the heart and skeletal muscle, with low contribution in the brain and gonad of either sex. Whether AG cells can pass through the germline is of great interest, but the neonatal fatality of all AG chimeras produced to date prevents this from being known. Most strikingly, severe abnormalities of skeletal elements, particularly of the rib cage cartilage, were noticed, which may explain the high incidence of early postnatal lethality (Barton *et al.*, 1991; Mann *et al.*, 1990; Mann and Stewart, 1991). Only chimeras with a low percentage of AG cells survived this early postnatal period. In contrast to PG cells, AG cells displayed an increased proliferation rate in all examined tissues of 13 dpc chimeras, but this difference diminished later in embryogenesis and was not reflected in an increased representation in the relevant tissues (Fundele *et al.*, 1995b).

A brain-specific difference in the relative distributions of AG and PG cells is of special interest. Reflecting trends already established by 10.5 dpc, PG cells in adult chimeras were completely excluded from the hypothalamus and enriched in the frontal cortex, striatum and hippocampus (Allen *et al.*, 1995). Just the opposite distribution was seen for AG cells in the brain (Keverne *et al.*, 1996). The authors also claimed that chimeric males made with PG cells were more aggressive than wild-type counterparts (Allen *et al.*, 1995).

Many authors have noted that PG cells in a chimera caused a decrease in body size that correlated with the extent of PG cell contribution (Fundele *et al.*, 1989; Fundele *et al.*, 1990; Pàldi *et al.*, 1989; Surani *et al.*, 1977). Whether this size difference can be explained by a decrease in proliferation potential is

controversial. Proliferation of PG cells, compared to wild-type cells in the same chimera, was seen to be only slightly impaired in a few tissues in pre- (Bender *et al.*, 1995) and post-natal chimeras (Jägerbauer *et al.*, 1992). As well, overall contribution of PG cells to chimeras was not seen to decrease with age postnatally (Nagy *et al.*, 1989). In contrast, AG cell-bearing chimeras were both larger and longer than normal litter mates (Barton *et al.*, 1991). Some evidence exists that AG cells do proliferate better than wild-type cells in the same chimera (Fundele *et al.*, 1995b). These findings are suggestive in that they may help explain the relative abilities of parthenogenotes and androgenotes to develop as embryos. They also suggest a fundamental role of genomic imprinting in influencing rates of growth during development, though the importance of this influence on growth rates to general development is under much debate (see below).

The expression and methylation patterns of imprinted genes are deregulated in PG and AG cells

Imprinted gene expression was examined in differentiated cultures of normal and UPD-derived ES cells, and results roughly paralleled that seen in embryos. In observations of later passage embryoid bodies (Allen *et al.*, 1994; Szabó and Mann, 1994), *Igf2* expression was low to undetectable in PG ES cells, while *Igf2r* expression wasn't observed in AG ES cells. *H19* was expressed in both cell types, however (Allen *et al.*, 1994; Szabó and Mann, 1994). Comparison of these results both within and between reports is difficult, however, because the various differentiation techniques delivered conflicting results. Regions known to be differentially methylated in *Igf2r*, *Igf2*, and *H19* were examined after *in vitro* differentiation (Feil *et al.*, 1994; Szabó and Mann, 1994). The methylation patterns were found to fall in a spectrum which

correlated with the parental derivation of the cells (e.g. paternally-methylated genes showed increased levels of methylation in AG ES cells, compared to wild-type and PG ES cells). The examined regions for the most part were ones which receive secondary imprints, and those imprinting boxes which were examined did not show a bimodal methylation pattern (i.e. biallelically methylated or unmethylated). Interestingly, co-expression of *Igf2* and *H19* was found in ES cells of all three types, both early passage embryoid bodies (Szabó and Mann, 1994) and teratomas (McKarney *et al.*, 1996). Once again, these results are hampered by an inability to distinguish methylation and expression levels of individual alleles. The results of these experiments indicate that in uniparentally-derived ES cells imprinting operates primarily as predicted from the uniparental origin of the cells, but other factors such as passage number, clonal variation and possible loss of imprinting patterns during ES cell derivation can influence methylation patterns and expression significantly.

In conclusion, the examination of specific imprinted genes has led to the hypothesis that aberrant expression levels of these genes cause the multiple developmental abnormalities seen in the embryonic cells derived from parthenogenotes and androgenotes.. The growth-promoting and -suppressing functions of many of these genes makes them attractive candidates to explain the growth-related abnormalities recorded in both mice and human displaying UPD for specific chromosomal segments. It must be kept in mind, however, that most imprinted genes discovered so far are clustered together in small imprinted gene domains. Thus, the expression of all imprinted genes on a given chromosomal segment is presumably deregulated in cases of UPD, so that it is very difficult to ascribe a phenotype to the over- or under-expression of any given gene. The precise function of

each imprinted gene in development will be best evaluated by gene disruption experiments.

Evolutionary origins of genomic imprinting

The evolutionary origin of genomic imprinting has caused much speculation. A simple explanation (Solter, 1988) views the necessity of equal paternal and maternal contribution to the zygote for normal development as a mechanism ensuring that parthenogenesis will not occur. The dangers of parthenogenesis include the uncovering of harmful recessive alleles and the production of dangerously invasive trophoblast (Hall, 1990). Along the same lines, it has been proposed that the monoallelic expression of a certain number of essential genes could act as a guard against recombination events that are associated with neoplastic transformation (Thomas, 1995). Loss of expression of an imprinted gene product, as a result of an aberrant recombination event, could either be sensed indirectly or lead directly to growth arrest. This argument would logically necessitate imprinted genes placed randomly through the genome, which is clearly not the case. Also, none of the three reported gene disruptions of imprinted genes is embryonic lethal (DeChiara *et al.*, 1990; Lau *et al.*, 1994; Leighton *et al.*, 1995a; Wang *et al.*, 1994). Finally, loss of imprinted gene expression is associated with many tumor types, including Wilms' tumor, rhabdomyosarcoma, neuroblastoma, and the Philadelphia-chromosome-positive leukemia (Reeve, 1995). In contrast to this "guardian of the genome" hypothesis, the deregulation of imprinted gene expression is at best a consequence of neoplastic transformation, and at worst a possible cause for tumor formation!

The Haig hypothesis: maternal and paternal genomes place conflicting demands upon fetal growth

A much more convincing theory to explain the evolutionary origin of genomic imprinting stems from the fact that the placental method of development used by mammals places an inordinate demand upon the female for resources expended in the nourishment of the growing fetus (Moore and Haig, 1991). In mammals such as mice, in which females have a high probability of multiple mates, it is to the advantage of females to conserve resources during a single pregnancy so that multiple litters can be generated in one lifetime. There are often multiple fathers within a single litter of mice, so embryos will compete with their siblings for scarce maternal resources. Males have an obvious advantage in maximizing the fitness of their offspring, in comparison to those of another male. This conflict in the interests of the maternal and paternal genomes could lead to alterations in the expression levels of genes that influence embryonic growth rates.

This theory predicts that paternally-inherited alleles would be up-regulated for growth-promoting genes and down-regulated for genes leading to growth suppression. This is precisely what one observes when the function of imprinted gene products are correlated with their expression. The paternally-expressed gene insulin-like growth factor 2 (*Igf2*) encodes a secreted growth factor (DeChiara *et al.*, 1990), while the maternally-expressed insulin-like growth factor 2 receptor (*Igf2r*) encodes a receptor which acts as a putative "sink" for the *Igf2* ligand, trapping the growth factor without transducing a signal (Haig and Graham, 1991). Three of the four reported imprinted-gene disruptions have shown an altered growth rate in the mutant embryos (DeChiara *et al.*, 1991; Lau *et al.*, 1994; Leighton *et al.*, 1995a; Wang *et al.*, 1994), and most of the products of the newly-discovered imprinted genes have

growth-related functions (Guillemot *et al.*, 1995; Hatada and Mukai, 1995; Plass *et al.*, 1996; Villar and Pedersen, 1994).

As pointed out by R. Jaenisch (1997), the specific genes which have imprinted gene expression would depend upon the reproductive behavior of the mammal in question, and could vary widely from species to species. Comparisons between mouse and humans reveals lack of conservation of imprinted gene expression, as one would expect for a function which is not essential to the development of an organism (Kalscheuer *et al.*, 1993; Pearsall *et al.*, 1996). It will be of great interest to examine the allelic expression of imprinted genes in other mammals, especially the monotremes and marsupials, in which the major contribution of maternal energy reserves to the fetus occurs before fertilization and after parturition, respectively (Graves, 1996).

Is genomic imprinting necessary for development?

According to the Haig model, imprinted gene expression would have evolved after a basic plan for vertebrate development had already been established. The origin of monoallelic expression of a pre-existing gene could occur, for example, through the use of a *de novo* DNA methylation event to inactivate this gene upon passage through the germline of one sex. If this gene were involved in growth enhancement, and were inactivated by methylation in the female germ line, compensatory paternal activation of a growth-enhancing gene may occur to counteract this initial silencing event. Precedent for the complementarity of oppositely-imprinted genes is seen on a gross scale in the reciprocal developmental phenotypes of parthenogenotes, which support extensive normal embryonic development but fail to form

extraembryonic tissue, and androgenotes, which support extensive proliferation of extraembryonic tissue but not of the embryo itself.

Complementarity can be seen on a gene-specific level by looking at the analysis of gene knockout experiments. Deletion of *Igf2* results in "proportionate dwarf" mice, which are 60% smaller than wild-type littermates (DeChiara *et al.*, 1991). In contrast, deletion of *Igf2r* results in mice which are 30% larger than normal, though they die at birth (Lau *et al.*, 1994; Wang *et al.*, 1994). Perinatal lethality of the *Igf2r* mutation is rescued in mice carrying both the *Igf2* and *Igf2r* mutations, which are normal in size (Wang *et al.*, 1994) and fertile (Ludwig *et al.*, 1996). Additionally, *H19*-deficient mice, which are 30% larger than their wild-type littermates, are normal-sized when crossed into the *Igf2* null background (Leighton *et al.*, 1995a). Though not every imprinted gene can be expected to have an oppositely-imprinted partner whose genetic interactions are directly antagonistic, as with *Igf2* and *Igf2r*, the sum effect of a loss of imprinted gene expression should allow normal development of the embryo.

The ES cells which have been generated by expression of the *Dnmt* cDNA in the background of *Dnmt* mutant cells provide an excellent reagent for addressing the importance of imprinted gene expression in embryonic development. As shown in Chapter Two, the imprinting boxes of both of the imprinted genes *H19* and *Igf2r* remained hypomethylated after expression of the *Dnmt* cDNA in the background of the *Dnmt* mutation. Both genes, as well as *Igf2*, assumed an abnormal, biallelic transcriptional profile as a consequence of this hypomethylation, as seen before in homozygous *Dnmt* mutant embryos (Li *et al.*, 1993).

According to the Haig hypothesis, genomic imprinting evolved in mammalian species after the basic developmental program of these species

was established. Making the crucial assumption that DNA methylation has been used during evolution to determine the monoallelic expression patterns of imprinted genes, the prediction is clear. The erasure of methylation at imprinted loci should restore their methylation status to one reflective of a time before the evolution of genomic imprinting, and this will result in a corresponding "pre-imprinting" transcriptional profile for every imprinted gene. Genes such as *Igf2* and *Igf2r* will be shut off, while genes such as *H19* will be expressed from both alleles. A mouse bearing such changes should show no developmental abnormalities, in distinct contrast to the developmentally-compromised parthenogenotes and androgenotes.

Unfortunately, this mouse is not available, but the rescued ES cells described in Chapter Two do have this genomic methylation profile. These ES cells have been used to make chimeras with wild-type embryos, and their tissue-specific distribution will be analyzed. As described above, PG and AG cells show restricted, complementary distribution in chimeras. Specifically, contribution to skeletal muscle is compromised in PG cells, while AG cells show extremely limited contribution to brain tissue. If normal development is restored after the erasure of genomic imprinting, rescued ES cells should be able to contribute to all of those tissues that AG and PG cells respectively cannot. Tissue-specific distribution of rescued ES cells in chimeras is being analyzed by two methods. Firstly, organs and tissues taken from adult chimeras are being analyzed by Southern blot to determine the fraction of contribution from rescued ES cells. Secondly, the CHIP protocol described in Chapter Two has been applied to embryonic germ (EG) cells which are homozygous for the $Dnmt^s$ mutation and also contain the ROSA 11 cell marker, a *lacZ* gene which is ubiquitously expressed in all tissues throughout development and early adulthood (Friedrich and Soriano, 1991). I have

generated rescued, *lacZ*-expressing EG cells which will be used to analyze contribution to fine structures such as the hypothalamus in the brain.

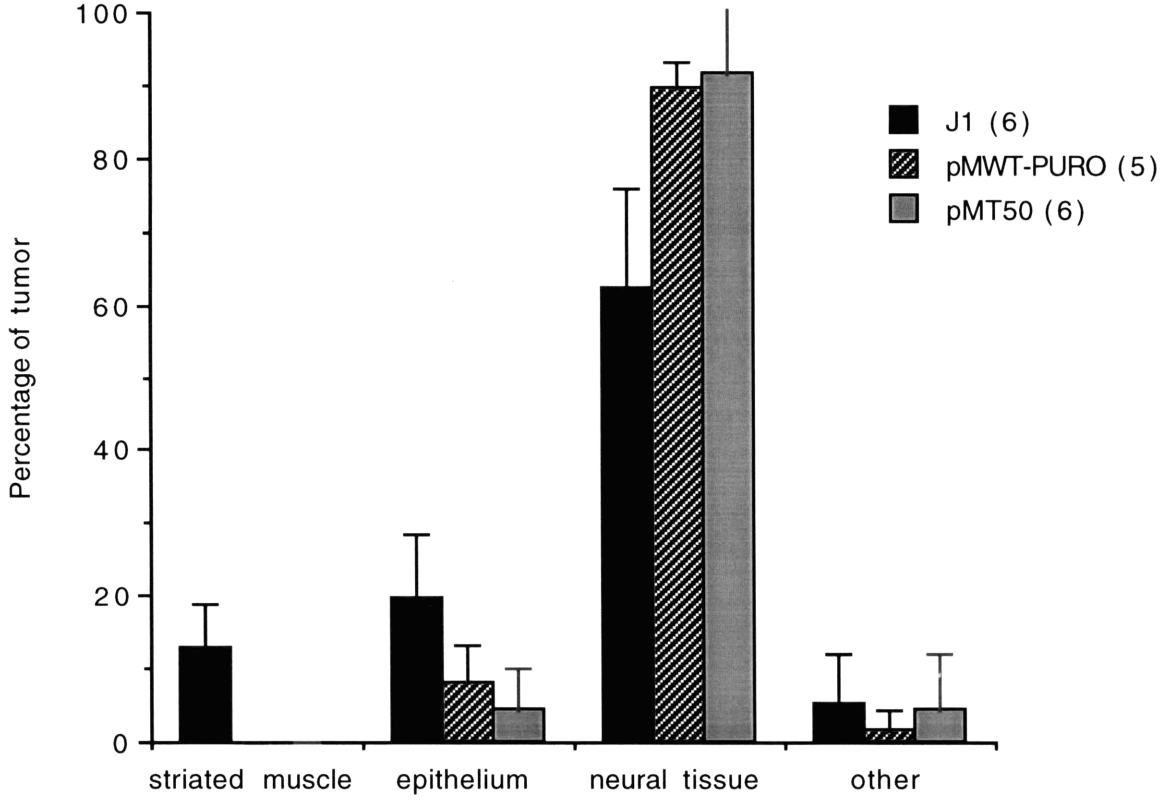
Preliminary results have been obtained from examination of teratomas formed by subcutaneous injection of rescued ES cells into syngeneic host animals. The variety of differentiated cell types in the resultant teratomas was tabulated and compared to that seen in wild-type ES cells[1]. No significant differences were observed, with the sole exception of a relative deficiency in striated muscle in teratomas formed by the rescued ES cells (Fig. 1). This resembles the phenotype shown by PG ES cells in teratomas (Allen *et al.*, 1994), in contrast to AG ES cells, which form exclusively striated muscle in teratomas (Mann *et al.*, 1990).

Mice which are genetically homogenous for a given mutation are normally made by having a mouse chimeric for ES cells containing the desired alteration to pass that mutation through the germline. However, this is not possible with the "epigenetic mutations" at imprinted loci observed in the rescued ES cells, because normal methylation patterns are restored to these genes upon passage through the germline. A technique has recently been developed which may allow the circumvention of this problem by the construction of a complete embryo from the rescued ES cells (Nagy *et al.*, 1990), allowing a definitive analysis of the developmental potential of these cells. In this technique, a wild-type 2 cell-stage embryo is electrofused to form a one-cell tetraploid embryo. This embryo develops *in vitro* to the morula stage, when it is aggregated with ES cells. The ability of tetraploid cells to donate to embryonic structures is extremely limited, while they have no difficulty forming extraembryonic structures. This developmental deficiency is complemented by the embryonic pluripotency of ES cells, which can go on

---

[1] Tabulation performed by H. Gardner

Fig. 1 - Differentiated cell types in teratomas made from wild-type vs. rescued ES cells



Teratomas were produced by injecting 3 x 10^6 ES cells subcutaneously into the flanks of adult male 129/Sv mice. Wild-type J1, *Dnmt*$^{s/s}$, and 3 clones each of rescued *Dnmt*$^{s/s}$ ES cells were used. *Dnmt*$^{s/s}$ cells failed to produce tumors (n = 20). pMWT-PURO refers to *Dnmt*$^{s/s}$ cells remethylated by the CHIP technique described in Chapter Two. pMT50 refers to *Dnmt*$^{s/s}$ cells remethylated by the *Dnmt* mini-gene vector described in Chapter One. The number of teratomas examined is indicated in parentheses. Tumors were excised at 2-3 weeks, fixed overnight in 4% paraformaldehyde, and prepared into haematoxylin- / eosin- stained paraffin sections. For each tumor, several 10 µm sections were examined by microscopy, and quantitation of cell-type distribution was made by noting the cell type found in consecutive ocular crosshair etchings, averaging 35 counts in 5 randomly-chosen fields over the length of the teratoma (0.5 - 1.5 cm in size). The category "other" includes bone, cartilage, endothelium, thyroid and fibrous tissue.

to form the complete embryo. The Haig hypothesis predicts that rescued ES cells should support embryonic development past the developmental block of parthenogenotes at the 25-somite stage, so that chimeras formed in this fashion can simply be examined for the appearance of standard morphological markers and for levels of imprinted gene expression.

In conclusion, the Haig hypothesis has emerged as the most sensible of the many theories put forth to explain the evolution of genomic imprinting. This theory argues that competing concerns between males and females for the viability of individual offspring in multi-father litters will inevitably lead to the sex-specific activation and repression of growth-related genes. The phenomenon of genomic imprinting can be seen as the result of a genetic "arms race" which was imposed recently upon a basic mammalian developmental program, and whose patterns of gene expression should vary from species to species. Although the deregulation of individual imprinted genes is of great consequence for both development and disease, the sum effect of the removal of all imprinted gene expression should be a cell or embryo with a normal developmental potential. I propose that I have created such a reagent with the construction of ES cell lines which lack methylation at imprinted loci, and various assays will be used to ascertain the developmental potential of ES cells lacking an imprinted genome.

References

Allen, N. A., Logan, K., Lally, G., Drage, D. J., Norris, M. L., and Keverne, E. B. 1995. Distribution of parthenogenetic cells in the mouse brain and their influence on brain development and behavior. *Proc Natl Acad Sci USA* **92:** 10782-10786.

Allen, N. D., Barton, S. C., Hilton, K., Norris, M. L., and Surani, M. A. 1994. A functional analysis of imprinting in parthenogenetic embryonic stem cells. *Development* **120:** 1473-1482.

Aniello, F., Locascio, A., Fucci, L., Geraci, G., and Branno, M. 1996. Isolation of cDNA clones encoding DNA methyltransferase of sea urchin *P. lividus*: expression during embryonic development. *Gene* **178:** 57-61.

Antequera, F., Boyes, J., and Bird, A. 1990. High levels of de novo methylation and altered chromatin structure at CpG islands in cell lines. *Cell* **62:** 503-14.

Ariel, M., Robinson, E., McCarrey, J. R., and Cedar, H. 1995. Gamete-specific methylation correlates with imprinting of the murine *Xist* gene. *Nat Genet* **9:** 312-5.

Barlow, D. P. 1995. Gametic Imprinting in Mammals. *Science* **270:** 1610-1613.

Barlow, D. P., Stoger, R., Herrmann, B. G., Saito, K., and Schweifer, N. 1991. The mouse insulin-like growth factor type-2 receptor is imprinted and closely linked to the *Tme* locus. *Nature* **349:** 84-7.

Bartolomei, M. S., Webber, A. L., Brunkow, M. E., and Tilghman, S. M. 1993. Epigenetic mechanisms underlying the imprinting of the mouse *H19* gene. *Genes Dev* **7:** 1663-1673.

Bartolomei, M. S., Zemel, S., and Tilghman, S. M. 1991. Parental imprinting of the mouse *H19* gene. *Nature* **351:** 153-5.

Barton, S. C., Ferguson-Smith, A. C., Fundele, R., and Surani, M. A. 1991. Influence of paternally imprinted genes on development. *Development* **113:** 679-688.

Barton, S. C., Surani, M. A., and Norris, M. L. 1984. Role of paternal and maternal genomes in mouse development. *Nature* **311:** 374-376.

Beard, C., Li, E., and Jaenisch, R. 1995. Loss of methylation activates *Xist* in somatic but not in embryonic cells. *Genes Dev* **9:** 2325-2334.

Beechey, C. V., and Cattanach, B. M. 1996. Genetic imprinting map. *Mouse Genome* **94:** 96-99.

Bender, R., Surani, M. A., Kothary, R., Li, L.-L., Fürst, D. O., Christ, B., and Fundele, R. 1995. Tissue specific loss of proliferative capacity of parthenogenetic cells in fetal mouse chimeras. *Roux's Arch Dev Biol* **204:** 436-443.

Bestor, T., Laudano, A., Mattaliano, R., and Ingram, V. 1988. Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases. *J Mol Biol* **203:** 971-83.

Bestor, T. H. 1992. Activation of mammalian DNA methyltransferase by cleavage of a Zn binding regulatory domain. *Embo J* **11:** 2611-7.

Bestor, T. H., Hellewell, S. B., and Ingram, V. M. 1984. Differentiation of two mouse cell lines is associated with hypomethylation of their genomes. *Mol Cell Biol* **4:** 1800-6.

Bestor, T. H., and Ingram, V. M. 1983. Two DNA methyltransferases from murine erythroleukemia cells: purification, sequence specificity, and mode of interaction with DNA. *Proc Natl Acad Sci U S A* **80:** 5559-63.

Bestor, T. H., and Ingram, V. M. 1985. Growth-dependent expression of multiple species of DNA methyltransferase in murine erythroleukemia cells. *Proc Natl Acad Sci U S A* **82:** 2674-8.

Bird, A. P. 1986. CpG-rich islands and the function of DNA methylation. *Nature* **321:** 209-13.

Bird, A. P. 1995. Gene number, noise reduction and biological complexity. *Trends Genet* **11:** 94-100.

Boyes, J., and Bird, A. 1991. DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein. *Cell* **64:** 1123-34.

Boyes, J., and Bird, A. 1992. Repression of genes by DNA methylation depends on CpG density and promoter strength: evidence for involvement of a methyl-CpG binding protein. *Embo J* **11:** 327-33.

Brandeis, M., Ariel, M., and Cedar, H. 1993. Dynamics of DNA methylation during development. *BioEssays* **15:** 709-713.

Brandeis, M., Kafri, T., Ariel, M., Chaillet, J. R., McCarrey, J., Razin, A., and Cedar, H. 1993. The ontogeny of allele-specific methylation associated with imprinted genes in the mouse. *Embo J* **12:** 3669-77.

Brannan, C. I., Dees, E. S., Ingram, R. S., and Tilghman, S. M. 1990. The product of the *H19* gene may function as an RNA. *Mol Cell Biol* **10:** 28-36.

Brooks, P. J., Marietta, C., and Goldman, D. 1996. DNA mismatch repair and DNA methylation in adult brain neurons. *J Neuroscience* **16:** 939-945.

Buschhausen, B., Wittig, B., Graessmann, M., and Graessmann, A. 1987. Chromatin structure is required to block transcription of the methylated herpes simplex virus thymidine kinase gene. *Proc Natl Acad Sci USA* **84:** 1177-1181.

Butler, M. G., and Palmer, C. G. 1983. Parental origin of chromosome 15 deletion in Prader-Willi syndrome. *Lancet* **1:** 1285-1286.

Carlson, L. L., Page, A. W., and Bestor, T. H. 1992. Properties and localization of DNA methyltransferase in preimplantation mouse embryos: implications for genomic imprinting. *Genes Dev* **6:** 2536-41.

Cattanach, B. M., Barr, J. A., Evans, E. P., Burtenshaw, M., Beechey, C. V., Leff, S. E., Brannan, C. I., Copeland, N. G., Jenkins, N. A., and Jones, J. 1992. A candidate mouse model for Prader-Willi syndrome which shows an absence of *Snrpn* expression. *Nat Genet* **2:** 270-274.

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

Chaillet, J. R., Vogt, T. F., Beier, D. R., and Leder, P. 1991. Parental-specific methylation of an imprinted transgene is established during gametogenesis and progressively changes during embryogenesis. *Cell* **66:** 77-83.

Chapman, V., Forrester, L., Sanford, J., Hastie, N., and Rossant, J. 1984. Cell lineage-specific undermethylation of mouse repetitive DNA. *Nature* **307:** 284-6.

Cheng, X., Kumar, S., Posfai, J., Pflugrath, J. W., and Roberts, R. J. 1993. Crystal structure of the HhaI DNA methyltransferase complexed with S-adenosyl-L-methionine. *Cell* **74:** 299-307.

Clarke, H. J., Varmuza, S., Prideaux, V. R., and Rossant, J. 1988. The developmental potential of parthenogenetically derived cells in chimeric mouse embryos: Implications for action of imprinted genes. *Development* **104:** 175-182.

Comb, M., and Goodman, M. 1990. CpG methylation inhibits proenkephalin gene expression and binding of the transcription factor AP-2. *Nuclei Acids Res* **18:** 3975-3982.

