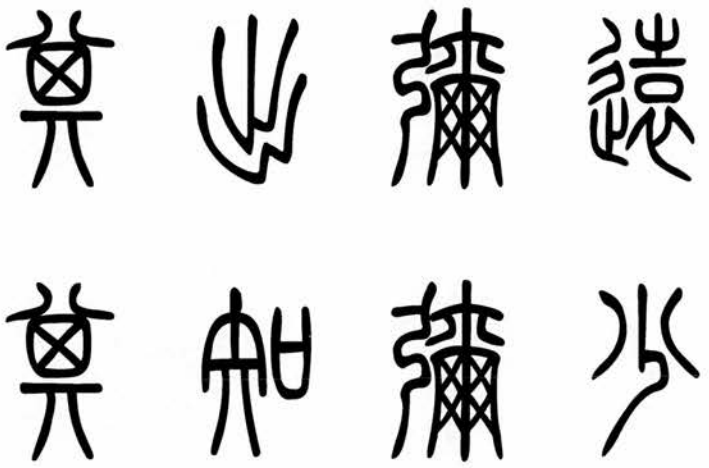
De novo methyltransferases, DNA methylation and cancer: a transgenic model

Tom Latham

PhD
The University of Edinburgh
2007



The further you travel
The less you understand

-Lao Zi

Declaration

I declare:

a) that the thesis has been composed by myself

(b) the work is my own unless explicitly indicated

(c) that the thesis has not been submitted for any other degree or professional qualification.

Signed:

22/05/07

Acknowledgements

Bernie Ramsahoye for being everything I could have hoped for as a supervisor, and his seemingly endless supply of ideas.

Lesley Forrester, Irena Stancheeva and David Melton for helping with my thesis committee.

John Davies and the consultant staff of the Haematology department for indulging my research habit and for generous financial support during some difficult periods.

David Brownstein for help with histology.

Jim Selfridge for help with blastocyst injections.

The staff of the Biomedical Research facility, Western General Hospital, for looking after the mice.

Charles Patek for providing the Apc mice and Adrian Bird for providing the Mbd2 mice.

Everyone in the John Hughes Bennett laboratory for providing friendship and a pleasant place to work.

The work was generously funded by grants from the Melville Trust for the Care and Cure of Cancer and Cancer Research UK.

Abstract

One of the most important controversies in the field of cancer epigenetics is the question of whether aberrant CpG island methylation of tumour suppressor genes can cause cancer or whether such abnormalities of DNA methylation are secondary to the malignant process. We address this question by prospectively causing abnormal methylation events *in vivo*. We have produced transgenic mice which over-express the *de novo* methyltransferase Dnmt3b under the widely expressed CAG promoter. *Tg(Dnmt3b)*+ mice devlop normally and are fertile, but die at 4-5 months, developing dilated cardiomyopathy. CpG island methylation is globally increased, and abnormal methylation of specific genes, including the tumour suppressor genes C*dkn1a*, *Cdkna2a* and *Hic1* is detectable. However, there is no spontaneous cancer incidence. Crossing the mice with the intestinal tumour prone $Apc^{+/min}$ mice does not lead to a significant increase in tumour number, methylation of the *Apc* promoter occurs with equal frequency in microdissected tumours from $Apc^{+/min}$ mice regardless of *Tg(Dnmt3b)*+ genotype and is not detectable in normal mucosa. However the proportion of tumours showing dysplasia is increased. The results show that increases of methylation can be well tolerated, suggesting that although methylation can be increased, active silencing by methylation is more unusual. In particular, the active silencing of tumour suppressor genes by *de novo* methylation is unlikely to be a primary event in the formation of tumours, although it may play a role in modulating tumour progression.

1. Introduction

Methylation of DNA is an important feature of the biology of higher eukaryotic organisms, appearing essential for normal mammalian development. Abnormalities of methylation are some of the most common abnormalities seen in cancer, and are widely viewed as having a causative role in the development of cancer. In this thesis, the role of DNA methylation in normal biology and cancer is reviewed, and the development of a transgenic mouse model for abnormal increases in methylation is described.

1.1 Essential biology of DNA methylation and CpG islands

Covalent modification of DNA by post-synthetic addition of a methyl group occurs in the full spectrum of living organisms. Its oldest function in evolutionary terms is as part of the restriction enzyme defence system against phage viruses in prokaryotic organisms, which show methylation of both adenine and cytosine bases. In eukaryotic organisms, methylation appears to be confined to the 5'- position of the cytosine ring. In evolution as a whole, DNA methylation is clearly not absolutely essential for life or the development of multicellular organisms, since the amount of methylation present in the genome varies widely among eukaryotic organisms. Yeasts and *Caenorrhabditis elegans (C.Elegans)* do not have detectable methylation and do not have detectable DNA methyltransferase genes in their genomes, and DNA methylation is detectable but rare in *Drosophila Melanogaster (D.Melanogaster)*, whereas methylation is common in plants and vertebrates (For review see Bird 2002).

Methylation in mice is almost entirely (but not completely) restricted to cytosine residues immediately 5' to guanosine in genomic DNA- the CpG dinucleotide (Ramsahoye, *et al* 2000). In organisms which have substantial DNA methylation, the CpG dinucleotide is under represented in the genome. The ratio CpG to GpC as measured by nearest neighbour analysis is 0.23, indicating that the frequency of CpG in the mammalian genome is approximately 4 times less than would be expected by

chance given the frequencies of C and G in the genome (Swartz, *et al* 1962). It is thought that the depletion of the CpG dinucleotide is likely to have occurred during evolution by conversion of methylated cytosines to thymidine. Unmethylated cytosines are converted to uracil by hydrolytic deamination which occurs spontaneously even at neutral pH (Lindahl and Nyberg 1974). This can be recognised by DNA repair mechanisms as incorrect and removed by base excision repair. However, methylcytosine is deaminated to form thymidine, creating a C to T mutation which is then heritable if not corrected (Ehrlich, *et al* 1986). The methyl binding protein Mbd4 acts as a DNA repair enzyme repairing C-T mismatches and may offer protection against such errors (Millar, *et al* 2002).

However, there are regions of the genome in which the expected frequency of CpG is preserved, which are referred to as CpG islands. CpG islands and their relationship to DNA methylation were initially identified by comparing discordant restriction digestion patterns between the restriction enzyme *HpaII* (CCGG) which is inhibited by CpG methylation and its isoschizomer *MspI* which is methylation insensitive. *MspI* digests genomic DNA into low molecular weight fragements, whereas most of the DNA digested with *HpaII* remains high molecular weight, indicating that the majority of CpG dinucleotides in *HpaII* sites are methylated (Singer, *et al* 1979). The unmethylated minority of the genome which was sensitive to *HpaII* digestion was found to be C+G and CpG rich. Based on sequence data from this analysis, sequence criteria for defining a CpG island were developed: G+C content greater than 0.5 and observed:expected CpG dinucleotide ratio of >0.6, both occurring within a window of 200bp or greater (Gardiner-Garden and Frommer 1987). These criteria are the most widely used for defining a CpG island, although alternative criteria, derived from modern genomic sequencing data have been proposed (increasing the required G+C content to >0.55 and the window size threshold to greater than 500bp), which decreases the identification of regions of repetitive sequence (Takai and Jones 2002).

CpG islands, then, are regions of the DNA which are enriched for the CpG dinucleotide and appear to be protected from DNA methylation, in contrast to the bulk of DNA which is CpG depleted and methylated. Classification of CpG islands according to the original definition shows that the promoter region lies within a CpG island in 50% of known genes, including all housekeeping genes, and that most CpG islands, excluding those within repetitive sequence, are associated with genes (Lindsay and Bird 1987). The chromatin structure of CpG islands is also different to bulk DNA with increased acetylation of histones H3 and H4, decreased levels of H1 and nucleosome-free regions (Tazi and Bird 1990). The majority of CpG island genes are unmethylated even if the gene is not expressed in adult life.

In normal physiology, there are three established exceptions to the principle that CpG island DNA is unmethylated: the phenomena of gene imprinting, X chromosome inactivation and silencing of transposable elements. These phenomena are associated with silencing of the methylated gene. X chromosome inactivation and imprinting share the phenomenon of monoallelic gene expression. X inactivation involves the random silencing of transcription from one female chromosome, whereas imprinting involves silencing of one allele according to parental origin. In these cases, methylation is associated with silencing of the gene on the inactive allele. Since both alleles are exposed to the same conditions and transactivating factors in the nucleus, this suggests that DNA methylation is part of a system to allow regulation of transcription that is independent of regulation by transcription factors. DNA methylation patterns can be copied on to the daughter strand during cell division, and therefore allow heritable control of gene expression at the methylated allele. Imprinting and X inactivation provide model systems for investigating the mechanisms of establishing epigenetic silencing and the mechanisms by which silencing is effected. These mechanisms are discussed in the following section.

1.2 Establishment of methylation patterns

1.2.1 The eukaryotic DNA methyltransferase system

There are three known families of DNA methyltransferases in eukaryotes of which two have an established role in producing DNA methylation. All three families share a conseved C-terminal catalytic domain which shows considerable homology to bacterial methyltransferases, but differ substantially in their N-terminal regions (For review see Chen and Li 2004).

The first to be discovered was Dnmt1, which was identified by protein purification from murine cells and peptide sequencing (Bestor, *et al* 1988). Murine Dnmt1 has a somatic cell specific isoform of 1620 amino acids and an oocyte specific isoform initiated within exon 4 of 1616 amino acids, both of which are catalytically active. The Dnmt1 N-terminal region includes nuclear localisation, replication focus targeting, methyl-binding, PCNA binding and bromo-adjacent homology domains (Figure 1.1). Dnmt1 is constitutively expressed in proliferating cells and is ubiquitous in somatic tissues during development. Dnmt1 is a nuclear protein and changes its nuclear distribution during cell cycle, associating with replication foci during S phase (Leonhardt, *et al* 1992). Dnmt1 has a 5- to 50 fold increased activity against hemi-methylated substrates compared to unmethylated substrates *in vitro* (Pradhan, *et al* 1999, Ruchirawat, *et al* 1987, Yoder, *et al* 1997), and fails to induce *de novo* methylation in ES cells (Chen, *et al* 2003a) or *D.Melanogaster* (Lyko, *et al* 1999). Deletion of *Dnmt1* in embryonic stem cells leads to progressive demethylation of the genome (Li, *et al* 1992) but does not affect *de novo* methylation of integrated MMLV provirus DNA (Lei, *et al* 1996). Dnmt1 is essential for development, with homozygous null mutant mice failing to survive past mid-gestation and developing genomic demethylation (Li, *et al* 1992). Mice with a hypomorphic *Dnmt1* allele which express Dnmt1 at approximately 10% normal level are

viable but runted and show reduced levels of global methylation, though not as severe as in the null mice (Gaudet, *et al* 2003). These results suggest that Dnmt1 is primarily concerned with the maintenance of DNA methylation patterns by copying the methylation signal following cell division, and has little if any involvement in the establishment of new DNA methylation patterns.

The second family of established importance is the Dnmt3 family which are the major enzymes concerned with *de novo* methyltransferase activity. The residual ability of *Dnmt1* null ES cells to methylate proviral DNA had suggested the presence of an independent methyltransferase mediating *de novo* methyltransferase activity. The Dnmt3 family was subsequently discovered by identification of EST clones from a human database showing homology to bacterial DNA methyltransferases (Okano, *et al* 1998a, Xie, *et al* 1999). There are two members of the family with active catalytic domains. The murine *Dnmt3a* gene encodes at least 2 proteins, Dnmt3a of 908 amino acids and Dnmt3a2 of 689 amino acids, by initiation from alternative promoters. Dnmt3a is localised to heterochromatin wheras Dnmt3a2 is diffusely distributed in the nucleus. The *Dnmt3b* gene encodes at least 4 isoforms in the mouse: Dnmt3b1, Dnmt3b2, Dnmt3b3 and Dnmt3b6. Dnmt3b1 and Dnmt3b2 are active enzymes, whereas the other isoforms lack motifs in the catalytic domain and are inactive (Chen and Li 2004). In addition to these enzymes, the Dnmt3 family also includes the Dnmt3L protein which lacks an active catalytic domain but which still appears to play a role in determining imprinting by interaction with other methyltransferases (Aapola, *et al* 2001, Bourc'his, *et al* 2001, Hata, *et al* 2002).

The Dnmt3a and 3b proteins are closely related with the N terminal region including 2 conserved regions: a cysteine rich region with a zinc finger and plant homeodomain- like sequence, which interacts with the histone deacetylase Hdac1 and a putative transcriptional repressor RP58, and a PWWP domain, which is of unknown function but may be involved in DNA binding and is essential for localisation to heterochromatin. Dnmt3a2 and Dnmt3b1 are highly expressed in ES cells, germ cells and early embryos,

but are almost undetectable in mature somatic tissues, whereas Dnmt3a and Dnmt3b3 are expressed at low levels in most somatic tissues (Chen, *et al* 2002). Dnmt3a and Dnmt3b have equal affinity *in vitro* for unmethylated and hemi-methylated DNA and were shown to be essential for *de novo* methylation by targeted deletion of the catalytic domain of either or both enzymes in embryonic stem cells (Okano, *et al* 1998a). Methylation of retroviral DNA following infection with MMLV was found to be absent in cells with both enzymes deleted, showing that the two enzymes together probably represent the entire set of *de novo* methyltransferases, and there is redundancy between the two enzymes in the function of methylation of retroviral insertions.

Both Dnmt3a and Dnmt3b are essential for normal development. *Dnmt3b*$^{-/-}$ embryos do not survive to term and have multiple defects including neural tube defects, ventricular septal defects, liver hypoplasia and growth impairment, although development appears normal until day E9.5 (Okano, *et al* 1998a, Ueda, *et al* 2006). *Dnmt3a*$^{-/-}$ embryos survive to term but are runted and die by 4 weeks of age (Okano, *et al* 1998a). Double mutant *Dnmt3b*$^{-/-}$ *Dnmt3a*$^{-/-}$ embryos, produced by intercrossing compound heterozygote mice have a much more severe phenotype, lacking somites and dying before E11.5. Double mutant embryos have impaired levels of methylation of endogenous retroviral DNA, which remains constant between blastocysts and E9.5 embryos. In contrast, *Dnmt1* null mice show loss of methylation between the blastocyst stage and E9.5, suggesting that *de novo* methylation is absent but maintenance methylation is intact. The results suggest that Dnmt3b and Dnmt3a have distinct functions during development but have overlapping functions in early embryogenesis. Further evidence for the importance of DNMT3B in humans comes from the rare inherited ICF syndrome which has been been shown to be caused by deletions involving the *DNMT3B* locus at 20q11 (OMIM reference #242860 http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM). The syndrome consists of variable immunodeficiency, centromeric instability involving chromosomes 1, 9 and 16 and facial dysmorphism. ICF patients show demethylation of DNA, with changes in

gene expression, decondensation of pericentromeric heterochromatin and impaired methylation of retroviral sequences.

The set of known mammalian DNA methyltransferases is completed by Dnmt2 which has a less well defined function. The *Dnmt2* gene encodes a protein of 415 amino acids in the mouse or 391 amino acids in the human (Van den Wyngaert, *et al* 1998, Yoder and Bestor 1998). In addition, Dnmt2 like proteins are present in *Saccharomyces pombe* (Wilkinson, *et al* 1995) and *D.Melanogaster* (Tweedie, *et al* 1999). Despite the presence of all regions of the catalytic domain, cloned Dnmt2 proteins do not have detectable catalytic activity *in vitro*. While there is some evidence that Dnmt2 proteins may have activity *in vivo* in *D.Melanogaster*, and may be involved with non-CpG methylation, no function has been detected in mammalian systems. Deletion of Dnmt2 has no effect on methylation patterns or ability to methylate DNA *de novo* in mouse ES cells, and has no detectable effect in homozygous knockout mice or on embryonic development in *D.Melanogaster* (Okano, *et al* 1998b). Interestingly a recent paper has shown that Dnmt2 has RNA methyltransferase activity, methylating a specific cytosine residue in the aspartate tRNA (Goll, *et al* 2006). The authors speculate that this might indicate evolution of the eukaroyotic DNA methyltransferases from prototypic RNA methyltransferases rather than prokaryotic DNA methyltransferases. However, the data do not support functional importance for Dnmt2, although the fact that it has remained highly conserved through evolution suggests that a vital role may still await discovery.

1.2.2 Factors determining methylation patterns

Although the presence of the appropriate DNA methyltransferases is clearly a necessary condition for the establishment of DNA methylation patterns, the pattern of enzyme activity is clearly an insufficient explanation for the determination of which regions in the genome are methylated, and in particular why CpG islands are normally unmethylated but are methylatable in certain circumstances.

There is evidence that targeting of *de novo* methyltransferases to specific regions can involve DNA sequence factors, protein-protein interactions, and RNA dependent mechanisms, and it is probable that different combinations of these factors are important in different contexts.

DNA sequence motifs involved in controlling methylation have been discovered for several imprinting genes including the *Igf2r* gene and the *H19* gene. The *Igf2r* gene is a maternally expressed gene, in which a region in intron 2 becomes methylated on the maternal chromososme. A 113 bp segment within this region appears to control methylation, with mutation at the 5'end of this segment leading to bi-allelic methylation, and mutation at the 3' end preventing methylation on both alleles (Birger, *et al* 1999). In the paternally methylated *H19* gene, point mutations in CTCF binding sites in the differentially methylated regions of the maternal chromosome allows methylation of the maternal chromosome to occur (Schoenherr, *et al* 2003), whereas deletion of a 1.6kb region of the parental differentially methylated region leads to mice being unable to maintain methylation and silencing of the promoter region (Thorvaldsen, *et al* 1998). These examples suggest that both repression and promotion of methylation by sequence motifs can be important for imprint establishment.

There has been considerable interest in the role of protein-protein interactions with other transcription repression systems in the recruitment of DNA methyltransferases and establishment of methylation. The largest body of evidence concerns interactions with histone proteins and their covalent modifications. As noted above, there is partial localisation of the *de novo* methyltransferases to pericentric heterochromatin, which is also characterised by methylation of lysine 9 of histone H3 (H3K9). The Suv39h histone methyltransferase co-immunoprecipitates with Dnmt3b, and knockout of Suv39h in embryonic stem cells leads to loss of H3K9 methylation, and loss of CpG methylation at pericentric heterochromatin (Lehnertz, *et al* 2003), suggesting that in at least some cases H3K9 methylation is required for DNA methylation. Other protein interactions that have been proposed to induce DNA methylation by recruitment of *de novo*

methyltransferases include polycomb group proteins (Mager, *et al* 2003, Vire, *et al* 2006), and recruitment by site-specific transcription factors- which has been demonstrated in a pathological context for the *Myc* oncogene (Brenner, *et al* 2005) and the PML-RARA fusion protein (Di Croce, *et al* 2002). However the significance of site-specific recruitment by transcription factors remains to be demonstrated in normal physiology.

Finally, an emerging area of interest is the role of RNA-dependent mechanisms in directing DNA methylation. In plants, transcriptional as well as translational silencing is induced by RNAi, and this is often associated with *de novo* methylation of the corresponding promoter. While this occurs much less commonly in mammalian cells, there are reports of *de novo* methylation occurring in mammalian cell lines following RNAi directed at the promoter region (Castanotto, *et al* 2005, Morris, *et al* 2004a) accompanied with transcriptional silencing. In one of these reports (Morris *et al* 2004a), there was robust evidence of transcriptional silencing of a reporter transgene by RNAi, which was reversed with 5'-azacytidine treatment. However the extent of methylation induced by RNAi in this study is not clear, since the only direct evidence for methylation was a PCR assay at a single restriction site. These studies have the weakness that they are performed in transformed cell lines in which epigenetic regulation may not reflect normal physiology. However, there is a case report of DNA methylation at the alpha globin locus occurring in an alpha-thalassaemia patient with a deletion which results in a truncated, widely expressed gene (*LUC7L*) ,which is transcribed from the antisense strand, becoming juxtaposed to a structurally normal α-globin gene and leads to the presence of antisense alpha globin transcripts (Tufarelli, *et al* 2003). The effects on DNA methylation of disruption of the RNAi machinery by deletion of the *Dicer* gene are conflicting (Kanellopoulou, *et al* 2005, Murchison, *et al* 2005) but raise the possibility of the involvement of endogenous miRNA or siRNA in the normal control of methylation at some loci. However, the fact that there are very few reports describing RNAi-induced DNA methylation in mammalian cells, even

though it is an attractive and obvious phenomenon to seek, would tend to suggest that it is of genuinely rare occurrence.

1.3 Mechanisms of transcriptional regulation

We have seen in the paragraph above that there is considerable interaction with other systems of transcriptional silencing in the establishment of methylation at specific loci. In this section, the complementary interactions are reviewed- the mechanism by which DNA methylation marks may promote transcriptional silencing. The simplest mechanism is that the methyl group may prevent access of transcription factors to their binding sites. The clearest example that this is the case is in the role of CTCF binding in the silencing of the imprinted *Igf2/H19* locus. Expression of the *Igf2* gene is dependent on a downstream enhancer element. Binding of the enhancer blocking CTCF protein at the imprinting control region between the *Igf2* gene and its enhancer thus prevents expression of the gene. The imprinting control region is methylated on the paternal allele, corresponding to its paternal- restricted expression pattern; CTCF binding can be shown to be abolished both by methylation *in vitro*, and by mutation of the binding sites *in vivo* (Bell and Felsenfeld 2000).

A second mechanism of transcriptional regulation by methylated DNA involves the interaction with methyl binding proteins. A family of methyl binding proteins (Fig 1.1) sharing a common methyl-CpG binding domain has been identified, which comprises the proteins Mbd1, Mbd2, Mbd3, Mbd4 (Hendrich and Bird 1998) and MeCP2 (Lewis, *et al* 1992). In addition there are further methyl-DNA binding proteins including Kaiso proteins, which do not posess a classical methyl binding domain but recognise methylated DNA through C-terminal zinc finger domains (Prokhortchouk, *et al* 2001) (for review see Hendrich and Tweedie 2003). The first described entity with methyl binding activity (Meehan, *et al* 1989) is now known to be a protein complex (MeCP1), which includes the Mbd2 protein (Ng, *et al* 1999) and the nuclear remodelling complex (NuRD) which consists of multiple proteins with transcriptional repressive activities

(Zhang, *et al* 1999), including the histone deactylases Hdac1 and 2. Mammalian Mbd3 does not have methyl DNA binding activity (Hendrich and Bird 1998) but also interacts with the NuRD complex separately to Mbd2. MeCP2 was the first distinct methyl binding protein to be isolated (Lewis, *et al* 1992, Meehan, *et al* 1992), has strong affinity for methyl DNA and also recruits histone deacetylases. Mbd1 also has transcription repression activity which is partially mediated by histone deactetylase recruitment, and has the ability to inhibit transcription from both methylated and unmethylated sequence (Cross, *et al* 1997, Fujita, *et al* 2000, Fujita, *et al* 1999). Finally Mbd4 appears to have a quite distinct function being more concerned with DNA repair in order to prevent C->T transitions caused by deamination of methylcytosine, as predicted by its increased affinity for mCpG:TpG mismatches and confirmed by the increased frequency of C->T transitions in $Mbd4^{-/-}$ knockout mice (Millar, *et al* 2002, Wong, *et al* 2002).

Despite the ability of methyl binding proteins to silence transcription upon binding to methylated DNA, the alterations in transcription and phenotypic changes of knockout mice for most individual methyl binding proteins are in fact fairly subtle, in contrast to the effects of deletion of the DNA methyltransferase enzymes. The exceptions are MeCP2 which shows neurological features similar to those seen in Rett syndrome (Guy, *et al* 2001), and Mbd3 (Hendrich and Bird 1998), which is even more surprising as it does not have methyl binding activity. The obvious explanation for this contrast is that there is redundancy in the methylation–reading machinery, with one of the remaining methyl binding proteins being able to compensate for the loss of one of the others. The evolutionary evidence that duplication of methyl binding protein genes appears to have occurred at the invertebrate-vertebrate transition, which corresponds to the increase of methylation in the genome, is perhaps supportive of this hypothesis.

An alternative, more radical possibility that is difficult to discount completely is that many of the regulatory functions of the DNA methyltransferases are mediated through mechanisms other than DNA methylation, and it is in fact the DNA methylation itself that is redundant. Since Dnmt3 proteins are able to interact directly with histone

deacetylases, and Dnmt1 also has multiple protein interactions, their role in transcriptional regulation may be primarily effected through this route. In addition, there is evidence in a number of models that gene silencing and histone modifications (in particular histone 3 lysine 9 methylation) precede, and are necessary for development of DNA methylation (Mutskov and Felsenfeld 2004). It has also been shown that silencing of X-inactivated genes (Sado, *et al* 2004), and establishment of imprinted gene expression (Lewis, *et al* 2004), can occur without *de novo* DNA methylation, suggesting that methylation is not necessary for epigenetic silencing. However, the accepted model is that histone modifications, chromatin structure and DNA methylation form an interrelated system mediating epigenetic silencing. In this model, recruitment of methyl binding proteins by methylated DNA and recruitment of DNA methyltransferases by histone modifications form a positive feedback loop to maintain the stability of transcriptional silencing. Since the heritability of DNA methylation is a well defined mechanism, it may be that the most important role of DNA methylation is to promote heritability of transcriptional silencing during cell replication, although experimental evidence that organisms that have DNA methylation have greater stability of epigenetic regulation during replication is lacking.

1.4 DNA methylation and genome stability

Although the majority of work on DNA methylation has focussed on its role in transcriptional silencing, it may have an equally important role in genome stability. As noted above, the DNA methylation system may have evolved as part of the bacterial defence system against viruses, and it retains this importance in eukaryotes. The methylation of integrated retroviral sequences (Harbers, *et al* 1981) and the reactivation of retroviruses on treatment with demethylating agents (Jahner, *et al* 1982) has long been established. In addition to this role, there is evidence that deficiencies in both *de novo* and maintenance DNA methylation lead to increased chromosomal instability. As noted above, in the ICF syndrome, partial loss of function of Dnmt3b leads to pericentromeric demethylation, and centromeric instability; these results have also been

shown in a mouse model (Ueda, *et al* 2006). Deletion of the maintenance enzyme Dnmt1 in embryonic stem cells (Chen, *et al* 1998) led to interstitial deletions; the importance of this effect was confirmed *in vivo* by the finding of increased chromosome alterations in lymphomas from Dnmt1 hypomorphic mice compared to that found in retroviral tumours (Gaudet, *et al* 2003). The balance between the effects of demethylation on genome stability and the effects of hypermethylation on gene silencing is a particularly important question in tumour biology and will be discussed at several points below.

1.5 DNA methylation in developmental biology

The establishment of methylation patterns is a dynamic process during development, with waves of genome-wide demethylation and remethylation occurring. Following fertilisation, the genome becomes demethylated, although imprinted genes are spared demethylation, retaining their parent-specific methylation pattern. Methylation patterns are subsequently re-established between the morula and blastocyst stages of development. There is a second wave of demethylation which occurs in primordial germ cells during which methylation patterns are reprogrammed in a sex specific manner to establish parental-specific imprinted marks in gametes to pass on the appropriate imprints to the next generation. Further *de novo* methylation rarely occurs in development subsequent to gastrulation.

Given the expansion of genomic DNA methylation frequency in higher organisms, it is tempting to suggest that DNA methylation has a role in maintaining tissue-specific gene expression patterns, enabling stability of tissue specification for the development of complex organs. While methylation patterns have long been shown to vary between tissues, a correlation between methylation and expression has been harder to establish. For tissue specific genes without a CpG island promoter, there are several examples where methylation of specific sites appears to regulate transcription, of which the β globin locus is probably the best studied example (Busslinger, *et al* 1983, Yisraeli, *et al*

1988). In contrast, the majority of CpG island genes are not regulated by methylation- the CpG island is unmethylated regardless of the transcriptional state. The strongest evidence that tissue specific expression can in some cases be regulated by methylation of CpG island genes comes from the maspin (*SERPINB5*) gene which shows a strong inverse correlation between expression and promoter methylation status in primary cell lines from various tissues (Futscher, *et al* 2002). Other strong candidates include the MAGE genes which are methylated in somatic tissues but unmethylated and expressed in the testis (De Smet, *et al* 1999). More recently, RLGS experiments from primary mouse tissues have identified as many as 150 CpG island regions which show tissue specific methylation (including CpG islands in intronic or exonic regions as well as promoter CpG islands) and correlation with tissue specific expression for a number of these genes (Song, *et al* 2005). While these experiments indeed suggest that there are tissue specific methylation patterns which correlate with tissue specific gene expression, it remains to be proven still what role these methylation marks play in tissue specification- whether they are necessary for tissue specification, promote stability of tissue specification or are simply footprints of gene silencing effected by other means.

1.6 Pathology of DNA methylation- role in cancer

It has long been recognised that the normal patterns of genomic DNA methylation described above are altered in cancer. The prototypical alteration seen is demethylation of the bulk DNA, but this is accompanied by aberrant CpG island methylation (for review see Herman and Baylin 2003). Both abnormalities have the potential to contribute to the pathogenesis of cancer, and it is a major area for research determining the mechanisms by which these abnormalities arise, and the exact role they play in the biology of cancer.

Hypomethylation may be involved with the pathogenesis in cancer by a number of mechanisms. Firstly, the demethylation of regions normally silenced by methylation may lead to reactivation of imprinted genes, X inactivated genes or retroviral elements.

Secondly, demethylation may lead to increased chromosome instability particularly in the pericentromeric regions. Lines of evidence that this mechanism indeed plays a role in cancer include the increase of chromosomal abnormalites and development of aggressive lymphomas in Dnmt1 hypomorphic mice (Gaudet, *et al* 2003).

Abnormal methylation of CpG islands has the potential to influence tumour development by the inactivation of tumour suppressor genes. This may provide an alternative mechanism of inactivation in addition to mutation or genomic deletion, which are well established mechanisms of carcinogenesis. Silencing by methylation could act as an early event in cancer, providing a second hit inactivating the normal allele after the opposite allele has been inactivated by somatic or inherited mutation. In addition, methylation could possibly inactivate both alleles at a tumour suppressor locus or several loci, which may be sufficient to cause cancer. Alternatively (or additionally), methylation increased as a secondary consequence of malignant transformation could promote the progression of already established tumours, leading to inactivation of further genes which could increase tumour growth or metastasis. Finally, it could act to stabilise epigenetic inactivation of tumour suppressor genes which have been epigenetically inactivated by other mechanisms such as histone modification.

That aberrant methylation of tumour suppressor genes can be identified in cancer, both in cells from primary tumours and cell lines, is not in doubt, and there are an ever increasing number of reports from many different tumour types describing the methylation of genes involved with such diverse functions as cell cycle regulation, cell-cell adhesion, drug resistance, genome stability. It appears that certain genes seem to be particularly prone to abnormal methylation in tumours, and it has also been suggested that specific patterns of methylation may be characteristic of the origin of the tumour (Esteller 2002 for review). A number of these genes have been shown to become expressed upon treatment with demethylating agents such as 5'-azacytidine. This has been taken as proof that these genes have been pathologically silenced by methylation and that pharmacological demethylation is restoring them to their normal state of

expression, with the implication that methylation is contributing to the behaviour of the tumour. However other observations show that these arguments are simplistic. Firstly, genomic experiments aimed to discover new hypermethylated genes in cancer in an unbiased manner show that there are in addition many aberrantly methylated genes for which it is difficult to predict a role in cancer pathogenesis (Dai, *et al* 2001, Rush, *et al* 2001, Rush, *et al* 2004). Since it is difficult to suggest that methylation of these genes would give a selective advantage to the cell such that a clone with methylation of these genes would become dominant, this observation favours the hypothesis that much of the aberrant CpG island methylation seen in cells is a bystander effect with no direct role in tumour behaviour. Secondly, 5'-azacytidine treatment clearly has substantial effects on gene expression independent of its demethylating action, with activation of silent but unmethylated genes demonstrable, since experiments using microarray analysis for genes upregulated by 5'-azacytidine as a method of identifying methylated genes have shown a variable rate of success in identification of genuinely methylated genes (Liang, *et al* 2002, Schmelz, *et al* 2005, Yamashita, *et al* 2006). This suggests that reactivation by 5'azacytidine is insufficient evidence to demonstrate that a gene would normally be active, and thus it is possible that methylated tumour suppressor genes may have been silent anyway with their methylation making no contribution to tumour behaviour.

More sophisticated arguments that DNA methylation does have an important role in cancer come from two lines of evidence. The first is the finding of an increased frequency of methylation at tumour suppressor gene loci in sporadic cases of tumours compared to hereditary cases in which a genomic abnormality of one allele is inherited. One of the oldest models is retinoblastoma, a retinal tumour caused by biallelic inactivation of the *Rb* gene, which occurs both sporadically and in a familial form in which a genomic abnormality of the *Rb* locus is inherited on one allele (OMIM reference +180200, http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM). The frequency of hypermethylation in unilateral, sporadic tumors was 9.3% (13/140), whereas the frequency was 1.0% in bilateral hereditary tumors (1/ 101) (Ohtani-Fujita, *et al* 1997), suggesting that there may be a background rate of inactivation by

methylation at this locus contributing to tumorigenesis independently. A second model is methylation of the mismatch repair gene *MLH1* in colonic tumours. This model is particularly instructive since the presence of the downstream consequence of *MLH1* inactivation, microsatellite instability, offers evidence that methylation has caused functionally significant inactivation (Herman, *et al* 1998). In one study, methylation of the *MLH1* locus was found in 7/10 sporadically occurring colon tumours with microsatellite instability compared to 0/10 in tumours from hereditary non-polyposis coli patients with inherited mutations of the *MLH1* gene (Wheeler, *et al* 2000), suggesting that DNA methylation is likely to be an early event in tumour formation. Finally, in colon cancer patients with inherited germline mutations of *MLH1* and breast cancer patients with inherited mutations of *BRCA1*, there is a greatly increased incidence of methylation in patients with retention of both alleles in the tumour compared to those with loss of heterozygosity at the mutated locus (Esteller, *et al* 2001).