Cuthbertson, K. S. R. 1983. Parthenogenetic activation of mouse oocytes in vitro with ethanol and benzyl alcohol. *J Exp Zool* **226:** 311-314.

Czank, A., Hauselmann, R., Page, A. W., Leonhardt, H., Bestor, T. H., Schaffner, W., and Hergersberg, M. 1991. Expression in mammalian cells of a cloned gene encoding murine DNA methyltransferase. *Gene* **109:** 259-63.

DeChiara, T. M., Efstratiadis, A., and Robertson, E. J. 1990. A growth-deficiency phenotype in heterozygous mice carrying an insulin-like growth factor II gene disrupted by targeting. *Nature* **345:** 78-80.

DeChiara, T. M., Robertson, E. J., and Efstratiadis, A. 1991. Parental imprinting of the mouse insulin-like growth factor II gene. *Cell* **64:** 849-859.

Doerfler, W. 1991. Patterns of DNA methylation - evolutionary vestiges of foreign DNA inactivation as a host defense mechanism. *Biol Chem (Hoppe-Seyler)* **372:** 557-564.

Duhl, D. M. J., Vrieling, H., Miller, K. A., Wolff, G. L., and Barsh, G. S. 1994. Neomorphic *agouti* mutations in obese yellow mice. *Nat Genet* **8:** 59-65.

Elson, D. A., and Bartolomei, M. S. 1997. A 5' differentially methylated sequence and the 3'-flanking region are necessary for H19 transgene imprinting. *Molecular and Cellular Biology* **17:** 309-317.

Fehér, Z., Kiss, A., and Venetianer, P. 1983. Expression of a bacterial modification methylase gene in yeast. *Nature* **302:** 266-8.

Feil, R., Walter, J., Allen, N. D., and Reik, W. 1994. Developmental control of allelic methylation in the imprinted mouse *Igf2* and *H19* genes. *Development* **120:** 2933-2943.

Ferguson-Smith, A. C., Cattanach, B. M., Barton, S. C., Beechey, C. V., and Surani, M. A. 1991. Embryological and molecular investigations of parental imprinting on mouse chromosome 7. *Nature* **351:** 367-370.

Ferguson-Smith, A. C., Sasaki, H., Cattanach, B. M., and Surani, M. A. 1993. Parental-origin-specific epigenetic modification of the mouse *H19* gene. *Nature* **362:** 751-5.

Finnegan, E. J., and Dennis, E. S. 1993. Isolation and identification by sequence homology of a putative cytosine methyltransferase from Arabidopsis thaliana. *Nucleic Acids Res* **21:** 2383-8.

Friedrich, G., and Soriano, P. 1991. Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice. *Genes Dev* **5:** 1513-23.

Frommer, M., McDonald, L. E., Millar, D. S., Collis, C. M., Watt, F., Grigg, G. W., Molloy, P. L., and Paul, C. L. 1992. A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc Natl Acad Sci U S A* **89:** 1827-31.

Fundele, R., Barton, S. C., Christ, B., Krause, R., and Surani, M. A. 1995a. Distribution of androgenetic cells in fetal mouse chimeras. *Roux's Arch Dev Biol* **204:** 484-493.

Fundele, R., Bober, E., Arnold, H. H., Grim, M., Bender, R., Wilting, J., and Christ, B. 1994. Early skeletal muscle development proceeds normally in parthenogenetic mouse embyros. *Dev Biol* **161:** 30-36.

Fundele, R., Li, L.-L., Herzfeld, A., Barton, S. C., and Surani, M. A. 1995b. Proliferation and differentiation of androgenetic cells in fetal mouse chimeras. *Roux's Arch Dev Biol* **204:** 494-501.

Fundele, R., Norris, M. L., Barton, S. C., Reik, W., and Surani, M. A. 1989. Systematic elimination of parthenogenetic cells in mouse chimeras. *Development* **106:** 29-35.

Fundele, R. H., Norris, M. L., Barton, S. C., Fehlau, M., Howlett, S. K., Mills, W. E., and Surani, M. A. 1990. Temporal and spatial selection against parthenogenetic cells during development of fetal chimeras. *Development* **108:** 203-11.

Fundele, R. H., and Surani, M. A. 1994. Experimental embryological analysis of genetic imprinting in mouse development. *Dev Genet* **15:** 515-22.

Gama-Sosa, M. A., Midgett, R. M., Slagel, V. A., Githens, S., Kuo, K. C., Gehrke, C. W., and Ehrlich, M. 1983. Tissue-specific differences in DNA methylation in various mammals. *Biochim Biophys Acta* **740:** 212-9.

Giddings, S. J., King, C. D., Harman, K. W., Flood, J. F., and Carnaghi, L. R. 1994. Allele specific inactivation of insulin 1 and 2, in the mouse yolk sac, indicates imprinting. *Nat Genet* **6:** 310-3.

Glenn, C. C., Saitoh, S., Jong, M. T. C., Filbrandt, M. M., Surti, U., Driscoll, D. J., and Nicholls, R. D. 1996. Gene structure, DNA methylation, and imprinted expression of the human *SNRPN* gene. *Am J Hum Genet* **58:** 335-346.

Glickman, J. F., and Reich, N. O. 1994. Baculovirus-mediated high level expression of a mammalian DNA methyltransferase. *Biochem Biophys Res Commun* **204:** 1003-8.

Goto, K., Numata, M., Komura, J.-I., Ono, T., Bestor, T. H., and Kondo, H. 1994. Expression of DNA methyltransferase gene in mature and immature neurons as well as proliferating cells in mice. *Differentiation* **56:** 39-44.

Graham, C. F. 1974. The production of parthenogenetic mammalian embryos and their use in biological research. *Biol Rev Cambridge Philos Soc* **49:** 399-422.

Graves, J. A. M. 1996. Mammals that break the rules: Genetics of marsupials and monotremes. *Annual Review of Genetics* **30:** 233-260.

Gruenbaum, Y., Cedar, H., and Razin, A. 1982. Substrate and sequence specificity of a eukaryotic DNA methylase. *Nature* **295:** 620-2.

Guillemot, F., Caspary, T., Tilghman, S. M., Copeland, N. G., Gilbert, D. J., Jenkins, N. A., Anderson, D. J., Joyner, A. L., Rossant, J., and Nagy, A. 1995. Genomic imprinting of *Mash2*, a mouse gene required for trophoblast development. *Nat Genet* **9:** 235-42.

Haig, D. 1996. Do imprinted genes have few and small introns? *BioEssays* **18:** 351-353.

Haig, D., and Graham, C. 1991. Genomic imprinting and the strange case of the insulin-like growth factor II receptor. *Cell* **64:** 1045-1046.

Hall, J. G. 1990. Genomic imprinting: review and relevance to human diseases. *Am J Hum Genet* **46:** 857-873.

Hatada, I., Kitagawa, K., Yamaoka, T., Wang, X., Arai, Y., Hashido, K., Ohishi, S., Masuda, J., Ogata, J., and Mukai, T. 1995a. Allele-specific methylation and expression of an imprinted U2af1-rs1 (SP2) gene. *Nucleic Acids Res* **23:** 36-41.

Hatada, I., and Mukai, T. 1995b. Genomic imprinting of $p57^{KIP2}$, a cyclin-dependent kinase inhibitor. *Nat. Genet* **11:** 204-6.

Hatada, I., Sugama, T., and Mukai, T. 1993. A new imprinted gene cloned by a methylation-sensitive genome scanning method. *Nucleic Acids Res* **21:** 5577-5582.

Hayashizaki, Y., *et al.* 1994. Identification of an imprinted U2af binding protein related sequence on mouse chromosome 11 using the RLGS method. *Nat Genet* **6:** 33-40.

Henry, I., Bonaiti-Pellie, C., Chehensse, V., Beldjord, C., Schwartz, C., Utermann, G., and Junien, C. 1991. Uniparental paternal disomy in a genetic cancer-predisposing syndrome. *Nature* **351:** 665-667.

Hoeller, M., Westin, G., Jiricny, J., and Schaffner, W. 1988. Sp1 transcription factor binds DNA and activates transcription even when the binding site is CpG methylated. *Genes Dev* **2:** 1127-35.

Howlett, S. K., and Reik, W. 1991. Methylation levels of maternal and paternal genomes during preimplantation development. *Development* **113:** 119-127.

Hurst, L. D., McVean, G. T., and Moore, T. 1996. Imprinted genes have few and small introns. *Nat Genet* **12:** 234-237.

Iguchi-Ariga, S. M., and Schaffner, W. 1989. CpG methylation of the cAMP-responsive enhancer/promoter sequence TGACGTCA abolishes specific factor binding as well as transcriptional activation. *Genes Dev* **3:** 612-9.

Jaenisch, R. 1997. DNA methylation and imprinting: Who needs it anyway? *Trends Genet*, in preparation.

Jägerbauer, E.-M., Fraser, A., Herbst, E. W., Kothary, R., and Fundele, R. 1992. Parthenogenetic stem cells in postnatal mouse chimeras. *Development* **116:** 95-102.

Jähner, D., and Jaenisch, R. 1985. Retrovirus-induced de novo methylation of flanking host sequences correlates with gene inactivity. *Nature* **315:** 594-7.

Jähner, D., Stuhlmann, H., Stewart, C. L., Harbers, K., Lohler, J., Simon, I., and Jaenisch, R. 1982. *De novo* methylation and expression of retroviral genomes during mouse embryogenesis. *Nature* **298:** 623-8.

Jinno, Y., Sengoku, K., Nakao, M., Tamate, K., Miyamoto, T., Matsuzaka, T., Sutcliffe, J. S., Anan, T., Takuma, N., Nishiwaki, K., Ikeda, Y., Ishimaru, T., Ishikawa, M., and Niikawa, N. 1996. Mouse/human sequence divergence in a region with a paternal-specific methylation imprint at the human *H19* locus. *Hum Mol Genet* **5:** 1155-1161.

Jones, P. A., Rideout, W. M., Shen, J. C., Spruck, C. H., and Tsai, Y. C. 1992. Methylation, mutation and cancer. *Bioessays* **82:** 9-12.

Jong, M. T. C., Glenn, C. C., Saitoh, S., Driscoll, D. J., and Nicholls, R. D. 1994. The putative imprinted locus *D15S9* within the common deletion region for

the Prader-Willi and Angelman syndromes encodes two overlapping mRNAs transcribed from opposite strands. *Am J Hum Genet Suppl* **55:** A29.

Jost, J. P., and Saluz, H. P. 1993. DNA methylation: molecular biology and biological significance. Birkhauser Verlag, Basel, Switzerland.

Jue, K., Benoit, G., Alcivar-Warren, A. A., and Trasler, J. M. 1995. Developmental and hormonal regulation of DNA methyltransferase in the rat testis. *Biol Reproduction* **52:** 1364-1371.

Junien, C. 1992. Beckwith-Wiedemann syndrome, tumourigenesis and imprinting. *Curr Opin Genet Development* **2:** 431-438.

Kafri, T., Ariel, M., Brandeis, M., Shemer, R., Urven, L., McCarrey, J., Cedar, H., and Razin, A. 1992. Developmental pattern of gene-specific DNA methylation in the mouse embryo and germ line. *Genes Dev* **6:** 705-14.

Kalscheuer, V. M., Mariman, E. C., Schepens, M. T., Rehder, H., and Ropers, H.-H. 1993. The insulin-like growth factor type-2 receptor gene is imprinted in the mouse but not in humans. *Nat Genet* **5:** 74-78.

Kaneko-Ishino, T., Kuroiwa, Y., Miyoshi, N., Kohda, T., Suzuki, R., Yokoyama, M., Viville, S., Barton, S. C., and Surani, M. A. 1995. *Peg1/Mest* imprinted gene on chromosome 6 identified by cDNA subtraction hybridization. *Nat Genet* **11:** 52-59.

Kawame, H., Gartler, S. M., and Hansen, R. S. 1995. Allele-specific replication timing in imprinted domains: absence of asynchrony at several loci. *Hum Mol Genet* **4:** 2287-93.

Kay, G. F., Penny, G. D., Patel, D., Ashworth, A., Brockdorff, N., and Rastan, S. 1993. Expression of *Xist* during mouse development suggests a role in the initiation of X chromosome inactivation. *Cell* **72:** 171-182.

Keverne, E. B., Fundele, R., Narasimha, M., Barton, S. C., and Surani, M. A. 1996. Genomic imprinting and the differential roles of parental genomes in brain development. *Dev Brain Res* **92:** 91-100.

Kimura, H., Ishihara, G., and Tajima, S. 1996. Isolation and expression of a Xenopus laevis DNA methyltransferase cDNA. *Journal of Biochemistry* **120:** 1182-1189.

Kitsberg, D., Selig, S., Brandeis, M., Simon, I., Keshet, I., Driscoll, D. J., Nicholls, R. D., and Cedar, H. 1993. Allele-specific replication timing of imprinted gene regions. *Nature* **364:** 459-463.

Klimasauskas, S., Kumar, S., Roberts, R. J., and Cheng, X. 1994. HhaI methyltransferase flips its target base out of the DNA helix. *Cell* **76:** 357-369.

Knoll, J. H., Nicholls, R. D., Magenis, R. E., Graham, J. M., Lalande, M., and Latt, S. A. 1989. Angelman and Prader-Willi syndromes share a common chromosome 15 deletion but differ in parental origin of the deletion. *Am J Med Genet* **32:** 385-390.

Knoll, J. H. M., Cheng, S.-D., and Lalande, M. 1994. Allele specificity of DNA replication timing in the Angelman/Prader-Willi syndrome imprinted chromosomal region. *Nat Genet* **6:** 41-46.

Kumar, S., Cheng, X., Klimasauskas, S., Mi, S., Posfai, J., Roberts, R. J., and Wilson, G. G. 1994. The DNA (cytosine-5) methyltransferases. *Nucleic Acids Res* **22:** 1-10.

Kuroiwa, Y., Kaneko-Ishino, T., Kagitani, F., Kohda, T., Li, L. L., Tada, M., Suzuki, R., Yokoyama, M., Shiroishi, T., Wakana, S., Barton, S. C., Ishino, F., and Surani, M. A. 1996. *Peg3* imprinted gene on proximal chromosome 7 encodes for a zinc finger protein. *Nat Genet* **12:** 186-90.

Laird, P. W., Jackson-Grusby, L., Fazeli, A., Dickinson, S., Jung, W. E., Li, E., Weinberg, R. A., and Jaenisch, R. 1995. Suppression of intestinal neoplasia by DNA hypomethylation. *Cell* **81:** 197-205.

Laird, P. W., and Jaenisch, R. 1994. DNA Methylation and Cancer. *Hum Molec Genet* **3:** 1487-1495.

Lalande, M. 1996. Parental imprinting and human disease. *Annual Review of Genetics* **30:** 173-195.

LaSalle, J. M., and Lalande, M. 1996. Homologous association of oppositely-imprinted chromosomal domains. *Science* **272:** 725-728.

Latham, K. 1995. Stage-specific and cell type-specific aspects of genomic imprinting effects in mammals. *Differentiation* **59:** 269-282.

Lau, M. M. H., Stewart, S. E. H., Liu, Z., Bhatt, H., Rotwein, P., and Stewart, C. L. 1994. Loss of the imprinted IGF2/cation-independent mannose 6-phosphate receptor results in fetal overgrowth and perinatal lethality. *Genes Dev* **8:** 2953-2963.

Lee, M. P., Hu, R.-J., Johnson, L. A., and Feinberg, A. P. 1997. Human *KVLQT1* gene shows tissue-specific imprinting and encompasses Beckwith-Wiedemann syndrome chromosomal rearrangements. *Nat Genet* **15:** 181-5.

Leff, S. E., Brannan, C. I., Reed, M. L., Ozcelik, T., Francke, U., Copeland, N. G., and Jenkins, N. A. 1992. Maternal imprinting of the mouse *Snrpn* gene and conserved linkage homology with the human Prader-Willi syndrome region. *Nat Genet* **2:** 259-64.

Lei, H., Oh, S. P., Okano, M., Jüttermann, R., Goss, K. A., Jaenisch, R., and Li, E. 1996. De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. *Development* **122:** 3195-3205.

Leighton, P. A., Ingram, R. S., Eggenschwiler, J., Efstratiadis, A., and Tilghman, S. M. 1995a. Disruption of imprinting caused by deletion of the *H19* gene region in mice. *Nature* **375:** 34-39.

Leighton, P. A., Saam, J. R., Ingram, R. S., Stewart, C. L., and Tilghman, S. M. 1995b. An enhancer deletion affects both *H19* and *Igf2* expression. *Genes Dev* **9:** 2079-89.

Leonhardt, H., and Bestor, T. H. 1993. Structure, function and regulation of mammalian DNA methyltransferase. In DNA methylation: molecular biology and biological significance (ed. J. P. Jost, and Saluz, H. P.), pp. 109-119. Birkhauser Verlag, Basel, Switzerland.

Leonhardt, H., Page, A. W., Weier, H. U., and Bestor, T. H. 1992. A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. *Cell* **71:** 865-73.

Lewis, J. D., Meehan, R. R., Henzel, W. J., Maurer, F. I., Jeppesen, P., Klein, F., and Bird, A. 1992. Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. *Cell* **69:** 905-14.

Li, E., Beard, C., and Jaenisch, R. 1993. Role for DNA methylation in genomic imprinting. *Nature* **366:** 362-5.

Li, E., Bestor, T. H., and Jaenisch, R. 1992. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69:** 915-26.

Liu, Y., Sun, L., and Jost, J.-P. 1996. In differentiating mouse myoblasts DNA methyltransferase is posttranscriptionally and posttranslationally regulated. *Nucl Acids Res* **24:** 2718-2722.

Ludwig, T., Eggenschwiler, J., Fisher, P., D'Ercole, A. J., Davenport, M. L., and Efstratiadis, A. 1996. Mouse mutants lacking the type 2 IGF receptor (IGF2r) are rescued from perinatal lethality in Igf2 and Igf1r null backgrounds. *Dev biol* **177:** 517-535.

Malcolm, S., Clayton-Smith, J., Nichols, M., Robb, S., Webb, T., Armour, J. A., Jeffreys, A. J., and Pembrey, M. E. 1991. Uniparental paternal disomy in Angelman's syndrome. *Lancet* **337:** 694-697.

Mann, J. R., Gadi, I., Harbison, M. L., Abbondanzo, S. J., and Stewart, C. L. 1990. Androgenetic mouse embryonic stem cells are pluripotent and cause skeletal defects in chimeras: implications for genetic imprinting. *Cell* **62:** 251-260.

Mann, J. R., and Lovell-Badge, R. H. 1984. Inviability of parthenogenones is determined by pronuclei, not egg cytoplasm. *Nature* **310:** 66-67.

Mann, J. R., and Stewart, C. L. 1991. Development to term of mouse androgenetic aggregation chimeras. *Development* **113:** 1325-1333.

McGrath, J., and Solter, D. 1983. Nuclear transplantation in the mouse embryo by microsurgery and cell fusion. *Science* **220:** 1300-1302.

McGrath, J., and Solter, D. 1984. Completion of mouse embryogenesis requires both the maternal and paternal genomes. *Cell* **37:** 179-183.

McKarney, L. A., Overall, M. L., and Dziadek, M. 1996. Expression of H19 and Igf2 genes in uniparental mouse ES cells during in vitro and in vivo differentiation. *Differentiation* **60:** 75-86.

McLaughlin, K. J., Szabó, P., Haegel, H., and Mann, J. R. 1996. Mouse embryos with paternal duplication of an imprinted chromosome 7 region die at midgestation and lack placental spongiotrophoblast. *Development* **122:** 265-270.

Meehan, R. R., Lewis, J. D., McKay, S., Kleiner, E. L., and Bird, A. P. 1989. Identification of a mammalian protein that binds specifically to DNA containing methylated CpGs. *Cell* **58:** 499-507.

Michaud, E. J., van Vugt, M. J., Bultman, S. J., Sweet, H. O., Davisson, M. T., and Woychik, R. P. 1994. Differential expression of a new dominant agouti allele (A^{iapy}) is correlated with methylation state and is influenced by parental lineage. *Genes Dev* **8:** 1463-1472.

Monk, M. 1990. Changes in DNA methylation during mouse embryonic development in relation to X-chromosome activity and imprinting. *Philos Trans R Soc Lond [Biol]* **326:** 299-312.

Monk, M., Adams, R. L., and Rinaldi, A. 1991. Decrease in DNA methylase activity during preimplantation development in the mouse. *Development* **112:** 189-92.

Monk, M., Boubelik, M., and Lehnert, S. 1987. Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development. *Development* **99:** 371-82.

Moore, T., and Haig, D. 1991. Genomic imprinting in mammalian development: a parental tug of war. *Trends Genet* **7:** 45-49.

Nagy, A., Gocza, E., Merentes Dias, D., Prideaux, V. R., Ivanyi, E., Markkula, M., and Rossant, J. 1990. Embryonic stem cells alone are able to suport fetal development in the mouse. *Development* **110:** 815-821.

Nagy, A., Paldi, A., Dezso, L., Varga, L., and Magyar, A. 1987. Prenatal fate of parthenogenetic cells in mouse aggregation chimaeras. *Development* **101:** 67-71.

Nagy, A., Sass, M., and Markkula, M. 1989. Systematic non-uniform distribution of parthenogenetic cells in adult mouse chimaeras. *Development* **106:** 321-4.

Nan, X., Campoy, J., and Bird, A. 1997. MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. *Cell* **88:** 471-481.

Neumann, B., Kubicka , P., and Barlow, D. P. 1995. Characteristics of imprinted genes. *Nat Genet* **9:** 12-13.

Nicholls, R. D., Knoll, J. H. M., Butler, M. G., Karam, S., and Lalande, M. 1989. Genetic imprinting suggested by maternal heterodisomy in non-deletion Prader-Willi syndrome. *Nature* **342:** 281-285.

Norris, D. P., Patel, D., Kay, G. F., Penny, G. D., Brockdorff, N., Sheardown, S. A., and Rastan, S. 1994. Evidence that random and imprinted *Xist* expression is controlled by preemptive methylation. *Cell* **77:** 41-51.

Noyer-Weidner, M., and Trautner, T. A. 1993. Methylation of DNA in prokaryotes. In DNA Methylation: Molecular Biology and Biological Significance (ed. J. P. Jost, and Saluz, H. P.), pp. 39-108. Birkhauser Verlag, Basel, Switzerland.

Pàldi, A., Gyapy, G., and Jami, J. 1995. Imprinted regions of the human genome display sex-specific meiotic recombination frequencies. *Curr Biol* **5:** 1030-1035.

Pàldi, A., Nagy, A., Markkula, M., Barna, I., and Dezso, L. 1989. Postnatal development of parthenogenetic <-> fertilized mouse aggregation chimeras. *Development* **105:** 115-8.

Pearsall, R. S., Shibata, H., Brozowska, A., Yoshino, K., Okuda, K., deJong, P. J., Plass, C., Chapman, V. M., Hayashizaki, Y., and Held, W. A. 1996. Absence of imprinting in U2AFBPL, a human homologue of the imprinted mouse gene *U2afbp-rs*. *Biochem Biophys Res Comm* **222:** 171-177.

Pfeifer, K., Leighton, P. A., and Tilghman, S. M. 1996. The structural H19 gene is required for transgene imprinting. *Proc Natl Acad Sci USA* **93:** 13876-13883.

Pinarbasi, E., Elliott, J., and Hornby, D. P. 1996. Activation of a yeast pseudo DNA methyltransferase by deletion of a single amino acid. *J Mol Biol* **257:** 804-813.

Plass, C., Shibata, H., Kalcheva, I., Mullins, L., Kotelevtseva, N., Mullins, J., Kato, R., Sasaki, H., Hirotsune, S., Okazaki, Y., Held, W. A., Hayashizaki, Y., and Chapman, V. M. 1996. Identification of Grf1 on mouse chromosome 9 as an imprinted gene by RLGS-M. *Nat Genet* **14:** 106-9.

Prendergast, G. C., Lawe, D., and Ziff, E. B. 1991. Association of myn, the murine homolog of max, with c_myc stimulates methylation -sensitive DNA binding and Ras contransformation. *Cell* **65:** 395-407.

Proffitt, J. H., Davie, J. R., Swinton, D., and Hattman, S. 1984. 5-Methylcytosine is not detectable in Saccharomyces cerevisiae DNA. *Mol Cell Biol* **4:** 985-8.

Reeve, A. E. 1995. Genomic imprinting in embryonal tumors and overgrowth disorders. In Genomic Imprinting: Causes and Consequences (ed. R. Ohlsson, Hall, K., and Ritzen, M.), pp. 209-223. Camridge University Press, Cambridge, England.

Reik, W., Collick, A., Norris, M. L., Barton, S. C., and Surani, M. A. 1987. Genomic imprinting determines methylation of parental alleles in transgenic mice. *Nature* **328:** 248-51.

Rouleau, J., Tanigawa, G., and Szyf, M. 1992. The mouse DNA methyltransferase 5'-region. A unique housekeeping gene promoter. *J Biol Chem* **267:** 7368-77.

Sanford, J. P., Clark, H. J., Chapman, V. M., and Rossant, J. 1987. Differences in DNA methylation during oogenesis and spermatogenesis in the mouse. *Genes Dev* **1:** 1039-1046.

Sapienza, C., Peterson, A. C., Rossant, J., and Balling, R. 1987. Degree of methylation of transgenes is dependent on gamete of origin. *Nature* **328:** 251-4.

Sasaki, H., Ferguson-Smith, A. C., Shum, A. S., Barton, S. C., and Surani, M. A. 1995. Temporal and spatial regulation of *H19* imprinting in normal and uniparental mouse embryos. *Development* **121:** 4195-202.

Sasaki, H., Jones, P. A., Chaillet, J. R., Ferguson-Smith, A. C., Barton, S. C., Reik, W., and Surani, M. A. 1992. Parental imprinting: potentially active chromatin of the repressed maternal allele of the mouse insulin-like growth factor II (*Igf2*) gene. *Genes Dev* **6:** 1843-1856.

Simon, D., Stuhlmann, H., Jahner, D., Wagner, H., Werner, E., and Jaenisch, R. 1983. Retrovirus genomes methylated by mammalian but not bacterial methylase are non-infectious. *Nature* **304:** 275-7.

Singh, J., and Klar, A. J. S. 1992. Active gene in budding yeast display enhanced in vivo accessibility to foreign DNA methylases: a novel in vivo probe for chromatin structure of yeast. *Genes Dev* **6:** 186-196.

Smrzka, O. W., Faé, I., Stöger, R., Kurzbauer, R., Fischer, G. F., Henn, T., Weith, A., and Barlow, D. P. 1995. Conservation of a maternal-specific methylation signal at the human IGF2R locus. *Hum Mol Genet* **4:** 1945-1952.

Solter, D. 1988. Differential imprinting and expression of maternal and paternal genomes. *Annu Rev Genetics* **22:** 127-146.

Spindle, A., Sturm, K. S., Flannery, M., Meneses, J. J., Wu, K., and Pedersen, R. A. 1996. Defective chorioallantoic fusion in mid-gestation lethality of parthenogenone <-> tetraploid chimeras. *Dev Biol* **173:** 447-458.

Stevens, L. C. 1978. Totipotent cells of parthenogenetic origin in a chimaeric mouse. *Nature* **276:** 266-7.

Stewart, C. L., Stuhlmann, H., Jähner, D., and Jaenisch, R. 1982. *De novo* methylation, expression, and infectivity of retroviral genomes introduced into embryonal carcinoma cells. *Proc Natl Acad Sci U S A* **79:** 4098-102.

Stöger, R., Kubicka, P., Liu, C. G., Kafri, T., Razin, A., Cedar, H., and Barlow, D. P. 1993. Maternal-specific methylation of the imprinted mouse *Igf2r* locus identifies the expressed locus as carrying the imprinting signal. *Cell* **73:** 61-71.

Sturm, K. S., Flannery, M. L., and Pedersen, R. A. 1994. Abnormal development of embryonic and extraembryonic cell lineages in parthenogenetic mouse embryos. *Dev Dynamics* **201:** 11-28.

Surani, M. A. 1994. Genomic imprinting: control of gene expression by epigenetic inheritance. *Curr Opin Cell Biol* **6:** 390-5.

Surani, M. A., and Barton, S. C. 1983. Development of gynogenetic eggs in the mouse: implications for parthenogenetic embryos. *Science* **222:** 1034-6.

Surani, M. A., Barton, S. C., Howlett, S. K., and Norris, M. L. 1988. Influence of chromosomal determinants on development of androgenetic and parthenogenetic cells. *Development* **103:** 171-8.

Surani, M. A., Barton, S. C., and Kaufman, M. H. 1977. Development to term of chimaeras between diploid parthenogenetic and fertilised embryos. *Nature* **270:** 601-3.

Surani, M. A., Barton, S. C., and Norris, M. L. 1987. Influence of parental chromosomes on spatial specificity in androgenetic<->parthenogenetic chimaeras in the mouse. *Nature* **326:** 395-397.

Surani, M. A. H., Barton, S. C., and Norris, M. L. 1984. Development of reconstituted mouse eggs suggests imprinting of the genome during gametogenesis. *Nature* **308:** 548-550.

Sutcliffe, J. S., Nakao, M., Christian, S., Örstavik, K. H., Tommerup, N., Ledbetter, D. H., and Beaudet, A. L. 1994. Deletions of a differentially methylated CpG island at the SNRPN gene define a putative imprinting control region. *Nat Genet* **8:** 52-58.

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.

Szabó, P., and Mann, J. R. 1994. Expression and methylation of imprinted genes during in vitro differentiation of mouse parthenogenetic and androgenetic embryonic stem cell lines. *Development* **120:** 1651-60.

Szabó, P. E., and Mann, J. R. 1995a. Biallelic expression of imprinted genes in the mouse germ line: implications for erasure, establishment, and mechanisms of genomic imprinting. *Genes Dev* **9:** 1857-1868.

Szabó, P. E., and Mann, J. R. 1995b. Allele-specific expression and total expression levels of imprinted genes during early mouse development: implications for imprinting mechanisms. *Genes Dev* **9:** 3097-3108.

Szabó, P. E., and Mann, J. R. 1996. Maternal and paternal genomes function independently in mouse ova in establishing expression of the imprinted genes Snrpn and Igf2r: No evidence for allelic trans-sensing and counting mechanisms. *EMBO Journal* **15:** 6018-6025.

Szyf, M., Bozovic, V., and Tanigawa, G. 1991. Growth regulation of mouse DNA methyltransferase gene expression. *J Biol Chem* **266:** 10027-30.

Szyf, M., Kaplan, F., Mann, V., Gioh, H., Kedar, E., and Razin, A. 1985. Cell cycle-dependent regulation of eukaryotic DNA methylase level. *J Biol Chem* **260:** 8653-8656.

Tajima, S., Tsuda, H., Wakabayashi, N., Asano, A., Mizuno, S., and Nishimori, K. 1995. Isolation and expression of a chicken DNA methyltransferase cDNA. *J. Biochem* **117:** 1050-1057.

Takagi, N., and Sasaki, M. 1975. Preferential inactivation of the paternally derived X chromosome in the extraembryonic membranes of the mouse. *Nature* **256:** 640-642.

Teubner, B., and Schulz, W. A. 1995. Regulation of DNA methyltransferase during differentiation of F9 mouse embryonal carcinoma cells. *J Cell Physiology* **165:** 284-290.

Thomas, J. H. 1995. Genomic imprinting proposed as a surveillance mechanism for chromosome loss. *Proc Natl Acad Sci U S A* **92:** 480-2.

Thomson, J. A., and Solter, D. 1988. The developmental fate of androgenetic, parthenogenetic, and gynogenetic cells in chimeric gastrulating mouse embryos. *Genes Dev* **2:** 1344-51.

Thomson, J. A., and Solter, D. 1989. Chimeras between parthenogenetic or androgenetic blastomeres and normal embryos: allocation to the inner cell mass and trophectoderm. *Dev Biol* **131:** 580-3.

Trasler, J. M., Alcivar, A. A., Hake, L. E., Bestor, T., and Hecht, N. B. 1992. DNA methyltransferase is developmentally expressed in replicating and non-replicating male germ cells. *Nucleic Acids Res* **20:** 2541-5.

Trasler, J. M., Trasler, D. G., Bestor, T. H., Li, E., and Ghibu, F. 1996. DNA methyltransferase in normal and $Dnmt^n/Dnmt^n$ mouse embryos. *Dev Dynamics* **206:** 239-247.

Tremblay, K. D., Saam, J. R., Ingram, R. S., Tilghman, S. M., and Bartolomei, M. S. 1995. A paternal-specific methylation imprint marks the alleles of the mouse *H19* gene. *Nat Genet* **9:** 407-13.

Tucker, K. L., Beard, C., Dausman, J., Jackson-Grusby, L., Laird, P. W., Lei, H., Li, E., and Jaenisch, R. 1996a. Germ-line passage is required for the establishment of methylation and expression patterns of imprinted but not of non-imprinted genes. *Genes Dev.* **10:** 1008-20.

Tucker, K. L., Talbot, D., Lee, M. A., Leonhardt, H., and Jaenisch, R. 1996b. Complementation of methylation deficiency in embryonic stem cells by a DNA methyltransferase minigene. *Proc Natl Acad Sci USA* **93:** 12920-12925.

Urieli-Shoval, S., Gruenbaum, Y., Sedat, J., and Razin, A. 1982. The absence of detectable methylated bases in Drosophila melanogaster DNA. *Febs Lett* **146:** 148-52.

Vanyushin, B. F., and Poirier, L. A. 1996. Drosophila melanogaster genomic DNA sequence homologous to mammalian cytosine DNA-methyltransferase gene. *Biochem Mol Biol International* **39:** 353-358.

Vardimon, L., Kressmann, A., Cedar, H., Maechler, M., and Doerfler, W. 1982. Expression of a cloned adenovirus gene is inhibited by in vitro methylation. *Proc Natl Acad Sci U S A* **79:** 1073-7.

Varmuza, S., Mann, M., and Rogers, I. 1993. Site of action of imprinted genes revealed by phenotypic analysis of parthenogenetic embryos. *Dev Genet* **14:** 239-248.

Vedel, M., Gomez-Garcia, M., Sala, M. A., and Sala-Trepat, J. M. 1983. Changes in methylation pattern of albumin and a-fetoprotien genes in developing rat liver and neoplasma. *Nucleic Acids Res* **11:** 4335-4354.

Vertino, P. M., Yen, R.-W. C., Gao, J., and Baylin, S. B. 1996. De novo methylation of CpG island sequences in human fibroblasts overexpressing DNA (cytosine-5-)-methyltransferase. *Mol Cell Biol* **16:** 4555-4565.

Villar, A. J., and Pedersen, R. A. 1994. Parental imprinting of the *Mas* protooncogene in mouse. *Nat Genet* **8:** 373-9.

Vogel, M. C., Papadopoulos, T., Mueller-Hermelink, H. K., Drahovsky, D., and Pfeifer, G. P. 1988. Intracellular distribution of DNA methyltransferase during the cell cycle. *FEBS Letters* **236:** 9-13.

Walsh, C., Glaser, A., Fundele, R., Ferguson-Smith, A., Barton, S., Surani, M. A., and Ohlsson, R. 1994. The non-viability of uniparental mouse conceptuses correlates with the loss of the products of imprinted genes. *Mech Dev* **46:** 55-62.

Wang, Z.-Q., Fung, M. R., Barlow, D. P., and Wagner, E. F. 1994. Regulation of embryonic growth and lysosomal targeting by the imprinted *Igf2/Mpr* gene. *Nature* **372:** 464-467.

Watt, F., and Molloy, P. L. 1988. Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter. *Genes Dev* **2:** 1136-43.

Wevrick, R., Kerns, J. A., and Francke, U. 1994. Identification of a novel paternally expressed gene in the Prader-Willi syndrome region. *Hum Mol Genet* **3:** 1877-82.

Wilkinson, C. R. M., Bartlett, R., Nurse, P., and Bird, A. P. 1995. The fission yeast gene *pmt1*$^+$ encodes a DNA methyltransferase homologue. *Nucl Acids Res* **23:** 203-210.

Wu, J., Herman, J. G., Wilson, G., Lee, R. Y., Yen, R.-W. C., Mabry, M., de Bustros, A., Nelkin, B. D., and Baylin, S. B. 1996. Expression of prokaryotic *Hha*I DNA methyltransferase is transforming and lethal to NIH 3T3 cells. *Cancer Res* **56:** 616-622.

Wu, J., Issa, J. P., Herman, J., Bassett, D. E., Jr., Nelkin, B. D., and Baylin, S. B. 1993. Expression of an exogenous eukaryotic DNA methyltransferase gene induces transformation of NIH 3T3 cells. *Proc Natl Acad Sci U S A* **90:** 8891-5.

Yeivin, A., and Razin, A. 1993. Gene methylation patterns and expression. In DNA Methylation: Molecular Biology and Biological Significance (ed. J. P. Jost, and Saluz, H. P.), pp. 523-568. Birkhauser Verlag, Basel, Switzerland.

Yen, R. W., Vertino, P. M., Nelkin, B. D., Yu, J. J., el-Deiry, W., Cumaraswamy, A., Lennon, G. G., Trask, B. J., Celano, P., and Baylin, S. B. 1992. Isolation and characterization of the cDNA encoding human DNA methyltransferase. *Nucleic Acids Res* **20:** 2287-91.

Zemel, S., Bartolomei, M. S., and Tilghman, S. M. 1992. Physical linkage of two mammalian imprinted genes, *H19* and insulin-like growth factor 2. *Nat Genet* **2:** 61-65.

Zuccotti, M., and Monk, M. 1995. Methylation of the mouse *Xist* gene in sperm and eggs correlates with imprinted *Xist* expression and paternal X-inactivation. *Nat Genet* **9:** 316-20.

Appendix A

Classification: Biochemistry

Complementation of methylation deficiency in embryonic stem cells by a DNA methyltransferase mini-gene

Kerry Lee Tucker*,†, Dale Talbot*, Min Ae Lee‡, Heinrich Leonhardt§, and Rudolf Jaenisch*,†,¶

*Whitehead Institute for Biomedical Research

†Massachusetts Institute of Technology and Department of Biology, Cambridge, MA 02142 USA

‡Brigham and Women's Hospital, Dept. of Medicine, Cardiovascular Division, Boston, MA 02115 USA

§Franz Volhard Clinic, Max Delbrück Center for Molecular Medicine, Dept. of Nephrology, Hypertension and Genetics, Wiltbergstr. 50, 13125 Berlin, Germany

¶To whom reprint requests should be addressed

Abbreviations: ES cells, Embryonic Stem cells; MTase, DNA (cytosine-5)-methyltransferase; ORF, open reading frame

ABSTRACT

Previous attempts to express functional DNA cytosine methyltransferase (MTase) in cells transfected with the available *Dnmt* cDNAs have met with little or no success. We show that the published *Dnmt* sequence encodes an amino terminal-truncated protein which is tolerated only at very low levels when stably expressed in ES cells. Normal expression levels were, however, obtained with constructs containing a continuation of an open reading frame with a coding capacity of up to 171 amino acids upstream of the previously defined start site. The protein encoded by these constructs comigrated in SDS-PAGE with the endogenous enzyme and restored methylation activity in transfected cells. This was shown by functional rescue of *Dnmt* mutant ES cells, which contain highly demethylated genomic DNA and fail to differentiate normally. When transfected with the minigene construct, the genomic DNA became remethylated and the cells regained the capacity to form teratomas, which displayed a wide variety of differentiated cell types. Our results define an amino-terminal domain of the mammalian MTase which is crucial for stable expression and function *in vivo*.

INTRODUCTION

Much evidence indicates that in vertebrates the methylation of DNA at cytosine residues affects gene transcription (1). Methylation is believed to be important in a variety of biological processes, including embryonic development, genomic imprinting, and cancer (2, 3). A single gene (*Dnmt*) encoding a DNA (cytosine-5)-methyltransferase (MTase) has been cloned from mammalian cells (4). MTase shows a marked preference for hemimethylated DNA (5), suggesting that its function is to maintain the methylation status of newly replicated DNA. The enzyme consists of an amino-terminal regulatory domain followed by a 500-amino acid domain assumed to contain the catalytic center, due to its homology to cytosine-methylating bacterial type II restriction methyltransferases (6).

The crucial importance of methylation in development, genomic imprinting, and cancer was demonstrated by the targeted mutation of the *Dnmt* gene in embryonic stem (ES) cells (7). Mice homozygous for the mutation die at midgestation; homozygous mutant ES cells proliferate normally with their genomic DNA highly demethylated but die upon differentiation (8). Imprinted genes display parent-specific monoallelic expression, and DNA methylation has been proposed to constitute the molecular mark that distinguishes the two alleles (3). Strong support for this notion was provided by the loss of monoallelic expression of these genes in homozygous *Dnmt* mutant embryos or ES cells (9). Furthermore, a large body of evidence links hypo- and hypermethylation of genomic DNA to cancer progression (10). A direct correlation between DNA methylation and intestinal

neoplasia was demonstrated in mice expressing different levels of MTase (11).

The present work used the complementation of the *Dnmt* mutation as an assay to define the sequences encoding a functional MTase. Previous reports of high expression of functional MTase in transient transfections (12, 13) contrasted with the inability to achieve high expression levels in stable transfection experiments. In fact, only one report described high levels of MTase expression in stably transfected clones which, however, displayed a transformed phenotype (14). We found that the published cDNA was incapable of expressing wild type levels of MTase and complementing the *Dnmt* mutation. A construct containing a previously reported (15) *Dnmt* promoter fused to the *Dnmt* cDNA also failed in this assay. In contrast, wild type levels of *Dnmt* expression were obtained after targeting the published cDNA to its cognate locus, where expression of the full-length cDNA from its cognate promoter fully restored MTase function to ES cells (16). This result suggested that sequences upstream of the previously identified initiation codon were crucial for stable and functional expression of the enzyme. In this paper we report a previously unidentified amino-terminal open reading frame which extends the coding capacity of the construct by up to 171 amino acids. Expression of the full-length *Dnmt* cDNA in mutant ES cells restored both normal methylation levels and the capacity to form teratomas, indicating a crucial role of the amino-terminal domain in normal enzymatic activity.

MATERIALS AND METHODS

Vectors. *Dnmt* constructs used were from pMG (T. Bestor), which is a 4934 bp *Dnmt* cDNA. pMT20 was made by inserting the cDNA into a vector which contains the *Pgk-1* promoter and polyadenylation sequences of p*pgk::hprt* (17) flanking a synthetic intron (IVS in Fig. 1A) (18). Removal of the IVS from pMT20 yielded pMT10. pMT30 was made by excising a 1357-bp *Nae*I fragment from pMT10, comprising the phage f1 ori, the 3' end of the *lacZ* gene, the *Pgk-1* promoter, and the first 180 bp of the *Dnmt* cDNA. A 4.7-kb *Dnmt* *Nae*I genomic fragment was cloned into this site, such that the *Dnmt* cDNA sequence removed from pMT10 was restored in its genomic context. pMT40 was constructed by inserting the *Dnmt* cDNA sequence into pEF-PGKhyg (gift of S. Orkin). This variant of pEF-PGKneo (19) included a 1188-bp fragment encompassing the promoter region of the human EF-1α gene, taken from pEF-BOS (20), an SV40-derived polyadenylation signal, and a *Pgk-1*-driven hyg^R gene. pMT42 was constructed from pMT40 by restriction at the unique *Nsi*I site in the *Dnmt* cDNA, destruction of the 3' overhang by T4 DNA polymerase, and insertion of a 12-bp *Nhe*I linker (NEB), containing amber stop codons in all three reading frames. pMT50 was made by inserting a 5.0-kb *Dnmt* *Eco*RI fragment upstream of the *Dnmt* genomic region in pMT30. The inserted sequence begins 3641 bp upstream of the first exon and ends 381 bp downstream of the testis-specific exon. The G418-selectable pPGK-RN (21) and the puromycin-selectable pPGK-Puro (16) were made as described.

Transfections. The J1 and the *Dnmt* mutant ES cell lines (7, 8) were cultured as described (7). Each linearized construct was mixed in five-fold molar excess with 250 ng of selectable marker and the cationic liposome DOTAP (Boehringer Mannheim). Cells were incubated with this mix according to manufacturer's protocol and plated on γ-irradiated murine fetal fibroblasts of the appropriate drug-resistance. Selection went as follows: G418 (BRL) @ 200 μg/ml (active), puromycin (Sigma) @ 2 μg/ml, and hygromycin (Boehringer Mannheim) @ 100 μg/ml. Isolated colonies were picked 9-12 days after lipofection and expanded. Nucleic acids were analyzed as described (16). Teratomas were made as described (16).

Antibody derivation and Western blot analysis. A synthetic peptide of the sequence NH2-CRSPRSRPKPRGPRRSK (Chiron Mimotopes) was coupled to Maleimide-activated KLH (Pierce) and injected into female NZW rabbits (HRP) with FCA. Boosts and bleeds were performed using standard protocols.

Protein was analyzed from clones grown for two passages without feeder cells. Confluent 25-cm^2 flasks of ES cells were lysed in 2X sample buffer followed by boiling for 5 minutes and sonication. SDS-PAGE on 8% gels were performed according to (22). Western blot analysis was performed according to (23) using antisera HM334 at 1 / 3000 dilution and chemiluminescence (ECL Amersham).

Cloning of 5' end of *Dnmt* cDNA. The first technique used polyA+ RNA purified from crude tissue extracts derived from a CD1-mouse kidney using the PolyATract System (Promega). The 5' RACE System

(BRL) was used with 300 ng of polyA+ RNA and a *Dnmt*-specific primer MMT2400AS (5'-AGGGTGTCACTGTCCGACTT, located in the fourth exon). Poly dC-tailed cDNA was subjected to 35 PCR cycles using an annealing temperature of 57° C in a 50 µl reaction containing 25 mM Tris-Cl (pH 8.4), 50 mM KCl, 1.5 mM $MgCl_2$, 40 µM of a nested *Dnmt*-specific primer MMT2363AS (Fig. 3A), 40 µM of a poly-d(GI)-anchor primer, and 10 U of Amplitaq polymerase (Perkin-Elmer). The PCR reaction was isolated and cloned into the pCRII vector (Invitrogen).

The second method used 400 cm^2 of confluent J1 ES cells. RNA was prepared by extraction with guanidine thiocyanate and centrifugation in a cesium chloride step gradient, as described (24). RNA was passed once over an oligo$(dT)_{12-18}$-Cellulose column (Pharmacia) as described (24). 12 µg polyA+ mRNA was hybridized with 1 µg of MMT2214AS (Fig. 3A), a *Dnmt*-specific primer. A cDNA library was constructed using the TimeSaver cDNA Synthesis Kit (Pharmacia LKB), and ligated with lambda ZAP II phage arms (Stratagene). Recombinant phage were packaged and plated out as described (24). 125, 000 independent clones were screened using a probe containing *Dnmt* cDNA sequences upstream of the *Eco*RI site in the Exon 3 (Fig. 3A). Standard procedures for hybridization, washing, and autoradiography were observed (24). After a secondary screen, isolated positives were picked and converted into plasmids (Stratagene).

The first additional exon (Exon 2) was located in the genome by identity to genomic sequence reported by Rouleau *et al.* (1992). The second upstream exon (Exon 1) and the testis-specific exon were

localized by using the cDNA fragment containing Exon 1 as a probe for Southern blots of plasmids containing 129/Sv-derived *Dnmt* genomic sequences (7).

RESULTS

***Dnmt* cDNA is inefficiently expressed in stable transfections of ES cells.** To establish stable wild type and *Dnmt* mutant cell lines expressing recombinant MTase, four different expression constructs carrying the *Dnmt* cDNA were created (Fig. 1A). The ubiquitously-expressed *Pgk-1* promoter (25, 26) was used in pMT10 and the human EF-1α promoter (20) in pMT40 to direct expression of the *Dnmt* cDNA. A synthetic intron (18) was inserted into pMT10 to increase gene expression levels (27), creating pMT20. pMT30 was derived from pMT10 by replacing the *Pgk-1* promoter and the first 180 bp of the *Dnmt* cDNA with a 4.7-kb *Dnmt* genomic fragment containing a previously reported *Dnmt* promoter (15) and the excised 180 bp of cDNA, restored in its genomic context.

The four constructs were separately co-transfected into wild type *Dnmt* J1 ES cells (7) and the homozygous mutant $Dnmt^{s/s}$ (8) and $Dnmt^{n/n}$ cell lines (7). The $Dnmt^{s}$ allele is completely inactivated whereas the $Dnmt^{n}$ allele is a partial loss-of-function mutation. Cells bearing either mutation in a homozygous state stably maintain a low level of DNA methylation (7, 8) (Fig. 1C, lanes 2, 3). Drug resistant colonies were picked after 9-12 days of selection. Genomic DNA was extracted and analyzed by Southern blot using *Dnmt* genomic and cDNA probes, which revealed high copy number integrants with no gross rearrangements of the constructs (data not shown). 18 positive clones were analyzed for mRNA expression by Northern analysis, using the *Dnmt* cDNA as a probe. The size of the endogenous mRNA is 5.2 kb (Fig. 1B, lane 1), while the expected size

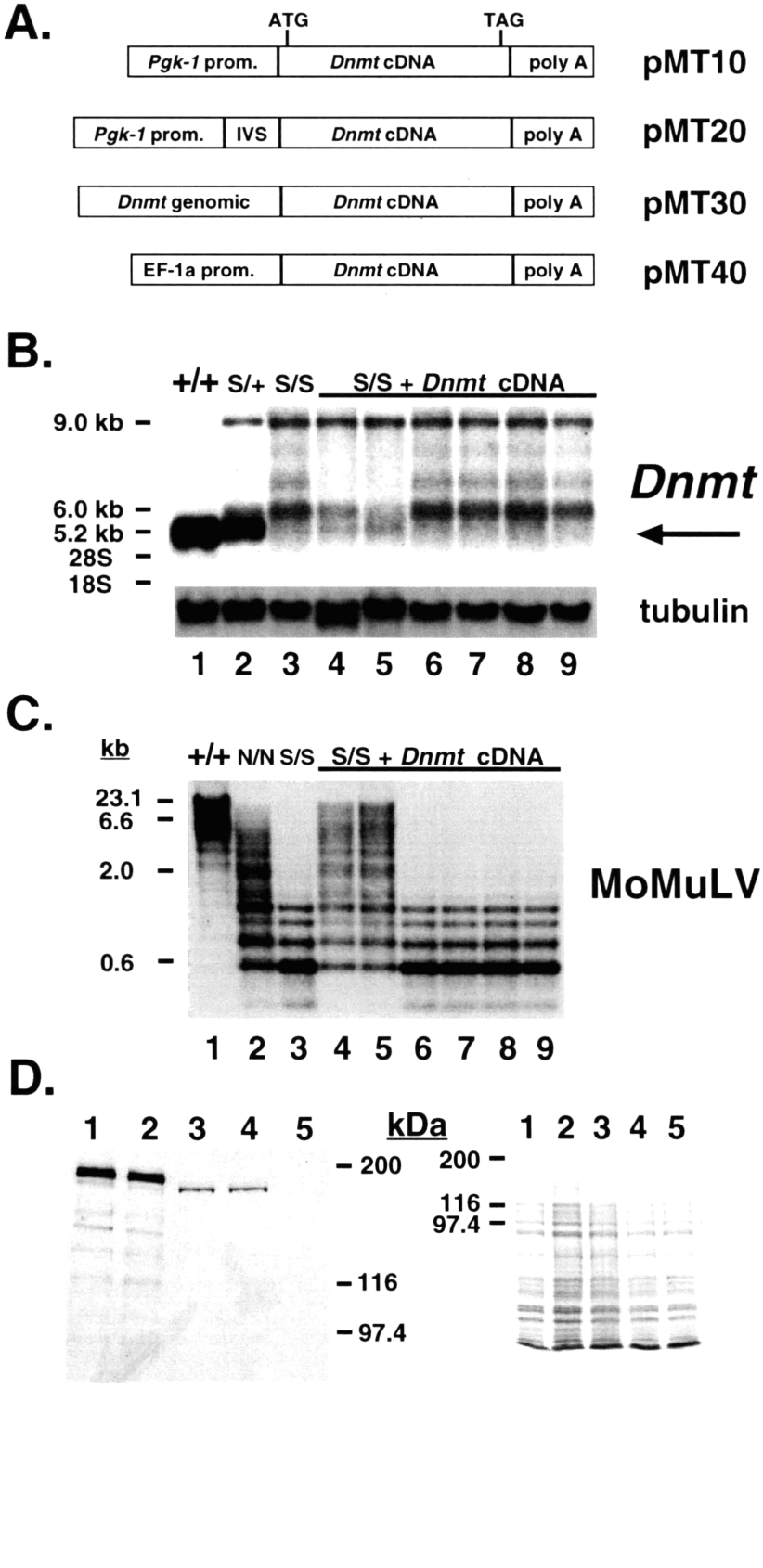

A.
ATG
TAG
Pgk-1 prom.
Dnmt cDNA
poly A
pMT10
Pgk-1 prom.
IVS
Dnmt cDNA
poly A
pMT20
Dnmt genomic
Dnmt cDNA
poly A
pMT30
EF-1a prom.
Dnmt cDNA
poly A
pMT40
B.
+/+ S/+ S/S S/S + Dnmt cDNA
9.0 kb
6.0 kb
5.2 kb
28S
18S
Dnmt
tubulin
1 2 3 4 5 6 7 8 9
C.
kb
+/+ N/N S/S S/S + Dnmt cDNA
23.1
6.6
2.0
0.6
MoMuLV
1 2 3 4 5 6 7 8 9
D.
1 2 3 4 5
kDa
200
116
97.4
1 2 3 4 5
200
116
97.4

Figure 1

from the transfected expression constructs was 5.1 kb. In the *Dnmt*$^{s/s}$ ES cells (lane 3) no 5.2 kb transcript was present, although aberrant transcripts of 9.0 and 6.0 kb representing abnormal splicing events caused by the insertion of the neomycin gene were observed (8). Figure 1B (lanes 4-9) shows a low level of expression in one-third of the *Dnmt*$^{s/s}$ clones transfected by pMT20 (lanes 4, 5), while the other transfectants (lanes 6-9) displayed no detectable expression. The decrease of the aberrant 6.0-kb transcript in clones expressing low levels of pMT20 cannot be explained at present (lanes 4, 5). Quantitation of band intensities using α-tubulin as a loading control (Fig. 1B) revealed that, in 5/18 clones, *Dnmt* cDNA expression levels were 5-10% of wild type *Dnmt*, whereas the remaining clones showed no detectable cDNA expression. Similarly, 23 J1 clones transfected with the same constructs showed no significant increase in levels of *Dnmt* expression over untransfected J1 ES cells (data not shown).