The second line of evidence comes from experimental animal models. The animal system which has been most widely studied in this context is the $Apc^{+/min}$ mouse. The $Apc^{+/min}$ mouse is one of the oldest established models of inherited tumour predisposition, and carries a mutation produced by chemical mutagenesis in the *Apc* gene which causes reduced expression. The phenotype is dominantly inherited, and mice develop multiple bowel tumours which appear to progress from microadenomas to adenomas to adenocarcinomas (Moser, *et al* 1990). The mouse *Apc* gene is the homologue of the human *APC* gene, which is mutated in the human inherited cancer syndrome of familial adenomatous polyposis coli (Su, *et al* 1992). The earliest evidence for an involvement of DNA methylation in the pathogenesis of $Apc^{+/min}$ mice came from demonstrating reduced tumour frequency with 5'-azacytidine treatment (Laird, *et al* 1995). More convincing is the demonstration of reduced frequency of tumours in mice carrying the *min* mutation on a background of various combinations of *Dnmt1* hypomorphic alleles (Cormier and Dove 2000, Eads, *et al* 2002). The number of polyps was reduced in proportion to the severity of Dnmt1 depletion, with complete suppression of polyps in the most severely depleted animals. More recently, these results have been

elaborated in $Apc^{+/min}$ mice with compound heterozygosity for *Dnmt1*-null and *Dnmt1* hypomorphic alleles (Yamada, *et al* 2005). This system showed an increased number of microadenomas as defined by abnormal nuclear beta-catenin staining but a decreased number of polyps, with increase in loss of heterozygosity in microadenomas from the Dnmt1 deficient mice. The results are interpreted as indicating an increased rate of tumour initiation as a result of chromosomal instability, but a reduction in survival to form established adenomas because of a requirement for maintenance methylation for tumour survival. In addition the role of maintenance methylation is clearly not simple, since as well as reduced intestinal tumour formation, there were increases in other types of cancer such as T cell lymphoma (Gaudet, *et al* 2003) or liver tumors (Yamada, *et al* 2005).

The role of *de novo* methyltransferases in $Apc^{+/min}$ mice has also been investigated by establishing $Apc^{+/min}$ mice with conditional deletion of *Dnmt3b* in the intestine (Lin, *et al* 2006). In contrast to the results in *Dnmt1* depleted mice, deletion of *Dnmt3b* was found to reduce the number of colonic adenmomas but not the formation of microadenomas. Importantly, Dnmt3b was found not to be essential for tumour formation since Dnmt3b negative tumours (both micro and macroadenomas) were found, although the proportion of macroadenomas which were Dnmt3b negative was reduced. These studies did not demonstrate the silencing by methylation of tumour suppressor genes, and thus strictly speaking demonstrate the necessity for methyltransferase enzymes rather than for DNA methylation itself. However, circumstantial evidence that DNA methylation itself may be important for tumour establishment in $Apc^{+/min}$ mice is provided by the reduced incidence of tumours in $Apc^{+/min}$ mice crossed with Mbd2 deficient mice (Sansom, *et al* 2003). Independent of the *Apc* model system, decreases in intestinal tumour formation have also been seen on an *Mlh1* deficient background (Trinh, *et al* 2002).

Looking at other tumour model systems, Dnmt1 deficiency has also been shown to reduce lung tumour formation in a tobacco-carcinogen induced lung tumour model

(Belinsky, *et al* 2003). The most direct demonstration that methylation can act in a causative manner to silence tumour suppressor genes is the study of tumours in *Hic1* heterozygous mice. The Hic1 transcription factor is a putative human tumour suppressor gene. Mice heterozygous for a *Hic1* mutation develop both solid tumours and lymphomas with high frequency from around 70 weeks of age (Chen, *et al* 2003b). Importantly, allele specific methylation specific PCR demonstrates methylation of the wild type allele but not the mutated allele in tumours from these mice, which corresponds to loss of Hic1 expression.

A major question is of the origin of abnormal methylation patterns in tumours. From an aetiological viewpoint, increased methylation of CpG islands have been described in association with aging (Issa, *et al* 1994, Waki, *et al* 2003), chronic inflammation (Hsieh, *et al* 1998, Issa, *et al* 2001) and chemical insults (Lee, *et al* 1995, Mass and Wang 1997). However, it remains unproven whether this corresponds to silencing of the gene, since the regions methylated in these models are mostly exonic or far upstream regions of the promoter, with the region immediately covering the transcription start site remaining unmethylated. Regarding the molecular mechanism of abnormal methylation induction, it has been suggested that dense, silencing-inducing methylation of core promoter regions may spread from sporadic methylation events or distant regions of methylation, perhaps caused by factors such as those described above (Stirzaker, *et al* 2004). Diminished transcription, including that caused by siRNA has also been shown to favour methylation (see above), and H3K9 methylation has been shown to precede DNA methylation in the remethylation of cancer cells following 5'-azacytidine treatment (Bachman, *et al* 2003). Finally DNA methyltransferases must be involved in development of abnormal methylation. Since little *de novo* methylation occurs after gastrulation and the *de novo* methyltransferases are downregulated, development of abnormal methylation events must be preceded by an increase in *de novo* methylation capacity. Accordingly, levels of DNA methyltransferases have been found to be upregulated in primary tumours (see Clark and Melki 2002 for review). As would be expected, the presence of both maintenance and *de novo* methyltransferase activity

seems necessary for maintaining cancer cell demethylation patterns (Leu, *et al* 2003, Rhee, *et al* 2002). More controversially it has been suggested that Dnmt1 may exhibit *de novo* methytransferase activity especially in the pathological context (Jair, *et al* 2006). The ability of DNA methyltransferases to cause abnormal methylation events was shown by the increased number of CpG islands methylated by over-expressing DNMT1 in human cancer cell lines (Feltus, *et al* 2003, Vertino, *et al* 1993), although levels of DNMT1 over-expression in this study far exceeded those seen in normal or pathological contexts. In addition the detection of abnormal CpG island methylation does not necessarily direct *de novo* methylation by Dnmt1; aberrant methylation events could still occur through improved maintenance of methylation established by proven *de novo* methyltransferases.

An important point about the above studies is that they are all conducted in cell lines, which are immortalised and in that sense may be argued to be already neoplastic. While they provide considerable insights into possible mechanisms, they tell us little about the actual course of events *in vivo*, and in particular they tell us nothing about the effects of disturbing DNA methylation mechanisms in normal cells, which is of critical importance if it is to be argued that methylation disturbance is an early, causative event in cancer.

Current controversy surrounds the question of whether certain tumours are more prone to abnormal methylation than others- the so called CpG island hypermethylator phenotype (CIMP). The existence of a discrete hypermethylator phenotype would be highly persuasive evidence that methylation in tumours has a distinct, independent cause which could be targeted, especially if a relationship to tumour prognosis or evolution can be shown to correlate with such a phenotype. As is often the case with this type of concept, there is no clear definition of what constitutes a hypermethylator phenotype. The largest study which claims the existence of the CIMP phenotype, an analysis of 864 colorectal tumours from patients with colonic cancer (Samowitz, *et al* 2005), examined a small panel of 5 genes for methylation. Undoubtedly there were tumours with different numbers of genes methylated, and multivariate analysis showed a relationship between

methylation of several genes (2/5) with age, stage and differentiation state. However, studies have given conflicting results on whether there is a bimodal distribution of methylated genes, which is surely an essential consequence if there is a distinct CIMP phenotype reflecting the presence or absence of a specific mechanism promoting methylation. A recent paper (Weisenberger, *et al* 2006) uses cluster analysis on a large set of cancer methylated genes in primary tumour samples from colorectal cancer patients. Following selection of genes which are aberrantly methylated in colorectal cancer, unsupervised clustering using these genes in an independent sample did appear to identify a subset of tumours with high numbers of methylated genes. Interestingly, the presence of mutation in the BRAF kinase appeared to be a necessary condition for membership of this group. However the cluster analysis also identified a set of genes which are commonly methylated in cancer but did not have a high frequency of methylation in the "hypermethylator" group. What this would seem to indicate is that specific methylation patterns can be consequent upon specific genomic abnormalites, and methylation patterns could therefore be characteristic of other abnormalities in the cell. While this paper does convincingly demonstrate a hypermethylator phenotype, in the sense of a distinct group of tumours which has extensive aberrant methylation, it remains to be seen whether there there is a truly independent hypermethylator concept. The identification of this will need to await the results of studies looking at unbiased selections of CpG islands and studies in other tumour types.

The attraction of DNA methylation in cancer as a subject for study is that methylation changes are potentially reversible with demethylating agents such as 5'-azacytidine and decitabine. If methylation induced silencing of genes essential for tumour behaviour can be reversed, this could be a powerful therapeutic strategy. Unfortunately, the results of clinical studies with demethylating agents do not really support a simple activity of widespread reactivation of tumour essential genes, as is often proposed by more enthusiastic advocates (Figure 1.2). The methylation inhibitor 5'-azacytidine has been available since the 1960s, and was initially used in clinical studies as a standard

cytotoxic nucleoside analogue. Despite its ability to prevent tumour development in animal models, it has had a disappointing record in human tumours. The most widespread experience clinically has been in the treatment of myelodysplasia, a set of clonal haematological stem cell disorders which are characterised by peripheral blood cytopenias and a variably aggressive risk of transformation to acute leukaemia. While 5'-azacytidine clearly does have activity in some patients with myelodysplasia, improving peripheral cell counts and improving survival, the results in real terms are disappointing, with an improvement in median survival of 7 months in the largest published study (Silverman, *et al* 2002b). It is unclear how much of this effect is due to methylation and how much due to the cytotoxic effect, though it might be predicted that demethylation effects would predominate at the low doses used, and also the clinical effects are probably stronger than in studies of treatment with low dose cytarabine, another cytosine analogue without demethylating ability. 5'-Azacytidine is a less than satisfactory drug for long-term administration due to its mutagenic effects and the need for administration by injection. In contrast there are no clinical reports of activity in solid tumours, even though much of the evidence for methylation in cancer comes from solid tumours, and it seems unlikely that a previously ineffective drug will become effective because of the discovery of an additional mode of action. A deeper understanding of the cause and effects of methylation in cancer could help to restrict clinical study to those patients and tumours in which demethylating agents are more likely to succeed.

Secondly, even if methylation changes do not always have a causative role in cancer they could still have a role in diagnosis and monitoring. The fundamental difficulty with the use of molecular techniques in cancer is the diversity of abnormalities found in tumours. The search for individual mutations in a tumour from a given patient is too labour intensive for clinical use, and so there is a need for abnormalities that are common across tumours of a given type between a significant number of patients. Such abnormalities are available for a small subset of tumours, for example the common chromosomal translocations in leukaemia, but are rare overall. The study of commonly

methylated genes as a molecular diagnostic technique has the attraction that a given methylated gene shows the same abnormality in tumours from different patients (or different types of tumour for that matter), enabling a higher throughput of testing. The use of methylation abnormalities for diagnosis and monitoring would have more validity if it can be shown that these abnormalities are essential to the behaviour of the tumour, although this can only be fully validated empirically in the clinical situation.

1.7 Aims of the project

To summarise the discussion above, there are a number of open questions regarding the development of CpG island methylation and its role in transcriptional regulation, in particular with reference to cancer. While abnormal CpG methylation is clearly a phenomenon associated with cancer, it remains unproven whether it can act as a primary cause of cancer. More generally, it also remains unproven whether DNA methylation can act as an initial mechanism to inactivate gene expression, or whether prior gene silencing and/ or histone modification is a necessary condition preceding methylation. Secondly, it is uncertain whether abnormal methylation events occur stochastically, essentially representing errors in determining correct methylation patterns, or whether they are targeted to specific loci. Finally, given the downregulation of *de novo* methyltransferases in adult life, it is uncertain whether the lack of the occurrence of further *de novo* methylation events post-gastrulation is mainly a consequence of insufficient *de novo* enzyme activity, allowing dysregulation of *de novo* methyltransferases to cause abnormal DNA methylation patterns, or whether there is a more coordinated direction of DNA methylation patterns.

In this project, an attempt is made to address some of these questions by prospectively creating abnormal patterns of DNA methylation *in vivo*. As noted earlier, it is necessary to work *in vivo* rather than in transformed cell lines to prove the role of abnormal

methylation events in the earliest stages of tumour formation, since cell lines have already in some respect undergone neoplastic transformation and also have increased levels of *de novo* methyltransferases. On the basis that over-expression of a *de novo* methyltransferase is a necessary condition for producing abnormal methylation events, the approach chosen is this is the development of transgenic mice which over-express Dnmt3b. Dnmt3b is particularly implicated in the development of aberrant methylation in cancer since it is more consistently increased in tumours than Dnmt3a or Dnmt1(Robertson, *et al* 1999). Dnmt3b has also been shown to be necessary for maintaining aberrant methylation patterns in tumour cell lines (Rhee, *et al* 2002), and for soft agar colony formation in mouse embryonic fibroblasts and human bronchial epithelial cells transformed with *ras* and SV40 large T antigen (Soejima, *et al* 2003).

The primary question that can be answered by this system is whether increases in *de novo* methyltransferase activity are sufficient to cause abnormal methylation events, either at CpG islands at specific loci or globally. If abnormal methylation events are indeed a consequence of DNA methyltransferase over-expression, it can be asked whether this has any consequences for normal development, and whether there is any spontaneous incidence of tumours. If it could be shown that there are differences in gene expression associated with methylation, or spontaneous tumour formation with silencing of tumour suppressor genes, this would be very strong evidence that DNA methylation can be a direct cause of gene silencing or tumour formation. If there is no spontaneous tumour formation, could Dnmt3b over-expression still play a role in tumour behaviour by influencing gene silencing later on in the natural history of tumour development? This question is addressed by studying the effects of Dnmt3b in a carcinogenesis model and also in the context of inherited cancer prone mutations, to study effects of DNA methylation in a system where the molecular mechanism of tumour formation is well defined.

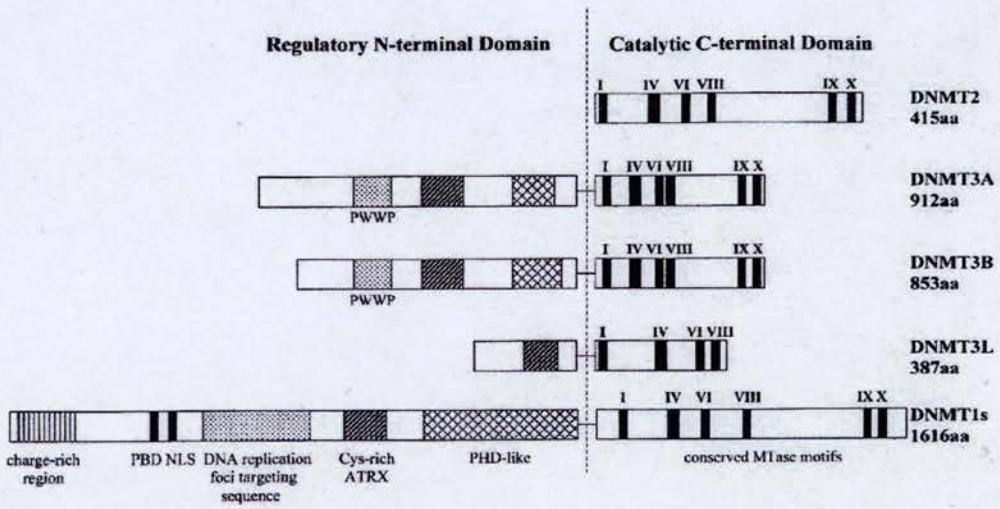
Subsidiary questions that may be answered by over-expressing Dnmt3b *in vivo* include whether changes in methylation in the context of active *de novo* methyltransferase

activity occur in a stochastic or pre-determined pattern, and to determine the relationship between DNA methylation and transcription- specifically, whether DNA methylation events precede transcriptional silencing and whether prior silencing is a necessary or sufficient condition to allow methylation in the context of active *de novo* methyltransferase activity.

In the following chapters the production of Dnmt3b over-expressing mice and analysis of their phenotype are described. The consequences of Dnmt3b over-expression on DNA methylation patterns *in vivo* are then demonstrated. Finally, the effects of increased Dnmt3b are investigated with respect to two specific phenomena which have attracted particular attention in the field of epigenetics- cardiac disease and cancer.

1.8 Figures

1.1a) DNA methyltransferases



1.1b) Methyl Binding Proteins

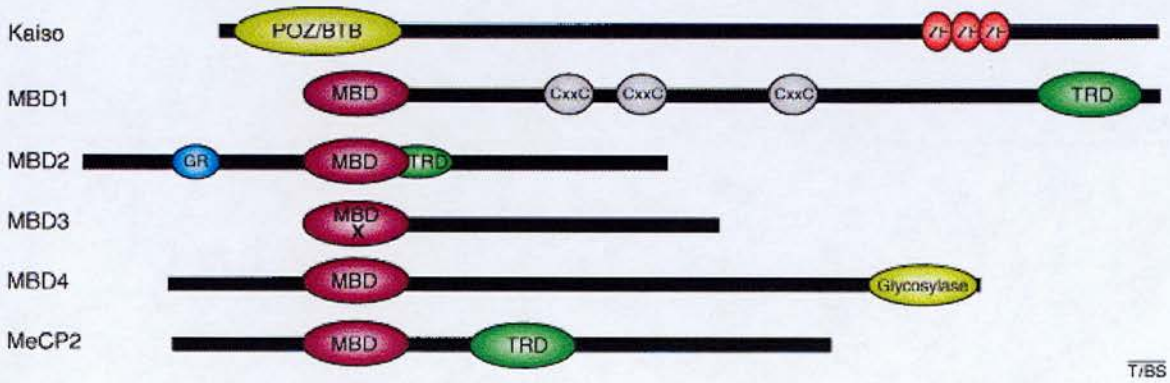


Figure 1.1 The mammalian DNA methyltransferases and methyl-binding proteins. Figure 1.1a) is reproduced from Turek-Plewa and Jagodzinski (2005) PBD- PCNA binding domain NLS– nuclear localisation signal. PHD– polybromo homology domain. PWWP– PWWP chromatin binding domain. ATRX– ATRX– homologous DNA binding domain. I-X conserved methyltransferase motifs. Figure 1.1b) is reproduced from Klose and Bird (2006). ZF=Zinc finger MBD– methyl binding domain . TRD transcriptional repressor domain. POZ/BTB– POZ/BTB domain. GR= GR repeat region . CxxC Zinc binding domain.

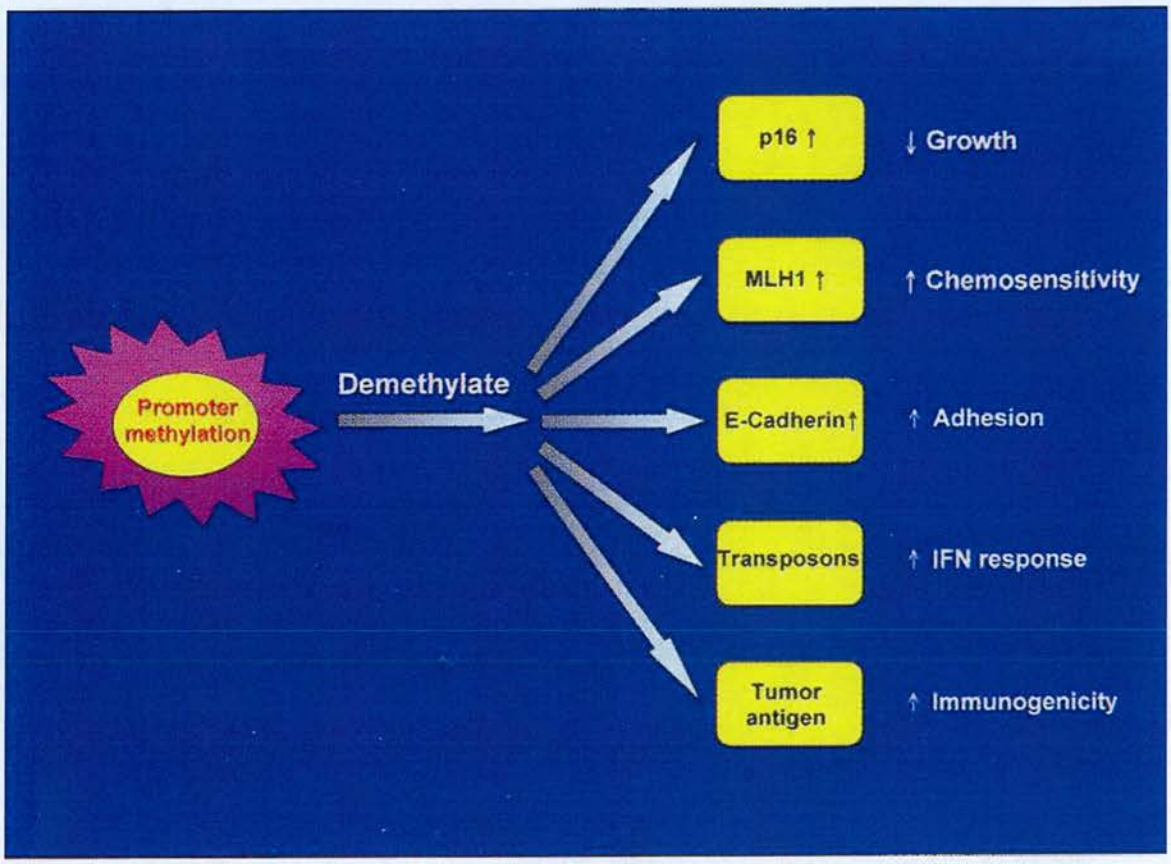


Figure 1.2 An optimistic view of demethylating agents in cancer. Reproduced from MGI Pharma booklet "Innovation in the treatment of haematological malignancies"

2. Production and characterisation of Dnmt3b over-expressing mice

Mice over expressing *Dnmt3* were produced by electroporating embryonic stem cells with a Dnmt3b expression construct consisting of full length *Dnmt3b1* cDNA driven by the ubiquitously expressed CAG promoter (shown in Figure 7.2). The construct has previously shown to be active, rescuing the global methylation levels in *Dnmt3b*$^{-/-}$ *Dnmt3a*$^{-/-}$ embryonic stem cells (Jackson, *et al* 2004). A total of 10 chimaeric mice were successfully produced by injection into C57BL6 blastocysts of transformed ES cells derived from a single ES cell clone. The male chimaeras were backcrossed and maintained on a C57BL6 background to form a transgenic strain Tg(Dnmt3b), and were genotyped by western blot.

2.1 *Tg(Dnmt3b)*+ mice are developmentally normal

Tg(Dnmt3b)+ mice develop to term and are born in the expected 50% transgenic – non transgenic ratio, with equal numbers of male and female mice born. They develop to term normally and appear to have no morphological abnormalities from birth to early adult life. Both males and females are fertile giving rise to viable young, although litter sizes are slightly below that expected for C57BL6 mice with a median litter size of 4.5 (mean 4.8+/-SD 2.3, n=57 litters) compared to expected size of 7 (Festing 1999). The effect may be on male fertility since litters derived from female transgenic mice are possibly larger with a median size of 7 (mean 5.9 SD 2.0 n=7 litters), although the difference is not statistically significant (95% confidence interval for difference in litter size between male and female transgenic parents=2.9 to -0.74).

Tg(Dnmt3b)+ mice are significantly smaller at all ages after weaning than *Tg(Dnmt3b)*- mice from the same strain although growth rates appear comparable (Figure 2.1, Table 2.1). The 95% confidence interval for ratio of weights is 0.74-0.92 for females and 0.66-0.75 for males, as estimated by linear regression of weights of *Tg(Dnmt3b)*+ against *Tg(Dnmt3b)*- mice (Figure 2.2). The outlying points on the regression graphs

represent a fall in weight of transgenic animals in the last 1-2 weeks of life as they become unwell.

2.2 *Tg(Dnmt3b)*+ mice over-express Dnmt3b in a wide range of tissues

Western blot analysis confirms that Dnmt3b is strongly expressed in the transgenic animals in a range of tissues in which it is weakly expressed if at all in adult wild type mice. Expression is seen in tissues including skin, heart, lung, kidney, gut and brain (Figure 2.3). The data are consistent with the tissue expression in other transgenic mouse models where expression is driven by the CAG promoter (Isoda, *et al* 2002, Okabe, *et al* 1997, Wiekowski, *et al* 2001). Dnmt3b is not detectable by Western blot in these tissues in wild type mice. While published data does suggest that *Dnmt3b* is transcribed at low levels in adult tissues (Xie, *et al* 1999) and in tissue stem cells (Hamra, *et al* 2004, Morris, *et al* 2004b), it can be concluded from this data that *Tg(Dnmt3b)*+ mice show expression of Dnmt3b protein which is many times higher in the above tissues than in wild type mice.

2.3 *Tg(Dnmt3b)*+ mice have reduced survival

A reproducible phenotype has emerged in *Tg(Dnmt3b)*+ mice. *Tg(Dnmt3b)*+ mice become unwell and die at a median age of 120 days (95% confidence interval 116-125 days). As can be seen from the survival curve in Figure 2.4, almost all deaths occur in a small range of ages, with very few animals surviving past 150 days. There is no significant sex difference in mortality (Figure 2.5). Figure 2.6 compares Kaplan-Meier estimates of survival between generations for all mice up to the F7 generation (including animals that were sacrificed early in the analysis as censored obsevations), and suggests there is a significant difference in survival between animals of different generations. There is no significant difference if F1 animals are excluded from the analysis, and there are no significant pairwise differences in survival between animals of generations other than F1 by post-hoc analysis of variance. This suggests that survival is particularly

reduced in the F1 generation but remains stable in subsequent generations. Possible explanations for this difference are that Dnmt3b over-expression is more toxic on a 129 strain background, and thus mice live longer with further outcrossing on to a C57BL6 background, or that transgene expression is stronger in the first generation of mice, possibly as a result of variegated transgene silencing in later generations (see Section 4.3).

2.4 *Tg(Dnmt3b)*+ mice have a reproducible phenotype which includes cardiomyopathy

There are three phenotypic abnormalites that appear to be characteristic of *Tg(Dnmt3b)*+ animals. All abnormalities are present in heterozygous dosage; there has been no attempt so far to create mice with homozygous expression of *Tg(Dnmt3b)*. The most striking abnormality is the development of cardiomyopathy, which is probably the most likely cause of premature death in the animals. Post-mortem analysis shows that hearts from *Tg(Dnmt3b)*+ animals are enlarged and have significantly increased mass (mean 0.3g +/- SD 0.07g vs 0.15+/- 0.03g for age matched non-transgenic control hearts from the same (F1) generation). Gross morphological analysis shows dilatation of both ventricular chambers and thinning of the ventricular walls (Figure 2.7). Atrial thrombus is often present at post mortem though this may be a post-mortem artefact. The liver commonly, though not invariably, shows congestion in keeping with the effects of heart failure. Histological analysis shows myofibre drop out and pallor, vacuolation and myocyte nuclear hypertrophy with prominent nucleoli (Figure 2.8a), and prominent interstitial fibrosis (Figure 2.8b). To determine whether the cardiac enlargement was due to hyperplasia or hypertrophy, BrdU immunostaining following *in vivo* labelling, and PCNA immunostaining were performed. These show that the myocardium remains non-proliferative (Figure 2.9). The histological findings are best described as those of a dilated cardiomyopathy, and appearances favour this being an inherited cardiomyopathy as there are no other features suggesting a haemodynamic cause such as hypertension. Further characterisation of the heart phenotype is discussed in chapter 4.

A second histological abnormality at post mortem is in the kidneys (Figure 2.10). *Tg(Dnmt3b)+* mice develop abnormalities in the juxtamedullary glomeruli with podocyte proliferation, and the presence of dilatation of the renal tubules with protein casts. Proteinuria is confirmed on testing of urine from *Tg(Dnmt3b)+* mice. The changes have some similarities to renal abnormalities seen in the aging mouse kidney, but are present at a much younger age than expected and are not present in littermate wild type animals.

A final abnormality is behavioural. As noted above, *Tg(Dnmt3b)+* females bred with wild type males produce morphologically normal offspring. However offspring are invariably killed and eaten by the mothers within 24 hours of birth. The finding is clearly a maternal effect, since offspring survive normally when fostered at birth to CD1 strain mothers. In addition, there is no increase in infanticide in offspring derived from male *Tg(Dnmt3b)+* males and wild type C57BL6 females. Anecdotally, *Tg(Dnmt3b)+* females also seem to have impaired nest-building activity with pregnant females building shallow nests similar to virgin females or male mice, although this finding has not been confirmed in a blinded manner.

2.5 Phenotypic changes are confirmed in a second line

Two possible confounding explanations are possible for the abnormal phenotypes seen in the initial *Tg(Dnmt3b)+* mouse line. The first possibility is that phenotypic changes are due to disruption of a gene by transgene integration, although the presence of the phenotype in heterozygous dosage would argue against this. A second, though more remote, possibility is that phenotypic changes may be influenced by the expression of the zeocin resistance gene (*zeo)* included in the construct as a selection marker. Although transgenic animals produced with *zeo* expression constructs reported in the literature do not show any of the phenotypic changes described here, it could be argued

that the effects of *zeo* expression in the context of active Dnmt3b expression are unknown.

In order to address these issues a second mouse strain was produced by electroporation of ES cells and blastocyst injection. A modified construct was produced with *frt* sequences inserted flanking the *zeo* gene to enable removal of the resistance marker using Flp recombinase either *in vitro* following selection or by crossing with Flp expressing mice. In order to answer the more pressing question of transgene integration more rapidly, blastocysts were injected with cells containing the non-recombined construct with a view to crossing with Flp expressing mice if the phenotype was confirmed.

A total of 6 male chimaeras, derived from a single ES cell clone survived to weaning age for establishment of a second line. One chimaera died at 4 months, and the remaining 5 chimaeras remain healthy at approximately 1 year of age. Chimaeras were backcrossed on to a C57BL6 background, as for the original line.

Successful transmission of the transgene to the F1 generation was detectable by PCR. There is a much higher rate of transgene silencing in the new line with only 3/9 F1 animals positive by PCR showing skin expression by western blot although transgene was detected by PCR in 5/9. However, all 3 animals with strong expression died early (93, 111, 88 days). In addition 2 animals which were found sufficiently soon after death for post-mortem analysis had enlarged hearts (heart weight/body weight ratios 9.3×10^{-3} and 14.2×10^{-3}), and had cardiomyopathic changes on histological evaluation. The confirmation of the phenotype, in particular the finding of early death only in animals with strong transgene expression leads us to conclude that the cardiomyopathy is indeed caused by over-expression of Dnmt3b. It has not been possible to exclude the possibility of a contribution due to the *zeo* gene as transgene expression has not been maintained past the first generation to enable crossing with *Flp* recombinase animals.

2.6 Discussion

In this chapter the successful production of mice which show strong over-expression of *Dnmt3b* is described. This is the simplest and least incontrovertible finding but in many ways it is also the most profound. The fact that strong over expression of a *de novo* methyltransferase does not lead to widespread developmental abnormalities suggests that epigenetic silencing is tightly regulated in early development and is not perturbed simply by over expression of a *de novo* methyltransferase. Dnmt3b clearly has an important role in early development since *Dnmt3b*$^{-/-}$ mice die between 13.5 and 16.5 days post-conception (Okano, *et al* 1998a). Since methylation patterns are determined early in normal development when Dnmt3b is strongly expressed, and little *de novo* methylation occurs after gastrulation when Dnmt3b expression is much weaker, it could have been hypothesised that continued strong expression of Dnmt3b would lead to abnormal methylation and gene silencing. While later chapters will demonstrate that this has indeed occurred to some extent, the essentially normal development of the mice suggests that the most important genes in development are protected from silencing by *de novo* methylation by other mechanisms. It also argues that *de novo* methyltransferase activity in early development is sufficient, in the sense that there is sufficient activity to create methylation patterns necessary for development, and so further increases in activity do not produce further gene silencing of genes essential for development. Again, this points to a coordinated regulation of epigenetic silencing and methylation determination early in development. The fact that animals are fertile suggests that *de novo* methylation in primordial germ cells may also be similarly sufficient.

However, the presence of reproducible phenotypic changes suggests that there are exceptions to this, and that over expression of a *de novo* methyltransferase is indeed able to perturb cellular biology in some cases. The fact that phenotypic changes occur reproducibly in all animals which show strong transgene expression rather than the

development of abnormalities at random suggests that this vulnerability is also pre-determined in some sense and that certain genes may be "primed for methylation" and remain unmethylated only because of the absence of *de novo* methyltransferase activity in adult life. It is equally possible as well that the perturbation of cell biology caused by Dnmt3b over-expression may be primarily mediated by functions of Dnmt3b other than DNA methylation such as direct recruitment of histone deacetylases. Evidence that abnormalities of DNA methylation indeed occur in *Tg(Dnmt3b)+* mice are discussed in the following chapter.

Along similar lines, it is an important question as to why the most prominent phenotypic changes are confined to particular organs. The most trivial explanation is that this is related to the level of transgene expression. While Figure 2.3 is not intended to be quantitative, it is clear that there is very strong expression in the heart. This is consistent with published data from transgenic mice which express spermine under the control of the CAG promoter, in which the highest spermine levels were achieved in heart and muscle (Ikeguchi, *et al* 2004). However it is clear that the presence of a phenotype is not solely related to the expression level of Dnmt3b, since the lung and skin have strong transgene expression but are phenotypically normal whereas the kidney has weak transgene expression but appears phenotypically abnormal. In the spermine over-expression model above, and in northern blots from a separate model where the tet-activator was overexpressed under control of the CAG promoter (Wiekowski, *et al* 2001) the relative expression of the transgene between these tissues was similar. Again, in the following chapter, the question of whether phenotypic changes are associated with methylation changes and whether these changes are causative or associative is explored further.

Moving on to the specific phenotypic alterations, the remainder of this project is concentrated on the cardiomyopathy phenotype since this is the most striking and easily defined abnormality. Detailed characterisation of the cardiac phenotype and the possible role of transcriptional regulation and epigenetic alterations in its pathogenesis are

described in the following 2 chapters. However it is worth commenting on other aspects of the *Tg(Dnmt3b)+* phenotype at this point as these suggest possible areas for further study.

The observation that *Tg(Dnmt3b)+* mice are slightly smaller than wild type animals is noteworthy. It is unlikely that this finding is due to strain differences since wild type animals in the analysis are littermates of the *Tg(Dnmt3b)+* animals and will have a similar genetic composition. A second possibility is that the smaller size represents chronic low grade ill health. While this is impossible to discount completely, there does not appear to be any clinically obvious difference in health between transgenic and wild type mice in early life. Furthermore, the reasonably linear correlation of weights of wild type with transgenic animals suggests that both sets of animals have similar rates of growth in early adult life, arguing against chronic ill health. A more speculative possibility is that over-expression of Dnmt3b may have altered imprinting behaviour, with post-meiotic germ cells positive for *Tg(Dnmt3b)* developing additional methylation and silencing of paternal-chromosome inherited genes (given that the mouse line is routinely propagated through the male germline). Since as a gross oversimplification paternal-expressed genes tend to favour growth whereas maternal-expressed genes tend to inhibit growth, this could be a mechanism for the observed size discrepancy. A simple way to investigate this possibility would be to compare weights of animals where the transgene is passed through the female germline, however given the difficulties with infanticide it would be non-trivial to produce sufficient offspring to compare statistically, especially since differences in weight are fairly subtle.