To investigate whether the low *Dnmt* cDNA expression levels would result in an increase in global DNA methylation, Southern blot analysis of genomic DNA digested with *Hpa*II, an enzyme sensitive to the methylation of internal cytosines in its CCGG recognition sequence, was performed. A Moloney Murine Leukemia Virus (MoMuLV) cDNA probe which cross-hybridizes with endogenous C-type retroviruses (28) that are heavily methylated in wild type cells was used (Fig. 1C, lane 1). DNA methylation in *Dnmt* mutant cells is greatly reduced, as reflected by the increasing intensity of lower molecular weight bands in this assay, with DNA methylation levels in the *Dnmt*$^{s/s}$ line (lane 3) lower than in the *Dnmt*$^{n/n}$ line (lane 2).

A slight increase in genomic methylation was detected only in those *Dnmt*$^{s/s}$ clones which had been transfected by either pMT20 (Fig. 1C, lanes 4,5) or pMT40 (data not shown) and showed plasmid expression of >5% of the wild type allele. The extent of remethylation was slightly higher than that seen in untransfected *Dnmt*$^{n/n}$ cell lines. By this assay, 8/19 *Dnmt*$^{s/s}$ clones containing a *Dnmt* cDNA displayed partial genomic remethylation, and this methylation level was stably maintained through five passages (data not shown). Also, inclusion into the expression construct pMT30 of the upstream genomic *Dnmt* sequences which had previously been reported to drive transcription (15) did not lead to genomic remethylation in *Dnmt*$^{n/n}$ cells (0/24 clones).

Western blot analysis, using an antibody directed against the amino terminus of the protein, indicated levels of MTase approximately 5-10% that of wild type in pMT20-transfected clones (Fig. 1D). Moreover, this recombinant MTase (lanes 3, 4) was shorter by approximately 8 kilodaltons than the protein seen in J1 cells (lane 1). We conclude that, using either heterologous promoters or the reported cognate *Dnmt* promoter (15), only low levels of the *Dnmt* cDNA are expressed in stably transfected wild type or mutant ES cells.

***Dnmt* cDNA is efficiently expressed from carboxy-terminal truncated constructs.** To test the hypothesis that the low mRNA expression was caused by the protein product of the *Dnmt* expression constructs being detrimental to cells, a construct was designed such that translation of the *Dnmt* cDNA would end

prematurely, producing truncated and possibly inactive peptides. A stop codon-containing linker was inserted into pMT40 at the unique *Nsi*I site 871 bp downstream of the initial ATG in the *Dnmt* cDNA, to create pMT42 (Fig. 2). This site is located in the middle of the domain which targets MTase to sites of DNA synthesis in S-phase nuclei (29). Translation of this message would be expected to produce a truncated protein of 292 amino acids with neither catalytic activity nor the ability to colocalize with replication foci. $Dnmt^{s/s}$ ES cells were transfected and individual colonies were examined by Northern blot analysis. In contrast to a typical clone carrying the wild type cDNA (Fig 2, lane 4), high levels of RNA were produced in 4 / 8 clones carrying the mutant cDNA construct (Fig. 2, lanes 7, 8, 11, and 12). As expected, none of the transfectants showed genomic remethylation (data not shown). These results are consistent with a detrimental effect of the protein produced from the original expression constructs resulting in selection against cells expressing high levels of the truncated *Dnmt* cDNA. This conclusion was corroborated by a reduced number of clones obtained with pMT 40 containing the ORF as compared to pMT 42 carrying the stop codon: transfection of $5x10^6$ cells with the latter clone yielded 500 colonies as compared to only 200 seen with pMT40.

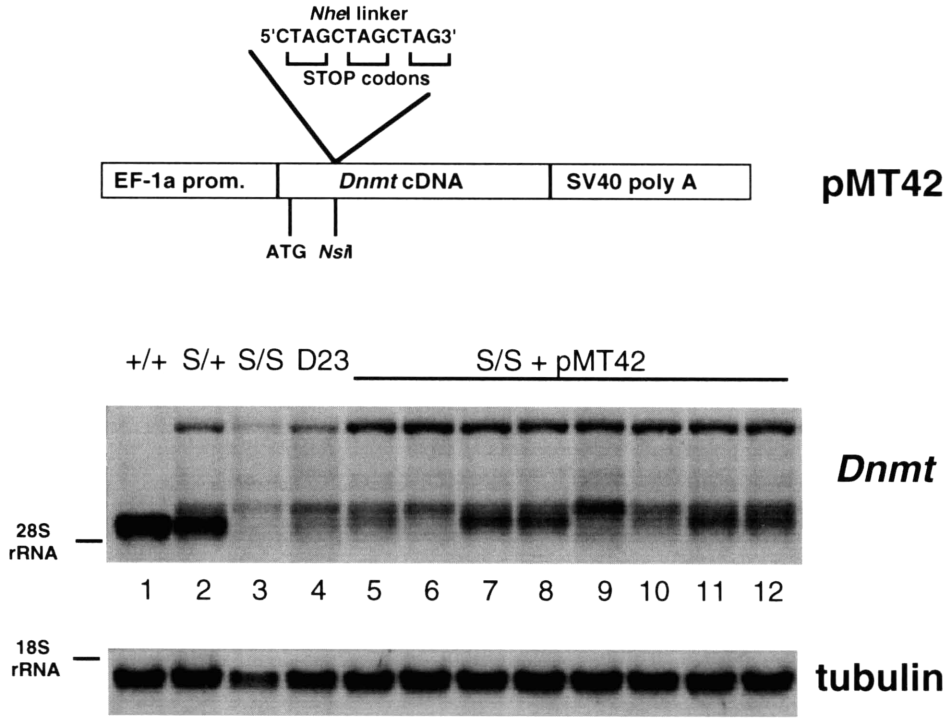

NheI linker
5'CTAGCTAGCTAG3'
STOP codons
EF-1a prom.
Dnmt cDNA
SV40 poly A
pMT42
ATG NsiI
+/+ S/+ S/S D23
S/S + pMT42
Dnmt
28S rRNA
1 2 3 4 5 6 7 8 9 10 11 12
18S rRNA
tubulin

Figure 2

The complete *Dnmt* cDNA contains two additional 5' exons which extend the open reading frame by at least 118 codons. The results described above suggested that high expression of the *Dnmt* cDNA was not tolerated in transfected cells. Given the results of Western blot analysis (Fig. 1D), it appeared that the cDNA was incomplete, encoding a truncated protein which was toxic to cells. Therefore, we attempted to derive further upstream sequences. First, CD1 kidney mRNA was used in 5' RACE experiments that employed the *Dnmt*-specific primer MMT2400AS for first strand cDNA synthesis. Subsequent PCR amplification of the cDNA product with the *Dnmt*-specific primer MMT2363AS (Fig 3A) yielded a product which included two additional exons.

Independently, polyA$^+$ mRNA from J1 ES cells was used for the construction of a cDNA library. The *Dnmt*-specific primer MMT2214AS (Fig. 3A) was employed for first strand synthesis, and 10^5 phage clones were screened with a probe containing *Dnmt* cDNA sequences upstream of the *Eco*RI site in the third exon (Fig. 3A). Examination of 12 clones revealed the same sequence cloned by 5' RACE, with the longest of the clones extending the 5' RACE product by 34 bp. In contrast to previous reports of *Dnmt* sequence (4, 15), close examination identified a 141-codon extension of the published open reading frame (ORF), including a new initial ATG 118 codons upstream of the previously-published initial ATG (Fig. 3A).

Comparison with the published *Dnmt* genomic DNA sequence (15) revealed that the new cDNA sequence comprises a 15-bp 5' extension of the published first exon (Exon 3) and two new exons (Exons 2 and 1, respectively; Fig. 3B). The 37-bp exon 2 is separated

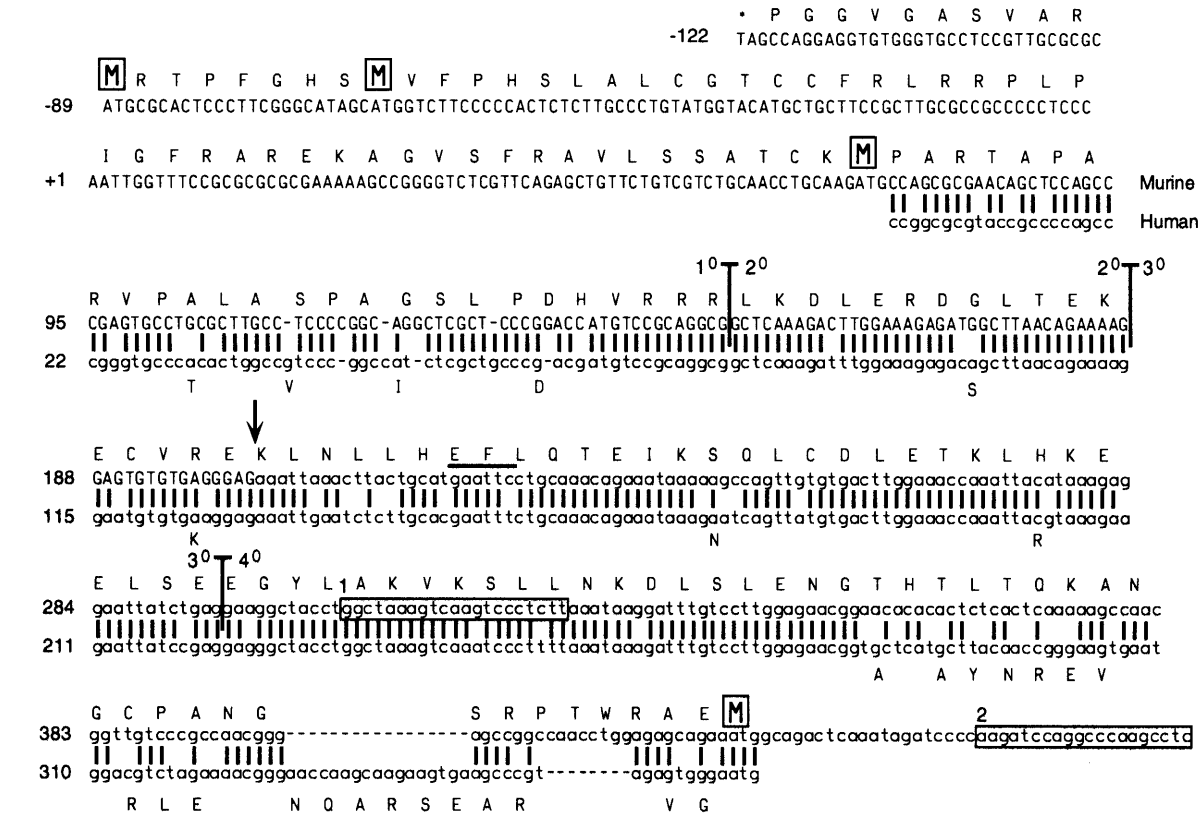

A.
Murine
Human
B.
Somatic splicing
Testis-specific splicing
ATG
ATG
10 kb
200 bp
CpG
GpC

Figure 3

from the 150-bp exon 3 by a 767-bp intron, while exon 1 lies 11.5 kb upstream of exon 2. Exon 1 was found in the center of a CpG island (Fig. 3B) and hence provides the *Dnmt* gene with the typical elements of a housekeeping gene. Examination of the genomic sequence around the new primary exon revealed a continuation of the ORF for another 122 bp upstream of the 5' limit of the current cDNA sequence. This additional genomic sequence begins with an in-frame stop codon and contains two in-frame ATGs (Fig. 3A). An alternatively-spliced upstream exon isolated both from mouse testis and HeLa cells (30) (EMBL Accession # X77486) was found to lie 800 bp downstream of exon 1 (Fig. 3B). These results show that tissue-specific and alternative transcriptional start sites in mouse and human tissues are localized upstream of the previously characterized 5' end of the gene (15).

A comparison of the new murine amino terminal cDNA sequence with the published human cDNA (31) showed that these sequences are highly conserved, especially at the beginning of the ORF (Fig. 3A). Similar to the murine sequence, the human sequence contains an ORF upstream of the published ATG initiation codon. In the first shared 282 bp, the respective gene sequences are 85% identical at the nucleotide level, and 91% identical at the amino acid level. 33 / 41 nucleotide mismatches between the two sequences result in silent mutations. This evolutionary conservation may indicate an important and essential function of this amino-terminal sequence, consistent with the data presented here.

Expression of complete *Dnmt* cDNA in *Dnmt* mutant ES cells restores normal genomic methylation levels. To test the possibility that the newly-discovered cDNA sequence would allow normal expression and MTase functioning *in vivo*, a mini-gene expression construct was made and stably transfected into *Dnmt* mutant ES cells. pMT50 was made by insertion of a 5.0-kb region of the *Dnmt* locus that included the first exon into pMT30 (Fig. 4A). pMT30 contained the previously reported *Dnmt* promoter (15) but had been shown not to function in *Dnmt* mutant ES cells (Fig. 1A). pMT50 resulted in the *Dnmt* cDNA fused 3' to a 9.7-kb piece of genomic *Dnmt* sequence that ended in the fourth exon and included 3.6 kb of genomic sequence upstream of the first exon, the two new exons reported above, three introns, and the beginning of the fourth exon.

Dnmt$^{s/s}$ cells were electroporated with the linearized construct and selected for puromycin resistance. Puromycin-resistant colonies were expanded and analyzed by Southern blot for vector integration and levels of genomic methylation. 7/12 clones examined showed random integration into the genome (data not shown) and normal levels of genomic methylation (Fig. 4B), as compared to wild type cells. 2/12 clones had undergone an homologous targeting of the construct to the endogenous *Dnmt* locus and the remaining 3 clones had no transgene (data not shown). All clones carrying an intact transgene displayed equivalent methylation profiles at repetitive gene sequences and *Xist*. The imprinted genes *H19* and *Igf2r*, however, were not remethylated (data not shown), as reported previously for a smaller, similar *Dnmt* cDNA-containing plasmid that

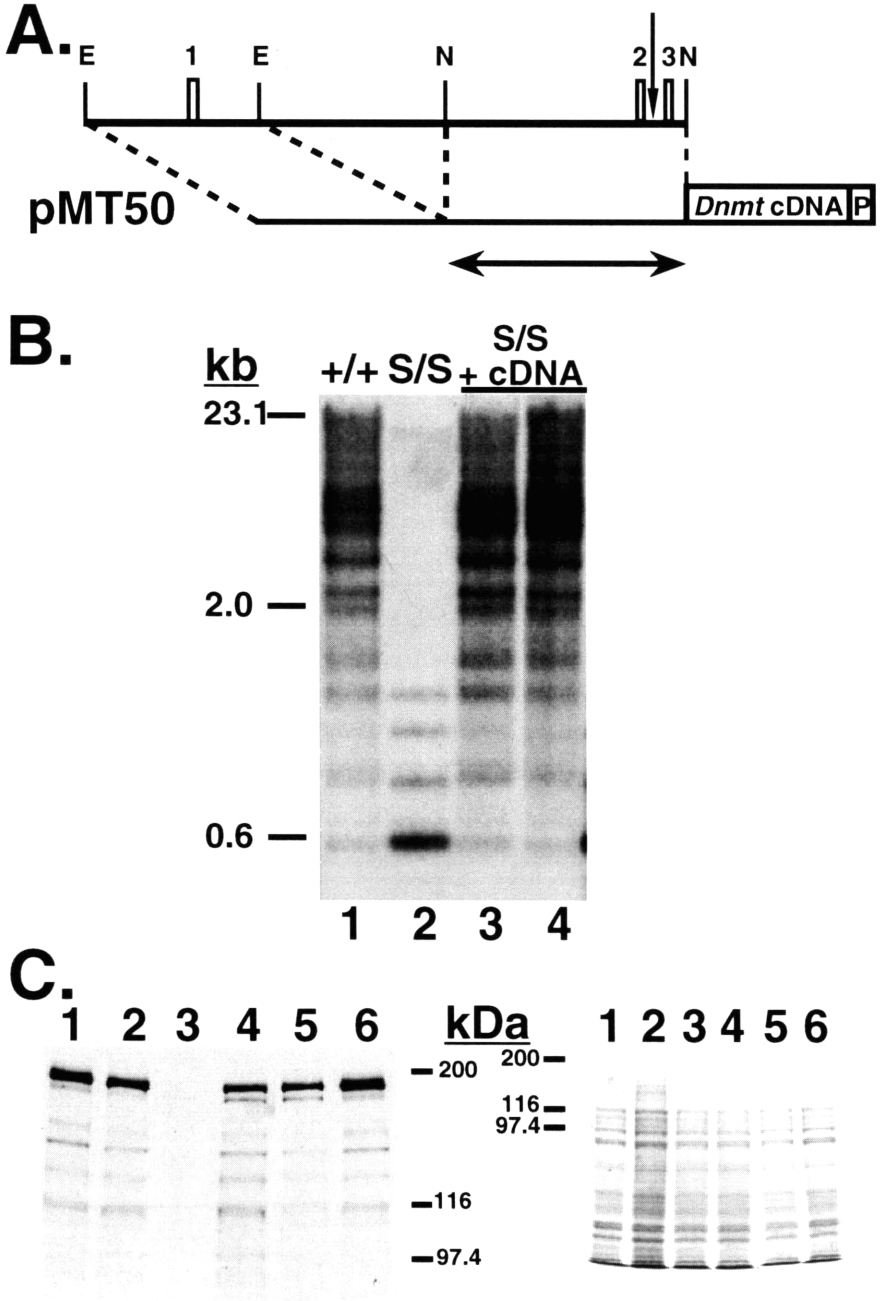

A.
E
1
E
N
2
3N
pMT50
Dnmt cDNA
P
B.
kb
+/+
S/S
S/S
+ cDNA
23.1
2.0
0.6
1
2
3
4
C.
1 2 3 4 5 6
kDa
200
116
97.4
1 2 3 4 5 6

Figure 4

was targeted to the endogenous *Dnmt* locus (16). Western blot analysis revealed a range of MTase levels from 10-100% that of a wild type allele, with the minigene-encoded enzyme co-migrating with wild type MTase (Fig. 4C).

Three of the rescued lines were used to induce teratomas by subcutaneous injection into syngeneic male host animals. While the parental mutant cells failed to form palpable teratomas within 3-4 weeks (n = 0 / 12), the rescued ES cells induced teratoma formation at a similar rate (n = 9 / 10) as the wild type J1 ES cells (n = 6 / 6). Histological analysis of these teratomas revealed a wide variety of differentiated cell types, consisting mostly of mature neural tissue and including discrete patches of epithelial tissue, cartilage, trophectoderm, osteoid, and a relative deficiency of striated muscle (data not shown). From these results we conclude that the additional sequences uncovered above are necessary for functional activity of MTase *in vivo*.

DISCUSSION

The experiments in this study were prompted by the failure of published *Dnmt* cDNA clones to functionally complement the hypomethylation phenotype of ES cells homozygous for various mutations of the *Dnmt* gene. Our results allow two main conclusions regarding the structure and function of DNA methyltransferase. (i) Stably transfected ES cells failed to express high levels of the previously described *Dnmt* cDNA clone, whereas expression of the full-length cDNA led to wild type expression levels. We conclude that the previously described cDNA is truncated, lacking an open reading frame (ORF) of up to 171 codons upstream of the ATG which was previously proposed to represent the initiation codon (4). (ii) Genomic methylation levels in ES cells, which were decreased substantially as a result of the targeted inactivation of the *Dnmt* gene, were restored to normal levels by expression of the *Dnmt* mini-gene. This restoration of methylation was accompanied by functional rescue of the transfected cells, as reflected in the ability to form teratomas. Rescued cells behave identically to those reported previously (16), in which the *Dnmt* cDNA was targeted to its cognate locus. We conclude that the inclusion of the newly-identified sequence is essential for stable expression and methyltransferase activity *in vivo*.

Functional rescue of the *Dnmt* mutant phenotype suggests that normal methylation patterns can by restored so as to assure appropriate gene expression in differentiating ES cells. De novo methylation of the demethylated genome is likely accomplished by

an independently encoded DNA methyltransferase, as discussed elsewhere (8, 16). Our results argue that demethylation of the mutant ES cell genome does not result in irreversible alteration in gene activity and / or chromatin structure. Rescue of the mutant phenotype seems to be the result of a default *de novo* methylation process which requires the expression of the hemi-methyltransferase encoded by *Dnmt,* which would maintain methylation levels set by the embryonic *de novo* methyltransferase.

Both the low expression levels in cells carrying the truncated cDNA in stably transfected ES cells, and the restoration of high levels of expression from constructs containing an in-frame stop codon in the ORF, were consistent with a truncated polypeptide being produced from the cDNA, translation of which may be detrimental to ES cells. Indeed, we isolated additional upstream cDNA sequence comprising two exons separated by an 11.5-kb intron. Examination of the genomic sequence surrounding this new primary exon revealed an uninterrupted open reading frame that begins upstream of and continues throughout the two new exons, extending the published coding sequence by 171 codons. Three new in-frame ATG codons were found in this ORF. Lack of these extra sequences could account for the size differences between MTase purified from cell lines and the product expressed from the truncated *Dnmt* cDNA (12). Comparison of the mobility in SDS-PAGE of wild type MTase with the recombinant MTase produced from the expression constructs (Fig. 1D, Fig. 4C) suggests the third ATG in the ORF most likely represents the translational start site. Although the results do not determine which of the three upstream ATGs is used as a translational start point *in*

vivo, they do show that the new ORF is necessary for proper functioning *in vivo*. Protein sequencing will resolve which ATG is used as a translational start.

The first exon was found to lie within a CpG island, as defined by Bird (32). Because of the GC-rich nature of the sequence, it is difficult to ascertain whether the 5' end heterogeneity of the 12 examined clones reflects pausing by the reverse transcriptase or a multiplicity of transcriptional start sites, as often observed in housekeeping promoters associated with a CpG island. Our results are not in agreement with the localization of the *Dnmt* promoter localized upstream of exon 3, as reported by Rouleau *et al.* (15). Expression from this promoter would produce a truncated MTase which is detrimental to stably transfected cells, and *Dnmt* expression constructs utilizing this promoter failed to effect complementation in *Dnmt* mutant cells. Instead we suggest that a promoter contained within the pMT50 construct represents the functionally relevant transcriptional start site.

Several reports have demonstrated the ability to express high levels of the truncated *Dnmt* cDNA in a transient fashion (12, 13). In contrast, only a few reports have described high levels of MTase expression in stably transfected clones of transformed cells (14, 33) or in a myoblast cell line (34)□. This raises the question of whether specific functional domains of the polypeptide may explain the low *Dnmt* expression and the apparent toxicity of the truncated cDNA when stably expressed in ES cells. It has been shown previously that the truncated *Dnmt* cDNA, when expressed in COS cells, generates an enzyme which methylates a hemimethylated DNA substrate with the

same kinetics as MTase purified from MEL cells (12). Proteolytic cleavage of the first 350 amino acids did not affect *in vitro* enzyme activity (35). Consistent with these observations, expression of the truncated *Dnmt* cDNA resulted in limited genomic remethylation in $Dnmt^{s/s}$ ES cells, raising the possibility that the detrimental effect of truncated MTase in ES cells can be separated from its methylating function. MTase is localized to nuclear DNA replication sites during S-phase by an amino-terminal targeting domain (29). Because insertion of stop codons within this targeting domain abolished the toxic effect of truncated cDNA expression, we speculate that toxicity may be mediated through this domain, possibly by targeting the truncated MTase to replication sites and thereby stalling the replication machinery. The isolation of a complete *Dnmt* cDNA allows for analysis of these and many other biological questions, which can finally be addressed through overexpression and tissue-specific expression of the enzyme. The identification of a new transcriptional start site also reopens the discussion on the regulation of *Dnmt* gene expression, which is known to be complex.

We thank Tim Bestor for pMG and Stuart Orkin for pEF-PGKhyg. K. T. was supported by the NIH / National Cancer Institute Training Grant T32-CA09541. D. T. was supported by the Human Frontier Science Program. H. L. was supported by the Council for Tobacco Research, U. S. A. R. J. was supported by the National Institutes of Health grant R35-CA44339.

REFERENCES

1. Yeivin, A. & Razin, A. (1993) in *DNA methylation: molecular biology and biological significance*, eds. Jost, J. P. & Saluz, H. P. (Birkhauser Verlag, Basel, Switzerland),

2. Jones, P. A., Rideout, W. M., Shen, J. C., Spruck, C. H. & Tsai, Y. C. (1992) *Bioessays* **82**, 9-12.

3. Barlow, D. P. (1995) *Science* **270**, 1610-1613.

4. Bestor, T., Laudano, A., Mattaliano, R. & Ingram, V. (1988) *J Mol Biol* **203**, 971-83.

5. Gruenbaum, Y., Cedar, H. & Razin, A. (1982) *Nature* **295**, 620-2.

6. Leonhardt, H. & Bestor, T. H. (1993) in *DNA methylation: molecular biology and biological significance*, eds. Jost, J. P. & Saluz, H. P. (Birkhauser Verlag, Basel, Switzerland),

7. Li, E., Bestor, T. H. & Jaenisch, R. (1992) *Cell* **69**, 915-26.

8. Lei, H., Oh, S. P., Jüttermann, R., Jaenisch, R., Goss, K. A. & Li, E. (in press) *Development*

9. Li, E., Beard, C. & Jaenisch, R. (1993) *Nature* **366**, 362-5.

10. Laird, P. W. & Jaenisch, R. (1994) *Hum Molec Genet* **3**, 1487-1495.

11. Laird, P. W., Jackson-Grusby, L., Fazeli, A., Dickinson, S., Jung, W. E., Li, E., Weinberg, R. A. & Jaenisch, R. (1995) *Cell* **81**, 197-205.

12. Czank, A., Hauselmann, R., Page, A. W., Leonhardt, H., Bestor, T. H., Schaffner, W. & Hergersberg, M. (1991) *Gene* **109**, 259-63.

13. Glickman, J. F. & Reich, N. O. (1994) *Biochem Biophys Res Commun* **204**, 1003-8.

14. Wu, J., Issa, J. P., Herman, J., Bassett, D. E., Jr., Nelkin, B. D. & Baylin, S. B. (1993) *Proc Natl Acad Sci U S A* **90**, 8891-5.

15. Rouleau, J., Tanigawa, G. & Szyf, M. (1992) *J Biol Chem* **267**, 7368-77.

16. Tucker, K. L., Beard, C., Dausman, J., Jackson-Grusby, L., Laird, P. W., Lei, H., Li, E. & Jaenisch, R. (1996) *Genes Dev.* **10**, 10008-20.

17. van der Lugt, N., Maandag, E. R., te Riele, H., Laird, P. W. & Berns, A. (1991) *Gene* **105**, 263-7.

18. Huang, M. & Gorman, C. (1990) *Nucleic Acids Res* **18**, 937-47.

19. Zhen, L., King, A. A. J., Xiao, Y., Chanock, S. J., Orkin, S. H. & Dinauer, M. C. (1993) *Proc Natl Acad Sci USA* **90**, 9832-9836.

20. Mizushima, S. & Nagata, S. (1990) *Nucleic Acids Res* **18**, 5322.

21. Rudnicki, M. A., Braun, T., Hinuma, S. & Jaenisch, R. (1992) *Cell* **71**, 383-90.

22. Laemmli, U. K. (1970) *Nature* **227**, 680-685.

23. Vassalli, A., Matzuk, M. M., Gardner, H. A. R., Lee, K.-F. & Jaenisch, R. (1994) *Genes Dev.* **8**, 414-427.

24. Sambrook, J., Fritsch, E. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.).

25. McBurney, M. W., Sutherland, A. C., Adra, C. N., Leclair, B., Rudnicki, M. A. & Jardine, K. (1991) *Nucleic Acids Res* **19**, 5755-61.

26. McBurney, M. W., Staines, W. A., Boekelheide, K., Parry, D., Jardine, K. & Pickavance, L. (1994) *Dev Dynamics* **200**, 278-293.

27. Choi, T., Huang, M., Gorman, C. & Jaenisch, R. (1991) *Mol Cell Biol* **11**, 3070-4.

28. Stuhlmann, H., Jahner, D. & Jaenisch, R. (1981) *Cell* **26**, 221-32.

29. Leonhardt, H., Page, A. W., Weier, H. U. & Bestor, T. H. (1992) *Cell* **71**, 865-73.

30. Kuemmerle, N. B., Hake, L. E. & Valerie, K. (submitted)

31. Yen, R. W., Vertino, P. M., Nelkin, B. D., Yu, J. J., el-Deiry, W., Cumaraswamy, A., Lennon, G. G., Trask, B. J., Celano, P. & Baylin, S. B. (1992) *Nucleic Acids Res* **20**, 2287-91.

32. Bird, A. P. (1986) *Nature* **321**, 209-13.

33. Vertino, P.M., Yen, R.-W.C., Gao, J. & Baylin, S.B. (1966) *Mol Cell Biol* **16**, 4555-4565.

34. Takagi, H., Tajima, S. & Asano, A. (1995) *Eur J Biochem* **231,** 282-291.

35. Bestor, T. H. & Ingram, V. M. (1985) *Proc Natl Acad Sci U S A* **82**, 2674-8.

FIGURE LEGENDS

FIG. 1. *Dnmt* cDNA is inefficiently expressed in stable transfections of ES cell lines. (*A*) Diagram of *Dnmt* expression constructs. The cDNA used in all four constructs is the same. See Materials and Methods for details of plasmid construction. (*B*) Northern blot analysis of *Dnmt* expression in $Dnmt^{s/s}$ lines transfected with pMT20 (lanes *4-9*). The wild type *Dnmt* transcript migrates at 5.2 kb and is indicated by an arrow to the right of the blot. Expression levels are <1% of the wild type allele in all clones shown, except for lanes *4* and *5*, which are 8 and 10 %, respectively. Wild type J1 (lane *1*), $Dnmt^{s/+}$ (lane *2*), and $Dnmt^{s/s}$ (lane *3*) cells are shown for comparison. The *Dnmt* cDNA was used as a probe. The lower figure shows the same blot re-hybridized with an α-tubulin cDNA (34) to control for amount of RNA loaded. The bands migrating at 9.0 and 6.0 kb are created by the disruption at the *Sal*I site by a *Pgk-Neo* cassette, and hybridize with a *Neo* probe (8). (*C*) Southern blot analysis of *Hpa*II-digested genomic DNA from $Dnmt^{s/s}$ lines transfected with pMT20 (lanes *4-9*, same order as above). The Moloney Murine Leukemia Virus (MoMuLV) cDNA (28) was used as a probe. Wild type J1 (lane *1*), $Dnmt^{n/n}$ (lane *2*), and $Dnmt^{s/s}$ (lane *3*) cells are shown for comparison. (*D*) Western blot analysis from $Dnmt^{s/s}$ lines transfected with pMT20, which show slight levels of genomic remethylation (lanes *3* and *4* correspond to lanes *4, 5* in Fig. 1B/C). An amino-terminal anti-MTase antibody was used in the left figure, and Coomassie-stained gel shown as a loading control to the right. Wild type J1 (lane *1*), $Dnmt^{s/+}$ (lane *2*), and $Dnmt^{s/s}$ (lane *5*) cells are

shown for comparison. Biorad Silver Stain molecular weight markers are shown to the side.

FIG. 2. *Dnmt* cDNA is efficiently expressed from carboxy-terminal truncated constructs. The top figure shows pMT42, made by inserting a stop-codon containing linker into the *Nsi*I site in pMT40, 871 bp downstream of the initial ATG. Shown beneath is a Northern blot analysis of *Dnmt* expression in $Dnmt^{s/s}$ lines transfected with pMT42. 4/8 of the examined transfectants (lanes *5-12*) showed an expression level approximately 50% that of the wild type cells. The following lines are shown for comparison: wild type J1 (lane *1*), $Dnmt^{s/+}$ (lane *2*), $Dnmt^{s/s}$ (lane *3*), and D23 (lane *4*), a pMT40-transfected $Dnmt^{s/s}$ cell line showing slight levels of *Dnmt* cDNA expression and genomic remethylation. The *Dnmt* cDNA was used as a probe. The lower figure represents the same blot re-hybridized with an α-tubulin cDNA to control for amount of RNA loaded. *S* allele-specific bands are as described in Fig. 1B.

FIG. 3. The complete *Dnmt* cDNA contains two additional 5' exons which extend the open reading frame by up to 171 codons. (*A*) Comparison between murine and human *Dnmt* sequences at the 5' end. New sequence reported in this paper is shown in uppercase letters, with previously reported sequence in lowercase. The first two rows of sequence are genomic sequence, while the seven rows beneath represent cDNA sequence. A previously reported human cDNA (31) is shown beneath the murine sequence with all matches represented by a vertical bar. A predicted translation of the 182-

codon extension of the ORF is shown above the murine nucleotide sequence, starting with the stop codon TAG and with the four methionine residues boxed. Differences in the human amino-acid sequence are represented beneath the human nucleotide sequence. Vertical bars running through the murine sequence denote exon-exon boundaries, with the numbers corresponding to the exon designations used in Fig. 3B. The two boxed sequences represent the primers MMT2214AS (1) and MMT2363AS (2) used in the cloning experiments. The *Eco*RI site that marks the beginning of the *Dnmt* cDNA used in the expression constructs in Fig. 1 is indicated by a bar over the recognition sequence. The major transcriptional start site found by Rouleau *et al.* (15) is indicated by a vertical arrow. (*B*) Map of exonic structure of the 5' end of *Dnmt*. The first four exons of *Dnmt* are represented to scale with respect to their size and relative position in the gene. Exon numbers described in text are given above. The splicing of the first four exons of *Dnmt* is diagrammed above, while the alternative splicing observed in mouse testis (30) is shown beneath. The numbering of the nucleotide scale starts with 0 as the 5' end of the new cDNA sequence. The two alternative exons are separated from the second, common exon by a 10-kb interval of genomic DNA, indicated by the broken lines in the splicing diagram. CpG and GpC incidence diagrams are plotted beneath, to scale with the rest of the diagram. The two 3'-proximal ATGs indicated by a boxed M in Fig. 3A are shown here in their genomic locations, while the first two ATGs are indicated by vertical arrows to the left of Exon 1. The major transcriptional start site found by Rouleau *et al.* (15) is indicated by a horizontal arrow beneath Exon 3.

FIG. 4. Restoration of normal genomic DNA methylation levels upon expression of the full-length *Dnmt* cDNA. (*A*) Map of pMT50, showing its relation to *Dnmt* genomic sequence and the exons therein (drawn to scale). The first three exons of *Dnmt* are represented by numbered boxes and correspond to those shown in Fig. 3B. The vertical arrow represents the transcriptional start site reported in Rouleau *et al.* (15). *Eco*RI (E) and *Nae*I (N) sites used in cloning are indicated. The *Dnmt* cDNA and *Pgk-1* polyadenylation signal (P) are both represented by boxes. The amount of *Dnmt* genomic sequence used in pMT30 is shown beneath by the arrow. (*B*) Southern blot analysis of *Hpa*II-digested genomic DNA, probed with the MoMuLV cDNA. Wild type J1 (lane *1,*) and $Dnmt^{s/s}$ cells (lane *2*) are shown, followed by pMT50-transfected $Dnmt^{s/s}$ cells (lanes *3, 4*). (*C*) Western blot analysis of the pMT50-transfected cells shown in Fig. 4B (lanes *4, 5*). An amino-terminal anti-MTase antibody was used in the left figure, and Coomassie-stained gel shown as a loading control to the right. Wild type J1 (lane *1, 6*), $Dnmt^{s/+}$ (lane *2*), and $Dnmt^{s/s}$ (lane *3*) cells are shown for comparison. Biorad Silver Stain molecular weight markers are shown to the side.

Appendix B

Germ line passage is required for establishment of methylation and expression patterns of imprinted but not of non-imprinted genes

Kerry L. Tucker[1,2], Caroline Beard[1], Jessica Dausman[1], Laurie Jackson-Grusby[1], Peter Laird[1], Hong Lei[3], En Li[3], and Rudolf Jaenisch[1,2]

[1]Whitehead Institute for Biomedical Research

[2]Massachusetts Institute of Technology, Dept. of Biology, Cambridge, MA 02142

[3]Massachusetts General Hospital East, Cardiovascular Research Center, 149 13th St., Charlestown, MA 02129

Running title: DNA methylation and the establishment of genomic imprinting

Keywords: DNA methylation, genomic imprinting, DNA methyltransferase, *H19*, *Igf2r*, *Xist*

ABSTRACT

Embryonic stem (ES) cells homozygous for a disruption of the DNA (cytosine-5)-methyltransferase gene (*Dnmt*) proliferate normally with their DNA highly demethylated but die upon differentiation. Expression of the wild type *Dnmt* cDNA in mutant male ES cells caused an increase in methylation levels of bulk DNA and of the *Xist* and *Igf2* genes to normal levels, but did not restore the methylation of the imprinted genes *H19* and *Igf2r*. These cells differentiated normally *in vitro* and contributed substantially to adult chimeras. While the *Xist* gene was not expressed in the remethylated male ES cells, no restoration of the normal expression profile was seen for *H19*, *Igf2r*, or *Igf2*. This indicates that ES cells can faithfully reestablish normal methylation and expression patterns of non-imprinted genes but lack the ability to restore those of imprinted genes. Full restoration of monoallelic methylation and expression was imposed on *H19*, *Igf2* and *Igf2r* upon germline transmission. These results are consistent with the presence of distinct *de novo* DNA methyltransferase activities during oogenesis and spermatogenesis which specifically recognize imprinted genes but which are absent in the post-implantation embryo and in ES cells.

INTRODUCTION

Much evidence indicates that in vertebrates the methylation of DNA at cytosine residues correlates with gene inactivity (Yeivin and Razin, 1993). Methylation patterns are generally stably maintained in somatic cells of the animal but change dramatically during early development when the genome of the mammalian embryo undergoes consecutive waves of demethylation and *de novo* methylation (Razin and Shemer, 1995). The methylation present in the DNA of the zygote is erased during cleavage so that the genome of the blastocyst is almost completely unmethylated (Monk et al., 1987; Kafri et al., 1992). After implantation, the embryo undergoes a wave of global *de novo* methylation which restores the genomic methylation levels of the gastrulating embryo to that seen in the adult (Jähner et al., 1982; Monk et al., 1987). During gametogenesis, the unmethylated DNA of the primordial germ cells becomes *de novo* methylated (Kafri et al., 1992; Monk et al., 1987), a process which ultimately leads to the high methylation levels of the zygote. *De novo* methylation of transduced retroviral genomes is also seen in pregastrulation embryos (Jähner et al., 1982) and in embryonic carcinoma (EC) and stem (ES) cells (Stewart et al., 1982; Lei et al., in preparation), consistent with the hypothesis that a *de novo* DNA methyltransferase activity is expressed in early stages of development.

The crucial importance of DNA methylation in development was demonstrated by the targeted mutation of the single known mammalian DNA methyltransferase gene (*Dnmt*) in embryonic stem cells (Li et al., 1992). *Dnmt* is believed to encode a maintenance

methyltransferase, because *in vitro* the enzyme (MTase) shows a substrate preference for hemimethylated DNA (Bestor and Ingram, 1983; Gruenbaum et al., 1982). Three mutant alleles of the *Dnmt* gene have been generated (Lei et al., in preparation; Li et al., 1993; Li et al., 1992). A partially inactivated allele ($Dnmt^n$) causes lethality of homozygous mutant embryos at the 25 somite stage, whereas embryos homozygous for completely inactivated alleles ($Dnmt^s$ and $Dnmt^c$) die at the 5 somite stage. ES cells homozygous for any of the alleles proliferate normally with their genomic DNA highly demethylated but die upon induction of differentiation. ES cells homozygous for any disruption, however, retain the ability to *de novo* methylate transduced DNA, supporting the notion of an independently-encoded *de novo* DNA methyltransferase being expressed in embryonic cells (Lei et al., in preparation).

Imprinted genes display parent-specific monoallelic expression (Barlow, 1995; Surani, 1994), although often not at all developmental stages (Szabo and Mann, 1995). Genes which are predominantly expressed from the maternal allele include *Igf2r* (Barlow et al., 1991), *H19* (Bartolomei et al., 1991), *Mash2* (Guillemot et al., 1995) and $p57^{KIP2}$ (Hatada and Mukai, 1995), while *Igf2* (DeChiara et al., 1991), *Mas* (Villar and Pedersen, 1994), *Snrpn* (Leff et al., 1992), *U2afbp-rs* (Hayashizaki et al., 1994), *ins1/ins2* (Giddings et al., 1994), *IPW* (Wevrick et al., 1994), and *Peg1/Mest* (Kaneko-Ishino et al., 1995) are expressed paternally. DNA methylation has been proposed to constitute the molecular mark which distinguishes the two alleles of imprinted genes. Evidence for DNA methylation constituting the imprinting mark includes differential methylation of

imprinted transgenes (Chaillet et al., 1991; Reik et al., 1987; Sapienza et al., 1987; Swain et al., 1987) and imprinted endogenous genes. The acquisition and developmental propagation of the methylation mark has been described in detail for the two imprinted genes *H19* and *Igf2r*, and for *Xist*, a gene encoded at the center for X chromosome inactivation which shows imprinted expression only in preimplantation embryos and extraembryonic tissues (Kay et al., 1993). The *H19* gene acquires paternal-specific methylation in a region upstream of its promoter during spermatogenesis (Tremblay et al., 1995), while *Igf2r* (Stöger et al., 1993) and *Xist* (Ariel et al., 1995; Zuccotti and Monk, 1995) acquire a maternal-specific methylation mark during oogenesis. Importantly, the allele-specific methylation differences in these "imprinting boxes" are resistant to the wave of global demethylation during cleavage and the subsequent wave of global *de novo* methylation before gastrulation (Bartolomei et al., 1993; Brandeis et al., 1993; Stöger et al., 1993; Tremblay et al., 1995). This assures that the methylation mark imposed on the two alleles of imprinted genes during gametogenesis remains unaltered during development and, therefore, distinguishes the two alleles in each cell of the organism. In contrast to *H19* and *Igf2r*, the methylation imprint of the maternal *Xist* allele is erased after implantation in cells of the embryonic lineage, and either the paternal or the maternal allele becomes methylated at the time of X inactivation (Norris et al., 1994).

Strong evidence for methylation being crucial for the maintenance of monoallelic expression of imprinted genes was provided by the analysis of *Dnmt* mutant embryos (Li et al., 1993; Beard et al., 1995).

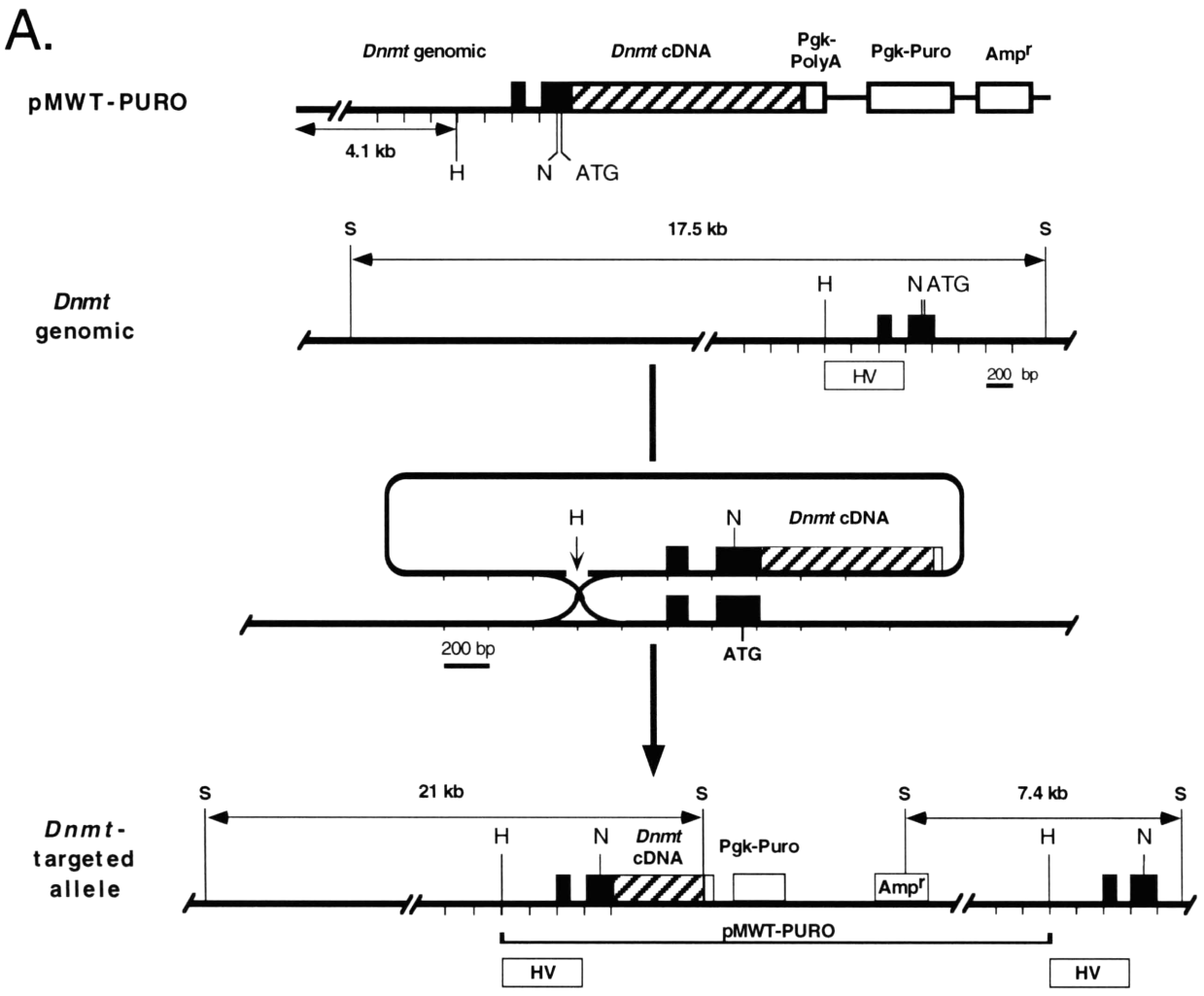
Hypomethylation of the DNA in mutant embryos or ES cells resulted in loss of monoallelic expression, either in the activation of the inactive *H19* and *Xist* alleles or in the inactivation of the active *Igf2r* and *Igf2* alleles. These experiments did not, however, provide insights into the role of methylation in the establishment of genomic imprinting during gametogenesis.

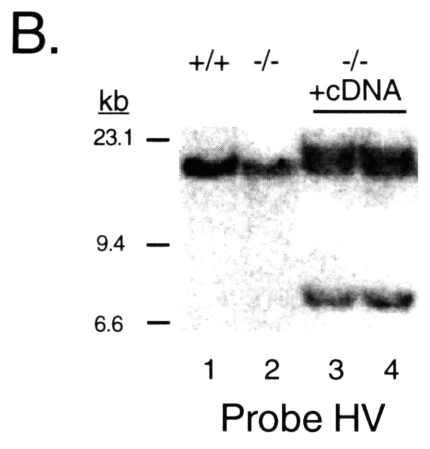
In this paper we address the process by which the allele-specific methylation of imprinted genes is acquired. We have expressed the wild type *Dnmt* cDNA in homozygous *Dnmt* mutant ES cells to assess whether this would restore the low methylation levels of the genomic DNA to normal. While normal methylation levels of non-imprinted DNA were reestablished, imprinted genes were neither remethylated nor appropriately expressed. However, passage of the remethylated ES cells through the germ line restored normal methylation and expression to all imprinted genes analyzed.

RESULTS

Expression of functional MTase in Dnmt mutant ES cells leads to restoration of genomic methylation

Previous work had demonstrated that the DNA of homozygous mutant *Dnmt* ES cells was highly hypomethylated (Lei et al., in preparation; Li et al., 1992). To assess whether expression of the wild type MTase would restore genomic methylation levels and rescue the lethal phenotype of differentiated mutant cells, we introduced the *Dnmt* cDNA into mutant ES and embryonic germ (EG) cells. The two cell types displayed identical molecular characteristics and functional capacity before and after expression of the *Dnmt* cDNA. For the sake of simplicity, they are collectively referred to as ES cells in the text. To assure correct quantitative and developmental expression of the construct, the cDNA was inserted by homologous recombination into the cognate *Dnmt* locus upstream of the mutation, using a strategy we refer to as CHIP (cDNA homologous insertion protocol; see Materials and Methods). The vector (Figure 1A), fused the *Dnmt* cDNA 3' to a 4.8-kb piece of genomic *Dnmt* sequence which ended in the second known exon (Rouleau et al., 1992). Homologous insertion of this hybrid vector was expected to leave all upstream control elements of the gene intact, leading to restoration of normally-regulated MTase expression. *Dnmt* mutant cells were electroporated with the linearized construct and selected for puromycin resistance. All puromycin-resistant colonies were expanded and analyzed by Southern blot for vector integration and levels of genomic


A.
Dnmt genomic
Dnmt cDNA
Pgk-PolyA
Pgk-Puro
Ampr
pMWT-PURO
4.1 kb
H
N
ATG
S
17.5 kb
S
Dnmt genomic
H
N ATG
HV
200 bp
H
N
Dnmt cDNA
200 bp
ATG
S
21 kb
S
S
7.4 kb
S
Dnmt-targeted allele
H
N
Dnmt cDNA
Pgk-Puro
Ampr
H
N
pMWT-PURO
HV
HV


B.
+/+
-/-
-/-
+cDNA
kb
23.1
9.4
6.6
1
2
3
4
Probe HV

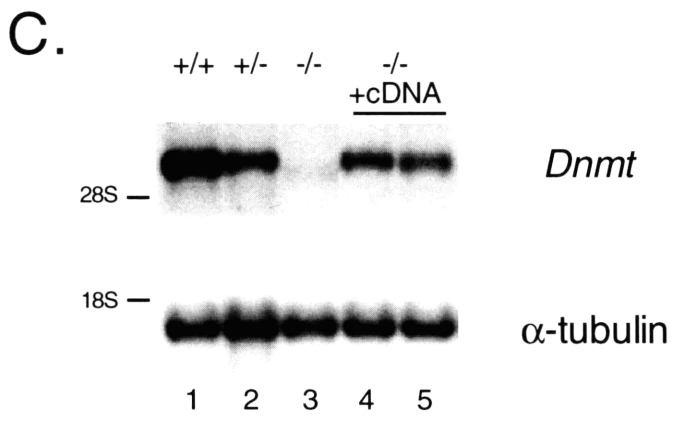

C.
+/+
+/-
-/-
-/-
+cDNA
28S
18S
Dnmt
α-tubulin
1
2
3
4
5

Figure 1

methylation. Genomic DNA was digested with *Sca*I and Southern blots hybridized with probe HV. This yields a 17.5-kb fragment in wild-type J1 and *Dnmt* mutant cell lines (Fig. 1B, Lanes 1, 2), but a 22-kb and a 7.4-kb fragment for the targeted allele of an homologous integrant (Lanes 3, 4). By this analysis, 73 / 182 puromycin-resistant clones had undergone homologous recombination events in which a single copy of the plasmid integrated at a single allele (data not shown). Northern blot analysis was used to assess the level of *Dnmt* transcription. The size of the endogenous mRNA is 5.2 kb (Fig. 1C, Lane 1). Virtually no mRNA of this size is observed in homozygous mutant *Dnmt* cells (Lane 3). Puromycin-resistant cells which underwent homologous recombination (Lanes 4, 5) show that the inserted cDNA was transcribed at a level and size comparable to that of the endogenous allele (lane 2).