The finding of infanticide by *Tg(Dnmt3b)+* mothers is particularly intriguing for a number of reasons, but difficult to investigate formally. Infanticide in mice is fairly common in response to distress and so the simplest possibility is that infanticide may occur because of maternal illness, perhaps as a result of early effects of the cardiomyopathy. However, since the infanticide seems to be universal in Tg(Dnmt3b)+ mothers and occurs in young animals which do not appear unwell, the possibility that

more direct effects of Dnmt3b over-expression are causing this could be considered. There are a number of phenomena where epigenetic changes have been shown to alter maternal behaviour. Mbd2 deficient mice appear to show impaired maternal instincts (Hendrich, *et al* 2001), although this manifests as reduced offspring weights prior to weaning rather than infanticide. In addition the paternally inherited imprinted gene *Peg3* has also been shown to regulate maternal behaviour (Li, *et al* 1999), with mothers inheriting a mutant paternal allele showing decreased nest building and infant retrieval. More fancifully, an interesting hypothesis is that the mothers may have abnormalities in olfaction. An intact olfactory system is necessary for maternal behaviour; mothers with surgical destruction of the olfactory bulb also demonstrate infanticide invariably (Gandelman, *et al* 1971). Since nuclear transfer experiments have demonstrated that olfactory receptor selection in a given olfactory cell is epigenetically determined (Eggan, *et al* 2004), it would be attractive to speculate that the olfactory repertoire may have been altered by over-expressing Dnmt3b. Finally the existence of neurological syndromes in association with mutations in *DNMT3B* (ICF syndrome) and the methyl binding protein MeCP2 (Rett syndrome) together with the presence of CNS abnormalities in *Dnmt3b*$^{-/-}$ and *Dnmt3a*$^{-/-}$ mice emphasise the importance of methylation changes in CNS development. Such methylation changes may also be susceptible to environmental manipulation post-natally with neuroendocrine and possibly also behavioural effects (Weaver, *et al* 2004). Since Dnmt3b is over-expressed in the brain of *Tg(Dnmt3b)+* mice, it would thus appear feasible that methylation changes may indeed have caused behavioural changes which may include abnormal maternal behaviour. Identification of the nature of these abnormalities could be extremely instructive but it will be challenging to find objective measures of behaviour and also to identify the significance of molecular changes in what is a relatively unexplored area for research.

The renal phenotype remains under-characterised. Given the histological similarities to age related renal changes, it is an interesting question of whether the effect of Dnmt3b expression in the kidney has been to produce premature aging. Further work could

certainly be done to seek the presence of other age related changes such as glomerular basement membrane thickening and immunoglobulin deposition- electron microscopy would probably give the largest amount of information not dependent on mechanistic speculation. It would also be useful to investigate changes in renal function *in vivo* by measurement of electrolytes, glomerular filtration rate and protein excretion and their changes with age in wild type and transgenic mice. Given the interest in age-related methylation changes and its biological significance, it would be an important result if it could be shown that increases in methyltransferases can cause premature aging. The difficulty of course is of defining in a rigorous way what constitutes a cellular or molecular definition of aging.

While there are still many unanswered questions regarding the phenotypic changes in *Tg(Dnmt3b)*+ mice, for the purposes of this project the generic effects on DNA methylation, and a few specific areas of the phenotype are studied in detail. In the next chapter the changes in DNA methylation found in *Tg (Dnmt3b)*+ mice are demonstrated.

2.7 Figures and Tables

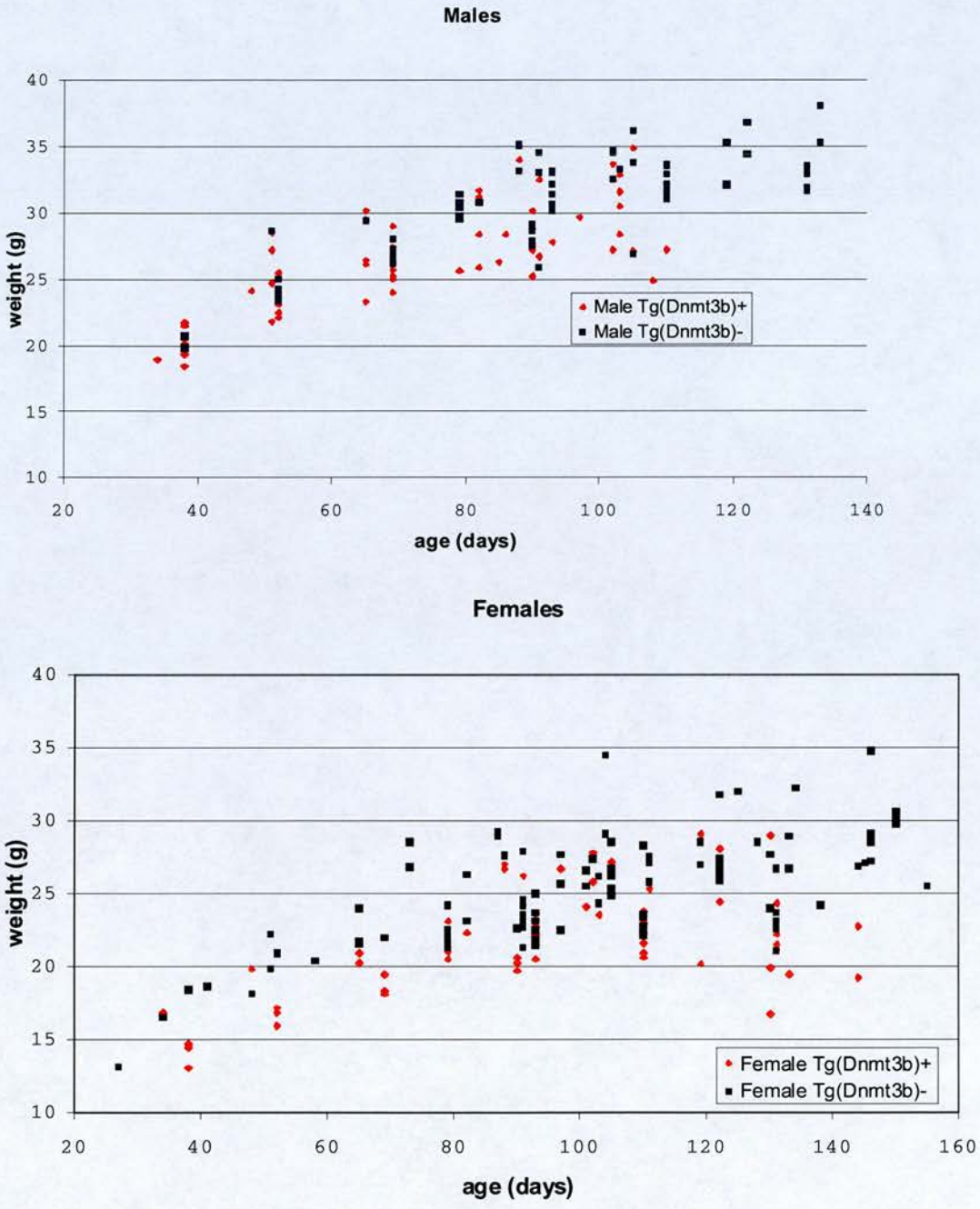


Figure 2.1 Weight of *Tg(Dnmt3b)* mice by age. Weight of *Tg(Dnmt3b)*+ mice compared with wild type littermate controls for the first 284 mice born cohort of mice, taken at a single timepoint.

A)

	Mean Weight by age: Females				
	Tg(Dnmt3b)-		Tg(Dnmt3b)+		
Age(d)	Mean Wt(g)	n=	Mean Wt(g)	n=	Ratio
30-39	17.5	2	14.7	4	0.84
40-49	18.4	2	19.8	1	1.08
50-59	20.8	4	17.4	4	0.84
60-69	22.3	4	19.4	5	0.87
70-79	23.3	9	21.5	7	0.92
80-89	27.0	5	25.3	3	0.94
90-99	23.6	18	22.5	13	0.95
100-109	26.8	14	25.6	6	0.95
110-119	25.2	12	23.0	10	0.91
120-129	28.0	9	26.3	2	0.94
130-139	25.2	14	21.7	9	0.86
140-149	28.9	6	21.0	2	0.72

B)

	Mean Weight by age: Males				
	Tg(Dnmt3b)-		Tg(Dnmt3b)+		
Age(d)	Mean Wt(g)	n=	Mean Wt(g)	n=	Ratio
30-39	20.2	5	20.1	8	0.99
40-49	None	0	24.1	1	
50-59	24.9	6	24.3	11	0.98
60-69	27.3	6	26.5	12	0.97
70-79	30.2	7	25.6	1	0.85
80-89	33.0	3	29.9	8	0.90
90-99	30.3	15	28.2	11	0.93
100-109	32.9	6	30.4	10	0.92
110-119	32.5	9	27.2	1	0.84

Tables 2.1A and 2.1B Weight of Tg(Dnmt3b) mice by age and genotype. Comparison of weight at different ages between *Tg(Dnmt3b)+* and *Tg(Dnmt3b)-*. Data are grouped by age, and shows the ratio (mean weight of *Tg(Dnmt3b)+* mice) / (mean weight of *Tg(Dnmt3b)-* mice) in each age group.

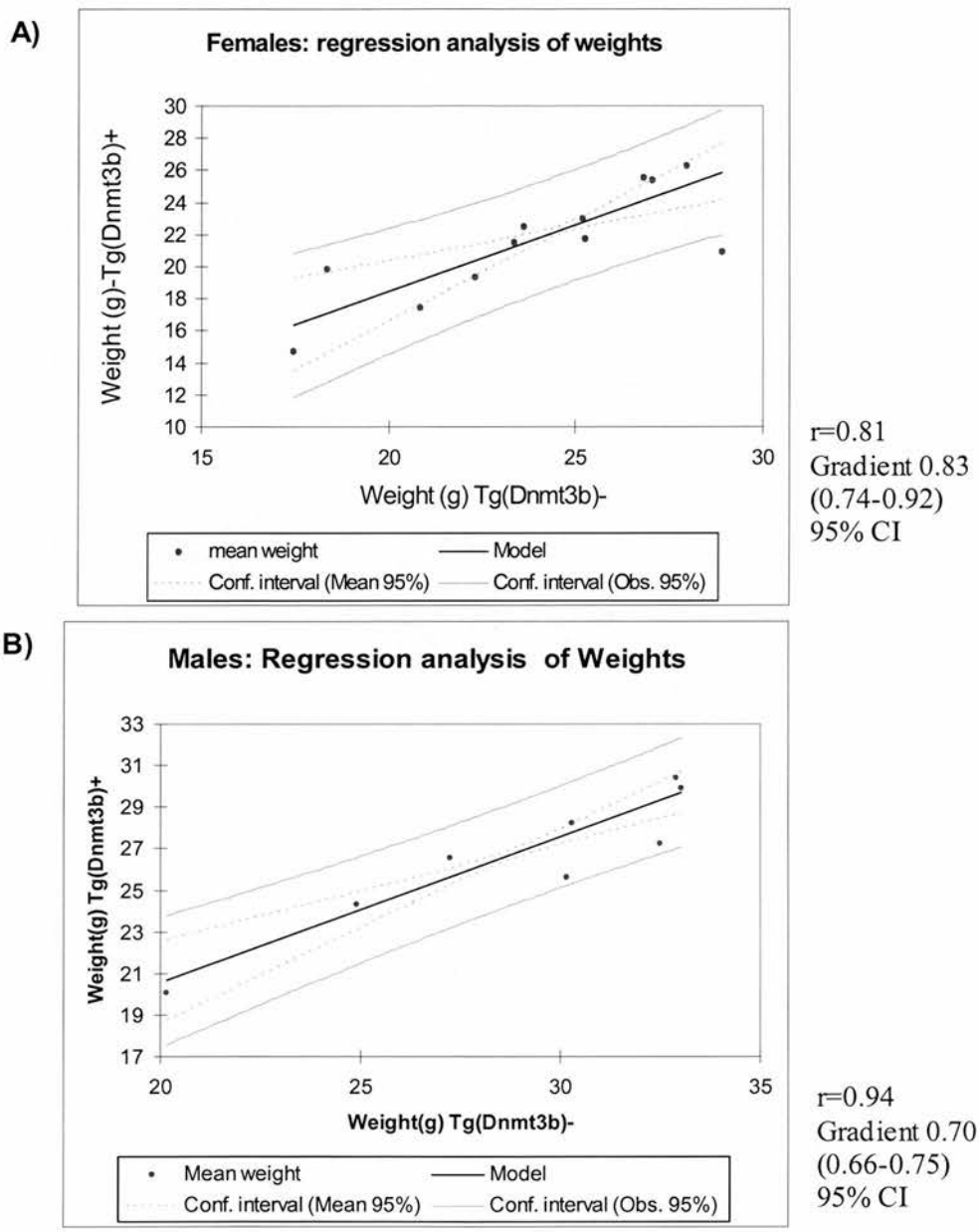


Figure 2.2 Regression of weight v age. Results of weighted linear regression of mean weights of *Tg(Dnmt3b)+* mice against mean weights of sibling *Tg(Dnmt3b)-* mice with data points at each age group as in Table 2.1. Data are weighted in the regression analysis according to the number of observations contributing to the mean at each data point. The dotted line shows 95% confidence limits for the regression line. The light shaded solid line gives 95% confidence limits for the mean predicted weight of transgenic animals at a given age given the weight of the non-transgenic animals at the same age.

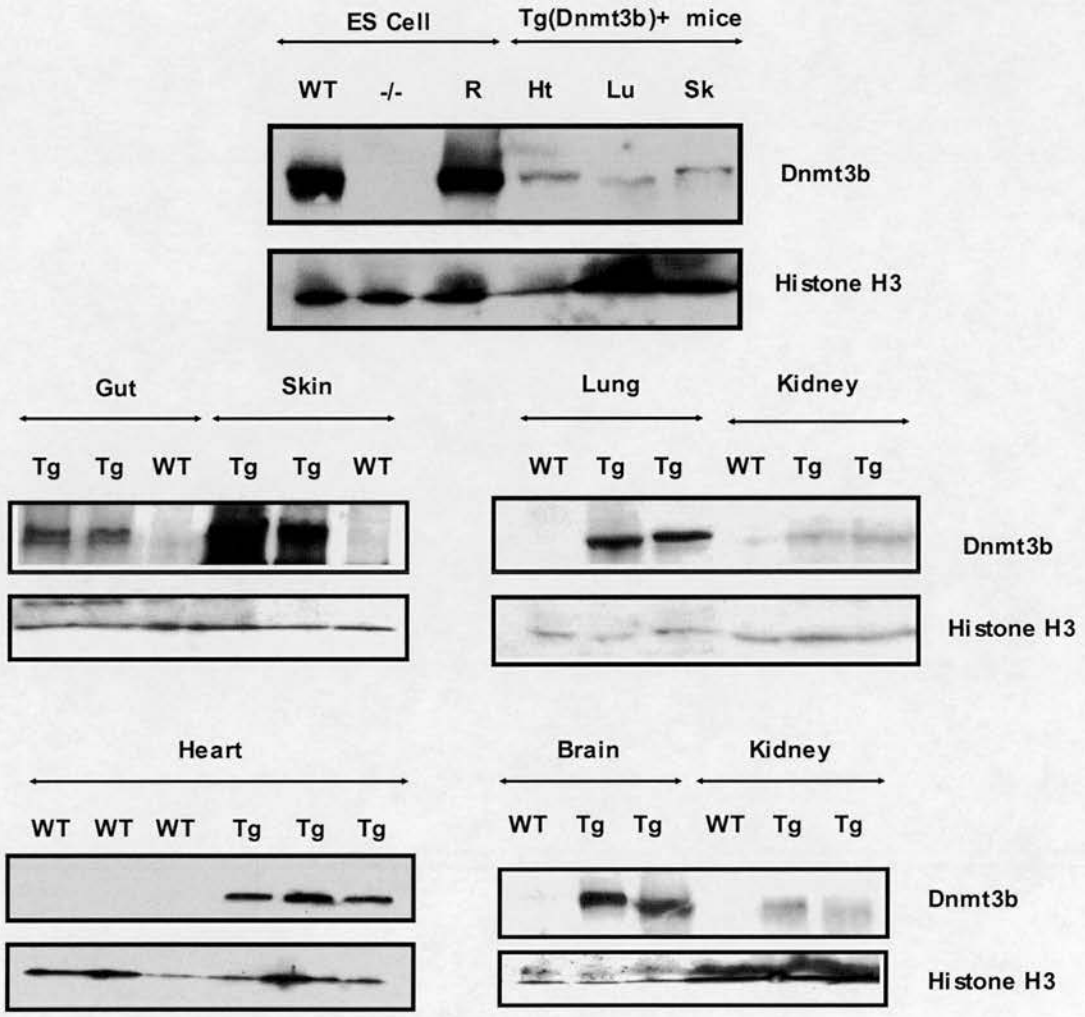
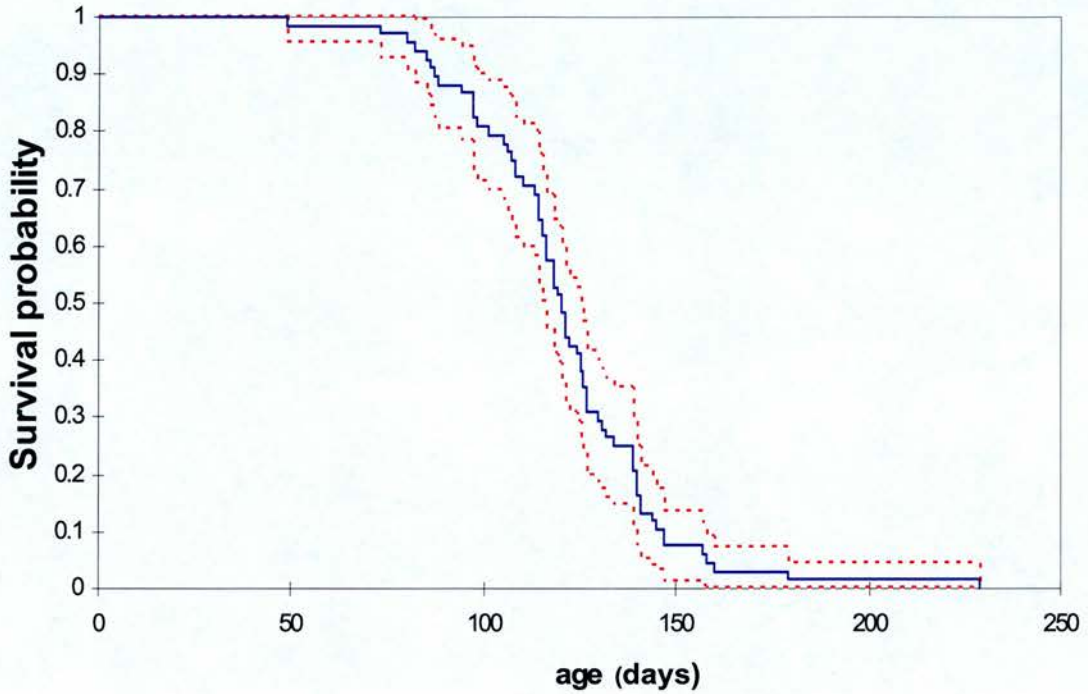


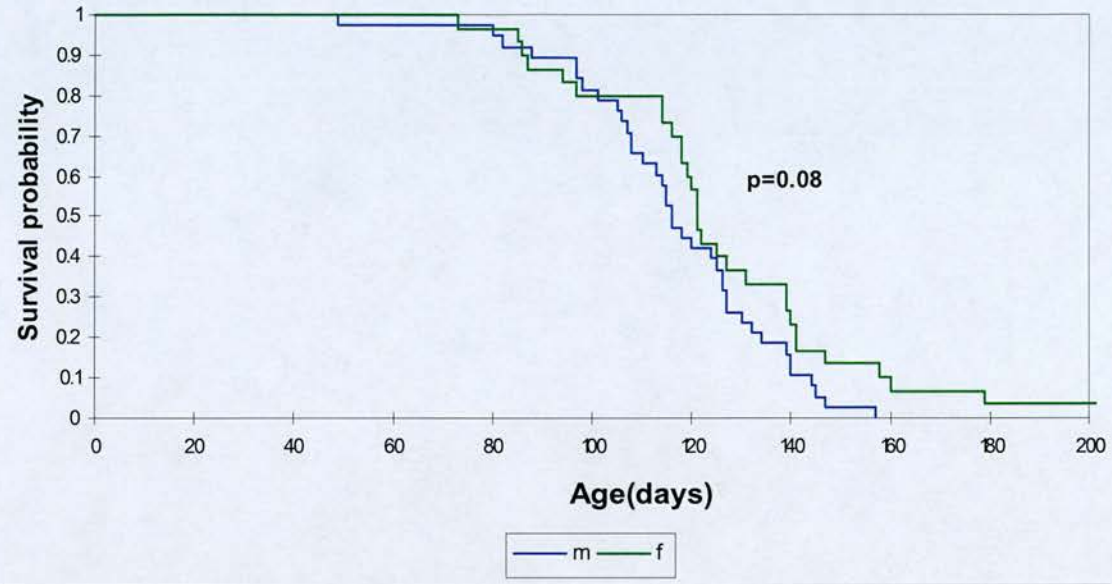
Figure 2.3 Tissue distribution of Dnmt3b expression . Western blots, using polyclonal anti Dnmt3b antiserum and poyclonal anti-histone H3 as loading control. Tg=Tg(Dnmt3b)+ mice(3 month old). WT=wild type mice. Br-Brain Sk-skin (ear) Ht– heart(ventricle) Ki-Kidney Lu-Lung Gu-Gut (distal small intestine). +ve= positive control from *Tg(Dnmt3b)+* ear clip. The top panel compares expression with ES cell protein extracts in order to demonstrate antibody specificity. WT= wild type ES cell. -/- = Dnmt3b-/- ES cell. R= Dnmt3b-/- ES cell rescued with Dnmt3b overexpression construct. The ES cell lines used are those described in Jackson et al (2004).



Survival quartiles

Quantile	Estimate (d)	(95% confidence interval)
75	134	126-140
50	120	116-125
25	108	97-115

Figure 2.4 Survival of *Tg(Dnmt3b)*+ mice. The proportion of Tg(Dnmt3b)+ mice surviving to a given age. The analysis includes all transgenic mice up to and including the f7 generation (n=68). Mice which were sacrificed early are excluded from the analysis. 95% confidence limits for the survival curve (dotted line) and the survival quartiles are also shown.



Survival quartiles

Females:		
Quantile	Estimate (d)	(95% confidence interval)
75	140	125-158
50	121	118-139
25	114	87-120

Males:		
Quantile	Estimate (d)	(95% confidence interval)
75	130	124-140
50	116	110-126
25	106	97-114

Figure 2.5 Survival of *Tg(Dnmt3b)+* mice by sex. Survival curves and 95% confidence limits for the quartiles of the survival function, comparing male and female *Tg(Dnmt3b)+* mice. Mice which were sacrificed early are excluded. The significance value for the survival curves based on the null hypothesis that they are equivalent is calculated using the log-rank test. Data includes mice up to the f7 generation as in figure 2.4

Generation	n observed	n censored	median survival (d)	95% conf. interval
f1	40	18	118	108-121
f2	7	1	115	114-158
f3	21	12	131	118-139
f4	7	3	147	140-not reached
f5	14	9	140	139-not reached
f6	21	9	140	127-157
f7	19	10	144	126-229

Figure 2.6 Survival of Tg(Dnmt3b)+ mice by generation. Kaplan-Meier estimate of survival and median survival for *Tg(Dnmt3b)+* mice from successive generations. All animals are included in the analysis with mice sacrificed early included as censored observations. Significance value by log-rank test for the survival curves, based on the null hypothesis that all curves are equivalent is shown for the whole data set. Significance value for the data set excluding the f1 generation is also shown, suggesting that shortened survival of the f1 mice is the major contributor to the non-equivalence of the set of survival curves.

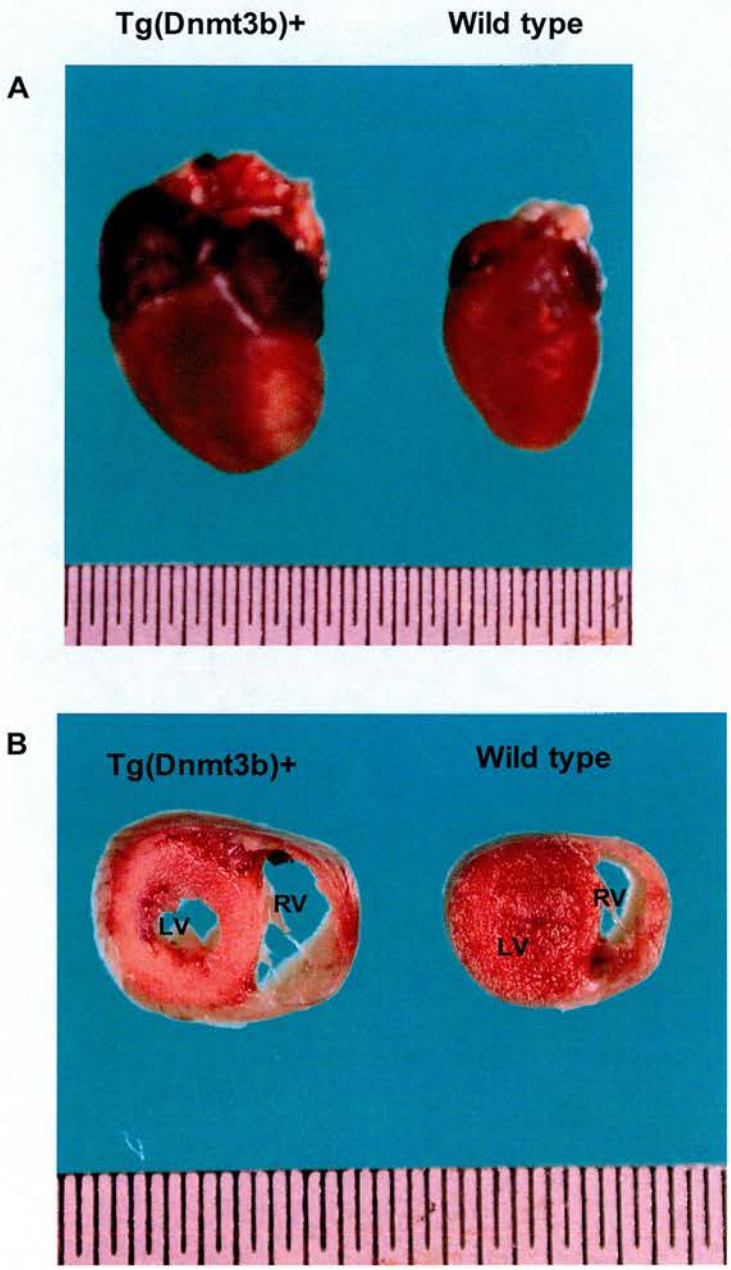
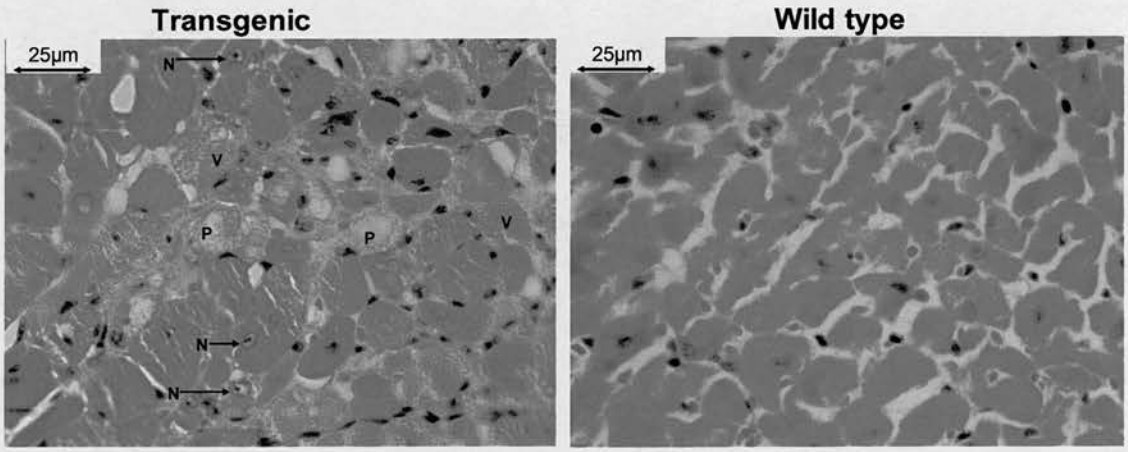


Figure 2.7 Gross anatomy of hearts from Tg(Dnmt3b) mice. Heart from a 4 month old F2 *Tg(Dnmt3b)+* mouse compared to a wild type control. The transverse section in B) shows dilatation of both left (LV) and right (RV) ventricles. Scale shows 1mm major graduations.

A) H+E stained sections from heart



B) Sirius red stained sections from heart

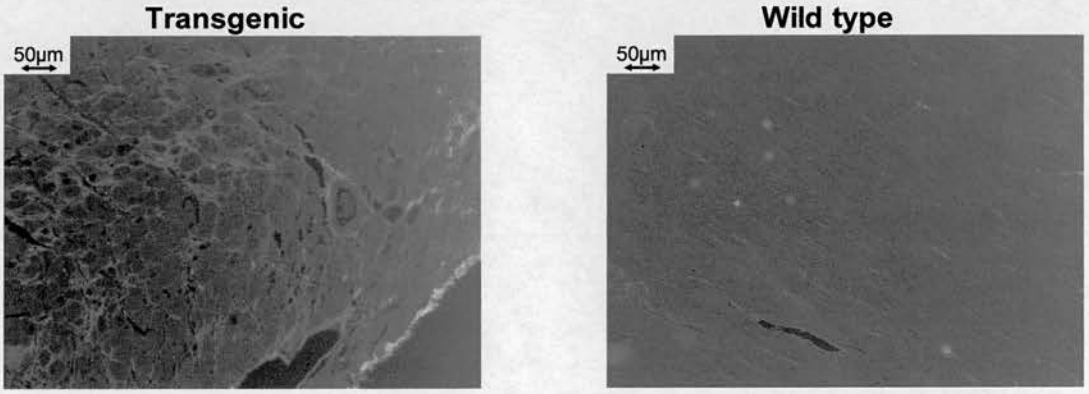


Figure 2.8 Histology of hearts from *Tg(Dnmt3b)*+ mice. A) H+E stained sections from 4 month old F10 mice, showing myofibre dropout with central pallor (P), cytoplasmic vacuolation (V) and prominent nucleoli (N) in the transgenic animals.
B) Sirius red stained section from the same animals showing extensive fibrosis in the transgenic mice (white staining in these photographs).

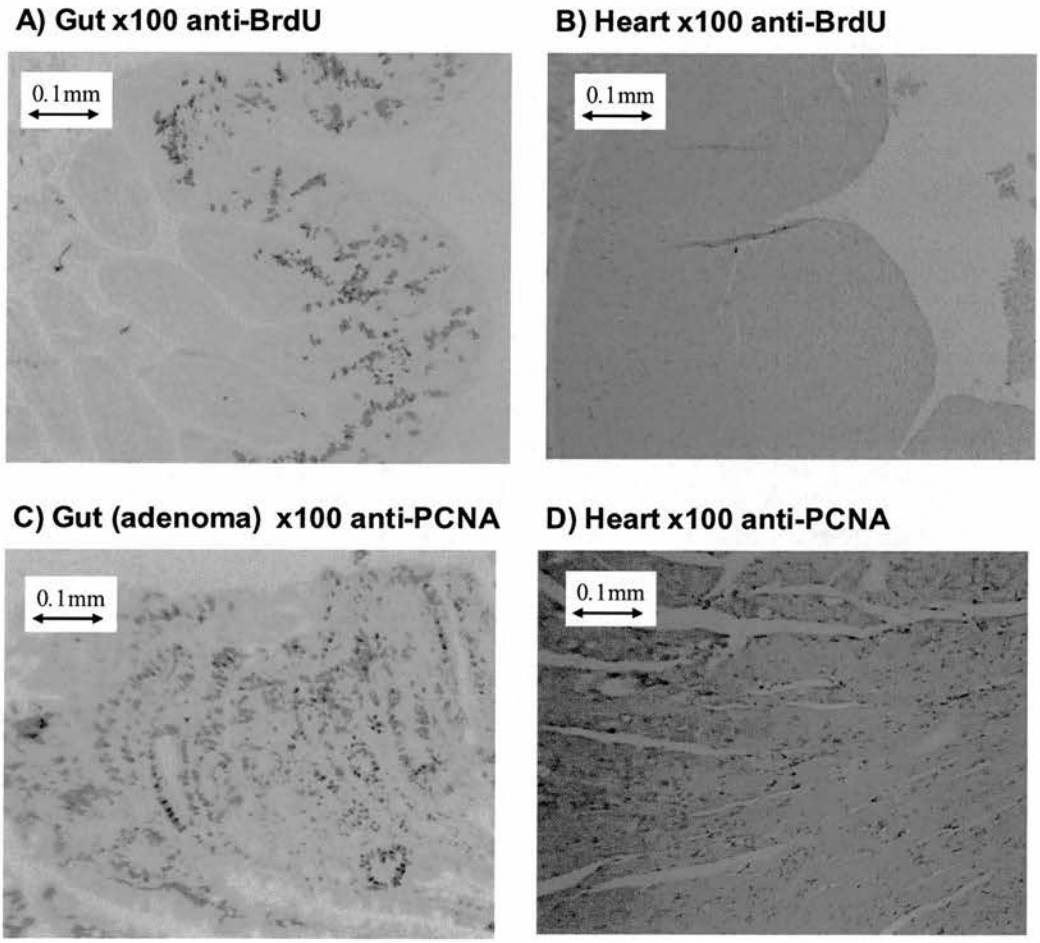


Figure 2.9 Proliferation in *Tg(Dnmt3b)+* hearts. A) shows gut tissue as positive control, showing nuclear BrdU staining. B) myocardium from 4 month old F11 *Tg(Dnmt3b)+* mouse following 6 hour in vivo BrdU labelling. showing absent staining. C) shows staining of an adenoma from an $Apc^{+/min}$ mouse as a positive control, demonstrating strong nuclear staining especially in the most cellular parts of the tumour. D) shows absence of PCNA staining in 4 month old *Tg(Dnmt3b)+* myocardium.