All 73 homologous recombinants examined showed near-complete genomic remethylation (Fig. 2), as assayed using two repetitive sequence probes: a Moloney Murine Leukemia Virus (MoMuLV) cDNA, which cross-hybridizes with endogenous C-type retroviruses scattered throughout the genome (Stuhlmann et al., 1981), and a minor satellite centromeric repeat (Chapman et al., 1984). Both of these repeated sequences are heavily methylated in J1 cells (Fig. 2, Lane 1), as seen by Southern blot analysis of genomic DNA digested with *Hpa*II, an enzyme sensitive to the methylation of internal cytosines in its CCGG recognition sequence. DNA methylation in *Dnmt* mutant cells was highly reduced (Lanes 2), as seen by the increasing intensity of lower molecular weight bands in *Dnmt* mutant cell lines. The methylation level of both sequences was restored to normal in all

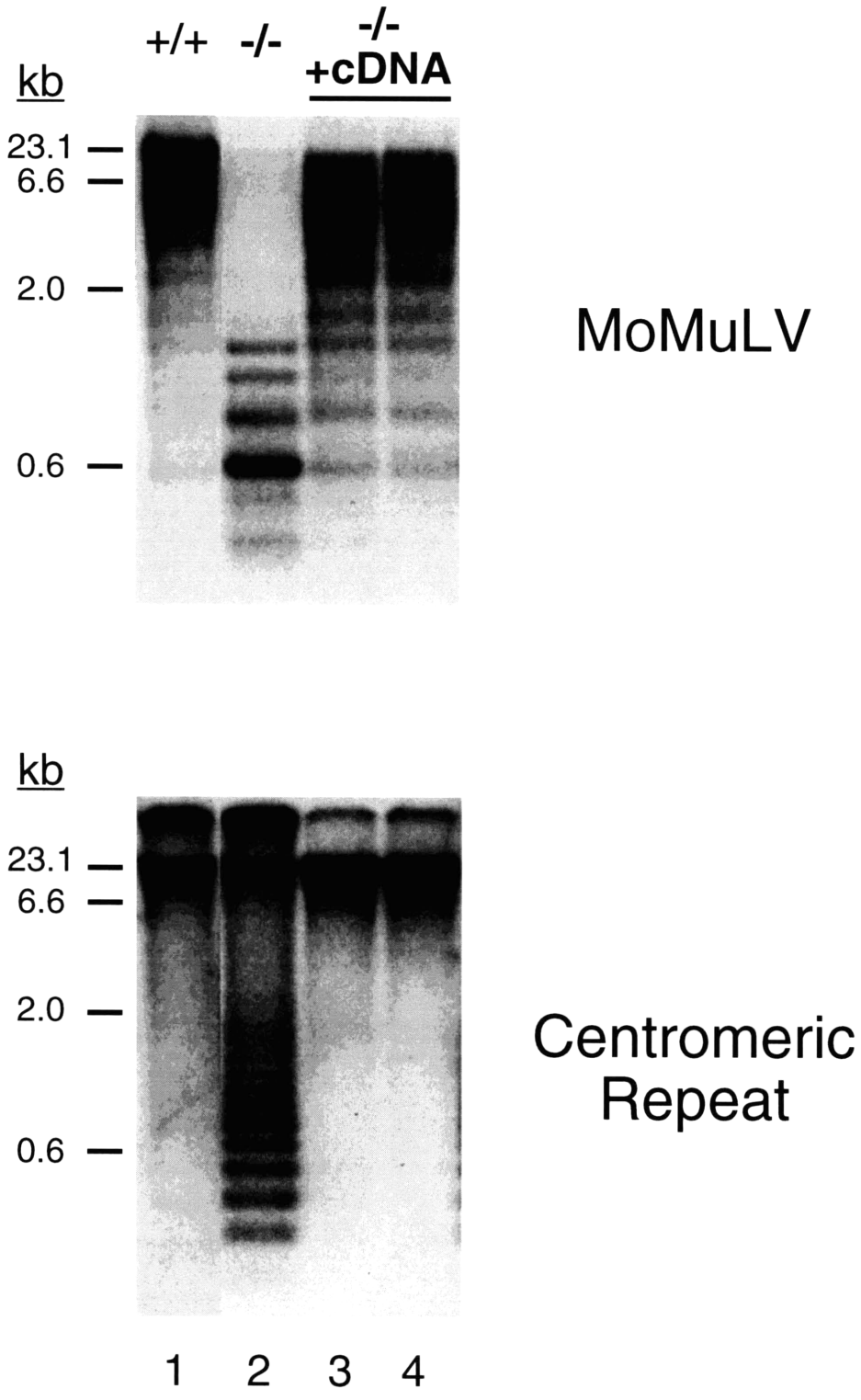

+/+
-/-
-/-
+cDNA
kb
23.1
6.6
2.0
0.6
MoMuLV
kb
23.1
6.6
2.0
0.6
Centromeric
Repeat
1
2
3
4

Figure 2

clones with a homologously integrated construct (Lanes 3, 4). No puromycin-resistant clones in which pMWT-PURO had integrated non-homologously displayed either *Dnmt* expression or genomic remethylation (data not shown). We conclude that remethylation of bulk DNA in the genome of homozygous mutant *Dnmt* cells is achieved by expression of the *Dnmt* cDNA from its endogenous promoter. *Dnmt* mutant cells which express the cDNA shall henceforth be referred to as "rescued" cells.

MTase expression complements lethal phenotype of differentiating mutant Dnmt ES cells

Previous observations indicated that homozygous mutant *Dnmt* ES cells, while proliferating normally when undifferentiated, die upon induction of differentiation (Beard et al., 1995; Lei et al., in preparation) and do not contribute to adult chimeras after injection into wild-type blastocysts (see below). We were interested to test whether restoration of genomic DNA methylation levels in mutant ES cells carrying the *Dnmt* cDNA would rescue the lethal mutant phenotype. To test *in vitro* differentiation, the rescued cells were induced to differentiate by plating on bacteriological petri dishes, in the absence of a fibroblast feeder layer and LIF. In contrast to the parental mutant cells (Lei et al., in preparation), the rescued cells differentiated and proliferated normally (data not shown). This indicates that *in vitro* differentiation potential was restored by genomic remethylation.

To more stringently test the developmental potential of the remethylated cells, *in vivo* experiments were conducted. First, the cells were injected subcutaneously into syngeneic male host animals. While the parental mutant cells failed to form palpable teratomas within 3-4 weeks (n = 0 / 24), the rescued ES cells induced teratoma formation at a similar rate (n = 21 / 24) as the wild-type J1 ES cells (n = 11 / 12). Second, the rescued cells were injected into wild-type BALB/c blastocysts to test their ability to contribute to chimeras. The parental *Dnmt* mutant cells were completely incapable of contributing to adult chimeras (n = 0 of 64 injected blastocysts). When rescued cells were used instead, eleven chimeras were generated with donation to coat color ranging from 10 to 70% (n = 11 of 92 injected blastocysts). The animals appeared normal in size and appearance, and contribution to the germline was observed in three chimeras derived from two rescued cell lines (see below).

The results described so far show, therefore, that expression of functional MTase in *Dnmt* mutant ES cells restores normal methylation levels and results in the rescue of the lethal mutant phenotype.

Methylation and monoallelic expression of imprinted genes is not restored in rescued ES cells

The methylation levels and expression of the imprinted genes *H19*, *Igf2r*, *Igf2*, and *Xist* were assessed by Southern blot and by RNase protection analysis. The analyses were performed on undifferentiated and *in vitro* differentiated ES cells. To investigate methylation and expression in terminally differentiated cells, fibroblasts were isolated from the chimeras described above. Primary fibroblast cultures were prepared from the carcasses of four chimeras. The primary cultures were immortalized with SV40 virus and then selected with puromycin to enrich for cells derived from the injected EG cells. Puromycin-selected populations were then subcloned at limiting dilution to prevent contamination from wild-type fibroblasts.

Methylation of H19: In normal ES cells and mice a 3.8-kb region just upstream of the promoter displays a monoallelic methylation pattern (Bartolomei et al., 1993), in which the paternally-derived allele is completely methylated (migrating at 3.8 kb, Fig. 3A, lanes 2, 3), while the maternally-derived allele is hypomethylated at the *Hha*I sites within the region (lower molecular weight bands, lanes 2, 3). Within this region lie three CpG sites that become methylated in spermatogenesis and remain methylated exclusively on the paternally-derived allele throughout development (Tremblay et al., 1995). This region becomes completely demethylated in *Dnmt* mutant ES cells (lanes 4, 5), indicating a complete loss of methylation at both alleles. In the rescued ES cells (lanes 6, 8) these sites

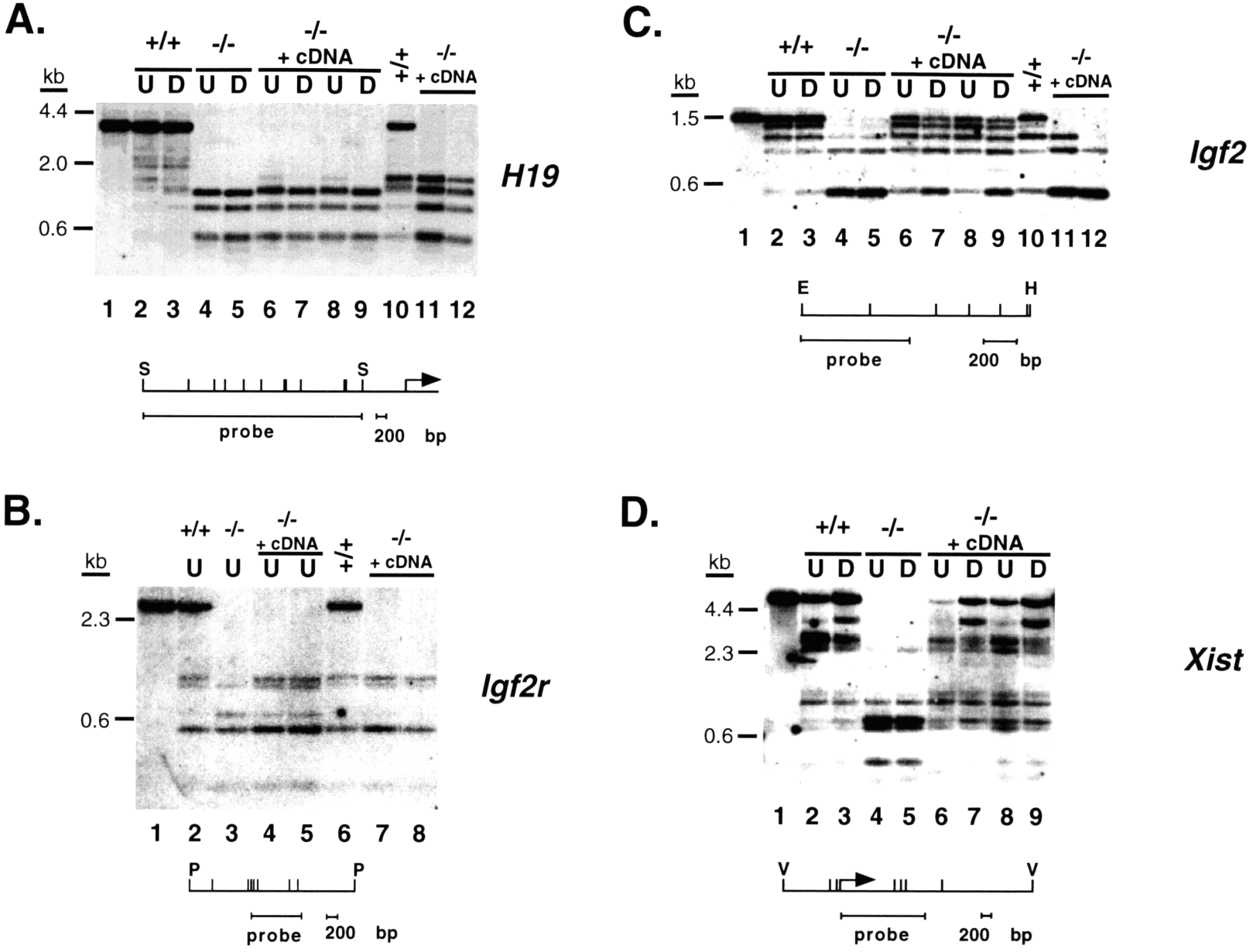

A.
+/+
-/-
-/- + cDNA
+/+
-/- + cDNA
kb
U D U D U D U D
4.4
2.0
0.6
H19
1 2 3 4 5 6 7 8 9 10 11 12
S
S
probe
200 bp
C.
+/+
-/-
-/- + cDNA
+/+
-/- + cDNA
kb
U D U D U D U D
1.5
0.6
Igf2
1 2 3 4 5 6 7 8 9 10 11 12
E
H
probe
200 bp
B.
+/+
-/-
-/- + cDNA
+/+
-/- + cDNA
kb
U U U U
2.3
0.6
Igf2r
1 2 3 4 5 6 7 8
P
P
probe
200 bp
D.
+/+
-/-
-/- + cDNA
kb
U D U D U D U D
4.4
2.3
0.6
Xist
1 2 3 4 5 6 7 8 9
V
V
probe
200 bp

Figure 3

remained hypomethylated at both alleles, and *in vitro* differentiation did not result in remethylation of this region (lanes 7, 9). To see if remethylation of this region occurred in terminally differentiated cell types *in vivo*, immortalized fibroblasts prepared from chimeras were examined. Puromycin-selected populations (lanes 11, 12), which were derived from the rescued ES cells, displayed no evidence of remethylation at either allele of *H19* [compare to fibroblasts derived from a wild-type mouse (lane 10), which display a fully-methylated allele (migrating at 3.8 kb) and a demethylated allele (lower molecular weight bands)].

Methylation of Igf2r: The methylation of the *Igf2r* gene was similarly determined. It has been shown that an intronic CpG-rich region (Stöger et al., 1993) displays a monoallelic methylation pattern in which the maternally-derived allele is completely methylated (migrating at 2.9 kb, Fig. 3B, lane 2), while the paternally-derived allele is hypomethylated at the *Hpa*II sites within the region (lower molecular weight bands, lane 2). Both alleles were completely demethylated in *Dnmt* mutant ES cells (lane 3), indicating a complete loss of methylation at both alleles. In the rescued ES cells (lanes 4, 5) these sites remained hypomethylated at both alleles. Neither *in vitro* differentiation of these cells nor twenty serial passages resulted in remethylation of this region (data not shown). The methylation pattern seen in puromycin-selected fibroblasts (lanes 7, 8) was similar to that seen in rescued ES cells [compare fibroblasts derived from a wild-type mouse (lane 6), which display a monoallelic methylation pattern, with rescued fibroblasts, which are hypomethylated at both alleles].

Methylation of Igf2: Unlike *H19* and *Igf2r*, *Igf2* does not possess a region which becomes differentially methylated in the gametes (Brandeis et al., 1993). However, in adult mice a 1.5-kb region upstream of the first *Igf2* promoter has been shown to display a differential methylation pattern (Brandeis et al., 1993; Feil et al., 1994; Sasaki et al., 1992). In contrast to *H19* and *Igf2r*, this region does not display absolute methylation differences, but the paternally-derived allele is on average more methylated. As reported before (Feil et al., 1994), examination of *Hpa*II-digested genomic DNA from wild-type ES cells revealed a pattern similar to that seen in embryonic and adult tissues, with a fully-methylated band (migrating at 1.5 kb, Fig. 3C, lane 2) and a ladder of fragments representing progressive demethylation, moving 3' to 5', of four 3'-proximal *Hpa*II sites (lower molecular weight bands, lane 2). This region became mostly demethylated in *Dnmt* mutant ES cells (lane 4), indicating a near-complete loss of methylation at both alleles. In the rescued ES cells (lanes 6, 8), a methylation pattern very similar to wild-type ES cells was established, though the fraction of demethylated 3'-proximal *Hpa*II sites remained higher than in wild-type cells. It is likely that this pattern reflected biallelic methylation. Upon *in vitro* differentiation, *Igf2* became slightly demethylated (lanes 7, 9), whereas puromycin-selected fibroblasts (lanes 11, 12) revealed a state of hypomethylation almost as extensive as that seen in *Dnmt* mutant ES cells (lanes 4, 5), in striking contrast to the pattern seen in wild-type fibroblasts (lane 10).

Expression of H19 , Igf2r, and Igf2: Expression was analyzed by RNase protection assay in wild type, *Dnmt* mutant and rescued ES

cells before and after induction of differentiation. All undifferentiated cells showed the same level of *Igf2r* expression and little or no expression of *H19* and *Igf2* (lanes 1, 3, and 5, in Fig. 4B, A, and D). Upon differentiation, however, the expression of *Igf2* increased significantly in wild-type ES cells (Fig. 4D, lane 2) with little or no increase seen in the mutant (lane 4) or rescued ES cells (lane 6). Although *Igf2r* expression appeared to increase only slightly upon differentiation of wild-type ES cells (Fig. 4B, lane 2), it showed a noticeable decrease in expression in the differentiated mutant (lane 4) and rescued ES cells (lane 6). Like *Igf2*, *H19* expression increased significantly in wild-type ES cells (Fig. 4A, lane 2) upon differentiation, and further increased by at least a factor of two in mutant (lane 4) and rescued ES cells (lane 6). In *Dnmt* mutant embryos, an increase in *H19* expression relative to wild-type embryos has been shown to be caused by biallelic expression of the gene (Li et al., 1993). The ES cell lines used for these studies were not marked by any polymorphism to distinguish the alleles of *H19*, but the overexpression seen in rescued ES cells is consistent with a biallelic transcriptional profile.

To determine if expression of these imprinted genes returned to wild-type levels in terminally differentiated cells *in vivo*, RNA was prepared from wild-type and rescued mutant fibroblasts derived from a P7 (P0 = day of birth) mouse. *Igf2r* expression, which was found at high levels in the wild-type fibroblasts (Fig. 4C, lanes 1, 2), was found to be completely extinguished in the rescued mutant fibroblasts (lanes 3-6). No significant expression of *H19* and *Igf2* was seen in either wild-type or rescued mutant fibroblast clones (data not

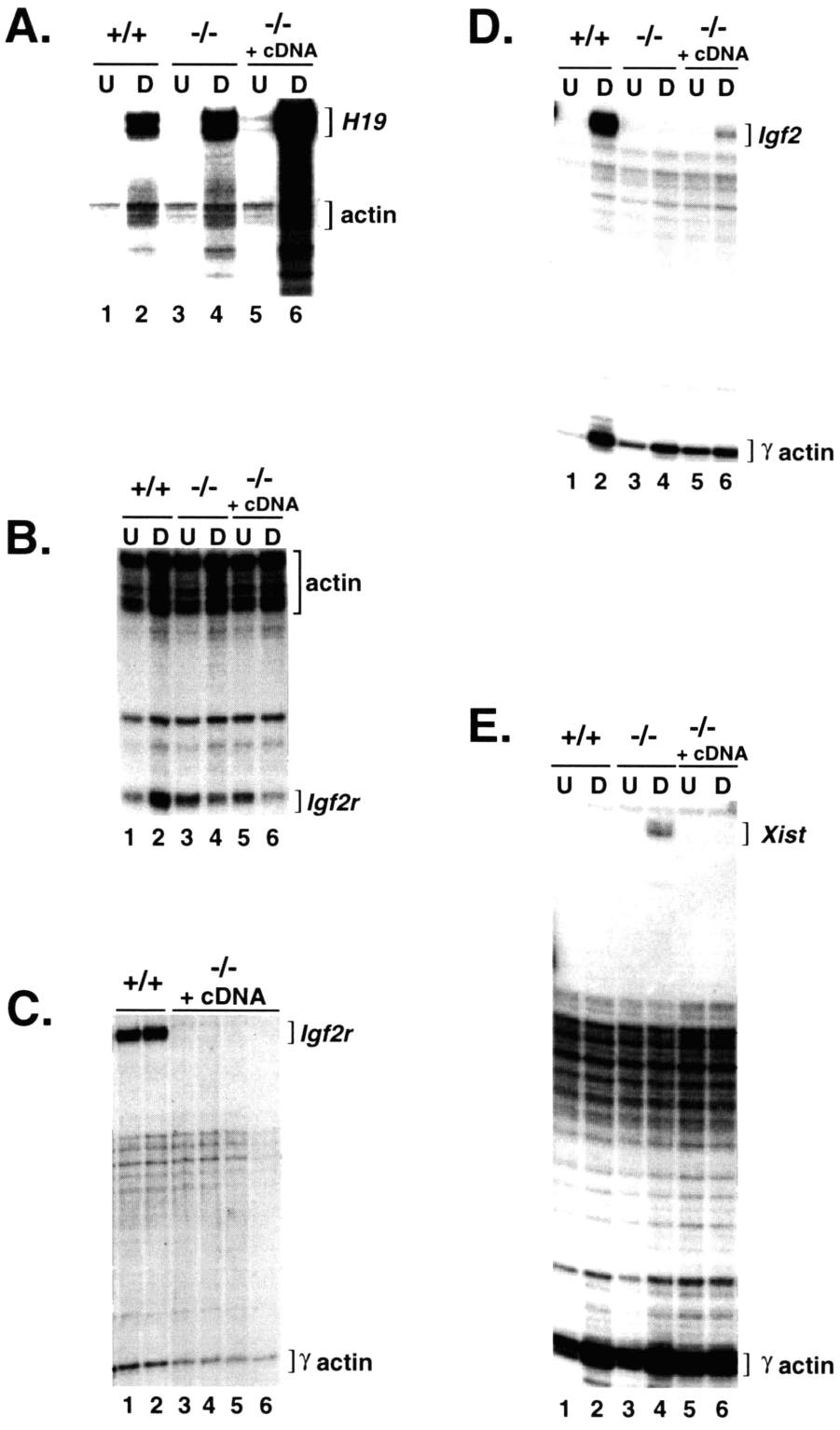

A.
+/+
-/-
-/- + cDNA
U D U D U D
H19
actin
1 2 3 4 5 6
B.
+/+
-/-
-/- + cDNA
U D U D U D
actin
Igf2r
1 2 3 4 5 6
C.
+/+
-/- + cDNA
Igf2r
γ actin
1 2 3 4 5 6
D.
+/+
-/-
-/- + cDNA
U D U D U D
Igf2
γ actin
1 2 3 4 5 6
E.
+/+
-/-
-/- + cDNA
U D U D U D
Xist
γ actin
1 2 3 4 5 6

Figure 4

shown). The lack of expression of *Igf2* and *H19* in wild-type cells was probably caused by a developmental down-regulation of these genes, for the following reasons: 1) expression of *Igf2* was observed in wild-type fibroblast cultures derived from a 16.5 day-old embryo (data not shown), and 2) *H19* is known to be switched off in most tissues soon after birth (Poirier et al., 1991).

Xist: This gene, in contrast to the imprinted genes described above, is imprinted only in extraembryonic tissues. The maternal allele is methylated and the paternal allele remains unmethylated throughout cleavage and in the extraembryonic lineage (Ariel et al., 1995; Zuccotti and Monk, 1995). Following implantation, however, the maternal imprint is erased and one of two alleles becomes randomly methylated and silent at the time of X chromosome inactivation (Norris et al., 1994). *Xist* becomes differentially methylated in a 1.5-kb region surrounding its promoter (Norris et al., 1994), with two clusters of CpG sites showing gamete-specific methylation, one cluster just upstream of the transcriptional start site (Zuccotti and Monk, 1995) and the other 600 bp downstream (Ariel et al., 1995). In male ES cells, with a single, maternally-derived X-chromosome, this region is almost completely methylated (Fig. 3D, lane 2). In *Dnmt* mutant ES cells, this methylation is completely erased (lane 4), including the gamete-specific methylation sites. The same degree of partial methylation seen in wild-type ES cells was observed in the rescued ES cells (lanes 6, 8). As shown before (Beard et al., 1995), *in vitro* differentiation of wild-type J1 cells results in an increase in methylation (lane 3), and this is also seen for the rescued ES cells (lanes 7, 9). Also as shown previously (Beard et al., 1995), this

demethylation results in the expression of *Xist* upon differentiation, of *Dnmt* mutant ES cells (Fig. 4E, lane 4) but not J1 (lane 2) or rescued ES cells (lane 6).

Our results show that methylation and expression of the imprinted genes *H19* and *Igf2r* were not restored to normal levels following expression of MTase in homozygous *Dnmt* mutant ES cells or their differentiated derivatives, in contrast to repetitive DNA sequences and *Xist*. These observations suggest, therefore, that once the genomic DNA has been demethylated by deletion of the maintenance MTase, the *de novo* methylation activity in ES cells can distinguish between two classes of genes: (i) the majority of genes, as evidenced by the repetitive probes in Figure 2 or *Xist*, which can be remethylated to normal levels so as to assure normal expression patterns during subsequent differentiation; (ii) imprinted genes such as *H19* and *Igf2r*, which cannot be remethylated and functionally restored once demethylated.

Normal methylation and expression patterns of imprinted genes are restored after passage through the germline

To test whether germline transmission would restore normal methylation and expression patterns to imprinted genes, five male chimeras, generated from a rescued ES cell line, were mated to BALB/c females. Three of these chimeras, derived from two independently-transfected ES cell lines, produced agouti offspring, indicating contribution of the rescued genome to the germline. The *Dnmt* allele which had undergone homologous recombination was transmitted at the expected frequency, and pups carrying the *Dnmt* cDNA appeared identical to their littermates. Paternal transmission of the imprinted genes contributed by the rescued ES cells makes clear predictions regarding the methylation and expression status of imprinted genes in the offspring. For example, after germline transmission the paternal *H19* allele should be methylated and the paternal *Igf2* allele should be expressed. *H19* methylation in tail preparations was examined as described above. The same methylation pattern was seen in pups derived from the rescued ES cells (Fig. 5D, lanes 2-4) as in pups derived from the host-blastocyst (lane 1). As the examined promoter-proximal region of *H19* is known to be unmethylated on the maternally-derived allele (Tremblay et al., 1995), we conclude that methylation of the paternally-derived allele had occurred in germline transmission. Methylation of *Igf2r* and *Igf2* was examined in the same regions described above. A wild-type methylation pattern was seen for *Igf2r* (Fig. 5F), which is consistent with the paternally-derived allele of this gene remaining

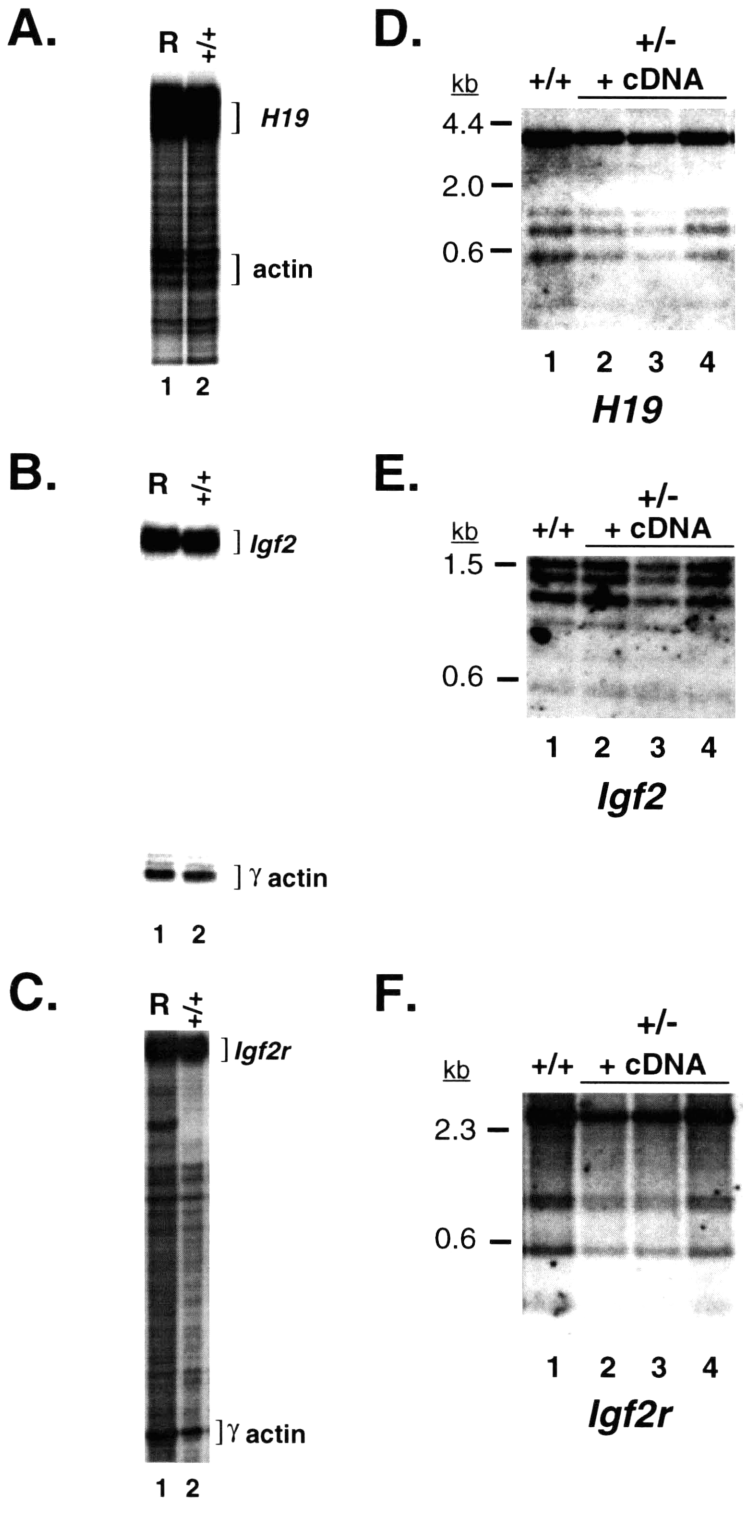

A.
R
+/+
H19
actin
1 2
B.
R
+/+
Igf2
γ actin
1 2
C.
R
+/+
Igf2r
γ actin
1 2
D.
+/-
kb
+/+
+ cDNA
4.4
2.0
0.6
1 2 3 4
H19
E.
+/-
kb
+/+
+ cDNA
1.5
0.6
1 2 3 4
Igf2
F.
+/-
kb
+/+
+ cDNA
2.3
0.6
1 2 3 4
Igf2r

Figure 5

hypomethylated after germline passage. Finally, a wild-type methylation pattern was observed for *Igf2* (Fig. 5E), indicating full methylation of the paternally-derived allele.

RNase protection analysis was performed on total RNA prepared from P0 pups. Identical levels of *Igf2* RNA were observed in pups derived from the rescued ES cells (Fig. 5B, lane 1) as in pups derived from the host-blastocyst (lane 2). Since *Igf2* is not expressed from the maternally-derived allele (DeChiara et al., 1991), we conclude that this expression was derived from the paternally-transmitted allele. Similar results were observed for expression levels of *H19* (Fig. 5A) and *Igf2r* (Fig. 5C).

These results indicate that normal methylation patterns were restored to both alleles of *H19* in the male gonad, resulting in the appropriate expression of both *H19* and *Igf2* from the paternally-derived alleles in the germline pups (see Discussion). The presence of an unmethylated *Igf2r* allele in these pups, and wild-type levels of *Igf2r* expression in germline-transmitted embryos, indicates that this gene was not remethylated in the male gonad. By extension, methylation of both alleles of *Igf2r* would presumably occur upon germline transmission through a female chimera, but none of the four female chimeras produced agouti offspring. In conclusion, our results suggest that *de novo* methylation of the imprinted genes *H19* and *Igf2r*, in contrast to "bulk" DNA and the *Xist* and *Igf2* genes, can only occur after passage through the germ line.

DISCUSSION

The experiments in this study were designed to examine the developmental specificity of DNA methylation activity in the mouse. Because gross genomic hypomethylation is tolerated by *Dnmt* mutant ES cells, we could start at a specific point in the murine life cycle with a near-complete deficiency of genomic methylation and determine the sites and times of methylation activity. Our results allow the two following conclusions regarding the function of the DNA methyltransferase: (i) Genomic methylation *levels* in ES cells, which were decreased substantially as a result of the targeted inactivation of the *Dnmt* gene, could be restored to normal by expression of a functional *Dnmt* gene. That this remethylation process has restored normal methylation *patterns* to the genome is inferred by the ability of these remethylated cells to undergo normal differentiation *in vitro*, to form teratomas, and contribute extensively to chimeras *in vivo*. In contrast, the parental mutant cells fail to differentiate normally in each of these assays. (ii) Upon *Dnmt* cDNA expression in ES cells, the imprinted genes *H19* and *Igf2r* were not remethylated at either allele, and as a consequence were not appropriately expressed. Despite this deficiency, the cells were able to contribute substantially to chimeras, including the gonad. Importantly, normal methylation patterns and expression of the two imprinted genes *Igf2* and *H19* were restored after passage through the male germline.

Expression of functional MTase in the fully hypomethylated *Dnmt* mutant ES cells complemented the mutant phenotype, leading to a

substantial restoration of normal genomic DNA methylation levels, as assayed by various repetitive sequence probes. *In vitro*, the enzyme prefers hemimethylated over unmethylated substrates (Gruenbaum et al., 1982), consistent with the notion that the *Dnmt* gene encodes the maintenance methyltransferase which is expressed in every cell. How might re-expression of a maintenance methyltransferase activity result in *de novo* methylation of the unmethylated genomic DNA in mutant ES cells? We have postulated previously that early embryos as well as undifferentiated ES cells express a *de novo* DNA methyltransferase that is not active in postgastrulation embryos (Jähner et al., 1982) or in differentiated ES cells (Stewart et al., 1982). The low but stable level of genomic DNA methylation in *Dnmt* mutant ES cells is likely caused by the activity of a separately-encoded enzyme (Lei et al., in preparation), which is insufficient to maintain normal methylation levels in the absence of the maintenance DNA methyltransferase. Re-expression of the maintenance methyltransferase would, however, be expected to lead to an increase in methylation levels as it would maintain the methylation levels set by the embryonic *de novo* methyltransferase activity.

Dnmt mutant ES cells, while proliferating normally when undifferentiated, grow poorly and die upon induction of differentiation, with the level of genomic DNA methylation decreasing (Lei et al., in preparation). This is consistent with the embryonic *de novo* DNA methyltransferase becoming inactive upon differentiation, leading to further hypomethylation. An important issue was whether re-expression of MTase not only would restore

genome-wide methylation levels but also lead to rescue of the mutant phenotype, which is a reduction in differentiation capacity *in vitro* and *in vivo* (Lei et al., in preparation; Li et al., 1992). Remethylated cells have clearly regained the ability to differentiate, as evidenced by three different assays: *in vitro* differentiation as embryoid bodies, teratoma formation, and contribution to the tissues of chimeras, which represents the most stringent functional test. The extensive coat color contribution to these normal chimeras suggests that single-copy genes had become appropriately remethylated and were correctly expressed, resulting in normal embryonic development. The restoration of wild-type methylation levels and consequent transcriptional inactivation was observed for the single-copy gene *Xist*. *Xist* is not imprinted in male ES cells and embryonic derivatives because the single copy of this gene is normally methylated and unexpressed (Norris et al., 1994). The partially methylated pattern seen in normal male ES cells (Beard et al., 1995) was restored to the same degree in the rescued ES cells. Importantly, the CpG sites showing gamete-specific methylation (Ariel et al., 1995; Zuccotti and Monk, 1995) were remethylated, supporting the notion that methylation of these sites leads to transcriptional inactivity of *Xist*. The remethylation of *Xist*, coupled with the functional complementation assays, suggests that the methylation of all or most other non-imprinted genes is fully restored. It is possible, however, that some genes, whose misexpression would not compromise normal development, are not becoming appropriately remethylated. A detailed study of remethylation at individual loci will address this issue and yield

information regarding the functional importance of methylation at specific sites as compared to overall methylation patterns.

The imprinted genes *H19* and *Igf2r* were not remethylated after introduction of the *Dnmt* cDNA into the mutant ES cells. The regions examined for methylation show a gamete-specific methylation difference and are believed to constitute the primary imprinting mark (Barlow, 1995). The methylation of these regions and expression status of these genes in rescued ES cells was identical to that seen previously in *Dnmt* mutant ES cells and embryos (Li et al., 1993). Furthermore, hypomethylation persisted in ES cell-derived fibroblasts isolated from chimeras, and methylation was restored to normal only after germline transmission. In contrast to *H19* and *Igf2r*, *Igf2* was extensively remethylated in an area which has been shown to be differentially methylated in the adult (Brandeis et al., 1993; Feil et al., 1994; Sasaki et al., 1992), but not in the gametes (Brandeis et al., 1993). Why methylation levels of *Igf2* decreased in fibroblasts derived from rescued ES cells is not clear at this time. It has been suggested that *Igf2* expression depends upon the availability of two enhancers shared with the closely-linked *H19* gene, *Igf2* being expressed only when the promoter of *H19* is methylated and inaccessible to its downstream enhancers (Bartolomei et al., 1993). In mice lacking the *H19* gene, a maternally-transmitted deletion results in both the methylation and the transcription of *Igf2* (Leighton et al., 1995a), while in mice lacking the downstream enhancers, a paternally-transmitted deletion results in an abrogation of *Igf2* transcription from that allele (Leighton et al., 1995b). In our study, methylation was restored to

Igf2, but little expression of the gene was seen in differentiated ES cells, similar to the *Dnmt* mutant ES cells and embryos. Our results support the idea that the methylation and expression status of *H19* is the primary determinant of *Igf2* expression, and that methylation of *Igf2* in the examined region does not affect its expression.

Our observations are of general significance in defining the stage in development for, and the mechanisms involved in, the establishment of genomic imprinting. Figure 6 contrasts the developmental changes in methylation of bulk DNA and *Xist* with that of the imprinted genes *H19* and *Igf2r*. While it is generally assumed that imprinted genes acquire their methylation mark during gametogenesis, no direct evidence of an enzyme activity involved in gamete-specific methylation has been demonstrated. Our results show conclusively that passage through the male germ line is necessary to obtain the allele-specific methylation mark on at least one imprinted gene, *H19*. This suggests the existence of a male-specific *de novo* methyltransferase activity which may specifically recognize all paternally-modified imprinted genes. Presumably a female-specific *de novo* methyltransferase activity exists, which would be active during oogenesis and would be expected to recognize maternally-imprinted genes such as *Igf2r*. While other results strongly suggest that the *de novo* methyltransferase expressed in ES cells is an independently-encoded enzyme (Lei et al., in preparation), we have no evidence whether the gamete-specific activities are encoded by separate genes, are due to modification of the maintenance MTase or the separately-encoded *de novo* methyltransferase activity (Lei et al., in preparation), or are

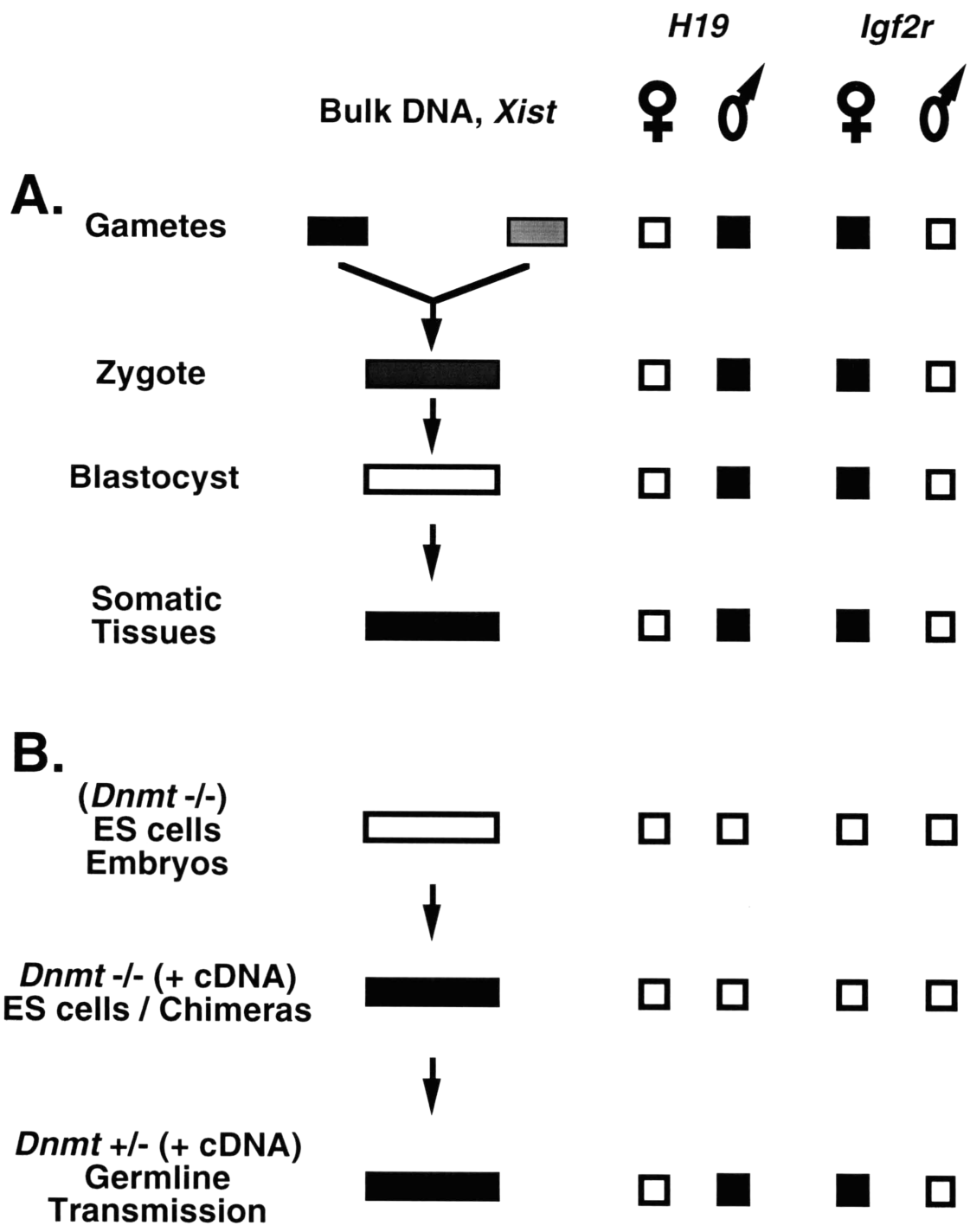

H19
Igf2r
Bulk DNA, Xist
A.
Gametes
Zygote
Blastocyst
Somatic Tissues
B.
(Dnmt -/-) ES cells Embryos
Dnmt -/- (+ cDNA) ES cells / Chimeras
Dnmt +/- (+ cDNA) Germline Transmission

Figure 6

reflective of a change in chromatin structure affecting substrate accessibility.

It is of interest to compare the rescued mutant ES cells with cells derived from androgenetic (AG) or parthenogenetic (PG) embryos (Fundele and Surani, 1994). Chimeras between wild type embryos and either PG- or AG-derived embryonic cells are generally inviable and show highly skewed and complementary distribution of the uniparental cells to different tissues of the chimera (Fundele et al., 1989; Fundele et al., 1990; Mann et al., 1990; Nagy et al., 1989; Thomson and Solter, 1988). The phenotype of the chimeras has been proposed to be caused by unbalanced expression of the imprinted genes in the uniparental component (Allen et al., 1994; Szabo and Mann, 1994). In contrast to PG- or AG-derived embryonic cells, the rescued ES cells described here have methylation erased at both paternally- and maternally-methylated loci, and the subsequent biallelic expression patterns of these genes resulted in a qualitatively different transcriptional profile. Chimeras produced from the rescued ES cells had extensive coat color contribution and were normal. Analysis of organ-specific contribution in chimeras and the possibility of producing embryos derived exclusively from the rescued ES cells (Nagy et al., 1990) will allow us to fully assess the developmental potential of a mouse lacking an imprinted genome.

MATERIALS AND METHODS

Tissue Culture

The wild-type J1 (Li et al., 1992), *Dnmt*$^{n/n}$ clone 10, (Li et al., 1992), *Dnmt*$^{s/s}$, and *Dnmt*$^{c/c}$ (Lei et al., in preparation; Li et al., 1993) ES cell lines were used. ES cell lines were cultured as described by Li *et al.* (1992). ES cells were induced to differentiate into embryoid bodies by suspension culture in bacteriological petri dishes, as described (Robertson, 1987), and were collected at day 15 for analysis. A *Dnmt*$^{s/s}$ EG cell line was derived from a cross between *Dnmt*$^{s/+}$ mice. Primordial germ cells were cultured from gestational day 8.5 embryos on SI^4-m220 feeder cells with 1000 U/ml LIF and 20 ng/ml bFGF, as described (Matsui et al., 1992). The three homozygous mutant ES cell lines and the EG cell line all showed equivalent molecular characteristics and functional capacity before and after transfection of the *Dnmt* cDNA, and they are collectively referred to as "rescued ES cells" in the text.

Fibroblasts were prepared by maceration and trypsinization of eviscerated, decapitated embryonic and post-natal carcasses and grown in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum, 0.1 mM nonessential amino acids (GIBCO), 0.1 mM β-mercaptoethanol, and antibiotics. SV40 virus infection was as described (Jat et al., 1986). To enrich for fibroblasts derived from the rescued ES cells, puromycin (Sigma) selection (2 μg/ml) was applied after infection. Puromycin-selected populations were plated at 0.5 - 1.0 cell / well density to produce subcloned lines. Subclones

were chosen from those plates with less than 1 colony / 3 wells. RNase protection experiments were performed on subcloned fibroblasts prepared from a P7 (P0 = day of birth) chimera made between a rescued *Dnmt*$^{s/s}$ EG cell line (described above) and a BALB/c host. Methylation data was obtained from these cells and fibroblasts prepared from two three month-old chimeras made with a rescued *Dnmt*$^{s/s}$ ES cell line.

CHIP (cDNA homologous insertion protocol)

All plasmids were prepared by alkali lysis from bacterial cultures and purified by centrifugation in continuous CsCl-ethidium bromide gradients (Sambrook et al., 1989). pMWT-PURO was made as follows. The *Dnmt* cDNA was derived from pMG (gift of T. Bestor), which comprises a 4934 bp *Dnmt* cDNA, starting at an *Eco*RI site 203 bp upstream of the first ATG and terminating at a *Bgl*II site 226 bp downstream of the stop codon TAG. This cDNA was inserted into pPGK-CAS, to utilize the *Pgk-1* polyadenylation tract, which comprises the 500 bp *Pvu*II-*Hind*III fragment at the 3' end of the *Pgk-1* cDNA (Adra et al., 1987). pPGK-CAS is described as the *Pgk-1*-driven expression vector in van der Lugt *et al.* (1991). A 4.8-kb *Dnmt Nae*I genomic fragment (Li et al., 1992) was cloned into a unique *Nae*I site in the open reading frame (ORF) of the *Dnmt* cDNA, such that the *Dnmt* cDNA sequence removed upstream of the *Nae*I site was restored in its genomic context. Finally, a *puro*^{R} cassette from pPgk-Puro was inserted downstream of the *Dnmt* cDNA. pPgk-Puro consists of the 660-bp *Hind*III-*Cla*I fragment of pBabePuro

(Morgenstern and Land, 1990) cloned into pPGK-CAS through several intermediate cloning steps.

This construct was designed to act as an insertion vector (Thomas and Capecchi, 1987) using a unique *Hind*III site (Fig. 1A), which lies 700 bp upstream of the *Nae*I site marking the genomic-cDNA junction and 4.1 kb downstream of the 5' end of the genomic *Dnmt* fragment. The construct was linearized at this site and then electroporated into ES cells. Upon homologous insertion at the *Dnmt* locus, the cDNA-genomic hybrid configuration of pMWT-PURO fused the *Dnmt* promoter and any missing upstream exons to the *Dnmt* cDNA. *Dnmt* mRNA transcripts were produced of the same size and amount as the endogenous gene, and they were spliced appropriately to allow production of the full-length protein. For electroporation, approximately 10^7 ES cells were trypsinized and washed and resuspended in 600 µl electroporation buffer, made as described (Thomas and Capecchi, 1987), except with 137 mM NaCl and 10 µg linearized pMWT-PURO. The cell-DNA mixture was electroporated in a BioRad Gene Pulser at 400 V and 25 µF. The cells were plated on 10-cm dishes pre-plated with γ-irradiated, puromycin-resistant embryonic fibroblasts prepared from pPgk-Puro transgenic mice (kind gift of E. Simpson). Media containing 2 µg/ml puromycin (Sigma) were changed every day, and isolated colonies picked 7-9 days after electroporation. DNA and RNA were analyzed by standard procedures (see below).

Southern blot analysis.

Genomic DNA was prepared from cell cultures and pup tails as described (Laird et al., 1991). 10 µg of DNA was digested with 40 units of restriction endonuclease, except for *Hha*I and *Hpa*II, which were used at double this amount. Digested DNA was electrophoresed in agarose gels, and blotted onto Zetabind nylon membranes (Cuno Laboratory Products). Usual Southern blot procedures were used (Sambrook et al., 1989), with hybridization buffer (Church and Gilbert, 1984) and α-^{32}P dCTP-labeled probe (Feinberg and Vogelstein, 1983) prepared as described. Blots were washed at 65° C with 0.1X SSC / 0.1% SDS, and exposed to AR-X film (Kodak). The following probes were used: HV is a 600-bp Hind*III*-Eco*RV* fragment of genomic DNA which includes the first exon of *Dnmt*; pMO contains a 5.5-kb MoMuLV cDNA fragment (Jähner et al., 1982); pMR150 contains a minor satellite centromeric repeat (Chapman *et al.*, 1984; gift of V. Chapman); pBSIISac contains a 3.8-kb genomic *Sac*I fragment from a region of the *H19* gene starting 750 bp upstream of the transcriptional start site, identical to the probe used in Tremblay *et al.*, (1995) (generously donated by M. Bartolomei); p1269 contains a 720-bp *Eco*RI-*Xba*I genomic *Igf2* fragment whose 5' end is approximately 4.5 kb from the first exon, corresponding to probe 2 in Feil *et al.*, (1994) (gracious gift of S. Tilghman); pPP4 contains a 1.1-kb *Bst*BI-*Mlu*I genomic *Igf2r* fragment located within the intronic region 2, identical to the probe used in Stöger *et al.*, (1993) (gift of D. Barlow); and pXistMc1 contains a 1.5-kb *Xist* cDNA fragment that includes the transcriptional start site, 1.5 kb of exon 1, and the gamete-specific methylation sites as reported in Ariel *et al.*, (1995) and Zuccotti and Monk (1995) (provided by H. Willard). A

mitochondrial probe was used to assay for completeness of *Hha*I and *Hpa*II digestions, because the mitochondrial genome does not become methylated. This was a gift of P. Laipis and comprises bp 4013 to 7720 of the published sequence from mouse L cells (Bibb et al., 1981).