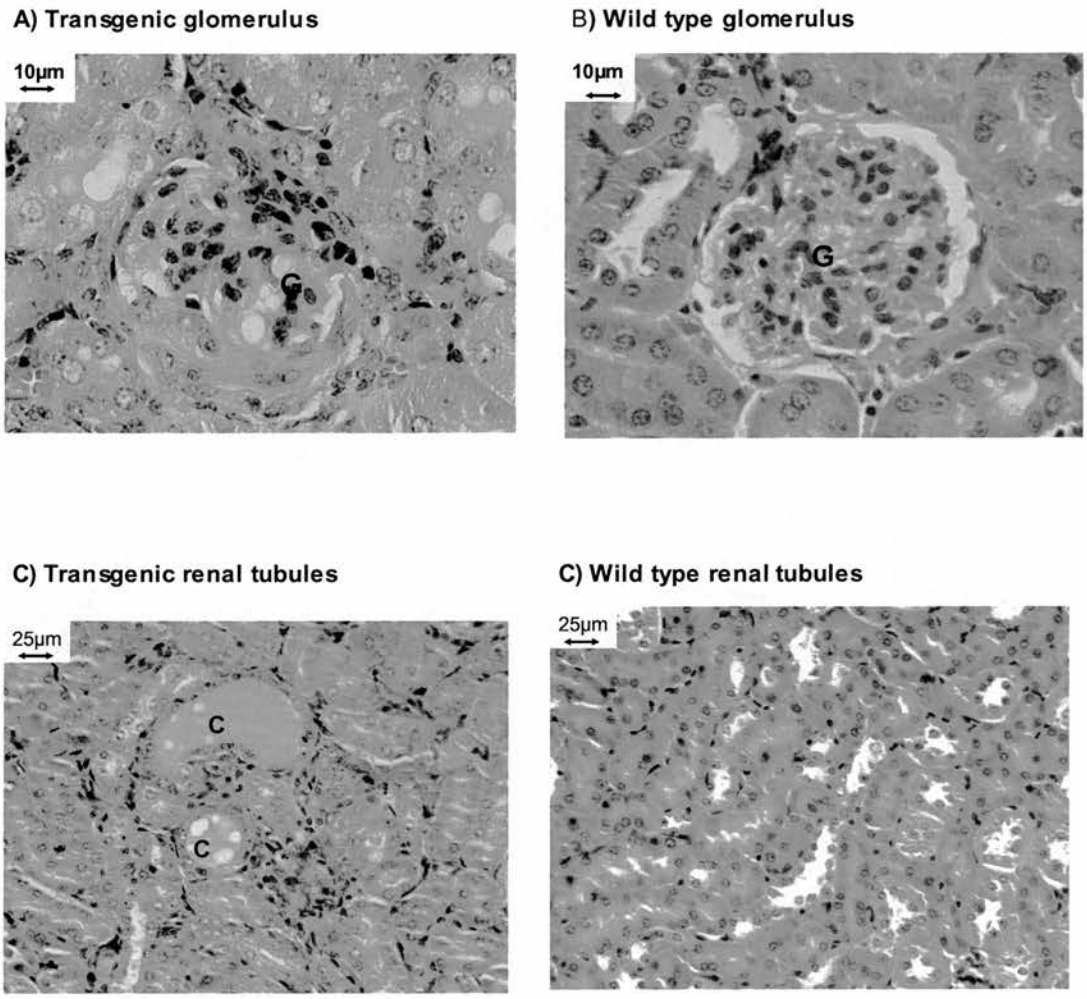


Figure 2.10 Kidney histology from *Tg(Dnmt3b)*+ mice. A and B) H+E stained sections showing a glomerulus (G) from the juxtamedullary region with abnormal podocyte proliferation in the transgenic sample A). C) and D) H+E stained section showing dilated renal tubules with protein casts (C) in the transgenic sample C). Samples are from 4 month old F10 mice.

3. Effects of Dnmt3b over-expression on methylation

3.1 Introduction

In the previous chapter, the production of transgenic mice which over-express Dnmt3b was described, and it was confirmed that the mice show increased Dnmt3b protein expression in a wide range of tissues, and show reproducible phenotypic changes in the heart, kidney and in maternal behaviour. In this chapter, the changes in DNA methylation consequent on this over-expression and the possible links between the methylation changes and phenotypic changes are described. Finally, what these changes tell us about the role of *de novo* methyltransferases in the determination of methylation patterns is discussed.

3.2 Global levels of methylation are increased in *Tg(Dnmt3b)*+ mice

Global changes in methylation in *Tg(Dnmt3b)*+ mice were measured by Nearest Neighbour analysis. Genomic DNA from heart and lung was extracted from 8 week old F13 *Tg(Dnmt3b)*+ mice and their wild type littermates. In order to examine changes in CpG methylation specifically in CpG island DNA, genomic DNA was digested with *MseI*, and the 2-4 kb fraction which is enriched for CpG islands was extracted for analysis.

In the heart there was a significant increase in methylation between *Tg(Dnmt3b)*+ and *Tg(Dnmt3b)*- animals (Figure 3.1) in this fraction of the genome. Mean CpG methylation in *Tg(Dnmt3b)*- hearts was 36.7% (SD 0.49%, n=3) compared to 46.0% (SD 3.2% n=3) giving a 95% confidence interval for increase in methylation of +1.7% to +16.9%. Replicate testing on one sample from each group shows the technique to have good precision with replicates within 2% of each other. In the lung, there was an

increase in methylation of a similar order of magnitude (7% increase between 1 pair of animals).

3.3 Aberrant, dense methylation of CpG islands occur in *Tg(Dnmt3b)*+ mice

It is not clear from the global methylation analysis whether the increase in methylation is the result of a diffuse increase in CpG methylation, or whether it has resulted in the dense methylation of CpG islands at specific loci. Methylation sensitive restriction fingerprinting (MSRF) was used as a screening technique to identify CpG islands which are aberrantly methylated in transgenic animal. In MSRF, genomic DNA is first digested with *MseI* to enrich for CpG islands and then with the methylation sensitive enzyme *BstUI*. The double-digested DNA is subjected to PCR with short, non-specific primers, and the patterns of PCR products can then be compared between transgenic and wild type animals.

Sample results from the MSRF experiment in are shown in Figure 3.2. Differences in restriction fingerprinting patterns were present in the heart, strongly suggesting that there are genes which are abnormally methylated in the hearts from transgenic animals, whereas there are also potentially methylatable sites which are resistant to methylation despite over-expression of the methyltransferase (see Figure 3.2 legend). In contrast to the results in heart, there were no differentially methylated bands identified in MSRF from lung or gut (Figure 3.3), despite the finding of increased global CpG island methylation in the lung.

DNA eluted from bands in the gel which suggest abnormal methylation was re amplified, cloned and sequenced. Clones were identified by BLAST search against the NCBI mouse genome database. A summary of genes thus identified and expression of the corresponding gene using data from microarray expression profiling in the heart (see

chapter 4) is given in Figure 3.4. Of the small number of genes identified, those showing differential methylation appear to have, at most, low level expression in both normal and transgenic animals.

In order to confirm the results from the MSRF experiment, the CpG islands which were aberrantly methylated according to the MSRF experiment were bisulphite sequenced. The CpG islands identified for the *Wnt3* and *Lrfn5* genes extend 5' to the transcriptional start site and extend in to the first exon and thus probably include the promoter region, whereas the *Viaat* CpG island is sited within an exon sequence, and is thus of more uncertain significance (Figures 3.5-3.7).

We obtained good quality sequence for a total of 17 clones from 3 normal and 13 clones from 3 transgenic mice (98 days old, F3 generation) for the *Viaat* CpG island, examining a total of 32 CpG dinucleotides in the PCR product from bisulphite modified DNA (Figure 3.9). The region appears to have sporadic methylation in the wild type animal at individual CpG dinucleotides which differ between clones, with a low level of dense methylation of the whole region. In the transgenic animals methylation is clearly increased with what appears to be increases both in the frequency of sporadic methylation and the proportion of clones showing dense methylation of the whole region. Since bisulphite conversion was not complete for most of these samples, it is not possible to exclude that apparent methylation at a particular CpG residue in a particular clone may be the result of conversion failure, but the difference in pattern is reasonably robust since the difference in conversion rates is not appreciably different between the wild type and transgenic clones.

More convincingly, there was a striking increase in methylation in the *Wnt3* and *Lrfn5* promoter regions. A summary of the methylation state of individual CpG residues in each clone sequenced is shown in Figures 3.8 and 3.10, showing that in several clones, the whole region has become densely methylated. There is a good rate of conversion of non-CpG cytosines to thymidine in both normal and transgenic animals, suggesting that

failure of deamination is unlikely to have caused the overall difference in methylation, although it is not possible to exclude the possibility that a given CpG in a given clone may be apparently methylated as a result of failure of deamination or as a result of PCR error. Of interest is the finding of CpA methylation in an increased number of clones from the transgenic animals at one CpA residue in both genes. While this is less definitive than the finding of increased CpG methylation in these clones, it is in keeping with the known activity of Dnmt3b in causing non-CpG methylation (Dodge, *et al* 2002).

3.4 Aberrant methylation of *Wnt3* and *Lrfn5* occurs in multiple tissues in *Tg(Dnmt3b)*+ mice

Quantitative real time PCR following restriction digestion with methylation sensitive restriction enzymes (M-QPR) was used to compare methylation of *Wnt3*and *Lrfn5* in several tissues from 3 normal and 3 transgenic mice (4 month old F8 generation). Methylation levels are expressed as a "methylation index": the ratio between the absolute quantity of DNA surviving digestion with *HpaII* and *HhaI* to the absolute quantity of starting DNA (as measured by a PCR across a region with no restriction sites). The results are shown Figures 3.11-3.12. Aberrant methylation is confirmed in the heart, providing some validation of the assay. The results suggest there is aberrant methylation of *Wnt3* in brain and kidney, though this may be at lower levels than in the heart (Figure 3.12). The results for *Lrfn5* are less consistent as a result of difficulties with the assay as noted in Section 7.3.4, but there is a consistent difference between the normal and transgenic group in heart, kidney (Figure 3.12). In addition the *Lrfn5* promoter is clearly either unmethylated or methylated at an insignificant level, in liver and lung of all normal and transgenic animals.

To compare the relationship of expression to methylation levels, expression of *Wnt3* and *Lrfn5* were measured by Taqman quantitative real time PCR. Gene expression levels

normalised relative to actin expression are shown in Table 3.1. There is consistent expression of both genes in brain and of *Wnt3* in gut, but it is at low levels in both tissues. There is inconsistent data suggesting there may be extremely low level expression in other tissues- the inconsistency in the data may result if expression is just at the threshold of sensitivity of the PCR technique. The results are consistent with the published data regarding the tissue expression of these genes in adult tissues (Morimura, *et al* 2006, Roelink, *et al* 1990).

Importantly, there is no expression of either gene in the tissues from wild type mice in which abnormal methylation or a histological change is seen in the transgenic mice. This suggests that methylation of *Wnt3* or *Lrfn5* is unlikely to have caused silencing of the gene, but has occurred in a gene that is already silent. In addition, the results show that the non-expression of a gene is not a sufficient condition to allow abnormal methylation in the context of active *de novo* methyltransferases, since there is no methylation of *Wnt3* or *Lrfn5* in tissue such as lung in which the gene is silent in the adult and in which Dnmt3b is strongly expressed.

3.5 Discussion

To summarise the results of this chapter, it has been shown that over-expression of a *de novo* DNA methyltransferase *in vivo* is sufficient to cause dense abnormal methylation of specific CpG island loci under certain conditions. However, the majority of CpG dinucleotides in the CpG island- rich fraction of the genome appear to be protected from methylation. This protection from methylation must include factors other than sequence or gene expression status, since there is a tissue-specific sensitivity to *Wnt3* and *Lrfn5* methylation regardless of the presence of strong *Dnmt3b* expression.

The discovery of abnormal *Wnt3* and *Lrfn5* methylation provides strong support for the notion that many of the abnormal methylation events seen in pathological contexts such as cancer have no causative role in the biology of the disease, since these genes are

abnormally methylated in the adult animal but were not expressed in the wild type animal either. A possible alternative hypothesis is that abnormal silencing had occurred due to methylation earlier in life. However this does not seem likely for the *Wnt3* gene since phenotypic changes seen in the mice to not correspond to the embryonic effects of *Wnt3* disruption. Two phenotypes have been described in association with *Wnt3* abnormalities. The first, $Wnt3^{-/-}$ mice, shows embryonic lethality with failure of mesoderm formation (Liu, *et al* 1999). The second, in humans, shows abnormal limb development (Niemann, *et al* 2004). While more subtle effects of a tissue specific interference with Wnt3 expression cannot be ruled out, neither of these phenotypes appears to show much similarity to the changes in *Tg(Dnmt3b)+* mice.

A further observation that has relevance for the role of DNA methylation in cancer is that abnormal methylation of *Wnt3* and *Lrfn5* have occurred in the heart, which does not undergo mitosis in adult life. It has been shown in the previous chapter that there is no increase in proliferation in the hearts of *Tg(Dnmt3b)+* mice. This indicates that abnormal methylation events are accumulating as a result of the transgene activity rather than becoming more prominent as a result of cellular proliferation. Thus, methylation events may accumulate in pathological contexts as a result of dysregulation of the DNA methyltransferase system in a polyclonal manner, rather than silencing of a tumour suppressor gene leading to clonal proliferation.

What role does abnormal methylation play in the pathogenesis of the phenotypic changes in *Tg(Dnmt3b)+* mice? It is instructive to note that it is easier to discover examples of dense CpG island methylation in tissues which are abnormal. The first possibility is that it is the presence of pathology which provides a suitable millieu for the development of methylation. In this model, the genome in the normal cell is protected from methylation. In the *Tg(Dnmt3b)+* animals, cellular biology is altered through mechanisms independent of methylation, such as protein-protein interactions involving the N-terminal domains of Dnmt3b, causing pathological change. As a result of the disturbed cellular biology, protection of genes from *de novo* methylation is disturbed,

allowing the development of methylation. This model has interesting translational consequences, as it would suggest that cellular damage induced by environmental or inherited insults could promote the development of abnormal methylation, with relevance to the finding of abnormal methylation events in cancer and other circumstances such as aging and inflammation. One difficulty with the extension of this hypothesis is that methylation changes common in cancer have also been described in histologically normal tissue surrounding tumours, tending to suggest that methylation changes occur earlier than cellular transformation. However it is still possible that the cells with methylation changes have been subject to molecular-level stressors which have not manifested at the histological level.

A second possibility is that the increase in global methylation indicates that DNA methylation has had a causative role in inducing the cardiac phenotype, most likely either by silencing genes which should be active, or preventing the activation of genes in appropriate contexts. If this could be robustly established, this would probably be the first direct demonstration that DNA methylation is in itself a sufficient factor to alter gene expression and alter cellular biology *in vivo*. The remainder of this this project concentrates on the heart on the assumption that the tissue with the most abnormal phenotype may be the most likely to show detectable abnormalities, and further efforts to prove the role of methylation in producing the heart phenotype are described in the following chapter.

3.6 Figures and Tables

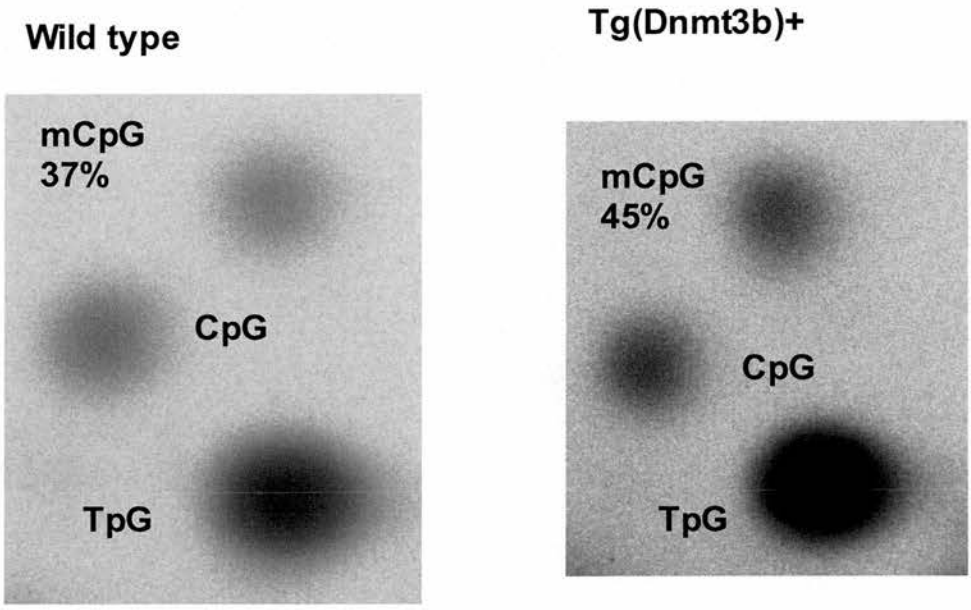


Figure 3.1 Global methylation in *Tg(Dnmt3b)* hearts. Sample autoradiographs from Nearest neighbour analysis comparing methyl CpG to CpG ratio between hearts of transgenic and wild type mice.

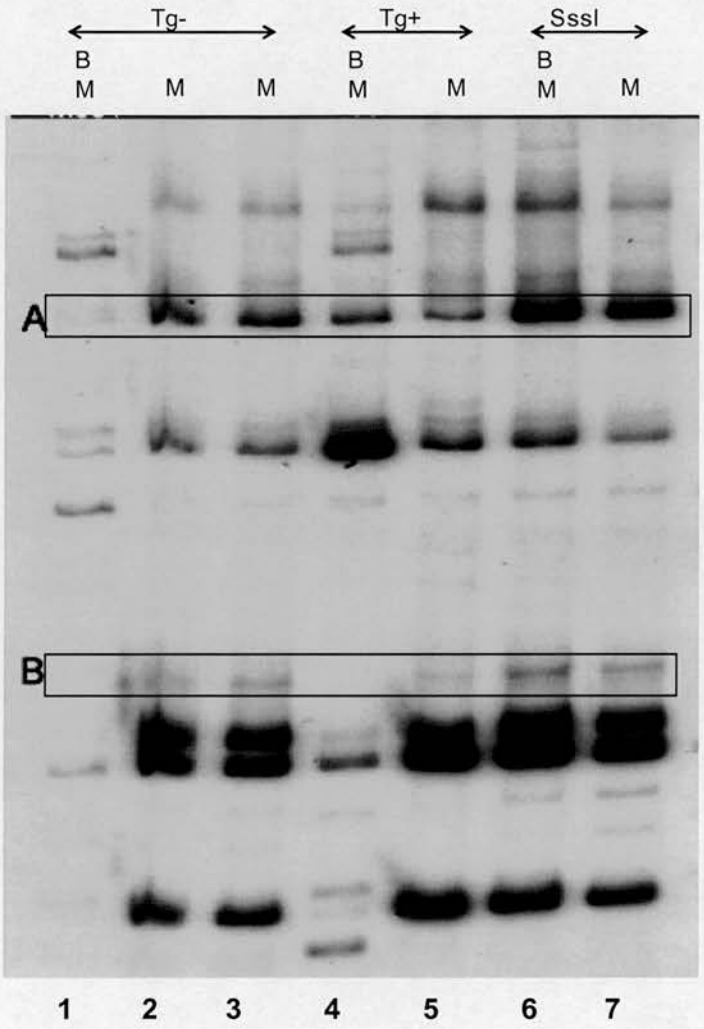


Figure 3.2. Methylation sensitive restriction fingerprinting from *Tg(Dnmt3b)* hearts. Example of methylation sensitive restriction fingerprint (MSRF) from heart (A1-A2 primer set), comparing transgenic (Tg+), wild type (Tg-) and *SssI* methylase treated DNA (*SssI*). B=*BstUI* digested; M=*MseI* digested. Band A shows a PCR product corresponding to a region methylated in the transgenic animal (band present in both lanes 4 (*BstUI* and *MseI* digested) and 5 (*MseI* digest only)) but not in the wild type animal (band present in *MseI* digested lanes 2 and 3 but not *BstUI* digested lane1). Band B shows a potentially methylatable site (Bands present in both *BstUI* and *MseI*-only lanes in *SssI* methylase treated sample) which remains unmethylated in both normal and transgenic animals (*BstUI* digested bands in lanes 1 and 4 absent).

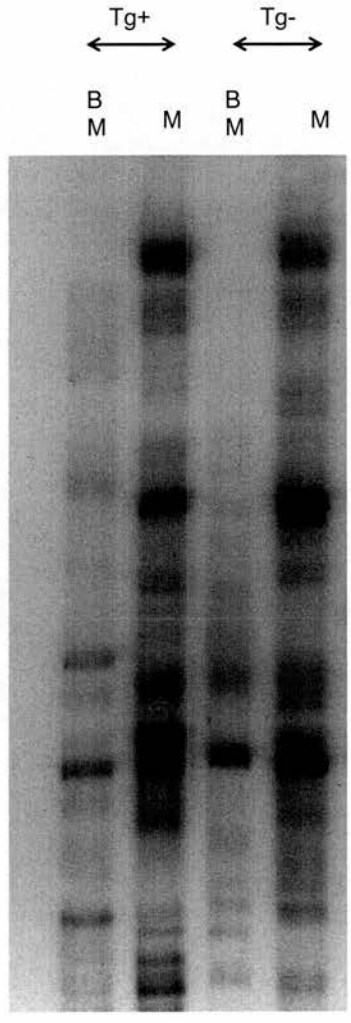
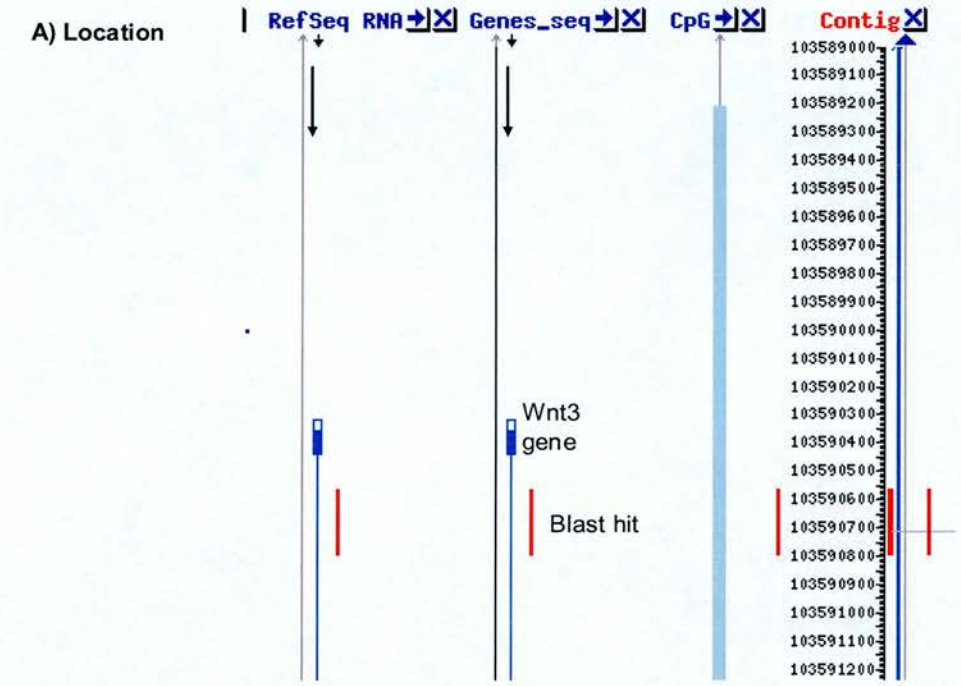


Figure 3.3. MSRF on lung from *Tg(Dnmt3b)* mice. Autoradiograph from methylation sensitive fingerprint comparing methylation patterns from lung between transgenic (Tg+) and wild type (Tg-) animals, showing no difference in *BstUI/MseI* digest patterns despite increased global CpG island methylation. B=*BstUI* digested; M=*MseI* digested.

Gene	Viaat	Wnt3	Hspa5	Dock4	Foxb2	Trim27	Lrfn5
Methylation Tg(Dnmt 3b)+	+	+	-	-	-	-	+
Methylation Tg(Dnmt3b)-	-	-	-	-	-	-	-
Sss1 methylase Methylatable	+	+	+	+	+	+	+
Expression in Tg(Dnmt3b)+	2.6	19	727*	?	6	193*	No^{+}
Expression in Tg(Dnmt3b)-	2.3	21	1382*	?	10	151*	No^{+}

Figure 3.4 Summary of genes identified by MSRF in heart. Methylation status of genes cloned from cardiac MSRF fragments as suggested by the presence of bands in the MSRF gel. Expression figures represent the arbitrary microarray probe signal strength (see section 4.7) ? - not represented on chip. * -signal is significantly increased above background (i.e. likely that gene is expressed). *Lrfn5* is not represented on array but is not expressed according to our own data (see section 3.4).



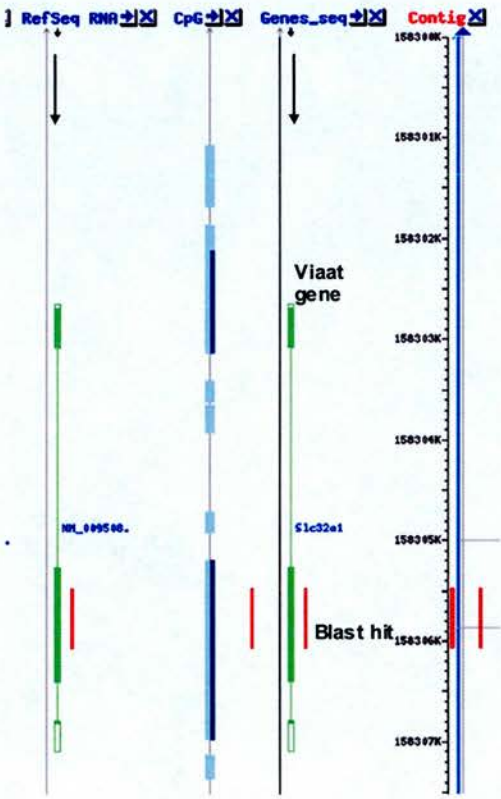
B) Sequence

```
Query  1         CCCCAGCCGTCCACCCCTCTTGGATCTTCTGCTGGTGTCGCCTTCCGAATCCAACTCCTG
60
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  15174952  CCCCAGCCGTCCACCCCTCTTGGATCTTCTGCTGGTGTCGCCTTCCGAATCCAACTCCTG
15175011

Query  61        GCTTCTCAGATAATCACCATTTTCCCTGCTACCCGCCGCAGAGGCGGGAGGGCTCTCCCT
120
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  15175012  GCTTCTCAGATAATCACCATTTTCCCTGCTACCCGCCGCAGAGGCGGGAGGGCTCTCCCT
15175071

Query  121       ACCCCTGAAGGTCTCTCGTTCTTAGTCTTCACCCTCTCCCACCCCAATTTGGGGGTCCCG
180
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  15175072  ACCCCTGAAGGTCTCTCGTTCTTAGTCTTCACCCTCTCCCACCCCAATTTGGGGGTCCCG
15175131

Query  181       AAGAACCTTCGTTTGCTCACACTTGCATTTCTGTGCTG  218
                 |||||||||||||||||||||||||||||||||||||
Sbjct  15175132  AAGAACCTTCGTTTGCTCACACTTGCATTTCTGTGCTG  15175169
```

Figure 3.5 BLAST results from MSRF: *Wnt3*

Figures 3.5-3.7 BLAST search results for 3 genes identified as abnormally methylated by MSRF in *Tg(Dnmt3b)+* hearts. A) Location of MSRF sequence relative to gene B) Sequence homology between sequence cloned from MSRF and the mouse genome. Mouse genome data is from NCBI website (see methods)

A) Location

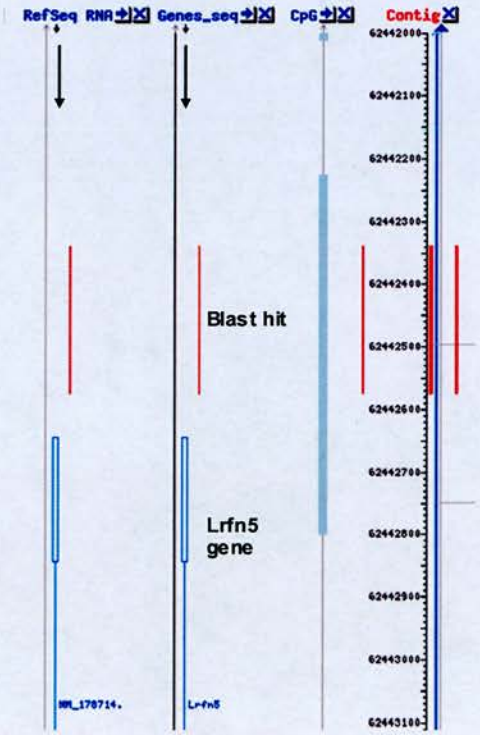


B) Sequence

Query	12	GCGGCCGCGTGGTCAATGTGGCGCAGATCATCGAGCTGGTGATGACGTGTATCTTGTACG	71
Sbjct	99383992	GCGGCCGCGTGGTCAATGTGGCGCAGATCATCGAGCTGGTGATGACGTGTATCTTGTACG	99384051
Query	72	TCGTGGTGAGCGGCAACCTCATGTACAACAGTTTCCCGGGGCTGCCCGTGTCGCAGAAGT	131
Sbjct	99384052	TCGTGGTGAGCGGCAACCTCATGTACAACAGTTTCCCGGGGCTGCCCGTGTCGCAGAAGT	99384111
Query	132	CCTGGTCCATCATAGCCACAGCGGTGCTGCTGCCCTGCGCCTTCCTGAAGAATCTCGAGG	191
Sbjct	99384112	CCTGGTCCATCATAGCCACAGCGGTGCTGCTGCCCTGCGCCTTCCTGAAGAATCTCAAGG	99384171
Query	192	CCGTGTCCAAGTTCAGTCTGCTGTGTACGCTGGCCCACTTCGTCATCAACATCCTGGTCA	251
Sbjct	99384172	CCGTGTCCAAGTTCAGTCTGCTGTGTACGCTGGCCCACTTCGTCATCAACATCCTGGTCA	99384231
Query	252	TCGCTTACTGTCTCTCTCGCGCGCGTGATTGGGCCTGGGAGAAGGTGAAGTTCTACATCG	311
Sbjct	99384232	TCGCTTACTGTCTCTCTCGCGCGCGTGATTGGGCCTGGGAGAAGGTGAAGTTCTACATCG	99384291
Query	312	ACGTCAAGAAGTTTCCCATCTCCATTGGCATCATCGTGTTCAGCTACACGTCGCAGATCT	371
Sbjct	99384292	ACGTCAAGAAGTTTCCCATCTCCATTGGCATCATCGTGTTCAGCTACACGTCGCAGATCT	99384351
Query	372	TCCTGCCCTCTCTCGAAGGCAACATGCAGCAGCCCAGCGAATTCCACTGCATGATGAACT	431
Sbjct	99384352	TCCTGCCCTCTCTCGAAGGCAACATGCAGCAGCCCAGCGAATTCCACTGCATGATGAACT	99384411
Query	432	GGACACACATCGCCGCCTGCGTGCTCAAGGGTCTCTTCGCGCTCGTCGCCTACCTCACCT	491
Sbjct	99384412	GGACACACATCGCCGCCTGCGTGCTCAAGGGTCTCTTCGCGCTCGTCGCCTACCTCACCT	99384471
Query	492	GGGCCGACGAGACCAAGGAAGTCATCACGGATAACCTGCCCGGCTCCATCCGCGCCGTGG	551
Sbjct	99384472	GGGCCGACGAGACCAAGGAAGTCATCACGGATAACCTGCCCGGCTCCATCCGCGCCGTGG	99384531
Query	552	TCAACCTCTTCCTGGTGGCCAAGGCGCTGCTGTCCTATCCGTTGCCCTTCTTCGCGGCCG	611
Sbjct	99384532	TCAACCTCTTCCTGGTGGCCAAGGCGCTGCTGTCCTATCCGTTGCCCTTCTTCGCGGCCG	99384591

Figure 3.6 BLAST results from MSRF: *Viaat* (legend is with Figure 3.5)

A) Location



B) Sequence

```
Query  1         CAAGGATTTGGCTGCAGGAGAACTTGAAGGAAACGTGATCAAAGGGCTATGTGAGAACGC
60
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  19985974  CAAGGATTTGGCTGCAGGAGAACTTGAAGGAAACGTGATCAAAGGGCTATGTGAGAACGC
19986033

Query  61        GACTGCTCTTTTTGCCGGTAGGCTGGTGGGCTTCCGGACCGGATCTTGAGAAGGACCAGA
120
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  19986034  GACTGCTCTTTTTGCCGGTAGGCTGGTGGGCTTCCGGACCGGATCTTGAGAAGGACCAGA
19986093

Query  121       ACTGTTTCTCCCTCCCGGTGCGGTTCTGGGagtgcttgtgcaagtgtgagtgcgcgcgcg
180
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  19986094  ACTGTTTCTCCCTCCCGGTGCGGTTCTGGGAGTGCTTGTGCAAGTGTGAGTGCGCGCGCG
19986153

Query  181       cccggctgagaatgtgtgagtgtgtgtgagcgtgtctctgtgtgtgCGCGGCCGC  235
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  19986154  CCCGGCTGAGAATGTGTGAGTGTGTGTGAGCGTGTCTCTGTGTGTGCGCGGCCGC
19986208
```

Figure 3.7: BLAST results from MSRF: ***Lrfn5*** (legend is with figure 3.5)

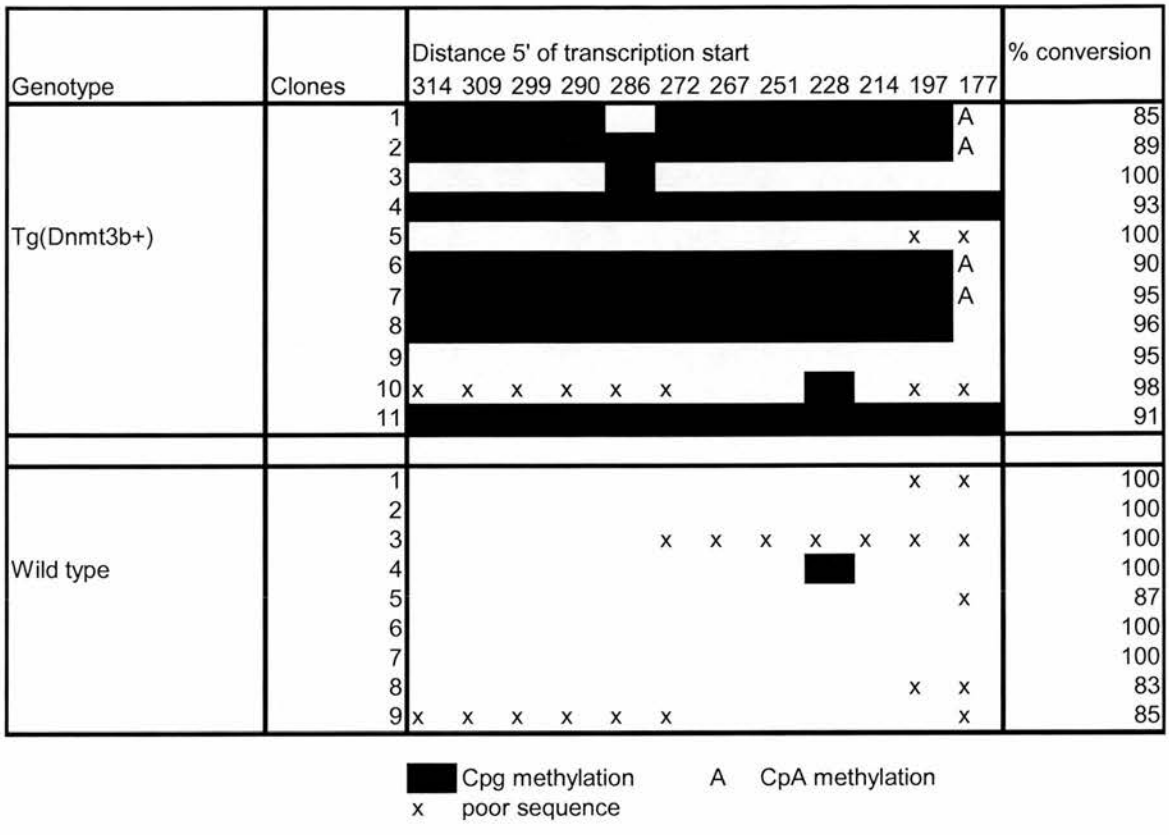


Figure 3.8 *Wnt3* bisulphite sequencing

Figures 3.8-3.10
Methylation status of genes identified by MSRF as abnormally methylated in heart confirmed by bisulphite sequencing.