Northern blot analysis

RNA was prepared as described (Chomczynski and Sacchi, 1987), from confluent ES cell cultures passaged without a fibroblast feeder layer. Denatured RNA was electrophoresed in formaldehyde-containing agarose gels and blotted onto Zetabind nylon membranes (Cuno) as described (Sambrook et al., 1989). Hybridization, washing, and exposure occurred as described above for DNA analysis. The *Dnmt* cDNA from pMG was used for a probe, and an α-tubulin cDNA probe (pILαT1) (Lemischka et al., 1981) was used to control for loading amounts.

RNase protection assays

Assays were performed as described for *H19*, *Igf2*, and *Igf2r* (Li et al., 1993), and for *Xist* (Beard et al., 1995).

Teratoma Formation and Generation of Germline Chimeras

Teratomas were produced by preplating 3 x 10^6 ES cells for 30' to remove embryonic fibroblasts, resuspending them in 300 µl of PBS,

and injecting them subcutaneously into the flanks of adult male 129/Sv mice. Wild-type J1, *Dnmt*$^{s/s}$, and rescued *Dnmt*$^{s/s}$ EG and *Dnmt*$^{c/c}$ ES cells were used. Tumors were excised at 3-4 weeks. Chimeras were made as described (Li et al., 1992), except that 2-3 ES or EG cells were injected into blastocysts. Male chimeras derived from the two rescued *Dnmt*$^{s/s}$ ES cell lines were mated to BALB/c females, and progeny were analyzed with Southern blot and RNase protection assays.

ACKNOWLEDGMENTS

We would like to thank Klaus Willecke and Edwin Cuppen for pMWT-PURO, Elizabeth Simpson for providing pPgk-Puro transgenic mice, Tim Bestor for pMG, Shirley Tilghman for the *H19* and *Igf2* probes, Denise Barlow for the *Igf2r* probes, Marisa Bartolomei for the *H19* probe, Huntington Willard for the 5' *Xist* probe, Argiris Efstratiadis for the *Igf2* probes, Verne Chapman for the centromeric repeat probe, Nathalie van der Lugt for pPGK-CAS, Michael Rudnicki for pILαT1, and Phil Laipis for the mitochondrial DNA probe. L. J.-G. was supported by a Cancer Research Fund of the Damon Runyon-Walter Winchell Foundation Fellowship (DRG-1223). P. W. L. was the recipient of a National Service Research Award (F32 CA 09097) from the National Cancer Institute. E.L. was supported by a grant from Bristol-Myers and Squibb. R. J. was supported by NIH grant #R35-CA44339.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

REFERENCES

Adra, C. N., Boer, P. H., and McBurney, M. W. 1987. Cloning and expression of the mouse *pgk-1* gene and the nucleotide sequence of its promoter. *Gene* **60:** 65-74.

Allen, N. D., Barton, S. C., Hilton, K., Norris, M. L., and Surani, M. A. 1994. A functional analysis of imprinting in parthenogenetic embryonic stem cells. *Development* **120:** 1473-82.

Ariel, M., Robinson, E., McCarrey, J. R., and Cedar, H. 1995. Gamete-specific methylation correlates with imprinting of the murine *Xist* gene. *Nat Genet* **9:** 312-5.

Barlow, D. P. 1995. Gametic imprinting in mammals. *Science* **270:** 1610-1613.

Barlow, D. P., Stöger, R., Herrmann, B. G., Saito, K., and Schweifer, N. 1991. The mouse insulin-like growth factor type-2 receptor is imprinted and closely linked to the *Tme* locus. *Nature* **349:** 84-7.

Bartolomei, M. S., Webber, A. L., Brunkow, M. E., and Tilghman, S. M. 1993. Epigenetic mechanisms underlying the imprinting of the mouse *H19* gene. *Genes Dev* **7:** 1663-1673.

Bartolomei, M. S., Zemel, S., and Tilghman, S. M. 1991. Parental imprinting of the mouse *H19* gene. *Nature* **351:** 153-5.

Beard, C., Li, E., and Jaenisch, R. 1995. Loss of methylation activates *Xist* in somatic but not in embryonic cells. *Genes Dev* **9:** 2325-2334.

Bestor, T. H., and Ingram, V. M. 1983. Two DNA methyltransferases from murine erythroleukemia cells: purification, sequence specificity, and mode of interaction with DNA. *Proc Natl Acad Sci U S A* **80:** 5559-63.

Bibb, M. J., Van Etten, R. A., Wright, C. T., Walberg, M. W., and Clayton, D. A. 1981. Sequence and gene organization of mouse mitochondrial DNA. *Cell* **26:** 167-80.

Brandeis, M., Kafri, T., Ariel, M., Chaillet, J. R., McCarrey, J., Razin, A., and Cedar, H. 1993. The ontogeny of allele-specific methylation associated with imprinted genes in the mouse. *Embo J* **12:** 3669-77.

Chaillet, J. R., Vogt, T. F., Beier, D. R., and Leder, P. 1991. Parental-specific methylation of an imprinted transgene is established during gametogenesis and progressively changes during embryogenesis. *Cell* **66:** 77-83.

Chapman, V., Forrester, L., Sanford, J., Hastie, N., and Rossant, J. 1984. Cell lineage-specific undermethylation of mouse repetitive DNA. *Nature* **307:** 284-6.

Chomczynski, P., and Sacchi, N. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem* **162:** 156-159.

Church, G., and Gilbert, W. 1984. Genomic sequencing. *Proc Natl Acad Sci U S A* **81:** 1991-1995.

DeChiara, T. M., Robertson, E. J., and Efstratiadis, A. 1991. Parental imprinting of the mouse insulin-like growth factor II gene. *Cell* **64:** 849-859.

Feil, R., Walter, J., Allen, N. D., and Reik, W. 1994. Developmental control of allelic methyation in the imprinted mouse *Igf2* and *H19* genes. *Development* **120:** 2933-2943.

Feinberg, A. P., and Vogelstein, B. 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal Biochem* **132:** 6-13.

Fundele, R., Norris, M. L., Barton, S. C., Reik, W., and Surani, M. A. 1989. Systematic elimination of parthenogenetic cells in mouse chimeras. *Development* **106:** 29-35.

Fundele, R. H., Norris, M. L., Barton, S. C., Fehlau, M., Howlett, S. K., Mills, W. E., and Surani, M. A. 1990. Temporal and spatial selection against parthenogenetic cells during development of fetal chimeras. *Development* **108:** 203-11.

Fundele, R. H., and Surani, M. A. 1994. Experimental embryological analysis of genetic imprinting in mouse development. *Dev Genet* **15:** 515-22.

Giddings, S. J., King, C. D., Harman, K. W., Flood, J. F., and Carnaghi, L. R. 1994. Allele specific inactivation of insulin 1 and 2, in the mouse yolk sac, indicates imprinting. *Nat Genet* **6:** 310-3.

Gruenbaum, Y., Cedar, H., and Razin, A. 1982. Substrate and sequence specificity of a eukaryotic DNA methylase. *Nature* **295:** 620-2.

Guillemot, F., Caspary, T., Tilghman, S. M., Copeland, N. G., Gilbert, D. J., Jenkins, N. A., Anderson, D. J., Joyner, A. L., Rossant, J., and Nagy, A. 1995. Genomic imprinting of *Mash2*, a mouse gene required for trophoblast development. *Nat Genet* **9:** 235-42.

Hatada, I., and Mukai, T. 1995. Genomic imprinting of $p57^{KIP2}$, a cyclin-dependent kinase inhibitor, in mouse. *Nat. Genet* **11:** 204-6.

Hayashizaki, Y., Shibata, H., Hirotsune, S., Sugino, H., Okazaki, Y., Sasaki, N., Hirose, K., Imoto, H., Okuizumi, H., Muramatsu, M., Komatsubara, H., Shiroishi, T., Moriwaki, K., Katsuki, M., Hatano, N., Sasaki, H., Ueda, T., Mise, N., Takagi, N., Plass, C., and Chapman, V. M. 1994. Identification of an imprinted U2af binding protein related sequence on mouse chromosome 11 using the RLGS method. *Nat Genet* **6:** 33-40.

Jähner, D., Stuhlmann, H., Stewart, C. L., Harbers, K., Löhler, J., Simon, I., and Jaenisch, R. 1982. *De novo* methylation and expression of retroviral genomes during mouse embryogenesis. *Nature* **298:** 623-8.

Jat, P. S., Cepko, C. L., Mulligan, R. C., and Sharp, P. A. 1986. Recombinant retroviruses encoding simian virus 40 large T antigen and polyomavirus large and middle T antigens. *Mol. Cell. Biol.* **6:** 1204-1217.

Kafri, T., Ariel, M., Brandeis, M., Shemer, R., Urven, L., McCarrey, J., Cedar, H., and Razin, A. 1992. Developmental pattern of gene-specific DNA methylation in the mouse embryo and germ line. *Genes Dev* **6:** 705-14.

Kaneko-Ishino, T., Kuroiwa, Y., Miyoshi, N., Kohda, T., Suzuki, R., Yokoyama, M., Viville, S., Barton, S. C., Ishino, F, and Surani, M. A. 1995. *Peg1/Mest* imprinted gene on chromosome 6 identified by cDNA subtraction hybridization. *Nat Genet* **11:** 52-59.

Kay, G. F., Penny, G. D., Patel, D., Ashworth, A., Brockdorff, N., and Rastan, S. 1993. Expression of *Xist* during mouse development suggests a role in the initiation of X chromosome inactivation. *Cell* **72:** 171-182.

Laird, P. W., Zijderveld, A., Linders, K., Rudnicki, M. A., Jaenisch, R., and Berns, A. 1991. Simplified mammalian DNA isolation procedure. *Nucleic Acids Res* **19:** 4293.

Leff, S. E., Brannan, C. I., Reed, M. L., Özçelik, T., Francke, U., Copeland, N. G., and Jenkins, N. A. 1992. Maternal imprinting of the mouse *Snrpn* gene and conserved linkage homology with the human Prader-Willi syndrome region. *Nat Genet* **2:** 259-64.

Lei, H., Oh, S. P., Jüttermann, R., Jaenisch, R., Goss, K. A., and Li, E. in preparation. An independently encoded mammalian *de novo* DNA methyltransferase is expressed in embryonic stem cells.

Leighton, P. A., Ingram, R. S., Eggenschwiler, J., Efstratiadis, A., and Tilghman, S. M. 1995a. Disruption of imprinting caused by deletion of the *H19* gene region in mice. *Nature* **375:** 34-39.

Leighton, P. A., Saam, J. R., Ingram, R. S., Stewart, C. L., and Tilghman, S. M. 1995b. An enhancer deletion affects both *H19* and *Igf2* expression. *Genes Dev* **9:** 2079-89.

Lemischka, I. R., Farmer, S., Racaniello, V. R., and Sharp, P. A. 1981. Nucleotide sequence and evolution of a mammalian α-tubulin messenger RNA. *J Mol Biol* **151:** 101-120.

Li, E., Beard, C., and Jaenisch, R. 1993. Role for DNA methylation in genomic imprinting. *Nature* **366:** 362-5.

Li, E., Bestor, T. H., and Jaenisch, R. 1992. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69:** 915-26.

Mann, J. R., Gadi, I., Harbison, M. L., Abbondanzo, S. J., and Stewart, C. L. 1990. Androgenetic mouse embryonic stem cells are pluripotent and cause skeletal defects in chimeras: implications for genetic imprinting. *Cell* **62:** 251-260.

Matsui, Y., Zsebo, K., and Hogan, B. L. M. 1992. Derivation of pluripotential embryonic stem cells from murine primordial germ cells in culture. *Cell* **70:** 841-7.

Monk, M., Boubelik, M., and Lehnert, S. 1987. Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development. *Development* **99:** 371-82.

Morgenstern, J. P., and Land, H. 1990. Advanced mammalian gene transfer: high titre retroviral vectors with multiple drug selection markers and a complementary helper-free packaging cell line. *Nucleic Acids Res* **18:** 3587-96.

Nagy, A., Gócza, E., Merentes Dias, E., Prideaux, V. R., Iványi, E., Markkula, M., and Rossant, J. 1990. Embryonic stem cells alone are able to support fetal development in the mouse. *Development* **110:** 815-821.

Nagy, A., Sass, M., and Markkula, M. 1989. Systematic non-uniform distribution of parthenogenetic cells in adult mouse chimaeras. *Development* **106:** 321-4.

Norris, D. P., Patel, D., Kay, G. F., Penny, G. D., Brockdorff, N., Sheardown, S. A., and Rastan, S. 1994. Evidence that random and imprinted *Xist* expression is controlled by preemptive methylation. *Cell* **77:** 41-51.

Poirier, F., Chan, C.-T. J., Timmons, P. M., Robertson, E. J., Evans, M. J., and Rigby, P. W. J. 1991. The murine *H19* gene is activated during embryonic stem cell differentiation *in vitro* and at the time of implantation in the developing embryo. *Development* **113:** 1105-14.

Razin, A., and Shemer, R. 1995. DNA methylation in early development. *Human Molec Genet* **4:** 1751-1755.

Reik, W., Collick, A., Norris, M. L., Barton, S. C., and Surani, M. A. 1987. Genomic imprinting determines methylation of parental alleles in transgenic mice. *Nature* **328:** 248-51.

Robertson, E. J. 1987. Embryo-derived stem cell lines. In Teratocarcinomas and embryonic stem cells: A practical approach (ed. E. J. Robertson), pp. 71-112. IRL Press, Oxford.

Rouleau, J., Tanigawa, G., and Szyf, M. 1992. The mouse DNA methyltransferase 5'-region. A unique housekeeping gene promoter. *J Biol Chem* **267:** 7368-77.

Sambrook, J., Fritsch, E., and Maniatis, T. 1989. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Sapienza, C., Peterson, A. C., Rossant, J., and Balling, R. 1987. Degree of methylation of transgenes is dependent on gamete of origin. *Nature* **328:** 251-4.

Sasaki, H., Jones, P. A., Chaillet, J. R., Ferguson-Smith, A. C., Barton, S. C., Reik, W., and Surani, M. A. 1992. Parental imprinting: potentially active chromatin of the repressed maternal allele of the mouse insulin-like growth factor II (*Igf2*) gene. *Genes Dev* **6:** 1843-1856.

Stewart, C. L., Stuhlmann, H., Jahner, D., and Jaenisch, R. 1982. *De novo* methylation, expression, and infectivity of retroviral genomes introduced into embryonal carcinoma cells. *Proc Natl Acad Sci U S A* **79:** 4098-102.

Stöger, R., Kubicka, P., Liu, C.-G., Kafri, T., Razin, A., Cedar, H., and Barlow, D. P. 1993. Maternal-specific methylation of the imprinted mouse *Igf2r* locus identifies the expressed locus as carrying the imprinting signal. *Cell* **73:** 61-71.

Stuhlmann, H., Jähner, D., and Jaenisch, R. 1981. Infectivity and methylation of retroviral genomes is correlated with expression in the animal. *Cell* **26:** 221-32.

Surani, M. A. 1994. Genomic imprinting: control of gene expression by epigenetic inheritance. *Curr Opin Cell Biol* **6:** 390-5.

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.

Szabo, P., and Mann, J. R. 1994. Expression and methylation of imprinted genes during in vitro differentiation of mouse parthenogenetic and androgenetic embryonic stem cell lines. *Development* **120:** 1651-60.

Szabo, P. E., and Mann, J. R. 1995. Biallelic expression of imprinted genes in the mouse germ line: implications for erasure, establishment, and mechanisms of genomic imprinting. *Genes Dev* **9:** 1857-1868.

Thomas, K. R., and Capecchi, M. R. 1987. Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. *Cell* **51:** 503-12.

Thomson, J. A., and Solter, D. 1988. The developmental fate of androgenetic, parthenogenetic, and gynogenetic cells in chimeric gastrulating mouse embryos. *Genes Dev* **2:** 1344-51.

Tremblay, K. D., Saam, J. R., Ingram, R. S., Tilghman, S. M., and Bartolomei, M. S. 1995. A paternal-specific methylation imprint marks the alleles of the mouse *H19* gene. *Nat Genet* **9:** 407-13.

van der Lugt, N., Maandag, E. R., te Riele, H., Laird, P. W., and Berns, A. 1991. A *pgk::hprt* fusion as a selectable marker for targeting of genes in mouse embryonic stem cells: disruption of the T-cell receptor δ-chain-encoding gene. *Gene* **105:** 263-7.

Villar, A. J., and Pedersen, R. A. 1994. Parental imprinting of the *Mas* protooncogene in mouse. *Nat Genet* **8:** 373-9.

Wevrick, R., Kerns, J. A., and Francke, U. 1994. Identification of a novel paternally expressed gene in the Prader-Willi syndrome region. *Hum Mol Genet* **3:** 1877-82.

Yeivin, A., and Razin, A. 1993. Gene methylation patterns and expression. In DNA methylation: molecular biology and biological significance (ed. J. P. Jost and H. P. Saluz), pp. 523-68. Birkhäuser Verlag, Basel, Switzerland.

Zuccotti, M., and Monk, M. 1995. Methylation of the mouse *Xist* gene in sperm and eggs correlates with imprinted *Xist* expression and paternal X-inactivation. *Nat Genet* **9:** 316-20.

FIGURE LEGENDS

Figure 1. CHIP (cDNA homologous insertion protocol) (*A*) Diagram representing the *Dnmt*-targeting construct pMWT-PURO and its method of integration into the *Dnmt* locus. pMWT-PURO is represented schematically. A 4.8-kb *Dnmt* genomic sequence that is shared between pMWT-PURO and the *Dnmt* locus is indicated by a broken line and two downstream exons (dark boxes) drawn to scale. This was fused to the remainder of the *Dnmt* cDNA (striped box) in the second exon at a *Nae*I (N) site, 20 bp upstream of the initial ATG. A *Pgk-1*-derived polyadenylation sequence (Pgk-PolyA) was inserted at the end of the cDNA, and the plasmid also contained a puromycin-selectable marker (Pgk-Puro) and pGEM-11zf(+) bacterial plasmid sequences (Amp^r). The *Hind*III linearization site is indicated (H). The *Dnmt* locus targeted with this construct is represented by a broken line upstream of the same two exons depicted in pMWT-PURO. The genomic probe (HV) for Southern blot analysis and the *Sca*I (S) digestion product expected from the locus are also displayed. *Hind*III (H) and *Nae*I (N) sites are as indicated above. Within the vertical arrow is a close-up, drawn to scale, of the crossover event in which pMWT-PURO was targeted to the *Dnmt* locus. The plasmid was cut at the unique *Hind*III (H) site indicated, and after electroporation the entire vector homologously inserted into the endogenous gene, yielding the genomic structure shown beneath. pMWT-PURO shares 4.1 and 0.7 kb of sequence identity with the target locus 5' and 3', respectively, of the *Hind*III linearization site. The bottom schematic illustrates the duplication of

the 4.8-kb *Dnmt* genomic sequence expected upon integration, seen with the exons (dark boxes) drawn to scale. It also displays the two products expected from a *Sca*I (S) digestion, using the HV probe. The scale is interrupted by several breaks in the genomic sequence, indicated by slashes through the line. The integrated pMWT-PURO plasmid is indicated by a line. (*B*) Southern blot analysis of *Dnmt*-targeted clones. *Sca*I-digested genomic DNA displayed a single 17.5-kb band from both untargeted *Dnmt* alleles in wild-type J1 (lane *1*) and homozygous *Dnmt* mutant (lane *2*) cells. Homologous insertion of a single copy of the pMWT-PURO plasmid at the *Dnmt* locus resulted in production of 21- and 7.4-kb bands (lanes *3* , *4*). The untargeted allele is also visible at 17.5 kb. *Hind*III-digested λ DNA size standards are indicated to the left. (*C*) Northern blot analysis of *Dnmt* expression in ES cell lines stably transfected with pMWT-PURO. Total RNA was prepared from the homologous integrants shown in Fig. 1B (lanes *4, 5*), and the *Dnmt* cDNA was used as a probe for the Northern blot (upper figure). The following cells are shown for comparison: wild-type J1 (lane *1*), a heterozygous (lane *2*), and homozygous (lane *3*) *Dnmt* mutant cell line. The lower figure represents the same blot re-hybridized with an α-tubulin cDNA, to control for amount of RNA loaded. The positions of the 28S and 18S ribosomal RNA species are indicated on the left.

Figure 2. Normal genomic DNA methylation levels are restored upon expression of the *Dnmt* cDNA. Southern blot analysis of *Hpa*II-digested genomic DNA, probed with either of two repetitive sequences: the Moloney Murine Leukemia Virus (MoMuLV) cDNA or

a centromeric minor satellite repeat. Wild-type J1 (lane *1*) and homozygous *Dnmt* mutant ES cells (lane *2*) are shown, followed by the rescued mutant cells seen in Fig. 1 (lanes *3, 4*). *Hind*III-digested λ DNA size standards are indicated to the left of each blot.

Figure 3. Methylation of the imprinted genes *H19* and *Igf2r* is not restored in rescued ES cells and their differentiated derivatives, but methylation of *Xist* and *Igf2* is restored to normal. Southern blot analysis of ES cells and fibroblast cultures derived from postnatal chimeras made with rescued ES cells. (*A, B, C, D*) Genomic DNA was isolated from undifferentiated (U) or differentiated (D) ES cells, cut with *Sac*I, *Pvu*II, Eco*RI*-Hind*III*, or *Eco*RV ((*A*), (*B*), (*C*), and (*D*), lane *1*, respectively), and probed with the gene sequence indicated beneath each respective blot. All other lanes were treated similarly but include the methylation-sensitive enzymes *Hha*I (*A, D*) or *Hpa*II (*B, C*). Wild-type J1 (+/+), homozygous *Dnmt* mutant (-/-), and two independent rescued cell lines (-/- + cDNA) are shown. In lanes *10-12* of (*A*) and (*C*), and lanes *6-8* of (*B*), SV40 virus-immortalized fibroblast cultures were examined. Fibroblasts were derived from an embryonic day 16.5 wild-type embryo (lane *10* of (*A*) and (*C*) and lane *6* of (*B*)) and were compared with puromycin-resistant fibroblasts derived from two postnatal chimeras (lanes *11, 12* of (*A*) and (*C*) and *7, 8* of (*B*)). The puromycin-resistant cells were shown to be derived exclusively from the injected rescued ES cells (data not shown). A map of the examined area is shown underneath each blot, with flanking *Sca*I (S), *Pvu*II (P), *Eco*RI (E), *Hind*III (H), and *Eco*R V (V) sites indicated. The probes used and the genomic regions

examined are described in Results and Materials and Methods. *Hha*I (*A*, *D*) and *Hpa*II (*B*, *C*) sites are represented by small vertical lines, while transcriptional start sites are shown in (*A*) and (*D*) by an arrow. *Hind*III-digested λ DNA size standards are indicated to the left of each blot.

Figure 4. Normal expression of imprinted genes is not restored in rescued ES cells. (*A, B, D, E*) Total RNA isolated from undifferentiated (U) and *in vitro* differentiated (D) ES cells was analyzed by RNase protection using antisense RNA probes for *H19* (*A*), *Igf2r* (*B*), *Igf2* (*D*), and *Xist* (*E*). Antisense RNA probes for cytoplasmic actin (*A, B*) or γ actin (*D, E*) were used as loading controls. Wild-type J1 (+/+), homozygous *Dnmt* mutant (-/-), and rescued ES cells (-/- + cDNA) are shown. (*C*) *Igf2r* RNase protection assay performed upon subcloned lines of immortalized fibroblasts prepared from a postnatal day 7 chimera made between rescued EG cells and a wild-type BALB/c host. Lanes *1,2* show two wild-type BALB/c lines, while lanes *3-6* show four puromycin-resistant lines derived from the injected rescued EG cells. Antisense RNA probes for *Igf2r* and γ actin (as a loading control) were used.

Figure 5. Normal methylation and expression patterns of imprinted genes are restored after passage through the germline. Southern blot and RNase protection analysis of germline embryos and pups produced by a chimera made from rescued ES cells. (*A, B, C*) RNase

protection analysis performed upon whole RNA from embryos, using probes indicated in Fig. 4. A male chimera was mated to BALB/c females, and embryos analyzed at embryonic day 19.0. Black-eyed embryos (R), derived by germline transmission of the rescued ES cells, were compared to albino wild-type littermates (+/+). (*D, E, F*) Southern blot analysis of genomic DNA from postnatal mice. DNA prepared from tails of three-week old mice was digested with methylation-sensitive enzymes and probed with *H19*, *Igf2*, or *Igf2r* sequences, as described in Fig. 3A, 3C, and 3B, respectively. A wild-type F1 (BALB/c x 129/Sv) mouse (lane *1*) was compared to three germline pups derived from a cross between a chimera generated with rescued ES cells and a BALB/c female (lanes *2-4*). *Hind*III-digested λ DNA size standards are indicated to the left of each blot.

Figure 6. (*A*) Developmental changes in methylation of bulk DNA and imprinted genes. Bulk DNA and non-imprinted genes undergo consecutive waves of demethylation and *de novo* methylation in early development before the adult somatic patterns are established (Razin and Shemer, 1995). *Xist* is depicted as non-imprinted as its expression is only imprinted in extraembryonic tissues. The monoallelic methylation patterns of the imprinted genes *H19* and *Igf2r* are established in the male and female gametes, respectively, and are not altered throughout development. (*B*) Methylation is almost completely erased for all examined DNA sequences in *Dnmt* mutant ES cells and embryos. Expression of the *Dnmt* cDNA in mutant ES cells restores the methylation of bulk DNA and single-copy genes such as *Xist*, but does not restore methylation to imprinted

genes such as *H19* and *Igf2r*. Methylation of these genes is restored only after transmission through the germline of male and female chimeras, respectively, implying the presence of gamete-specific DNA methylation processes for imprinted genes. Degree of shading in boxes indicates extent of DNA methylation. Maternally- and paternally-derived alleles are both displayed for *H19* and *Igf2r*.

Acknowledgements

I consider it a blessing that I have been able to attend college and university just thirty miles from my parents, my five siblings, their spouses, and their children. They have encouraged both scholarship and dissipation, and in their proper measures. The marshy tidewaters of the North Shore, where they dwell, shall never be surpassed in my esteem.

Of course, I thank Rudolf Jaenisch for allowing me to work in his laboratory and encouraging the research so enthusiastically.