The % conversion figure gives frequency of CpH->TpH conversion as a measure of bisulphite conversion efficiency.

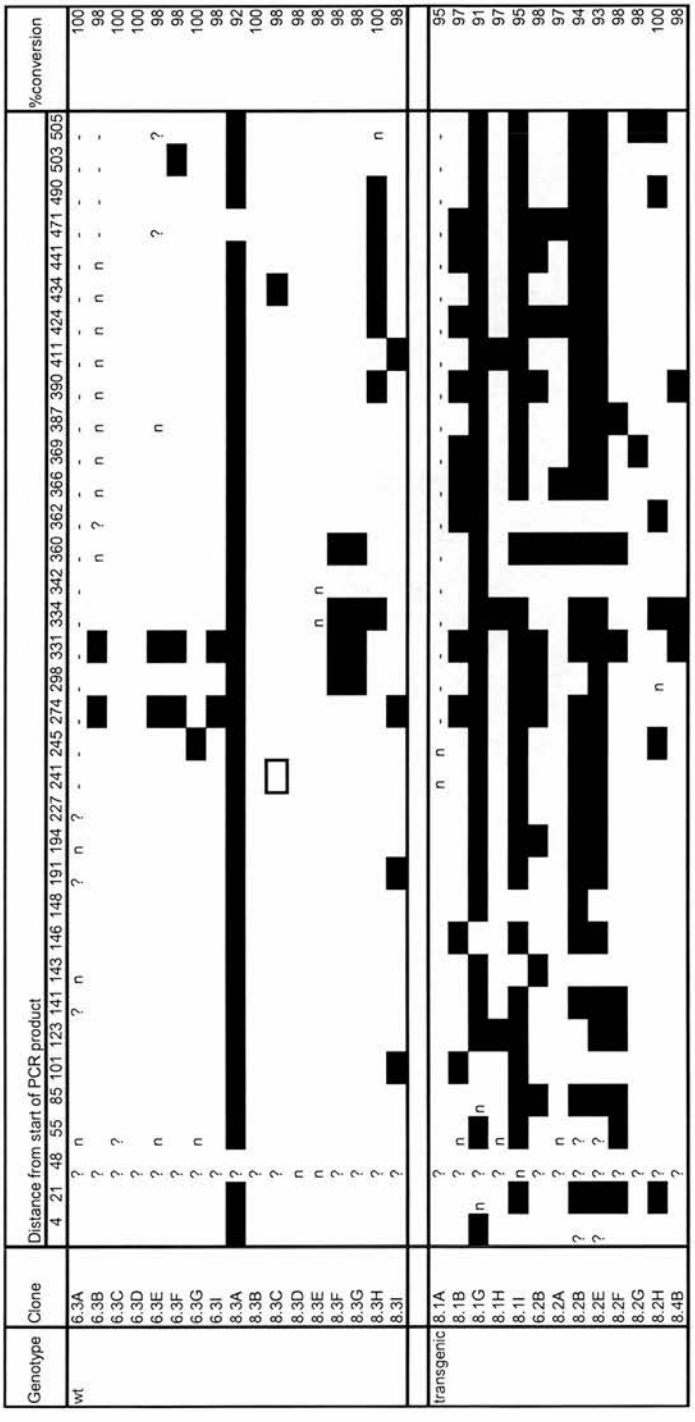


Figure 3.9 *Viaat* bisulphite sequencing. Legend is with Figure 3.8

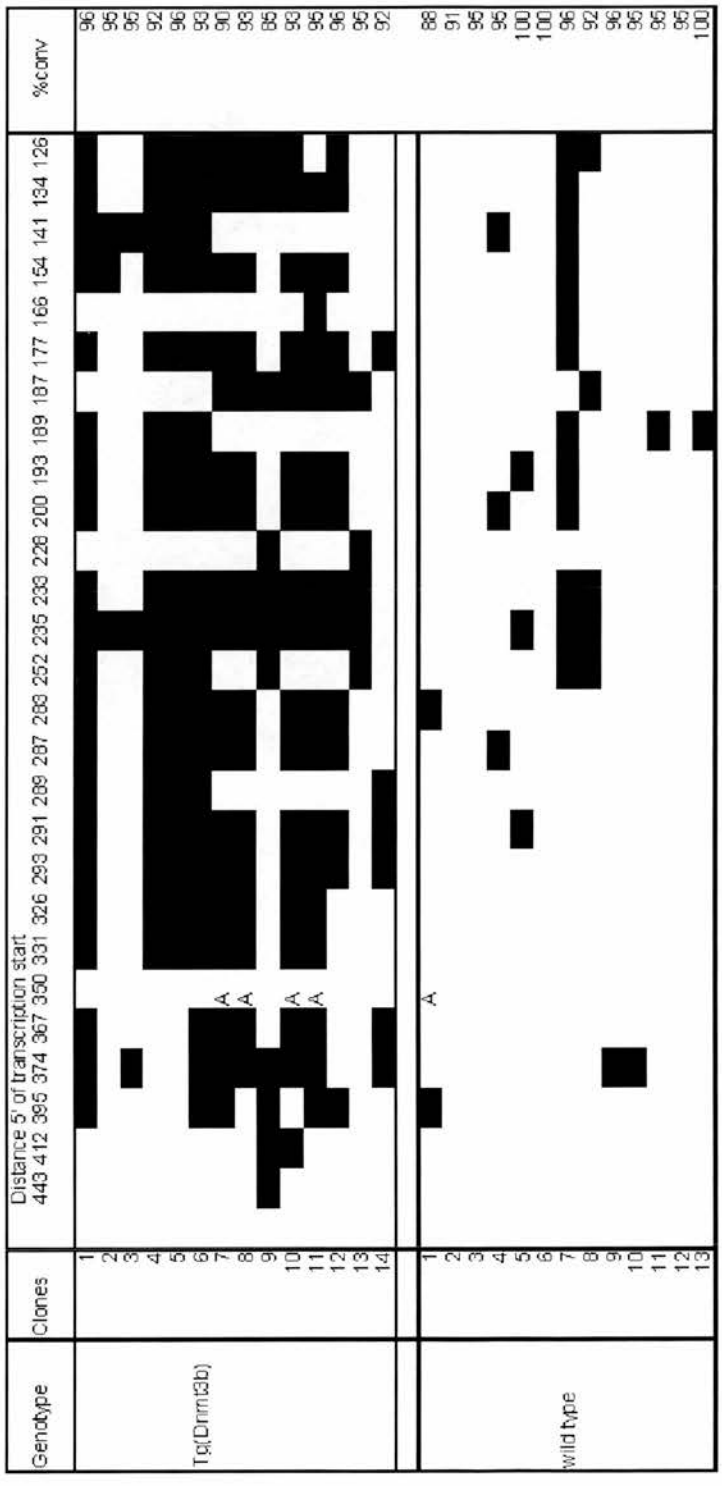
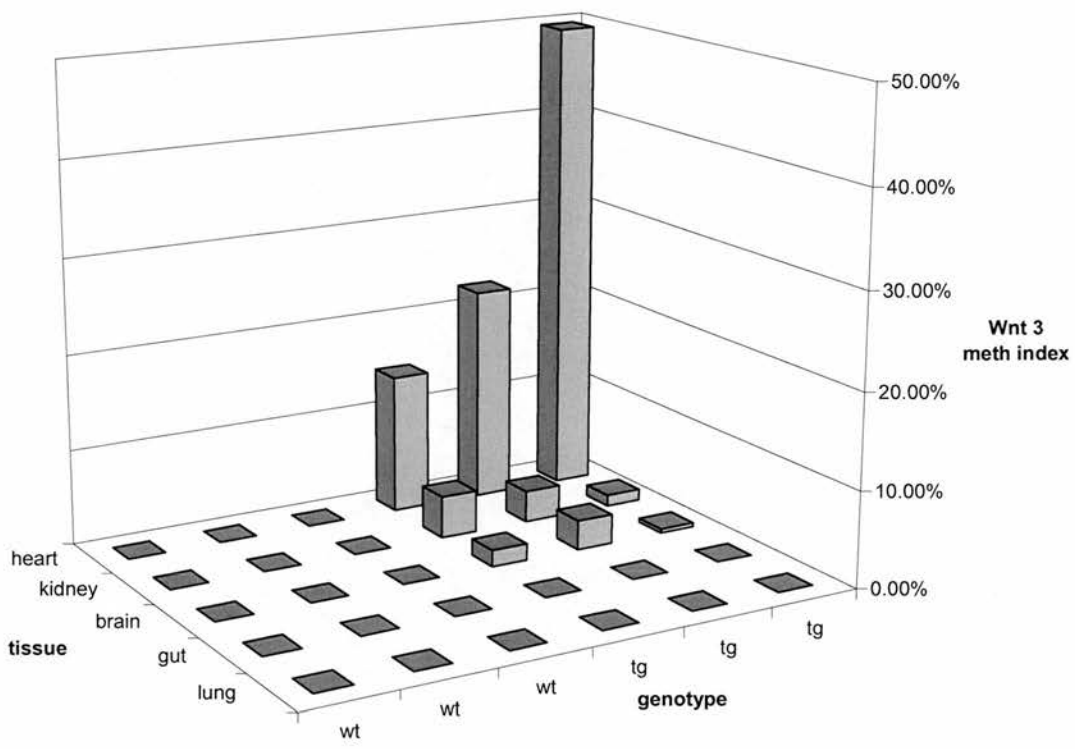


Figure 3.10 *Lrfn5* bisulphite sequencing. (Legend is with Figure 3.8)

A)



B)

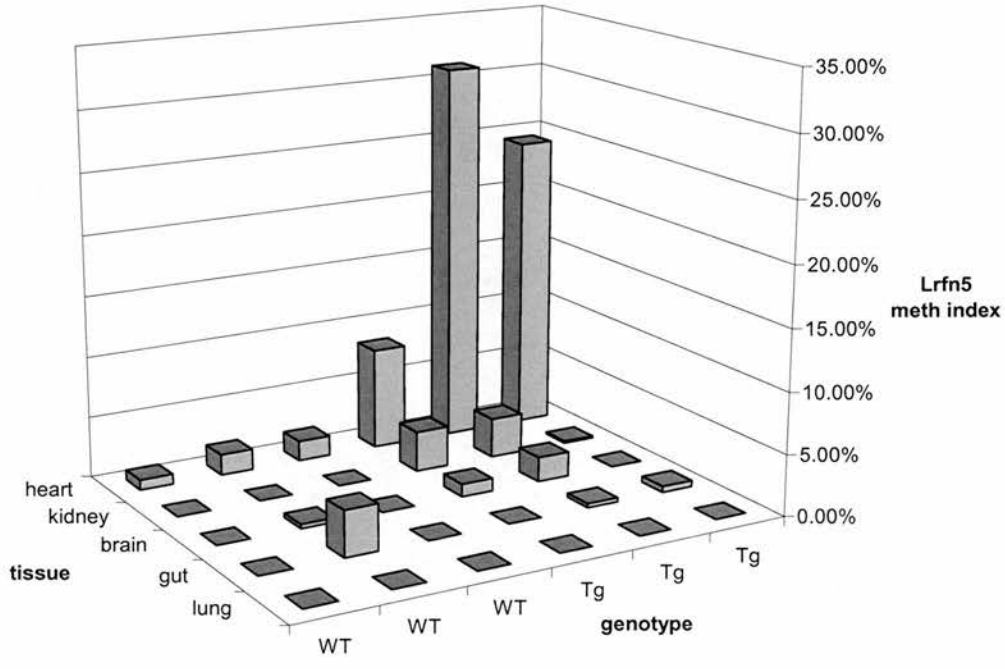
Wnt3 methylation index

	Wild Type			Transgenic		
	C_1	C_2	C_3	$3b_1$	$3b_2$	$3b_3$
Brain (+/-)	**0**	**0**	**0**	**1.7%** (1.1-2.6%)	**3.1%** (1.5-6.2%)	**0.42%** (0.31-0.58%)
Gut (+/-)	**0**	**0**	**0**	**0**	**0**	**0**
Heart (+/-)	**0**	**0**	**0**	**15%** (11-19%)	**22%** (16-31%)	**50%** (30-81%)
Kidney (+/-)	**0**	**0**	**0**	**4.4%** (3.1-6.3%)	**3.3%** (2.6-4.1%)	**1.2%** (0.97-1.5%)
Lung	**0**	**0**	**0**	**0**	**0**	**0**

Figure 3.11 *Wnt3* methylation by tissue

Figures 3.11 and 3.12 Graphs showing methylation status of *Wnt3* and *Lrfn5* as measured by M-QPR in different tissues from *Tg(Dnmt3b)*+ and wild type mice. The tables show methylation index in each tissue and range +/- 1 transformed standard error.

A)



B)

	Wild type			Transgenic		
	C1	C2	C3	3b1	3b2	3b3
Brain (+/-)	0.00%	0.34% (0.18-0.65%)	0.00%	1.0% (0.63-1.7%)	2.0% (1.7-2.4%)	0.027% 0.01-0.11%
Gut (+/-)	0.00%	3.7% (2.9-10%)	0.00%	0.00%	0.35% (0.019 -6.3%)	0.48% (0.023-10%)
Heart (+/-)	0.81% (0.60-1.1%)	1.7% (1.3-2.3%)	1.7% (1.0-2.8%)	8.4% (6.2-11%)	31% (28-35%)	24% (21-28%)
Kidney (+/-)	0.00%	0.00%	0.00%	3.3% (2.1-5.2%)	3.3% (2.2-4.7%)	0.14% 0.02-1.3%)
Lung (+/-)	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%

Figure 3.12 ***Lrfn5*** **methylation by tissue.** Legend is with Figure 3.11

A) *Wnt3* expression

	Wild Type 1	Wild Type 2	Tg 1	Tg2
Brain	**4.5E-03** (4.4-4.5)E-3	**2.5E-03** (2.2-2.8)E-3	**1.6E-03** (1.5-1.7)E-3	**7.9E-04** (6.9-9.2)E-4
Gut	**7.2E-05** (6.5-7.9)E-5	**9.8E-05** (9.7-1.0)E-5	**2.0E-03** (1.9-2.1)E-3	**1.0E-05** (0.89-1.2)E-5
Heart	ND	ND	ND	ND
Kidney	ND	ND	ND	ND
Liver	ND	**3.0E-05** (2.7-3.4)E-5	ND	ND
Lung	**1.6E-05** (0.61-4.3)E-5	ND	ND	ND
Ovary	ND	ND	ND	ND
Spleen	ND	ND	ND	ND
Thymus	ND	ND	ND	ND

B) *Lrfn5* expression

	Wild Type 1	Wild Type 2	Tg1	Tg 2
Brain	**1.0E-02** (0.96-1.0)E-2	**5.6E-03** (4.9-6.3)E-3	**3.2E-03** (3.1-3.3)E-3	**1.1E-03** (1.0-1.2)E-3
Gut	ND	ND	**3.6E-03** (3.2-4.0)E-3	ND
Heart	ND	ND	ND	ND
Kidney	ND	ND	ND	ND
Liver	ND	ND	ND	ND
Lung	ND	ND	ND	ND
Ovary	ND	ND	ND	ND
Spleen	ND	ND	ND	ND
Thymus	ND	ND	ND	ND

Table 3.1 *Wnt3* and *Lrfn5* expression. Expression levels of A)*Wnt3* and B) *Lrfn5* relative to actin expression as measured by quantitative PCR. ND= not detected (limit of detection $1x10^{-5}$ actin level). Error intervals represent transformed standard errors (see methods). Tg1 and Tg2 are the *Tg(Dnmt3b)+* animals.

4. Effects of Dnmt3b over-expression in the heart

4.1 Introduction

In this chapter the cardiac pathology of *Tg(Dnmt3b)+* mice is explored in more detail. The development of cardiomyopathy in *Tg(Dnmt3b)+* mice is probably the most striking and unexpected aspect of the phenotype. It is shown in the previous chapters that cardiomyopathy is not caused by an increase in hyperplastic myocardial growth. It has also demonstrated the presence of global increases in methylation and abnormal methylation of specific CpG islands in the hearts of *Tg(Dnmt3b)+* mice. Finally it has been shown that the phenotype can be recapitulated in a second mouse line, thus excluding an integration event as the cause for the cardiomyopathy. In this chapter, what is known about the role of epigenetic factors in cardiac development and disease is discussed, and there is an attempt to establish causative links between Dnmt3b expression, abnormal methylation and cardiomyopathy.

Cardiomyopathy could be defined as intrinsic disease of the myocardium which leads to contractile dysfunction, and can be primary (inherited) or secondary to a wide range of insults. Of these, pressure overload due to valvular disease or hypertension, toxic insults such as alcohol excess or anthracycline cytotoxic agents, hyperthyroidism, haemochromatosis, and inflammation due to viral myocarditis or autoimmune disease are the most important. The natural history of disease is progressive, with initially compensated disease progressing to the syndrome of heart failure, where the contractile capability of the heart is no longer sufficient to meet circulatory demands. The cardiomyopathies are classified morphologically as hypertrophic, dilated or restrictive. Genetic abnormalities causing dilated cardiomyopathy (DCM), which is the classification which best describes the pathology of *Tg(Dnmt3b)+* mice, have been identified in a number of inherited cases and provide some background to what is known about the molecular pathogenesis of this condition. All modes of inheritance including maternal (mitochondrial) have been described in DCM, although the predominant mode

of inheritance is autosomal dominant. The majority of mutations discovered so far involve cardiac structural and cytoskeletal genes (for review see Fatkin and Graham 2002), and transgenic mouse models have confirmed the importance of a number of these. Dilated cardiomyopathies also result from mutations in a number of genes involved in mitochondrial energy metabolism (both nuclear and mitochondrial DNA encoded genes), usually as part of a multisystem disorder. Cardiomyopathy has also been caused in transgenic mice with mutations in genes involved with other cellular processes, including calcium homeostasis, mitochondrial DNA replication fidelity, and cell signalling.

The most important question to address in our model is whether the phenotypic changes in *Tg(Dnmt3b)+* mice are directly caused by methylation, thus indicating that methylation is a sufficient condition on its own to perturb cell biology. In this chapter these questions are addressed by taking three separate approaches.

Firstly, a dose-response relationship between the level of Dnmt3b and the severity of cardiomyopathy is established, and the chronology of development of the histological and methylation abnormalities previously discovered in the adult hearts is investigated. It would support the hypothesis that methylation events are causative if abnormal methylation can be discovered before the onset of overt cardiac disease.

Secondly, an attempt is made to reverse the phenotype by treatment with demethylating agents and by breeding with mice deficient for the methyl binding protein Mbd2, in order to prove the role of DNA methylation in the pathogenesis of the cardiac disease. Several pharmacological inhibitors of DNA cytosine methyltransferases have been developed which consist of cytosine molecules modified at the 5- position of the pyrimidine ring. The most widely used of these are 5-azacytidine (AzC) and 5-aza 2'deoxycytidine/ decitabine (AdC) (Figure 4.1). Both molecules prevent methylation by becoming incorporated in to DNA and covalently linked to DNA methyltransferase enzymes. As well as their demethylating actions they also have cytotoxic activities as a

result of their cytosine analogue actions, inhibiting DNA synthesis and causing structural instability at their site of incorporation. Decitabine appears to have lower toxicity, possibly as a result of its predominant incorporation into DNA rather than RNA, leading to greater specificity of action. Both agents have been widely shown to be effective demethylating agents both *in vivo* and *in vitro* (for review see Goffin and Eisenhauer 2002). A complementary though indirect experiment which would support the hypothesis that the phenotype is mediated through DNA methylation would be to investigate whether the phenotype could be ameliorated by deficiency of downstream mediators of the methylation signal. As discussed in Section 1.3, the methyl binding protein Mbd2 is strongly implicated as such a mediator. *Mbd2*$^{-/-}$ mice have been produced by targeted insertion of a β geo cassette into exon 2 of the *Mbd2* gene, leading to a truncated protein with an intact N-terminus but no methyl-binding domain (Hendrich, *et al* 2001). *Mbd2* deletion leads to deficient silencing of methylated reporter constructs in cell lines, but does not appear to cause re-expression of imprinted genes or methylated retroviral elements *in vivo*. This perhaps suggests that Mdb2 is important for silencing of *de novo* methylated elements, but that gene silencing established during development is maintained by multiple mechanisms, thus leading to redundancy.

Finally we use gene expression profiling to identify genes which are downregulated in *Tg(Dnmt3b)*+ hearts and study the methylation status of the most strongly downregulated genes.

There is a limited but increasing literature on gene expression profiling in various models of cardiomyopathy and heart failure, which is somewhat contradictory. In the mouse, the most extensive data set (Blaxall, *et al* 2003) examines gene expression profiles from two different transgenic models. Profiles from one structural protein knockout, *Mlp*$^{-/-}$, and one calcium binding protein over-expression model (Calsequestrin) are compared with those from normal mice and with mice which have been rescued from heart failure. In that study, a total of 73 genes were found to be differently expressed beween the normal and heart failure models and restored in the rescued

phenotype. A complementary study examines expression profiles in 4 transgenic models of cardiac hypertrophy in young mice which have developed hypertrophy but have not yet developed heart failure (Aronow, *et al* 2001). In this study, there was no set of genes which was consistently associated with hypertrophy, with genes differing from normal instead appearing to reflect the specific models.

In studies from humans with cardiomyopathy and heart failure, many genes have been found to be dysregulated. Since the subjects for study are necessarily much more diverse than in mouse studies, and effects may well be confounded by the results of drug treatment, there is considerable variation between studies as regards dysregulation of individual genes. However, a number of common patterns do appear to emerge, including a predominant downregulation of genes involved in calcium homeostasis and energy metabolism, and an upregulation of sarcomeric, cytoskeletal and extracellular matrix genes in dilated cardiomyopathy (for review see Sanoudou, *et al* 2005). The simplistic message from these studies, as would be expected, is that on the one hand there do appear to be common changes in transcription which occur in heart failure from diverse causes, and on the other hand there are changes which are specific to the cause. It may therefore be possible to differentiate specific changes associated with Dnmt3b over-expression from non- specific changes as a result of heart failure and preliminary attempts to do this are described below.

4.2 Cardiomyopathy is absent up to 6 weeks of age

In order to determine the age of onset of cardiac histological changes and methylation abnormalities, H+E sections from hearts from F10 generation mice sacrificed at 2 weeks, 6 weeks, and 4 months of age were examined by microscopy. The microscopist was blinded to the age and genotype of the mice. A summary of the animals examined and cardiac weights is shown in Table 4.1. Histological abnormalities were detected only in the transgenic mice at 4 months of age.

In order to investigate whether methylation changes precede histological changes, genomic DNA extracted from the 2-week and 6-week mice was analysed for methylation of *Wnt3*, and the cyclin-dependent kinase inhibitor genes *Cdkn1b* (p21cip1) and *Cdkn2a* (p19arf) by M-QPR. The cell cycle regulatory genes were chosen on a candidate gene basis since these have been found to be methylated in cancer, and are thus potentially prone to aberrant methylation; if these were aberrantly silenced this could lead to cell proliferation and provide a plausible mechanism for the increase in heart size. No methylation was detected for any of these genes in either the 2 week or 6 week animals. Standard curves and quantification of DNA by real time PCR showed 100% efficiency of amplification, and detectable PCR product was present in the control (no restriction sites) reaction, showing that the lack of a PCR product was not due to failure of the PCR reaction or poor quality of starting DNA.

We conclude that cardiomyopathy occurs late, between 6 weeks and 4 months of age. The presence of global increases in cardiac CpG methylation at 8 weeks of age (see Section 3.2), when the mice are also apparently healthy is strong evidence that abnormal methylation precedes the development of cardiomyopathy and is not simply secondary to disease. However the late development of methylation at specific loci suggests that at some loci methylation abnormalities are not present in early life but may be secondary to either age or disease.

4.3 The presence of cardiomyopathy correlates with the level of Dnmt3b expression

In the later generations (from F11) of *Tg(Dnmt3b)+* mice there a decrease in the proportion of mice showing expression in the skin by western blot has been noted. This has been particularly marked in a few individual litters of mice, especially in the F12 generation. The phenomenon appears to be due to silencing of the transgene in skin, since mice negative by western blot can still be positive by PCR genotyping. Silencing does not appear to be passed to subsequent generations in any manner which is simply

predictable, since offspring of mice which are PCR positive for the transgene but have no expression in the skin can show strong skin expression. Despite the occurrence of silencing in the skin, early death has still been noted in mice which are transgene positive by PCR negative by western blot.

A possible explanation for this is that the transgene may still be expressed in other tissues despite silencing in the skin. In order to determine whether transgene silencing was occurring in the heart, Dnmt3b expression in the heart at 5 months of age was analysed in six mice from the F12 generationwhich were PCR positive for transgene but negative by Western blot in skin (Figure 4.2). The results demonstrate that transgene silencing can occur in the heart, but that mice showing silencing in the skin can still have strong expression in heart.

In order to determine whether transgene expression is correlated with phenotypic changes, H+E and Sirius Red (collagen) stained sections from the hearts were examined by microscopy. The microscopist was not aware of the cardiac expression results.

Table 4.2 compares the genotype and cardiac Dnmt3b expression with cardiac weights and the presence or absence of abnormal histology, as judged by the microscopist. In the small number of mice studied, the strength of transgene expression does appear to predict the presence or absence of abnormal histology. The correlation is most clearly seen in the Sirius red stained sections (Figure 4.2) which show a marked increase in fibrosis only in the animals with strong transgene expression. In addition, heart weight: body weight ratio is significantly increased in the mice showing transgene expression ($p<0.025$ by Mann Whitney U test). There is no significant increase in heart weight to body ratio for mice which are transgene positive but do not show cardiac expression ($p<0.40$ by Mann Whitney U test).

The results show that the cardiomyopathy is caused directly by the action of Dnmt3b protein as opposed to being an effect of random transgene integration. It is difficult to

prove that the mice with silent expression in the heart may not have developed cardiomyopathy at a later age, since it is necessary to sacrifice the mice to investigate the heart histology and expression level. It is also difficult to prove that silencing of Dnmt3b in the heart leads to improved survival for similar reasons (and it is also difficult to prove that early death is caused by cardiac disease). However supportive evidence that it is transgene expression and not simply inheritance which predicts survival can be obtained from the second line of mice (Section 2.5) where there is a much higher frequency of transgene silencing. Out of the 7 from this line so far which have shown transgene expression in the skin, none have survived past 166 days, whereas there are currently 14 mice which are positive for transgene transmission by PCR but not by western blot which remain alive at over 320 days (range 322-389).

The presence of silencing could have implications for the interpretation of other results, raising the question of whether effects on methylation are really comparable in experiments performed on mice of different generations. However, the fact that skin transgene expression inevitably predicts early death, even in later generation mice, suggests that skin expression is a good predictor of strong expression in other tissues, especially the heart, although the converse may not always be the case.

4.4 Aberrant CpG island methylation is secondary to cardiomyopathy at some loci

To investigate whether methylation events at specific loci may be secondary to disease, methylation status of *Wnt3*, *Cdkn1b* (P21cip1) and *Cdkn2a* (P19arf) were measured by M-QPR in the 5 month old animals with variable transgene silencing studied in Section 4.3. Results are shown in Figure 4.3. There is a consistent increase in *Cdkn1b* methylation, albeit at low level, in the animals showing strong cardiac transgene expression and cardiomyopathy. However, *Wnt3* and *Cdkn1b* methylation is only detectable in one of the affected animals. Although the number of animals studied is small, the results confirm that cardiomyopathy can occur before methylation changes at these loci are detectable. The results suggest that methylation of *Wnt3* or *Cdkn1b* are

not necessary for the development of cardiomyopathy. This also supports the hypothesis proposed in Section 4.2 that many of the methylation events occurring in the context of methyltransferase expression may be secondary to or facilitated by the presence of disease.

4.5 The phenotype of Dnmt3b over-expression is not reversed by 5'aza 2-deoxycytidine treatment

It would provide strong evidence for the hypothesis that DNA methylation is involved in the pathogenesis of the heart disease if the phenotype could be prevented or reversed by treatment with demethylating agents. We therefore attempted to reverse the phenotype by treating mice with 5'deoxy-2azacytidine (AdC/ decitabine). A total of 9 transgenic mice from the F5 and F6 generations were administered 1mg/kg AdC weekly for 11 weeks by subcutaneous injection starting from 6 weeks of age. The dose of AdC was chosen as this dose was sufficient to reduce tumour formation and reactivate silenced genes in $Apc^{+/min}$ mice (Laird, *et al* 1995). Wild type littermates were also treated to control for AdC toxicity.

Treatment with AdC produced no evidence of a difference in survival when compared against transgenic mice from the F5 and F6 generations ($p=0.44$ for difference in survival using Kaplan Meier estimate). The control animals remained healthy with all non-transgenic animals surviving beyond the longest living transgenic animal. See Figure 4.4.

Thus, AdC treatment in a dosage shown in previous literature to be sufficient to cause demethylation is unable to reverse the early mortality in *Tg(Dnmt3b)+* mice. While it has not strictly speaking been proven that the early mortality is caused by cardiomyopathy, the results are strongly suggestive that a period of demethylating agent therapy is unable to reverse the cardiomyopathy.

In order to demonstrate that demethylation in the myocardium is achievable using AdC treatment, a pilot experiment was carried out to investigate whether an imprinted gene can be demethylated in the heart by AdC. A separate group of 3 normal and 3 transgenic mice from the F13 generation were administered 1mg/kg AdC, and a further 3 transgenic mice were administered a higher dose of 2.5mg/kg. Mice were sacrificed 24 hours following the final injection.

Demethylation at was studied by measurement of methylation at the imprinted *Igf2r locus* by M-QPR, and compared to age-matched mice which had not received AdC treatment (Figure 4.5). As expected, the methylation index underestimates the proportion of methylated alleles (expected to be 50% in the wild type control animals). This is because only the most densely methylated alleles which are methylated at all 5 restriction sites within the amplicon are quantified as methylated. There were significant differences in methylation between the treatment groups ($p=0.02$ by analysis of variance). Post-hoc analysis suggests that this is due to a combination of demethylation being achieved in the AdC treated wild type animals, and hypermethylation in both AdC treated and untreated transgenic animals. The higher dose of AdC caused more demethylation in the transgenics, reducing *Igf2r* to normal levels. We can conclude that it is likely that AdC treatment failed to reduce mortality as demethylation could not be achieved in the transgenic animals with the dose administered.

While the higher dose of AdC may have caused more demethylation, this dose appears too toxic for long term administration. A further three transgenic animals received ongoing treatment twice weekly with 2.5 mg/kg AdC. Two animals died at 114 and 119 days of age (earlier than would be expected for untreated transgenic animals), and tissues were too decomposed for analysis; the final animal was sacrificed at 125 days of age after a total of 12 injections, as it looked unhealthy. *Igf2r* methylation analysis on this animal showed evidence of continued demethylation (methylation index. 4.5% , transformed standard error 0.5%). There were no histological changes of cardiomyopathy in this animal, but it is not possible to conclude that AdC treatment

prevented these changes since the cardiac histology may not yet be detectable at this age. It would appear more likely that the animals receiving ongoing treatment have died as a result of AdC toxicity.

4.6 Mbd2 deficiency does not prevent cardiomyopathy in *Tg(Dnmt3b)+* mice

Dnmt3b over-expressing mice which are Mbd2 deficient were produced by mating *Tg(Dnmt3b)+* mice (F8 generation) to $Mbd2^{-/-}$ knockout mice (previously reported in Hendrich et al 2001) to produce compound heterozygotes. The *Tg(Dnmt3b)+*, $Mbd2^{+/-}$ compound heterozygotes were then crossed with $Mbd2^{-/-}$ mice to produce litters with the expected Mendelian ration of genotypes *Tg(Dnmt3b)+*, $Mbd2^{-/-}$ 25%: *Tg(Dnmt3b)-*, $Mbd2^{-/-}$ 25%: *Tg(Dnmt3b)+*, $Mbd2^{+/-}$ 25%: *Tg(Dnmt3b)-*, $Mbd2^{-/-}$ 25% .

Survival curves for *Tg(Dnmt3b)+*, $Mbd2^{-/-}$ and *Tg(Dnmt3b)+*, $Mbd2^{+/-}$ mice are shown in Figure 4.6 and are compared with *Tg(Dnmt3b)+*, $Mbd2^{+/+}$ mice derived from males of the same generation used for intercrossing. The *Tg(Dnmt3b)+*, $Mbd2^{+/-}$ mice include both F1 and F2 compound heterozygotes. There is no evidence of any difference in survival in either *Tg(Dnmt3b)+*, $Mbd2^{-/-}$ or *Tg(Dnmt3b)+*, $Mbd2^{+/-}$ genotypes. In addition, hearts from mice at post mortem show histological changes identical to *Tg(Dnmt3b)+*, $Mbd^{+/+}$ mice. Although numbers are small, it can be concluded that Mbd2 is not necessary for the pathogenesis of the cardiac pathology.

4.7 Downregulation of genes in Tg(Dnmt3b)+ can occur by methylation-independent mechanisms.

On the basis that phenotypic changes caused by Dnmt3b over-expression are likely to result from the silencing of gene expression by methylation, gene expression profiles in hearts from a single 130 day old *Tg(Dnmt3b)+* mouse and a wild type littermate control were measured by microarray, as a screening process in order to find downregulated genes which may have been aberrantly methylated. The Affymetrix U74v2 array

contains oligonucleotides representing approximately 12000 genes. Of the transcripts for which the signal from at least one probe from either animal wass detectable, there were 274 transcripts increased > 2 fold in the transgenic and 204 genes decreased >2 fold (data not shown).

We investigated the promoter methylation state of the two characterised genes which showed the greatest downregulation in themicroarray analysis. Bisulphite sequencing of the promoter region of the *Hop* and *Ces3* genes was performed on DNA extracted from hearts from normal and transgenic mice. The downregulation of Hop was also confirmed by western blotting (Figure 4.7b). The *Hop* gene (6.4 fold reduced in the Affymetrix data) has a recognised CpG island promoter. There is also a second CpG island around 20kb 3' to this region, which is immediately 5' to a second region showing homology to expressed sequence tags corresponding to *Hop*, which probably represent an alternative promoter (Figure 4.7a). Both regions were examined by bisulphite sequencing. There was no evidence of a difference in CpG island methylation examining the 13 CpG sites present in the product from the most 5' CpG island (HopP1), or the 17 CpG sites in the 3' island (HopP2) comparing 10 normal to 10 transgenic clones in either CpG island (Fig 4.8). The promoter of the *Ces3* gene (17.5-fold reduced) does not contain a CpG island but does contain a CpG dinucleotide in an Sp1 transcription factor binding site which has shown to be important in transcriptional regulation (Douglas, *et al* 2001). 5/28 clones showed methylation at this site in the normal animal compared to 5/16 in the transgenic animal (p=0.16 by Chi-square test). The lack of change in methylation suggests that these gene expression changes may be secondary to the cardiac pathology, or may be downstream effects of silencing of regulatory factors.

4.8 Discussion

The presence of a cardiac phenotype in the *Tg(Dnmt3b)+* mice is perhaps the most unexpected result but one which could have widespread implications. If it could be shown that the cardiac pathology is the result of epigenetic silencing caused by Dnmt3b over-expression, this could open the door to a potentially new area of investigation into the role of epigenetic regulation and dysregulation in cardiac disease. At the least it may provide insights in to new mechanisms of cardiomyopathy, which would have relevance not only for inherited myopathies but also end stage cardiac muscle changes that occur in heart failure. More interestingly, it is possible that maladaptive epigenetic responses are involved in the pathogenesis of heart failure and of remodelling of the heart in response to stresses such as hypertension and following myocardial infarction, which are major causes of morbidity in the developed world.

Interpretation of the findings of this study is made more difficult because of the limited understanding of both epigenetic processes in the heart and the transcriptional effects of heart failure and cardiomyopathy, making it difficult to distinguish between cause and effect. However some inference can be drawn from the limited number of published studies in both areas, which are both becoming more prominent areas for research.

While I am unaware of any studies describing functions of DNA methylation in cardiac development or pathology other than the development of ventricular septal defects in *Dnmt3b*$^{-/-}$ embryos (Ueda, *et al* 2006) , there is an emerging literature of the importance of histone modifications and histone modifying enzymes in the control of cardiac hypertrophy.

The earliest suggestion that there may be a link between chromatin structure and cardiac regulation came from the discovery of increased transcription of fetal gene products, such as the embryonic beta-myosin heavy chain, with exposure of nuclease hypersensitive sites (Huang and Liew 1998). Class II histone deacetylases were

subsequently shown to play a role in this shift in transcription by being exported from the nucleus in response to phosphorylation by a calcium-calmodulin dependent kinase (McKinsey, *et al* 2000).

The classical histone deacetylases (HDAC) cause deacetylation by simple hydrolysis of the acetyl group and are classified in to two classes: Class I HDACs, which include proteins Hdac1,2,3 and 8 and have close homology to the yeast transcriptional regulator RPD3; and class II HDACs, which include Hdac 4,5,6,7,9 and have homology to the yeast protein HDA3 (for review see de Ruijter, *et al* 2003). Class I histone deacetylases are generally ubiquitous whereas class II HDACs tend to have more tissue specific distributions and in addition to the catalytic regions have other domains that are involved with protein-protein interactions. In addition there is a third class of HDACs, the sirtuins or class III HDACs which have homology to the yeast protein SIR2, which are NAD^+ dependent and release acetyl groups as 2'-o-acetyl-ADP-ribose.

The importance of histone deacetylases was confirmed by deleting the Class II histone deacetylase Hdac9 in transgenic mice. *Hdac9*$^{-/-}$ mice are essentially normal at birth but gradually develop cardiac hypertrophy by 8 months of age, which is accelerated by aortic banding pressure overload and increased calcium dependent signalling (Zhang, *et al* 2002), suggesting that Hdac9 has an anti-hypertrophic function. Paradoxically in the light of these findings, treatment with histone deacetylase inhibitors has an inhibitory effect on cardiac hypertrophy caused by pressure overload (Kong, *et al* 2006) and pharmacological manipulation (Kee, *et al* 2006). It has been suggested that this effect could be explained if class I histone deactetylases have the converse action, promoting hypertrophy. Indirect evidence that this may be the case comes from study of the cardiac specific nuclear factor Homeodomain only protein (Hop), which causes cardiac hypertrophy when over-expressed in transgenic mice by binding to and repressing the anti hypertrophic transcription factor Serum response factor (Srf). Hop- induced hypertrophy appears to be dependent on interaction with the class I histone deacetylase Hdac2, since the phenotype is not produced by a mutant Hop protein with a mutated

Hdac2 binding site, and is reversed by treatment with the histone deacetylase inhibitor trichostatin (Kook, *et al* 2003). Histone acetyltransferases have also been implicated in cardiac remodelling, with mice with deletion of the p300 acetyltransferase showing defective remodelling following experimental myocardial infarction (Miyamoto, *et al* 2006).

The basic message from these studies is that the epigenetic regulation of cardiac transcription is likely to be a dynamic process in the developed animal, with transcriptional responses to stimuli being regulated by epigenetic mechanisms as well as transcription factors. Given the close interactions between DNA methyltransferases and the histone modification system, it is perhaps not surprising that altering the levels of DNA methyltransferases may have perturbed the balance of cardiac gene regulation, either through inducing abnormal methylation or by direct interaction with histone modifying enzymes. Some features of the current study perhaps favour the latter possibility, such as the late development of methylation changes and the failure of 5'-azacytidine treatment or Mbd2 deficiency to reverse the phenotype. Along similar lines, while it might be thought more than a coincidence that Hop appears to be downregulated in the *Tg(Dnmt3b)+* mice given the close association of Hop with histone deacetylases, downregulation did not seem to be caused by promoter methylation. However, it is still quite possible that Dnmt3b over-expression and Hop are linked in the pathogenesis of the cardiomyopathy, either through methylation independent silencing of *Hop* or interference with the function of Hop by altering levels of histone deacetylases.

One of the major questions arising from *Tg(Dnmt3b)+* animals is the question of why abnormalities occur is specific tissues only, and most severely in the heart. One feature from the histone deacetylase studies which may be instructive is the timecourse of events in the Hdac9 knockout animals. While the histological changes in the Hdac9-/- animals is distinct from the changes in the *Tg(Dnmt3b)+* mice, and it might furthermore be expected that opposite effects would be seen given the recruitment of histone

deacetylases by Dnmt3b and methyl binding proteins, a similar phenomenon of late onset disease confined to the heart occurs in both models. One could argue that this may help to explain why the heart is the most severely affected organ in the *Tg(Dnmt3b)+* mice despite widespread expression of the transgene, since Hdac9 is also widely expressed but depletion only gives a cardiac phenotype. This may indicate a particular sensitivity to alterations in the balance of epigenetic modifications in the heart.

One can speculate what factors might lead to an increased sensitivity to epigenetic regulation in the heart. Thinking at the most general level, the two most unusual characteristics of the heart are the fact that it is non-mitotic, and its heavy metabolic workload. Since the essential nature of epigenetic modifications is that their control of transcription is heritable during cell division, this provides a possible mechanism. It is possible that error-checking processes are important for maintaining the fidelity of epigenetic replication during mitosis (for example it is known that Dnmt1 has higher specificity for CpG than either of the established *de novo* methyltransferases), and the absence of this error checking in the non- mitotic cell may lead to accumulation of errors in transcriptional regulation. Arguing against this however is the proliferation of epigenetic abnormalities in cancer where cells are undergoing extensive mitosis.

Could the high metabolic requirements of the heart cause sensitivity to epigenetic changes? Trivially, the substrates for methylation and acetylation are products of metabolism. Single carbon groups are not used as energy sources and so interference with methylation seems unlikely. This is supported by the lack of cardiac effects of nutritional deficiencies affecting single carbon metabolism such as deficiencies of vitamin B12, folate or methionine. However, the substrate for histone acetylation is acetyl-coenzymeA, which is central to energy metabolism. It is an attractive hypothesis that histone deacetylation could occur under conditions of increased aerobic metabolic demand or decreased acetyl supply from diet or lipid catabolism, as a result of decreased acetyl supply to histone acetyltransferases. This could lead to a metabolic autoregulation of transcription, promoting histone deacetylation and transcriptional

inactivation under periods of metabolic stress. The hypothesis would be consistent with the amelioration of pressure induced hypertrophy by histone deacetylases, and would also be consistent with increased histone deacetylation recruitment in *Tg(Dnmt3b)+* mice causing a phenotypic changes in the heart. The hypothesis would be difficult to prove, not least since there are few molecular biologists with detailed experience of "old fashioned" methods of manipulating metabolism, but perhaps the simplest counterexample to seek would be to investigate whether cellular levels of acetyl coenzyme A fluctuate in response to metabolic demand sufficiently to affect histone acetylation rates meaningfully.

An alternative, and not mutually exclusive, possibility for the predominant phenotype being cardiac in nature is that Dnmt3b is involved with maintaining fetal gene expression patterns, since both Dnmt3b expression and cardiac growth by hyperplasia, as opposed to hypertrophy, are reduced after birth. The continued expression of Dnmt3b could result in the failure to terminate fetal expression patterns leading to maladaptive responses to stimuli similar to those seen in the *Hdac9*$^{-/-}$ animals.

While the development of cardiomyopathy as a result of Dnmt3b over-expression was unexpected, it ties in well with the growing interest in the importance of epigenetic regulation in the heart, and the idea that the heart may be particularly sensitive to epigenetic changes. Given the widespread interest of epigenetic regulation in development and in mitotic phenomena such as cancer, it is ironic that the clearest phenotype seen in *Tg(Dnmt3b)+* mice is seen in an organ which is end-differentiated and has an extremely low cancer incidence. Although it may seem that the over-expression of Dnmt3b in an organ which it is not normally expressed would at first seem to have limited relevance to actual physical phenomena, perhaps it raises deeper questions about the role of epigenetic factors in the end-differentiated cell.

4.9 Figures and Tables

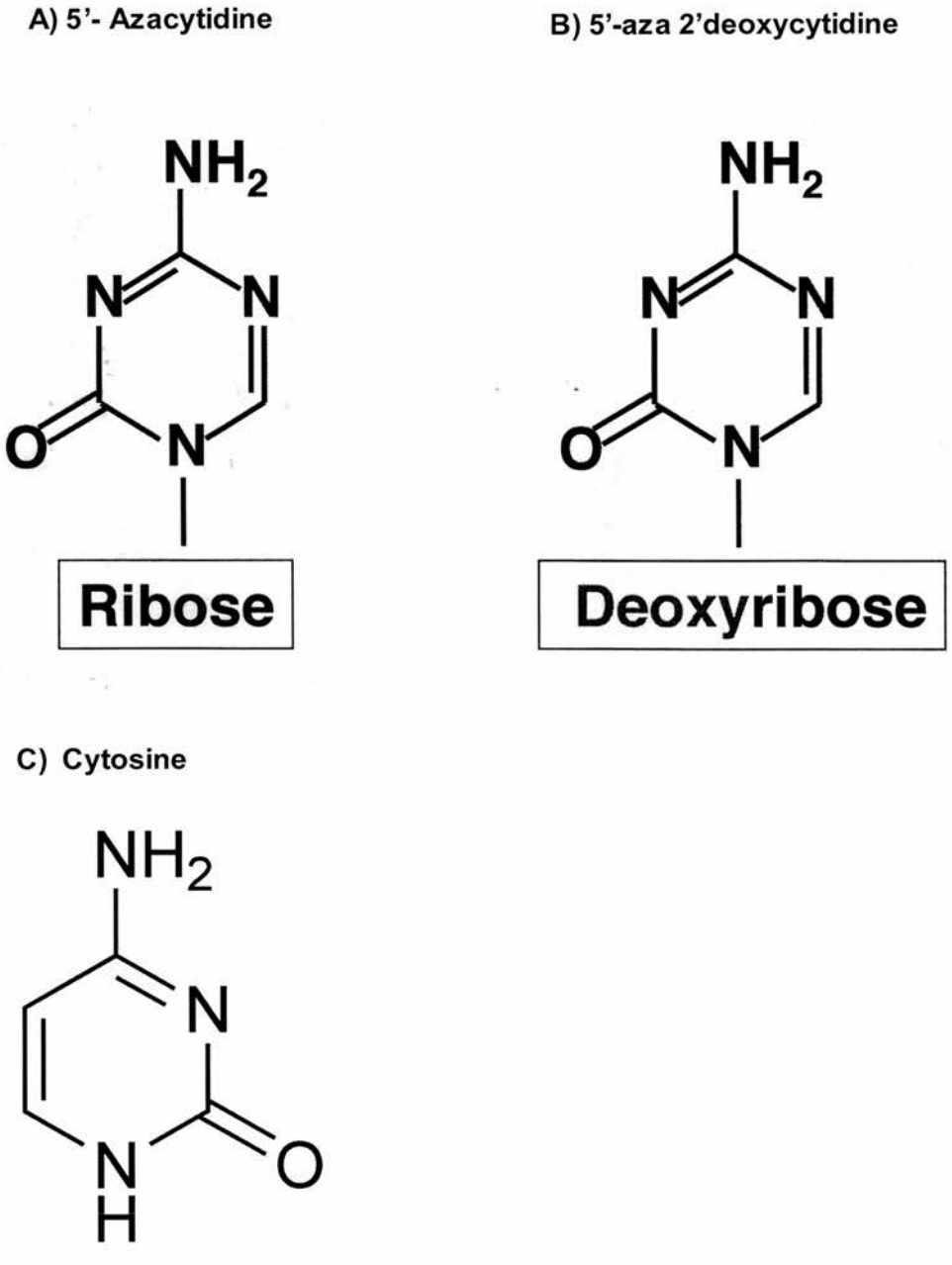


Figure 4.1 Molecular structures of demethylating drugs. Structure of 5'-azacytidine and 5'-Aza 2'deoxycytidine (decitabine), with structure of cytosine for comparison.

Age	Genotype	Body Weight (g)	Heart weight (mg)	Heart:body ratio (x1E-3)	Histology
4 months	pos	21.9	170	7.8	abnormal
4 months	neg	21.9	165	7.5	normal
4 months	neg	24.2	130	5.4	normal
4 months	pos	25.0	190	7.6	abnormal
6 weeks	neg	23.8	180	7.6	normal
6 weeks	pos	21.4	170	7.9	normal
6 weeks	pos	22.1	130	5.9	normal
6 weeks	neg	19.4	140	7.2	normal
6 weeks	pos	18.1	120	6.6	normal
6 weeks	neg	18.2	140	7.7	normal
2 weeks	pos	9.4	69	7.3	normal
2 weeks	pos	10.0	76	7.6	normal
2 weeks	pos	9.2	64	7.0	normal
2 weeks	neg	9.2	67	7.3	normal
2 weeks	neg	9.8	67	6.8	normal
2 weeks	neg	10.4	73	7.0	normal
2 weeks	neg	8.4	57	6.8	normal

Table 4.1 Presence of cardiomyopathy in *Tg(Dnmt3b)* mice by age. Relationship between genotype (by western blot), age and the presence or absence of histological changes characteristic of cardiomyopathy in F10 *Tg(Dnmt3b)* mice.

A)

mouse	genotype	heart expression	histology	body wt(g)	heart(mg)	heart:body
795	-	-	normal	37.4	171	4.57
796	+	-	normal	27.7	166	5.99
797	-	-	normal	26.2	147	5.61
798	-	-	normal	34.6	205	5.92
799	+	-	normal	26.9	195	7.25
801	+	+	abnormal	22.1	167	7.56
802	-	-	normal	33.2	217	6.54
803	+	-	normal	25.7	132	5.14
829	+	-	normal	28.1	214	7.62
837	+	+	abnormal	20.4	239	11.72
838	+	+	abnormal	18.25	185	10.14

B)

heart expression	n	Rank sum	Mean rank	U
no	8	37.0	4.63	23.0
yes	3	29.0	9.67	1.0

p=0.02

C)

genotype	n	Rank sum	Mean rank	U
pos	4	22.0	5.50	4.0
neg	4	14.0	3.50	12.0

p=0.34

Table 4.2 Relationship of heart mass to Dnmt3b expression. A) Comparison of cardiac weights and histology according to genotype (as determined by PCR), and cardiac transgene expression in 5 month old F12 mice. Heart weight: body weight ratios are $x10^{-3}$. B) Nonparametric statistics for heart:body weight ratio according to cardiac transgene expression status. C) Nonparametric statistics for heart:body weight ratio according to genotype for animals which do not show detectable transgene expression. Significance values are calculated by the 2-tailed Mann-Whitney U test. The results show an increase in heart mass which is significantly related to expression of the transgene in the heart but not to PCR genotype.

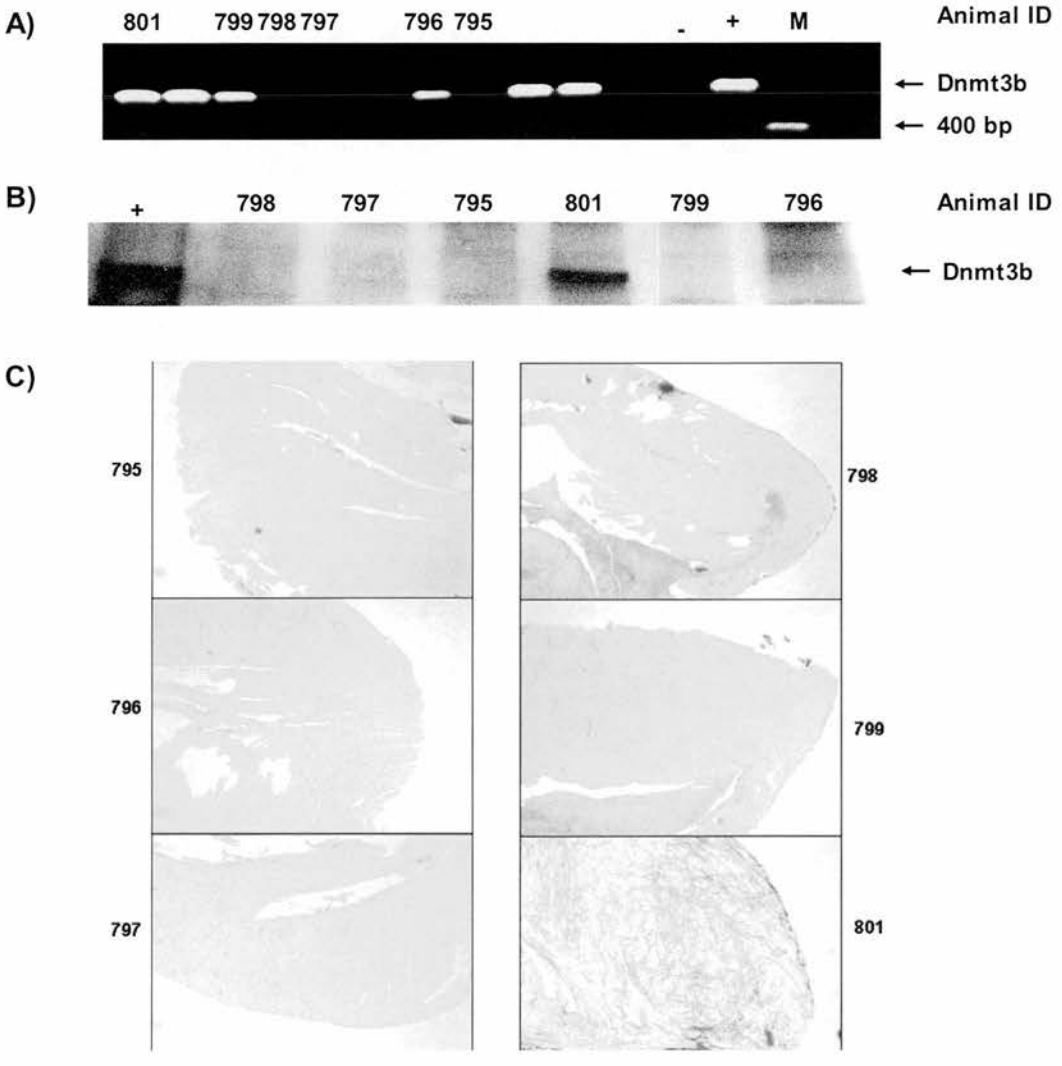


Figure 4.2 Relationship of cardiac phenotype to Dnmt3b protein expression level. A) PCR genotyping of mice from F12 litter. M=size marker. B) Western blot of cardiac Dnmt3b expression from same animals. C) Sirius red stained sections from hearts of above animals demosnstrating extensive fibrosis only in the animal with strong cardiac transgene expression. Animals were sacrificed at 5 months.

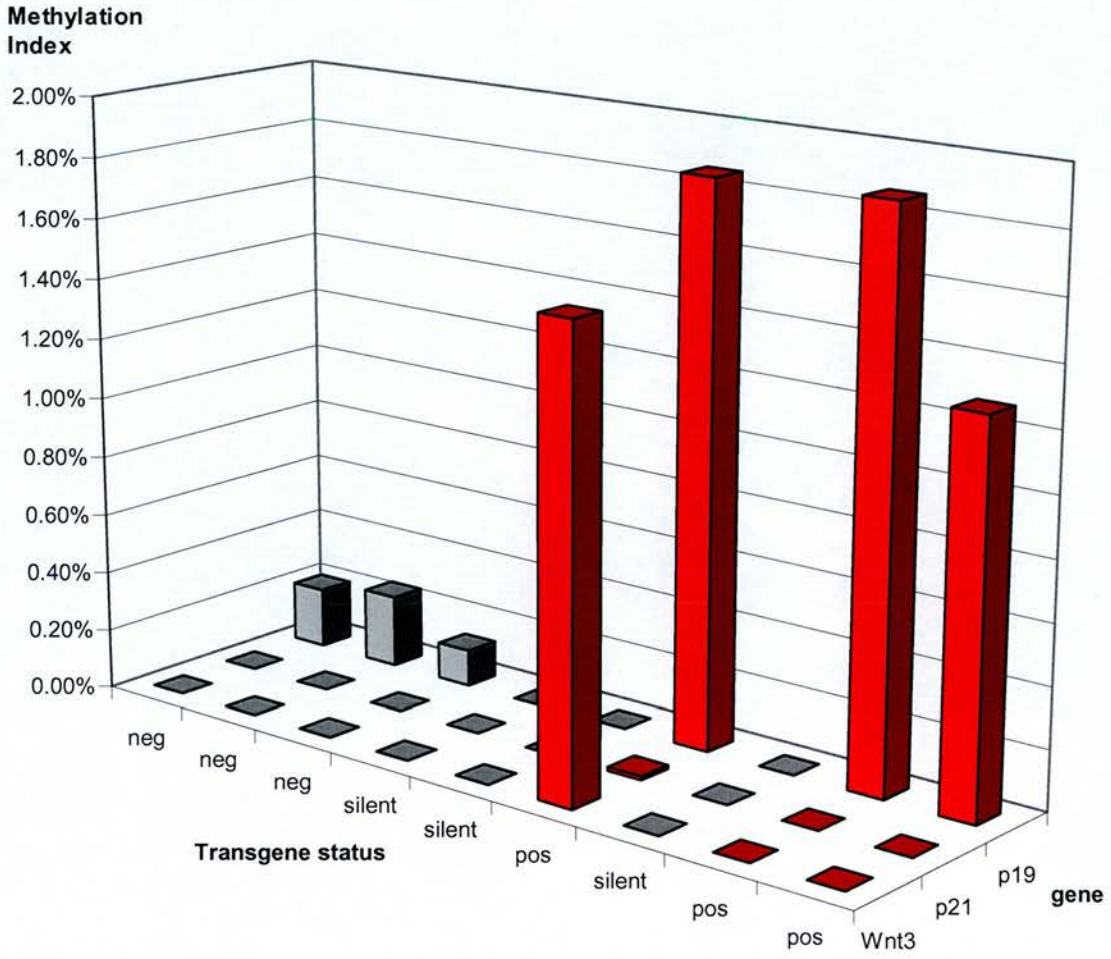
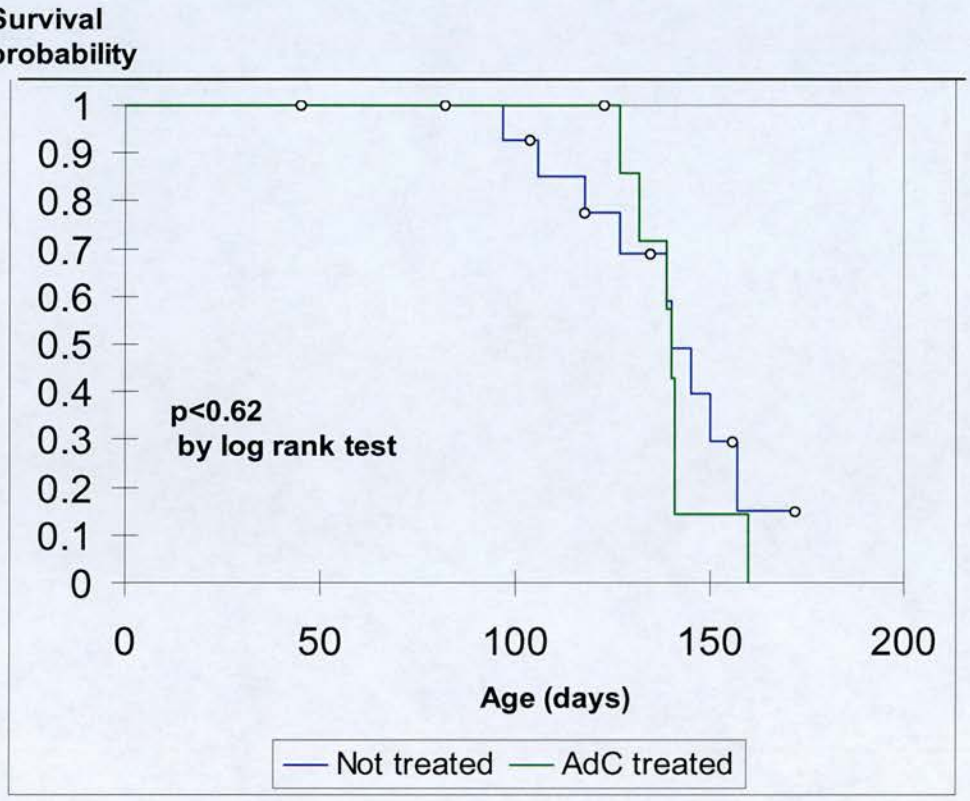


Figure 4.3 Aberrant methylation levels by Dnmt3b expression level. Methylation status of *Wnt3*, *Cdkn1a* (p21), *Cdkn2a* (p19) in heart, classified according to Tg(Dnmt3b) genotype and cardiac Dnmt3b expression. neg= *Tg(Dnmt3b)-;* silent= *Tg(Dnmt3b)*+ by PCR but undetectable expression in heart and skin; pos= *Tg(Dnmt3b)*+ by PCR, Dnmt3b detectable in heart but not skin.

A)



B)

	Total n=	censored	median survival (d)	95% confidence interval
untreated	17	8	140	127-157
treated	8	1	140	132-147

Figure 4.4 Effect of decitabine treatment on survival of *Tg(Dnmt3b)*+ mice. A) Kaplan-Meier survival curves and B) estimated median survival, comparing survival of *Tg(Dnmt3b)*+ mice treated with 1 mg/kg decitabine with untreated *Tg(Dnmt3b)*+ mice from the same generations (F5/6). Mice which were sacrificed early are included in the analysis as censored observations (shown as circles).

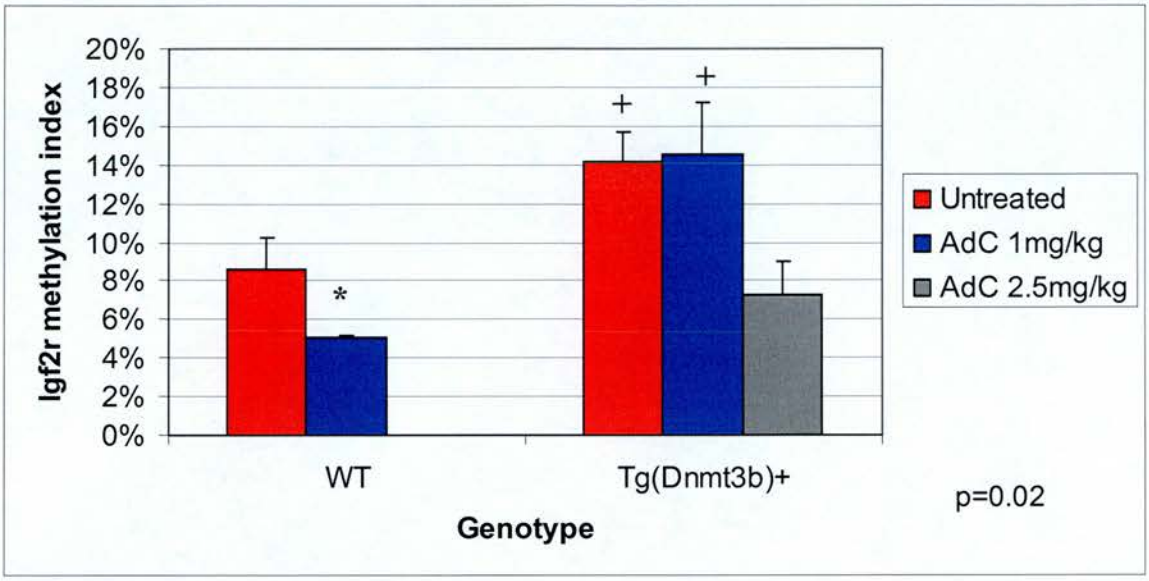
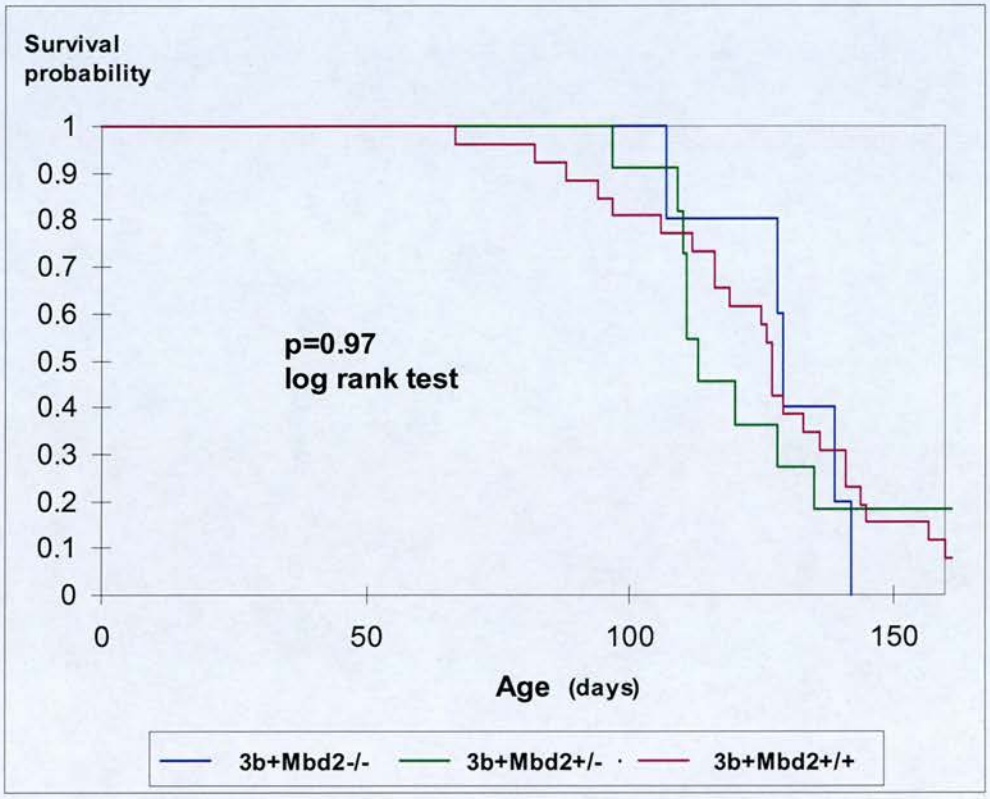


Figure 4.5 Effect of decitabine treatment on *Igf2r* methylation. Graph comparing cardiac *Igf2r* methylation as measured by M-QPR following decitabine treatment. Animals were treated for 2 weeks (5 injections) with the dose shown in the legend. Error bars represent standard error from 3 separate animals. Significance value is for the existence of difference in means in the experiment as a whole by one way analysis of variance. The group marked as * is significantly different from groups marked + at the 5% significance level (Bonferonni correction for multiple testing). Wt=Wild type.

A)

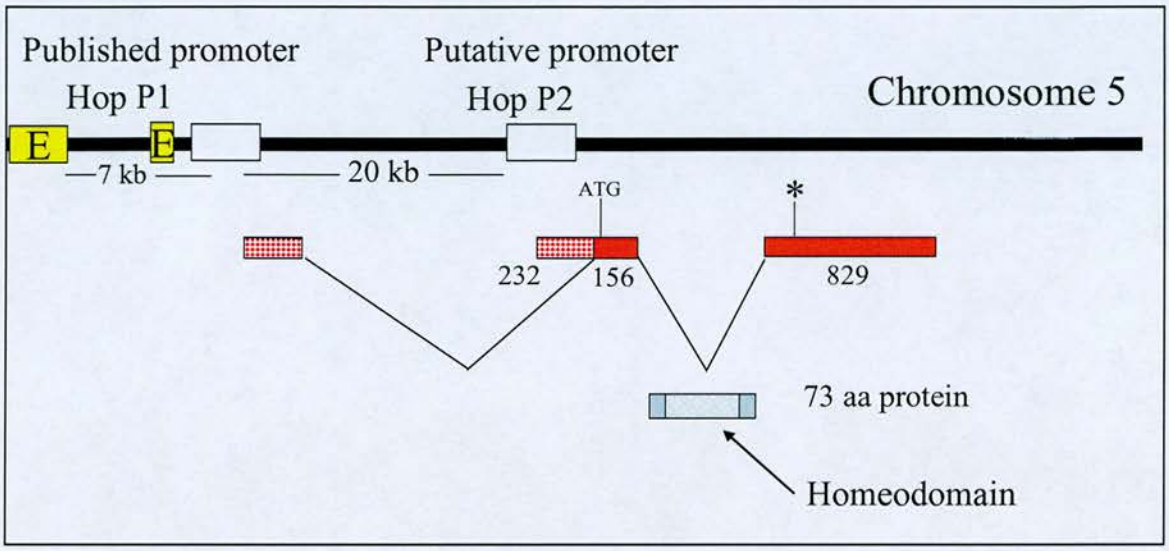


B)

Genotype	n	median survival (d)	(95% conf. interval)	
3b+Mbd2-/-	5	128	107	142
3b+Mbd2+/-	11	113	110	135
3b+Mbd2+/+	26	127	116	136

Figure 4.6 Effect of Mbd2 deficiency on survival in *Tg(Dnmt3b)+ mice.* A) Survival curves and B) median survival data for *Tg(Dnmt3b)+* x *Mbd2*$^{-/-}$ mice, classified by *Mbd2* genotype. 3b+ indicates *Tg(Dnmt3b)+*. Significance value quoted in figure 4.6a is based on null hypothesis that the survival curves are the same.

A) *Hop* promoter structure



B) Hop expression

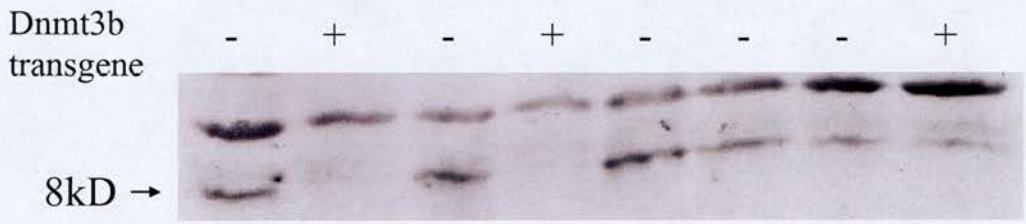


Figure 4.7 *Hop* expression and promoter structure. A) promoter structure of *Hop* B) Expression of Hop (8Kda band) by western blot in hearts from *Tg(Dnmt3b)*+ and wild type mice.

Fig 4.8 (overleaf) A) Methylation status of P1 promoter in wild type and *Tg(Dnmt3b)*+ mice by bisulphite sequencing. B) Methylation status of individual CpG nucleotides in the 3' *Hop* promoter (HopP2) by bisulphite sequencing. The number of clones examined varies at individual CpG dinucleotides since clones where sequence was unclear at a given CpG were excluded from the analysis. Data are from F3 mice.

A) P1 promoter

genotype	clone #	CpG=distance from start of PCR product 38 65 69 94 100 191 205 207 214 262 277 353 362	conversion%
Tg(Dnmt3b)+	01	?	96
	02		100
	03		100
	05		100
	06		100
	07		100
	08		100
	09	- -	96
	10	n ? n	97
Tg(Dnmt3b)-	22		98
	23		100
	24		96
	25		96
	26		96
	27		100
	28		100
	29		100
	30		100

CpG methylation
n ambiguous sequence
? Sequence does not correspond to reference
- Unreadable sequence

B) P2 promoter

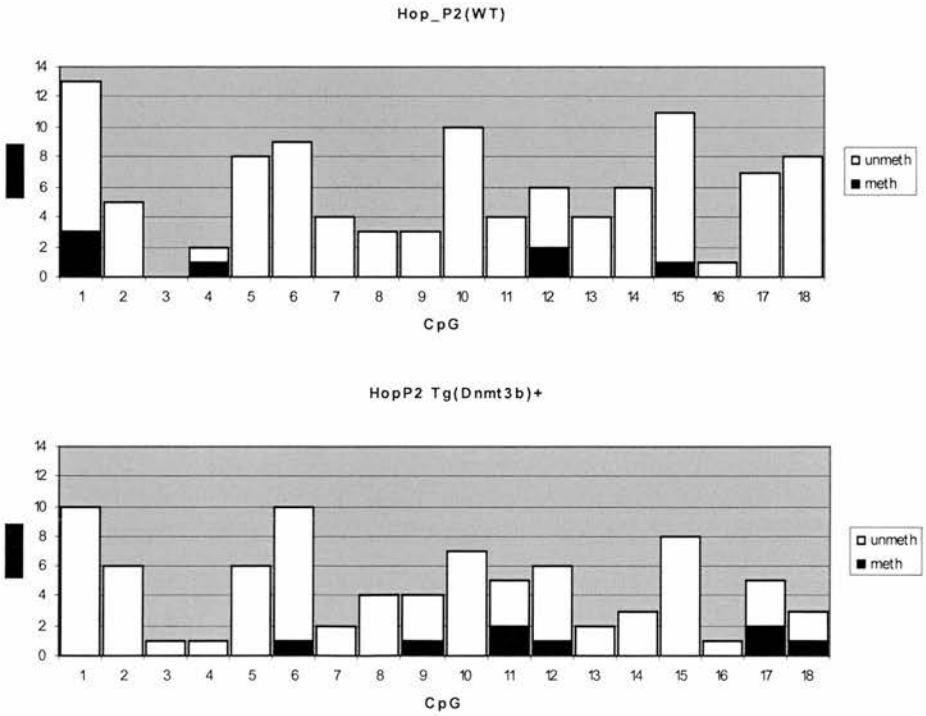


Figure 4.8 *Hop* promoter methylation in heart. (Legend on preceding page)

5. Effects of Dnmt3b over-expression on cancer

5.1 Introduction

In this chapter the effects of Dnmt3b over-expression on tumour formation are examined. Three complementary models are studied: firstly the spontaneous incidence of tumours in *Tg(Dnmt3b)*+ mice, secondly the effects of Dnmt3b over-expression on multi-stage skin carcinogenesis, and finally the effects of Dnmt3b over-expression in $Apc^{+/min}$ mice, a hereditary cancer predisposition model in which the mechanism of tumour formation is well defined. One of the difficulties of investigating tumour development in *Tg(Dnmt3b)*+ mice is that their reduced lifespan may not give sufficient time for tumours to form. We have therefore chosen the two induced cancer models such that tumours develop with short latency, allowing any effect to be seen within the lifespan of the *Tg(Dnmt3b)*+ mice.

Multi stage skin carcinogenesis is a well established model system for studying the effect of factors which influence susceptibility to cancer. Mice are treated topically with a carcinogen, 7,12-dimethylbenzanthracene, (DMBA) which causes DNA damage, in what is referred to as the initiation phase of treatment. This is followed by regular topical treatment with a mitogen 12-O-tetradecanoylphorbol-13-acetate (TPA) which is termed the promotion phase of treatment. There are several different treatment regiments which have different periods of tumour latency. Using an aggressive regimen of 2 initiation treatments 6 weeks apart and twice weekly promotion treatments, animals develop tumours at around 10 weeks after starting treatment (Owens, *et al* 1999). The mice initially develop benign squamous-cell papillomas, which progress to form malignant squamous cell carcinomas. A number of molecular abnormalities have been described in these tumours, especially in the transformation to malignant carcinoma, but a common feature which occurs in almost all papillomas is the development of mutations in Ha-*ras* (Zoumpourlis, *et al* 2003) for review. In addition, the archetypal

changes in DNA methylation have been described in papillomas produced by this regimen, with decreases in global levels of methylation, aberrant methylation of a number of tumour suppressor genes, and increases in Dnmt1 and Dnmt3b (Fraga, *et al* 2004). The authors of this study imply that levels of DNA methyltransferases may be correlated with tumour aggressiveness, since the highest levels of methyltransferase activity are found in the cells with the most malignant behaviour. The system therefore would appear to meet our requirements for an induced carcinogenesis model of short tumour latency and a possible role for methylation in modulating tumour behaviour, and also has the advantage of a simple outcome measure (counting skin tumours).

As an inherited tumour model, the $Apc^{+/min}$ mouse has been chosen for study, partly in view of its short tumour latency, but also since there is substantial evidence for the importance of DNA methyltransferases in its pathogenesis (see Section 1.6). The Apc protein binds directly to β catenin and promotes its phosphorylation and subsequent degradation by the proteasome (For review see Brembeck, *et al* 2006). In the absence of Apc, β catenin levels increase and leads to abnormal accumulation in the nucleus in addition to its normal expression at the cell membrane. Nuclear β catenin acts as a transcription factor, and is the main downstream mediator of Wnt signalling. The aberrant expression of β catenin in the nucleus in the absence of Apc thus mimics Wnt signalling, promoting the transcription of proliferation and survival associated genes including *c-myc* and *Cyclin D1,* which can promote malignancy. In humans, mutations in the *APC* gene can be found in 60% (Powell, *et al* 1992) of colorectal tumours; in addition inherited mutations is the *APC* gene lead to the syndrome of familial adenomatous polyposis coli, in which subjects develop multiple colonic adenomas early in life which subsequently progress to carcinoma (OMIM reference +175100; http://www.ncbi.nlm.nih.gov/entrez/dispomim.cgi?cmd=entry&id=175100).

The mouse *Apc* gene is transcribed from two known promoters (Karagianni, *et al* 2005). The first promoter is immediately upstream of exon 1 and is not formally a CpG island promoter although it contains a 50 bp segment which meets all CpG island criteria other

than length. This promoter is the predominant promoter in the gut. A second more recently discovered promoter is the predominant promoter in ES cells. This promoter region is 40 kb upstream of the initiating ATG, is a CpG island promoter, and leads to transcripts with the addition of a 55 nt untranslated sequence at the 5' end. The internal promoter has been shown to be hypermethylated in human tumours; however there are differences in the human promoter structure, most importantly that the human internal promoter includes a CpG island. Loss of Apc in the intestine in an inducible transgenic mouse model (Sansom, *et al* 2004) leads to rapid (within 5 days) increases in proliferation, alterations in cell migratory behaviour and up-regulation of *Wnt* target genes.

As mentioned in Section 1.6 , $Apc^{+/min}$ mice are heterozygous for a mutation in the *Apc* gene which leads to premature termination at amino acid 850 (Su, *et al* 1992). This mutation arose as part of a chemical mutagenesis experiment. The $Apc^{min/min}$ phenotype is embryonically lethal, however $Apc^{+/min}$ mice develop multiple adenomas, predominantly in the small intestine with smaller numbers of colonic tumours. This leads to rectal bleeding and anaemia, normally causing death at around 7 months of age. Tumour development proceeds from the development of microadenomas, which can be detected from as early as 5 weeks of age, with subsequent development of adenomatous polyps. Analogous to the human condition, these adenomas also progress to form invasive adenocarcinomas. Loss of *Apc* appears both necessary and sufficient for microadenoma formation in $Apc^{+/min}$ mice (Yamada, *et al* 2005), although evolution of polyps is known to be modulated by other loci (Gould, *et al* 1996, Silverman, *et al* 2002a).

5.2 Over-expression of Dnmt3b does not cause spontaneous tumours

The most simplistic hypothesis for the role of DNA methylation in carcinogenesis is that abnormal methylation events occur essentially at random. If these random errors happen to involve tumour suppressor genes, this could cause their silencing and lead to cancer.

According to this hypothesis, it would be expected that expressing a *de novo* methyltransferase in cells in which it is not normally expressed would enable the development of abnormal events, and lead to an increased incidence of cancer.

There have been no macroscopically detectable tumours in any of the *Tg(Dnmt3b)+* mice so far in any of the first 220 observed, based on gross post mortem dissection of animals that had died spontaneously and were found sufficiently soon after death not to have decomposed, and the lack of spontaneous externally visible tumours in any animals. In addition there was no evidence of microscopic tumours in histology from two *Tg(Dnmt3b)+* F1 animals examined at 139 days, or in any other animals which have had had tissues taken for histology for other reasons. It can be concluded that any effect of Dnmt3b over-expression on spontaneous tumour formation is absent or extremely small.

It could validly be argued that an increase in cancer predisposition is not detected in these animals because of their reduced lifespan due to non-malignant disease, and that spontaneous cancer formation may have been seen if the animals had a normal lifespan. While this argument cannot be refuted, it can partially be addressed by the chimaeric founder animals which had a normal lifespan. Out of 10 chimaeric animals, there was one stomach tumour occurring at approximately 2 years of age. This level of incidence is not significantly different to the published lifetime cancer incidence of 7% in 129 strain mice (Festing 1979). The findings support the hypothesis that the effect of Dnmt3b on spontaneous tumour formation is small, if present at all.

5.3 Aberrant methylation of known tumour suppressor genes is detectable in *Tg(Dnmt3b)+* mice

In order to address the possibility that no cancers had occurred in the *Tg(Dnmt3b)+* mice because no genes relevant to cancer had become methylated, the methylation status of a number of known tumour suppressor genes were examined by M-QPR. A candidate

gene approach was employed, concentrating on genes that have been shown to be methylated in tumours from mice, as there are almost certainly differences in the occurrence of methylation in tumours between mice and humans, since humans have more CpG island genes than mice.

Figures 5.1-5.4 show the distribution of abnormal methylation in tissues from wild type and transgenic animals (F8 generation) of a number of known tumour suppressor genes (*Hic1*- 5' promoter, *Cdkn1b* (p21), *Cdkn2a* (p19) and *Mlh-1*), as measured by M-QPR. The salient points of these data are as follows. There is clear evidence of substantially increased methylation of hic1 in heart and kidney in the transgenic mice. There is also a consistent increase in *Cdkn2a* (p19) and *Cdkn1b* methylation in heart and kidney though at a lower level, possibly indicating methylation of a sub population of cells in the tissue. There is no convincing difference between methylation patterns for *Mlh-1*, with methylation detectable in both wild type and transgenic mice. In addition, low levels of *p19* and *Hic1* methylation appear to be detectable in tissues from wild type mice although there is variation between samples.

The results indicate that an increased level of a *de novo* methyltransferase is a sufficient condition to cause dense methylation of known tumour suppressor genes, at least at low level. However this has not translated in to an increased spontaneous tumour incidence. The functional significance of such low level methylation is uncertain, since it is not clear whether this relates to methylation of a minority subpopulation of cells or to methylation occurring in a small proportion of cells from the predominant cell type. It is also unclear whether the methylation observed with this technique is mono-allelic or bi-allelic.

5.4 Sensitivity to skin carcinogenesis is not greatly increased

In order to investigate whether Dnmt3b over-expression causes increased susceptibility to carcinogenesis, A preliminary study using the model of multistage skin

carcinogenesis (Owens, *et al* 1999) was conducted. 4 transgenic and 6 non-transgenic littermate mice (F7 generation) were exposed topically to the carcinogen dimethylbenzanthracene (DMBA) at 0, 2, 4 weeks with promotion using the mitogen 2-o-tetradecanoylphorbol-13-acetate (TPA) applied topically twice weekly for 10 weeks. Treatment started at 4-6 weeks of age. No macroscopic skin tumours were seen in any of the treated transgenic mice. Transgenic mice became unwell at the expected age, with autopsy studies revealing cardiomegaly as expected but no gross tumour formation. The remaining four control animals were observed to identify the background rate of tumour formation. One control animal developed a skin tumour 14 weeks after the start of treatment. The remaining animals were sacrificed at 240 days and remained tumour free.

The results are not conclusive, due to the lower than expected incidence of tumours in the control animals. This may be due to the relative resistance of the C57BL6 strain of mice to multistage carcinogenesis (Reiners, *et al* 1984). In addition, the reproducibility of this method of dosing is somewhat uncertain, since there is likely to be much variation in the amount of carcinogen actually applied to the skin in each topical treatment. Thus there was no attempt to repeat the experiment with larger numbers or alternative tumour promoting agents. Nevertheless, the experiments suggest that susceptibility to tumour formation caused by multistage chemical carcinogenesis is not massively increased by Dnmt3b over-expression.

5.5 Tumour progression, but not *Apc* methylation, is increased in $Apc^{+/min}$ mice crossed with *Tg(Dnmt3b)*+ mice

The most direct evidence for a causative role of methylation in the development of cancer would be to demonstrate the methylation of tumour suppressor genes in a model in which the mechanism of tumour formation by silencing of these genes is well defined. *Tg(Dnmt3b)*+ mice (from F8 to F11 generations) were therefore crossed with $Apc^{+/min}$ mice to produce mice heterozygous for both mutations (see Section 7.2.5).

Apc$^{+/min}$, *Tg(Dnmt3b)+* and *Apc*$^{+/min}$, *Tg(Dnmt3b)-* mice were sacrificed at 125 days and tumour burden was estimated by taking transverse sections at approximately 30 levels through the small intestine and calculating the mean number of tumours per section- referred to as mean tumour index (MTI) (see Section 7.2.7). This statistic cannot be viewed as a simple estimator of total tumour number, and changes in MTI could result from changes in either number or tumour size distribution.

A total of 7 *Apc*$^{+/ min}$, *Tg(Dnmt3b)+* mice and 17 *Apc*$^{+/min}$, *Tg(Dnmt3b)-* mice survived to 125 days. The number of *Tg(Dnmt3b)+* mice is smaller since some mice died before 125 days as a result of cardiac disease. As expected, multiple microadenomas, adenomatous polyps and adenocarcinomas were seen in intestines of both groups of animals (Figure 5.5 for examples of histology). Figure 5.6 shows the mean tumour index for each animal studied, classified by genotype. In both *Tg(Dnmt3b)+* and *Tg(Dnmt3b)-* animals, there appear to be a small number of animals with a particularly high tumour burden with the majority of animals having a smaller tumour burden. The size of the effect of Dnmt3b over-expression on tumour number was estimated by calculating the ratio of the median MTIs between the two groups. Based on this statistic, the tumour burden was modestly increased in *Tg(Dnmt3b)+* mice but this was not statistically significant, with a ratio of MTI *Tg(Dnmt3b)+ / Tg(Dnmt3b)-* of 1.3 (95% confidence interval 0.84 to 2.68).

There are regional differences ($p<0.0001$ by Kruskal-Wallis test) within the gut as regards tumour burden, with an increased tumour burden in the distal gut regardless of *Tg(Dnmt3b)* genotype. This distal predominance is significantly more marked in *Tg(Dnmt3b)+* animals (Figure 5.8). Hence there may be a difference in the distribution of tumours between *Tg(Dnmt3b)-* and *Tg(Dnmt3b)+* animals, either as a result of increased tumour initiation or growth which is confined to the distal gut. Since the effect seems to amplify the normal distal predominance, it could be argued that it is most likely that Dnmt3b is interacting with pre-existing condition that give rise to tumours rather than acting as an independent contributor to tumour formation.

In order to study the effects of Dnmt3b over-expression on tumour evolution from benign adenoma to malignant adenocarcinoma, polyps were also classified according to the presence or absence features of malignant transformation- the presence of invasion or dysplasia (Figure 5.7). The proportion of dysplastic polyps was increased in *Tg(Dnmt3b)+* with a relative risk of dysplasia (i.e. probability of *Tg(Dnmt3b)+* polyp being dysplastic / probability of *Tg(Dnmt3b)-* polyp being dysplastic) of 1.52 ($p<0.003$ by Fisher's exact test, 95% confidence interval 1.2-2.0). The relative risk of invasion was also increased but not significantly ($p=0.54$ by Fisher's exact test, relative risk 1.51, 95% confidence interval 0.60-3.8).

To investigate whether any differences in tumour numbers between $Apc^{+/min}$, *Tg(Dnmt3b)+* and $Apc^{+/min}$, *Tg(Dnmt3b)-* mice may be mediated through methylation at the *Apc* locus, the promoter methylation status of the *Apc* gene was investigated. To examine the methylation status of the Apc promoters in normal gut tissue, DNA from whole gut from both $Apc^{+/+}$, *Tg (Dnmt3b)-* and $Apc^{+/+}$, *Tg(Dnmt3b)+* mice was examined by bisulphite sequencing. No evidence of methylation was found in either promoter in either *Tg(Dnmt3b)+* or *Tg(Dnmt3b)-* gut samples (Figure 5.9). In addition the methylation status of the external CpG island promoter was examined by M-QPR in order to detect low levels of methylation. The level of methylation in normal gut is very small if present (of the order of 0.1%), and there was no significant difference between the *Tg(Dnmt3b)-* and *Tg(Dnmt3b)+* animals. It can be concluded that the *Apc* gene is not susceptible to spontaneous methylation as a consequence of Dnmt3b over-expression.

To examine the *Apc* methylation status in tumours, individual tumours were microdissected from the paraffin embedded gut samples. DNA was extracted from these, and pooled DNA from 20 $Apc^{+/min}$, *Tg(Dnmt3b)+* tumours and 20 $Apc^{+/min}$, *Tg(Dnmt3b)-* tumours was analysed for methylation by bisulphite sequencing.

From the internal promoter, good quality sequence was obtained from 22 clones from *Apc*$^{+/min}$, *Tg(Dnmt3b)*+ tumours and 20 clones from *Apc*$^{+/min}$, *Tg(Dnmt3b)*- tumours. The results show dense methylation of several adjacent CpG residues in 4 clones from both wild type and transgenic animals. In addition, sporadic methylation of single CpG residues in 3 clones from the *Tg(Dnmt3b)*+ tumours and 1 clone from *Tg(Dnmt3b)*- tumours (Figure 5.10) was present. From the external promoter, there was no evidence of methylation in 20 clones from from *Apc*$^{+/min}$, *Tg(Dnmt3b)*- mice (Figure 5.11). There was no evidence of dense methylation in 21 clones from *Apc*$^{+/min}$, *Tg(Dnmt3b)*+ tumours. However 7 clones showed methylation at individual CpG residues.

It can be concluded that the internal *Apc* promoter is densely methylated with detectable frequency in intestinal tumours from *Apc*$^{+/min}$ animals regardless of Tg(Dnmt3b) expression. However there is no evidence for any increase in the frequency of dense methylation caused by Dnmt3b over-expression. Since *Apc* inactivation is a necessary condition for polyp formation, it follows that even if the small changes in tumour burden noted above were genuine, they must be caused not by increased tumour initiation as a result of methylation of the wild type *Apc* allele, but by effects on tumour growth mediated at other loci.

To investigate this point further, two approaches were taken to examine whether there are differences in silencing at the *Apc* locus. Firstly, DNA from individual tumours was also assayed for loss of heterozygosity at the *Apc* locus by PCR and *HindIII* digestion. Secondly, immunohistochemical staining was performed for beta-catenin as a surrogate marker of Apc inactivation.

Results of the loss of heterozygosity analysis for 8 tumours from each genotype are shown in Figure 5.12. The most important finding from this experiment is that DNA of sufficient quality to be successfully amplified by PCR has been successfully extracted from the majority of microdissected tumours, and so the bisulphite sequencing data is likely to be representative of several independent tumours. The analysis shows that the

microdissection has been successful in enriching for tumour DNA since samples from the tumours show a much stronger *min* band than wild type band. It is not possible to determine from this analysis whether the presence of a weaker wild type band indicates that the dissection of tumours has included more normal tissue or whether there is genuinely some loss of heterozygosity at the *Apc* locus in a small proportion of cells within the tumour. An independent observer, blinded to the genotype of the mice was unable to find any difference in nuclear beta-catenin staining between *Tg(Dnmt3b)*+ or *Tg(Dnmt3b)*- sections (Figure 5.13), viewing 5-6 sections from distal intestine from 5 mice in each group. This is in keeping with the hypothesis that there is no increase in silencing of the *Apc* locus in *Tg(Dnmt3b)*+ mice.

The increase in sporadic methylation in both promoter regions raises the possibility of an effect of Dnmt3b on causing sporadic methylation events. It is not possible to discount the possibility of individual events being artefacts as a result of incomplete bisulphite modification. However the observation that conversion rates are not appreciably different between the two groups would argue against this explaining the differences in the experiment as a whole. Importantly the increase in sporadic methylation appears to have occurred in the tumour samples and not the samples from histologically normal $Apc^{+/+}$ tissue. This again supports the hypothesis that changes in the malignant cell other than increases in methyltransferase level provide a milleu which favours the development of aberrant methylation. Again differences in bisulphite conversion between paraffin embedded tissues and fresh prepared tissues cannot be excluded, but if anything the conversion rates are better in the tumour samples.

To summarise, there is no evidence that over-expression of Dnmt3b causes tumours in $Apc^{+/min}$ mice by methylating the wild type *Apc* allele. However, on detailed analysis, Dnmt3b over-expression does appear to alter aspects of tumour behaviour such as the distribution of tumours and the rate of progression to dysplastic carcinoma in the context of a genetically tumour prone background.

5.6 Discussion

The results in this chapter show that Dnmt3b over-expression does not lead to a large increase in tumour susceptibility. This is despite the fact that abnormal methylation of known tumour suppressor genes can be detected, and the fact that the Dnmt3b levels achieved in *Tg(Dnmt3b)*+ mice are sufficient to cause both changes in methylation and phenotypic changes. Nevertheless, Dnmt3b over-expression does appear to influence tumour biology in a well defined tumour prone model, although this does not appear to be mediated by direct silencing of the locus which causes cancer.

One possible explanation for the lack of a strong cancer phenotype might be that in the transgenic animals, Dnmt3b is not expressed in cells which give rise to cancer. While this argument cannot be directly refuted, because attempts to investigate the detailed tissue distribution of Dnmt3b expression of Dnmt3b by immunohistochemistry on paraffin sections were unsuccessful (data not shown), data from other models where the transgene is expressed under control of the CAG promoter have shown strong expression in relevant cells (Wiekowski, *et al* 2001). Specifically, there is strong expression in the intestinal epithelium, including the intestinal crypts in which tissue stem cells reside and microadenoma formation is thought to occur in $Apc^{+/min}$ mice.

What does this lack of a clear cancer phenotype tell us about the role of methylation and methyltransferases in cancer? The results argue strongly against the possibility that dysregulation of methyltransferases can act as the primary initiating events in cancer. This is perhaps not very illuminating, since there are no known reports in cancers of translocations or mutations causing overactivity of *de novo* methyltransferases. However given the appearance of abnormal methylation events provoked by non-malignant phenomena such as aging (see Section 1.6), *de novo* methyltransferase activity must be induced under certain conditions in adult life. Nevertheless, the fact that no definite effect on tumour induction is seen even despite expressing Dnmt3b to supra-physiological levels would tend to suggest that aberrant DNA methylation as the

initial event in gene silencing is likely to be very rare at physiological levels of Dnmt3b expression.

Since silencing of tumour suppressor genes by methylation is the most plausible mechanism for a role of DNA methylation in causing cancer, the finding of abnormal tumour suppressor gene methylation in *Tg(Dnmt3b)+* mice is instructive. It shows that, in principle, dysregulation of a *de novo* methyltransferase is sufficient to cause methylation of tumour suppressor genes. However, the proportion of cells within the whole tissue which are susceptible to methylation is very small. The factors which allow methylation in the context of methyltransferase over-expression remain undefined; it is unclear whether the low level of methylation relates to methylation occurring in a minority cell type within the tissue, or whether methylation is occurring in a subpopulation of cells which have other factors permissive of methylation. The functional significance of such low levels of methylation also is unclear. Clearly methylation of tumour suppressor genes has not translated into an increased spontaneous cancer incidence. There are many possible explanations for this. Firstly it is not known from our PCR based assay whether methylation is mono-or bi-allelic. It would not be surprising if mono-allelic methylation has no functionally detectable consequences in the absence of genomic deletion. Secondly, the effects of bi-allelic loss of a single tumour suppressor gene are often subtle. Mice with null mutations of many of the genes studied here suffer only small increases in tumour susceptibility (Chen, *et al* 2003b, Deng, *et al* 1995, Sharpless, *et al* 2004), and so it is not surprising that of methylation of these genes in small numbers of cells has not led to spontaneous tumours. Finally, and more radically, it could be argued that the methylation seen in *Tg(Dnmt3b)+* mice has only occurred in cells in which it is irrelevant, for example if the gene is already silenced. If this were the case, it would support the hypothesis that methylation events seen in tumours and cell lines may not necessarily reflect functional changes in gene expression.

Nevertheless, the detection of methylation shows that under certain (undefined) circumstances, tumour suppressor genes which have been found to be commonly methylated in cancer are able to be methylated simply by inducing *de novo* methyltransferase activity. Since the genes chosen for study here were selected using a candidate gene approach rather than as a result of a screening assay, it raises the possibility that such genes are methylated in cancer since they are especially prone to methylation. This hypothesis could imply that these genes may be a common factor in the pathogenesis of many different types of cancer. On the other hand, it could also be taken to indicate that these genes are especially prone to methylation as a bystander effect of the malignant process, whether this has any functional consequence or not. In addition, the lack of spontaneous tumour development argues that there has not been widespread silencing of tumour suppressor genes, as there are many genes (including the *Apc* gene) which when deleted give rise to an aggressive tumour prone phenotype which would be seen within the lifespan of the *Tg(Dnmt3b)+* mice. Taken together, these results suggest that although tumour suppressor genes which are methylatable are not uncommon, the silencing of an active gene in a functionally important manner is much rarer.

How do our findings tie in with other experiments on the role of methyltransferases in the development and progression of tumours in $Apc^{+/min}$ mice? The most important study to compare with is the inducible deletion of *Dnmt3b* in the intestine (Lin et al 2006). To recap, this study showed that Dnmt3b is upregulated in established macroadenomas from $Apc^{+/min}$ mice. Conditional deletion of *Dnmt3b* in the intestine was achieved by producing mice with loxP sites flanking exons 16 to 19 of the *Dnmt3b* gene ($Dnmt3b^{2lox/2lox}$), and breeding these mice with transgenic mice with Cre recombinase under the control of transcriptional regulatory elements from the fatty acid binding protein gene (*Fabp*), which is intestine specific. Comparing ($Apc^{+/min}$, $Dnmt3b^{2lox/2lox}$, $Fabp^{Cre+}$) and ($Apc^{+/min}$, $Dnmt3b^{2lox/2lox}$, $Fabp^{Cre-}$) mice, the authors found a 40% reduction in the number of macroscopic colonic tumours, with no difference seen in the small intestine, possibly as a result of low recombination rates in

the small intestine. No differences in the number microadenomas, as defined by abnormal β-catenin staining were found. Dnmt3b was not essential for tumour growth, since tumours not expressing Dnmt3b protein were found. However *Dnmt3b* expression did seem to provide a selective advantage since in the colon, a higher proportion of established macroadenomas expressed *Dnmt3b*, whereas approximately 50% of microadenomas expressed Dnmt3b. The authors concluded that the effect of *Dnmt3b* deletion must act to prevent the transition between microadenomas and macroscopic adenomas, since there was no difference in tumour volume between ($Apc^{+/min}$, $Dnmt3b^{2lox/2lox}$, $Fabp^{Cre+}$) and ($Apc^{+/min}$, $Dnmt3b^{2lox/2lox}$, $Fabp^{Cre-}$) mice. This is consistent with the findings of our study suggesting that Dnmt3b over-expression appears to have no effect on tumour initiation at the *Apc* locus, but appears to affect subsequent tumour behaviour.

A key difference between the induction of Dnmt3b seen in the above study and in *Tg(Dnmt3b)+* mice is that the increase in Dnmt3b in *Tg(Dnmt3b)+* mice is present in the normal tissue rather than being increased only in tumours. Therefore, one might not expect to add anything to the effects of Dnmt3b on macroadenoma formation outlined above, since the enzyme becomes upregulated in polyps anyway. Similarly, the fact that dense *Apc* methylation levels are found in both $Apc^{+/min}$, *Tg(Dnmt3b)+* and $Apc^{+/min}$, *Tg(Dnmt3b)-* animals, but that there is no great difference between methylation frequencies between the two groups suggests that Apc methylation is a late event, occurring after tumour formation and the induction of Dnmt3b. Furthermore since dense *Apc* methylation is not detectable in normal mucosa from *Tg(Dnmt3b)+* mice, it seems that tumour formation is the necessary condition rather than *de novo* methyltransferase expression in determining *Apc* methylation. Since *Apc* inactivation is necessary for the formation of intestinal tumours in $Apc^{+/min}$ mice, this would imply that dense methylation of *Apc* promoters is occurring in a functionally inactive gene (either methylation of the mutated allele, or an allele previously silenced by other mechanism such as histone modifications). This suggests that any effect of Dnmt3b over-expression on tumour number is mediated at loci other than the *Apc* locus. Since *Apc* inactivation is

widely accepted as an early event in tumour formation, this hypothesis is consistent with the published findings that Dnmt3b deletion affects polyp formation but not microadenoma formation.

To develop this point further, since Dnmt3b is increased in tumours in $Apc^{+/min}$ mice without the necessity for it being over-expressed by a trangene, it follows that differences between $Apc^{+/min}$, *Tg(Dnmt3b)+* and $Apc^{+/min}$, *Tg(Dnmt3b)-* are likely to arise from methylation events (or other effects of Dnmt3b) occurring prior to tumour formation. In this model, over-expression of Dnmt3b has led to methylation events in histologically normal cells, which does not disturb the behaviour of the cell, perhaps because it has occurred in genes which are silent in the normal state. Upon inactivation of the *Apc* gene, for example by mutational loss of the wild type allele, the cell undergoes malignant transformation. In the *Tg(Dnmt3b)+* mice, genes (for example tumour suppressor genes) which would normally become reactivated in response to transformation may have been methylated, thus preventing their re-expression. This may alter the natural history of the tumours, and although it would have no effect on tumour initiation it could affect their subsequent progression. The findings of low- level tumour suppressor gene methylation in histologically normal tissue may be relevant in this regard, promoting tumour growth once tumours are initiated by other means such as genomic deletion.

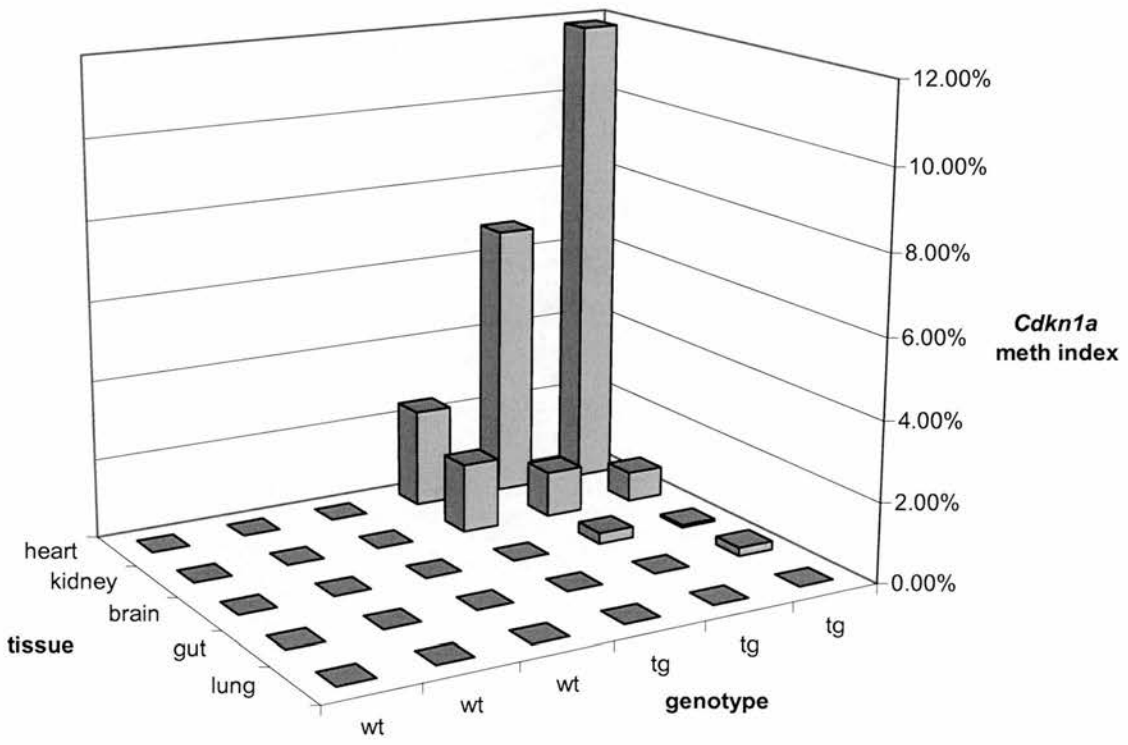
A problem arising from this hypothesis, which also underpins the whole field of gene silencing in cancer, is why increased methylation should increase rather than decrease cancer. There is no *prima facie* reason why silencing should not affect genes which promote cancer such as oncogenes; in fact it might be argued, given the role of methylation in retrovirus silencing, that this would be more likely. The finding of methylation of *Wnt3,* a putative oncogene, in *Tg(Dnmt3b)+* mice (see Section 3.4) also demonstrates that oncogene methylation can indeed occur. The problem is especially critical in the *Apc* model, since the immediate effect of *Apc* inactivation is the promotion of transcription of β catenin target genes, and if these genes had been

methylated it might have protected against cancer. The obvious answer is that it is easy to find examples where methylation events have promoted cancer progression since proliferation will be favoured, whereas it would be difficult to find (and much harder to prove) cells that would have become malignant had certain genes not been silenced, unless this phenomenon was very common. An interesting possibility is that both effects occur and counterbalance each other, which is why the differences seen in *Tg(Dnmt3b)+* mice are fairly subtle, and also might explain the paradox that loss of maintenance methylation increases some tumours but decreases others (Yamada, *et al* 2005). A more radical explanation is either that tumour suppressor genes are intrinsically more prone to silencing than oncogenes or that the set of potentially methylatable genes includes more tumour suppressor genes than oncogenes, thus tipping the balance in favour of oncogenesis in a methylation-prone environment.

The idea that methylation events in tumours occurs both prior to, and as a consequence of an oncogenic insult, with pre-existing methylation patterns altering subsequent tumour behaviour, has an interesting consequence. It indicates that the set of genes methylated in a given tumour may be divided into those methylated subsequent to malignant transformation, which may be common to tumours of similar origin, and those which precede malignant transformation, which may differ substantially between different tumours. In theory, this might make it possible to classify tumours into those with pre-transformation methylation and those in which methylation changes are simply secondary to malignancy. This would have interesting and in some ways paradoxical therapeutic implications. Demethylating agent therapy would be more likely to be an independent contributor to a treatment regimen in tumours where methylation abnormalities are not simply secondary to the malignant process. Secondly, it poses a diagnostic problem since the search for genes commonly methylated in given cancers is most likely to discover those which are secondary to the cancer, whereas what would be more useful is to discover those which differ between apparently similar cancers and thus may have discriminating value. This may not be feasible in clinical practice, unless

there are genes that offer discriminating value but are sufficiently frequently methylated to warrant measurement in routine practice.

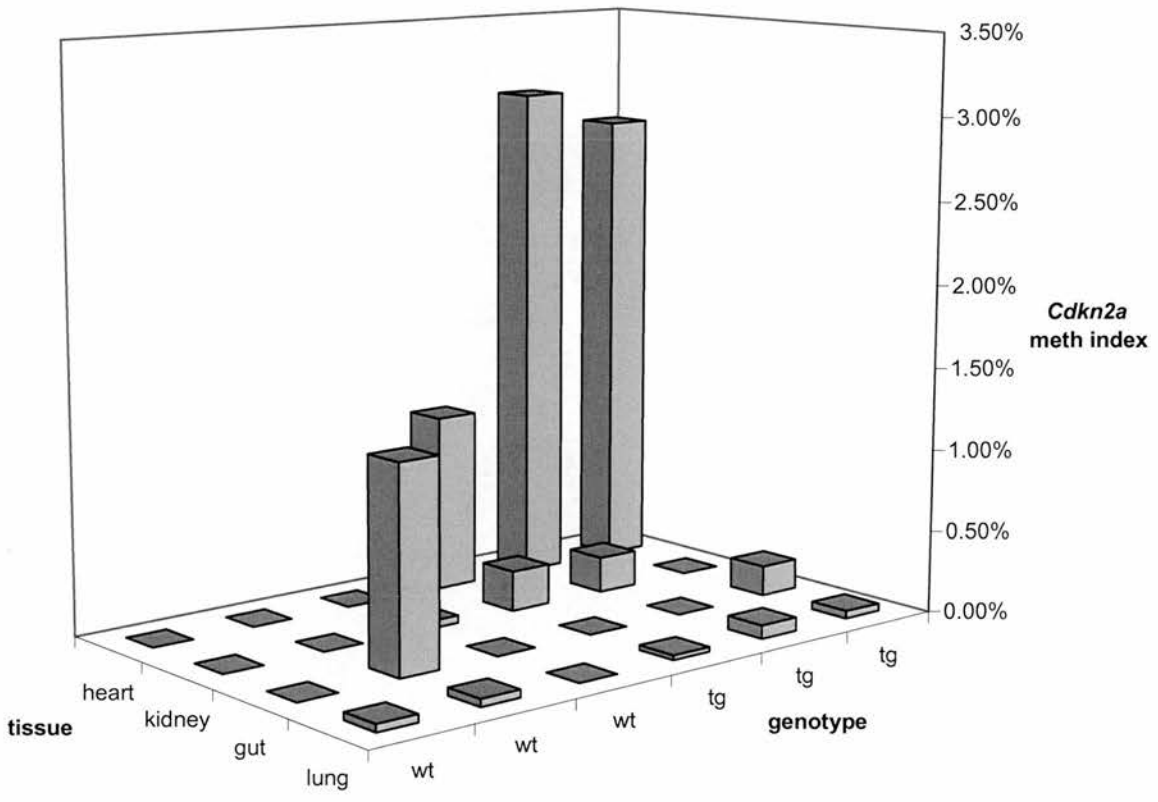
5.7 Figures and Tables



	Wild Type			Transgenic		
	wt1	wt2	wt3	3b1	3b2	3b3
Brain	0	0	0	0	0.27%	0.06%
					(0.19-0.38%)	(0.05-0.08%)
Gut	0	0	0	0	0	0.21%
						(0.20-0.22%)
Heart	0	0	0	2.5%	6.9%	12%
				(2.2-2.7%)	(6.0-7.8%)	(11-13%)
Kidney	0	0	0	1.73%	1.14%	0.74%
				(1.1-2.6%)	(0.64-2.0%)	(0.30-1.9%)
Lung	0	0	0	0	0	0

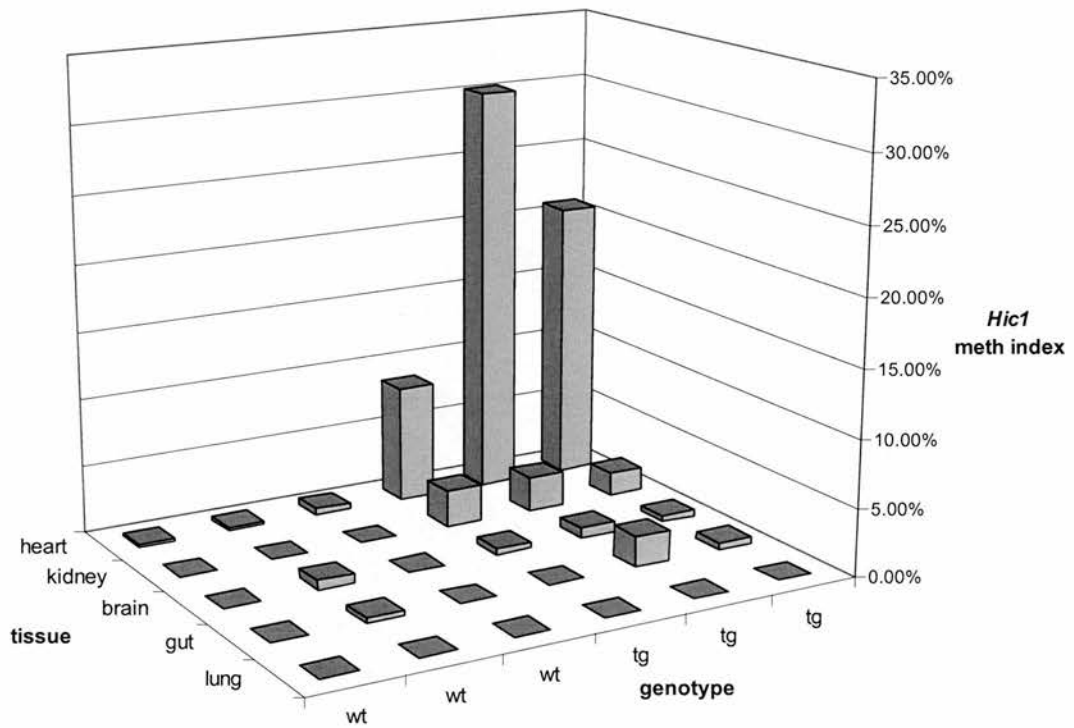
Figure 5.1 *Cdkn1a* methylation

Figures 5.1-5.3 Methylation of 5.1) *Cdkn1a* (p21cip1), 5.2) *Cdkn2a* (P19arf) and 5.3) *Hic1* (5' promoter) as measured by M-QPR. Tables show methylation index in various tissues (range +/- transformed standard error).
wt= wild type. tg =transgenic *Tg(Dnmt3b)+*. Data are from F8 mice.



	Wild Type			Transgenic		
	wt1	wt2	wt3	3b1	3b2	3b3
Gut	0.00%	1.24%	0.00%	0.00%	0.00%	0.18%
		(0.53%-2.91%)				(0.12%-0.28%)
Kidney	0.00%	0.00%	0.05%	0.24%	0.22%	0.00%
			(0.01%-0.17%)	(0.20%-0.29%)	(0.13%-0.36%)	0.00%
Lung	0.05%	0.05%	0.00%	0.03%	0.08%	0.05%
	(0.03%-0.08%)	(0.03%-0.09%)		(0.02%-0.05%)	(0.05%-0.13%)	(0.04%-0.05%)
Heart	0.00%	0.00%	0.00%	1.08%	3.03%	2.82%
				(0.79%-1.47%)	(2.41-3.82%)	(1.84-4.33%)

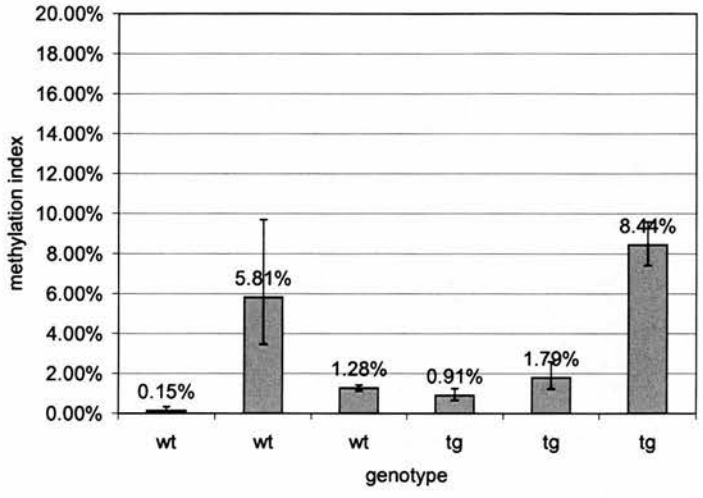
Figure 5.2 *Cdkn2a* methylation (Legend with Figure 5.1)



	Wild type			Transgenic		
	wt1	wt2	wt3	3b1	3b2	3b3
Brain	**0**	**0.69%**	**0**	**0.47%**	**0.76%**	**0.44%**
		(0.51%-0.95%)		(0.37%-0.59%)	(0.64%-0.92%)	(0.37%-0.51%)
Gut	**0**	**0.34%**	**0**	**0**	**2.23%**	**0.49%**
		(0.25%-0.47%)			(0.76%-6.6%)	(0.38%-0.63%)
Heart	**0.18%**	**0.24%**	**0.47%**	**8.6%**	**30%**	**21%**
	(0.11%-0.30%)	(0.20%-0.28%)	(0.43%-0.51%)	(5.7%-13%)	(28%-33%)	(18%-23%)
Kidney	**0**	**0**	**0**	**2.7%**	**2.5%**	**1.7%**
				(2.2%-3.3%)	(1.9%-3.3%)	(1.2%-2.7%)
Lung	**0**	**0**	**0**	**0**	**0**	**0**

Figure 5.3 *Hic1* methylation (Legend with Figure 5.1)

A) *Mlh1* methylation-gut



B) *Mlh1* methylation-heart

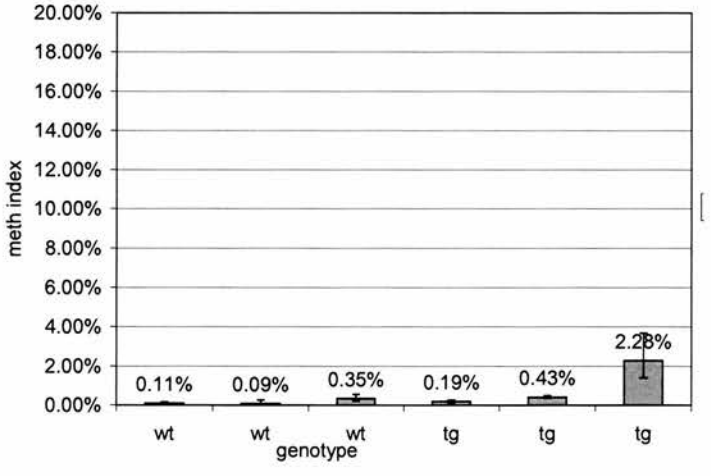


Figure 5.4 Mlh1 methylation. Methylation of *Mlh1* in heart and gut as measured by M-QPR. The error bars represent transformed standard errors. Data are from F8 mice.

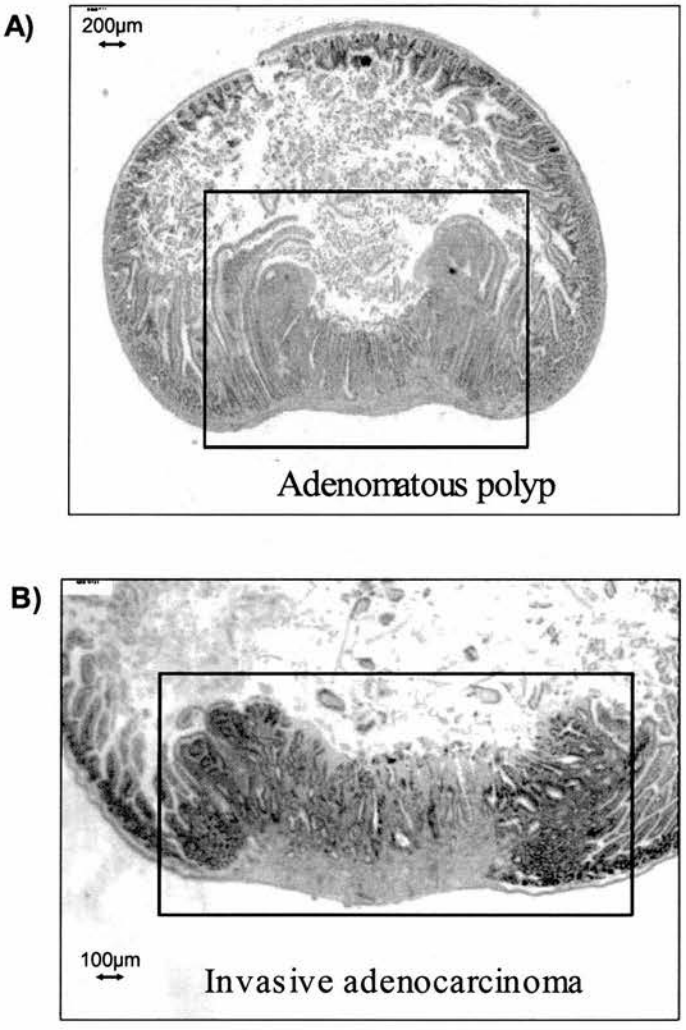
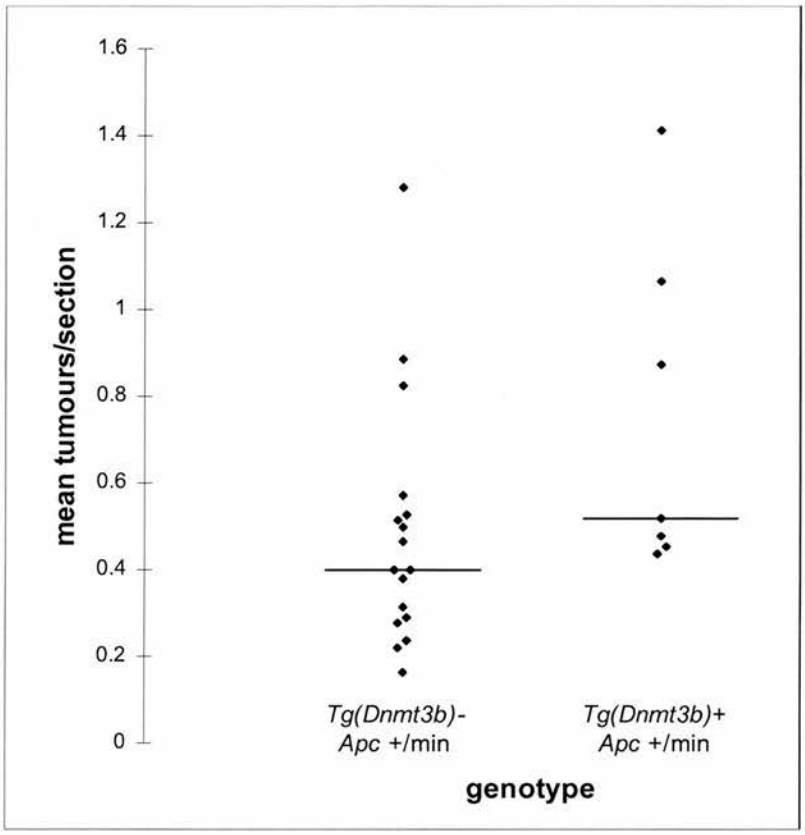


Figure 5.5 Histology of tumours from *Apc*$^{+/min}$ mice. H+E stained transverse sections of small intestine from intestines of *Apc* $^{+/min}$ mice showing A) benign adenomatous polyp B) malignant adenocarcinoma with necrotic centre and invasion in to the mucosa.



genotype	n	Mean	SD	SE	95% CI of Mean
Tg(Dnmt3b)-	17	0.49	0.28	0.07	0.34 0.63
Tg(Dnmt3b)+	7	0.75	0.38	0.14	0.40 1.10
Ratio of means(pos:neg)		1.54 (95%CI 0.99 to 2.44)*			

genotype	n	Median	IQR	95% CI of Median
Tg(Dnmt3b)-	17	0.40	0.24	0.29 0.53
Tg(Dnmt3b)+	7	0.52	0.51	0.43 1.41
Ratio of medians (pos:neg)		1.30 (95% CI 0.84 to 2.68) *		

Figure 5.6 Effect of Dnmt3b overexpression on tumour burden in *Apc*$^{+/min}$ mice. Scatterplot and summary statistics comparing tumour burden (as measured by mean tumours/section) between *Tg(Dnmt3b)+*, *Apc*$^{+/min}$ and *Tg(Dnmt3b)-* , *Apc* $^{+/min}$ mice.
* Confidence intervals for ratio of medians and ratio of means are estimated by bootstrapping with 2000 resamples of the data using nonparametric resampling and bias-corrected and accelerated percentile method for estimating confidence intervals.

A)

genotype	invasive	not invasive	Total
Tg(Dnmt3b)+ Apc +/min	7	87	94
Tg(Dnmt3b)- Apc +/min	10	193	203
Total	17	280	297

P=0.54

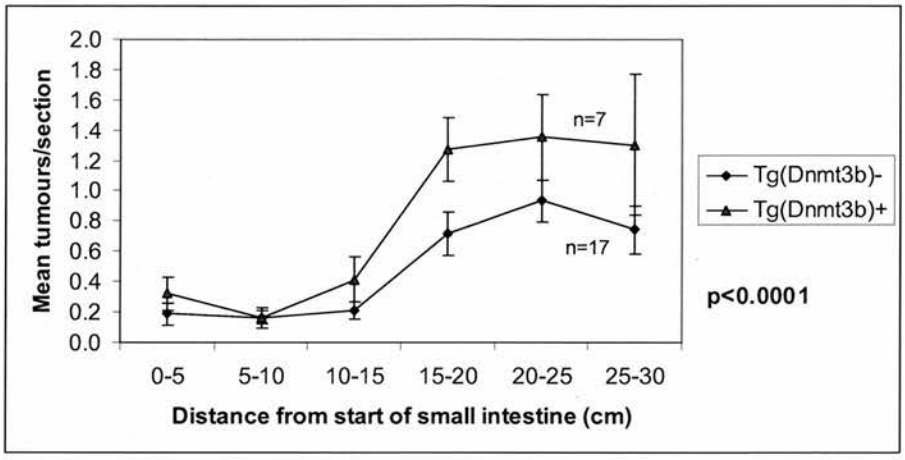
B)

genotype	dysplastic	not dysplastic	Total
Tg(Dnmt3b)+ Apc +/min	52	44	96
Tg(Dnmt3b)- Apc +/min	71	132	203
Total	123	176	299

P=0.003

Figure 5.7 Effect of Dnmt3b overexpression on tumour progression in *Apc* $^{+/min}$ mice. Contingency tables comparing proportion of polyps from *Apc* $^{+/min}$ mice comparing features of malignant transformation between *Tg(Dnmt3b)* genotypes. The number of polyps differ between the two groups since polyps in which the base of the tumour was not clearly seen were excluded from the analysis. Significance values are by Fisher exact test.

A)



B)

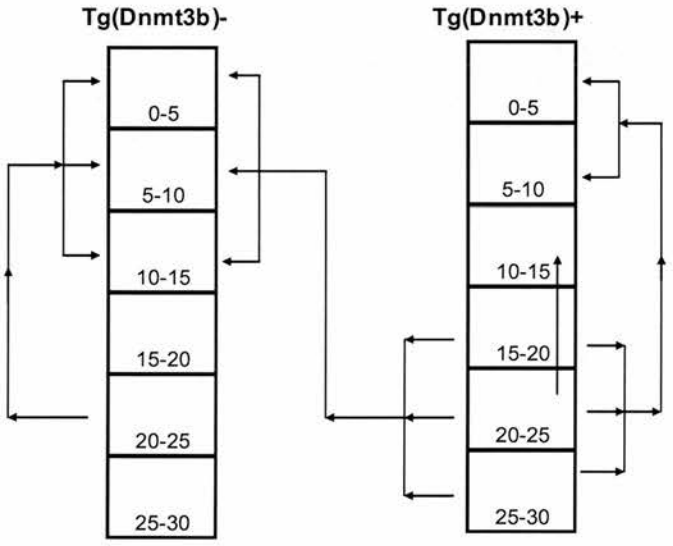


Figure 5.8 Regional differences in tumour burden. A) Mean tumours per section in each segment of the gut at given distance from the stomach in *Apc*$^{+/min}$ mice, classified according to Tg(Dnmt3b) status. Error bars represent standard error. The significance value is for the existence of a regional difference by Kruskal-Wallis test (non parametric analysis of variance).

Panel B) shows which segments are significantly different from each other at the 5% significance level (post-hoc analysis of variance with Bonferonni correction for multiple testing). There is a significant difference in tumours between 2 regions if an arrow can be followed between the boxes representing the 2 regions. The direction of the arrow denotes that there is a greater number of tumours in the region at the start of the arrow than in the region at the end of the arrow. For example, the 25-30 cm region from the Tg(Dnmt3b)+ animals has significantly more tumours than each of the 0-5cm and 5-10cm regions from Tg(Dnmt3b)+mice, and also has significantly more tumours than each of the 0-5cm , 5-10cm and 10-15cm regions from Tg(Dnmt3b)- mice.

A) Internal promoter

Genotype	Clone	Distance upstream from transcript start: 210	192	190	188	171	144	142	121	%Conversion
Apc +/+	A1								n	100
Tg(Dnmt3b)-	A2								n	100
	B1									93
	B2									100
	C1									100
	E1								n	100
	F1								-	100
	G1									100
	H1									94
	A6					n	-	-	-	100
	A7						-	-	-	100
	B6						-	-	-	100
	D6							?	n	100
	E6						-	-	-	100
	G6								n	100
Apc +/+	A4					-	-	-	-	100
Tg(Dnmt3b)+	B4						-	-	-	100
	C4								-	100
	D4								-	100
	E4		n	n	n		-	-	-	100
	F3						n			100
	F4								?	100
	G3								?	100
	H3									100
	C7								-	100
	E7								-	100

B) External promoter

Genotype	Clone	Distance upstream from transcript start: 397	391	371	363	358	355	conversion%
Apc +/+	03E	n	n		n			100
Tg(Dnmt3b)-	03F							100
	03G	n	?		n			100
	03H		?		n	n		100
	04A	n	n		n	n		100
	04B	n	?		n			100
	04C							100
	04D	n			n			100
	04E							100
	04F							100
	08F	n	?		n			100
	08G	n	?		n	n	n	100
	08H	n					n	100
	09A	n						100
	09B	n						100
	09C	n			n	n		100
	09D	n			n	n		100
	09E	n			n		n	100
	09F	n			n			100
	09G							100
	09H	n			n			100
	10A	n						100
	10B					n		100
	10C	n			n			100
	10D	n						100
	10E	n	n		n			100
	10F	n	?		n			100
	10G	n						100
	10H	n			n		n	100
Apc +/+	02C							100
Tg(Dnmt3b)+	02D							100
	02F							100
	02G	n	?		n	n	n	100
	03A	n						100
	03B							100
	03C	n			n			100
	03D	n	?		n			100
	06B	n			n	n	n	100
	06C							100
	06D	n						100
	06F	n			n			100
	06G	n			n		n	100
	06H	n	?		n			100
	07B	n	n		n	n	n	100

Figure 5.9. Methylation status of *Apc* promoter regions. Methylation of A) Internal *Apc* promoter B) External *Apc* promoter as measured by bisulphite sequencing. DNA was extracted from freshly harvested distal small intestine from *Apc* $^{+/+}$, *Tg(Dnmt3b)+* and *Apc* $^{+/+}$, *Tg(Dnmt3b)-* mice. □=unmethylated. n =ambiguous sequence — =unreadable sequence. Frequency of CpH-CpT conversion in each clone is given as a measure of completeness of bisulphite conversion

		Distance upstream of transcript start								Conversion
Genotype	Clone	210	192	190	188	171	144	142	121	%
Apc +/min	01D	.	.	.	.	.	.	.	.	100
Tg(Dnmt3b)+	01G	.	.	.	.	.	.	.	.	100
	01H	■	■	.	■	■	■	■	n	100
	02A	.	.	.	.	.	.	.	.	100
	02C	.	.	.	.	.	■	.	.	100
	02E	.	.	.	.	.	.	.	.	100
	02F	.	.	.	.	.	.	.	.	92
	03A	.	.	.	.	.	.	.	.	100
	03C	.	.	.	.	.	.	.	.	100
	03D	.	.	.	.	.	n	n	-	100
	03E	■	.	.	.	.	.	.	.	100
	03F	.	.	.	.	.	.	.	.	100
	03G	.	.	.	.	.	■	.	.	100
	03H	■	■	■	■	.	.	.	.	100
	01B	.	.	.	.	.	.	.	.	100
	01F	.	.	.	.	.	.	.	.	92
	02B	.	.	.	.	.	.	.	.	100
	02G	■	■	■	■	■	■	■	■	100
	02H	■	■	■	■	■	■	■	■	100
	03B	.	.	.	.	.	.	.	.	100
Apc +/min	05G	.	.	.	.	.	.	.	.	100
Tg(Dnmt3b)-	06A	.	.	.	■	.	.	.	.	100
	04D	.	.	.	.	.	.	.	.	100
	04E	.	.	.	.	.	.	.	.	100
	04F	.	.	.	.	.	.	.	.	100
	04H	■	■	.	■	■	■	■	■	100
	05E	.	.	.	.	■	■	■	■	100
	05H	.	■	.	■	■	■	■	■	93
	06F	.	.	.	.	.	.	.	.	100
	06G	.	.	.	.	.	.	.	.	100
	04A	.	.	.	.	.	.	.	.	100
	04B	n	.	n	n	n	.	n	-	100
	04C	.	.	.	.	.	.	.	.	100
	04G	■	■	■	■	■	■	■	.	100
	05B	.	.	.	.	.	.	.	.	100
	05C	.	.	.	.	.	.	.	.	100
	05D	.	.	.	.	.	.	.	.	100
	05F	.	.	.	.	.	n	.	n	100
	06B	.	.	.	.	.	n	.	n	100
	06C	.	n	.	.	.	.	.	.	100
	06E	.	.	.	.	.	.	n	n	100
	06H	.	.	.	.	.	.	.	.	100

Figure 5.10 *Apc* internal promoter methylation in microdissected tumours.

Figures 5.10 and 5.11. Methylation status of 5.10) Internal *Apc* promoter and 5.11) External *Apc* promoter as measured by bisulphite sequencing . DNA was pooled from 20 tumours microdissected from paraffin blocks of intestines from $Apc^{+/min}$. *Tg(Dnmt3b)+* and $Apc^{+/min}$, *Tg(Dnmt3b)-* animals. ■= methylated. . =unmethylated. n=ambiguous sequence — =unreadable sequence. CpH-TpH conversion frequency is given for each clone in the right hand column as a measure of bisulphite conversion completeness.

Genotype	Clone	Distance from transcript start																		Conversion
		397	391	371	363	358	355	347	338	336	334	330	321	314	297	275	261	253	249	%
Apc +/min	07A	.	.	.	.	.	.	.	■	.	.	.	.	.	n	.	-	-	-	100
Tg(Dnmt3b)+	07F	.	.	.	.	.	.	.	.	.	.	.	.	.	n	.	n	n	.	94
	07G	.	.	.	.	n	.	.	.	.	n	.	.	.	.	.	.	n	.	92
	07H	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	n	.	100
	08B	n	.	.	.	n	.	n	.	.	.	n	.	.	.	-	-	-	-	100
	08C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	■	.	.	95
	08D	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	100
	08E	.	.	.	.	.	.	n	.	.	.	.	.	.	.	.	.	n	.	100
	08G	n	.	.	.	.	.	n	.	.	.	.	.	.	.	.	n	.	.	100
	08H	.	.	.	.	.	.	.	.	.	.	.	n	.	.	n	.	.	.	100
	09D	.	.	.	.	.	.	n	.	.	.	.	■	.	.	n	n	■	n	95
	09E	.	.	.	.	n	.	.	.	.	.	.	.	.	n	n	n	-	-	100
	09F	.	.	.	.	.	.	n	■	.	.	.	.	.	.	.	n	-	-	100
	09G	.	.	.	.	n	n	n	.	.	n	.	.	.	-	-	-	-	-	100
	09H	.	.	.	.	.	.	.	.	.	n	.	.	.	n	-	-	-	-	100
	07B	n	n	n	n	n	n	n	n	n	n	n	n	.	.	.	.	.	.	100
	07C	.	.	.	.	.	.	.	■	.	.	.	.	.	.	.	.	.	.	95
	07D	.	.	.	.	.	.	.	.	.	.	.	■	.	.	.	.	.	.	95
	07E	.	.	.	.	.	.	.	.	.	.	.	■	.	.	.	.	.	.	100
	09B	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	95
	09C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	95
Apc +/min	10A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	95
Tg(Dnmt3b)-	10C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	91
	10D	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	100
	10E	n	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	100
	10H	n	.	.	.	.	.	n	.	.	.	.	.	.	.	.	.	.	.	100
	11B	.	.	.	.	n	.	.	.	.	.	.	.	.	n	.	n	.	.	100
	11C	.	.	.	.	.	.	n	.	n	n	n	n	.	-	-	-	-	-	100
	11H	.	.	.	.	.	.	.	.	n	n	.	.	n	n	n	-	-	-	100
	12A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	100
	12D	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	n	n	.	100
	12G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	100
	10B	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	91
	10F	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	100
	10G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	100
	11D	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	100
	11E	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	100
	11G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	100
	12B	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	100
	12C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	n	.	100
	12H	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	100

Figure 5.11 *Apc* internal promoter methylation in microdissected tumours. Legend on preceding page.

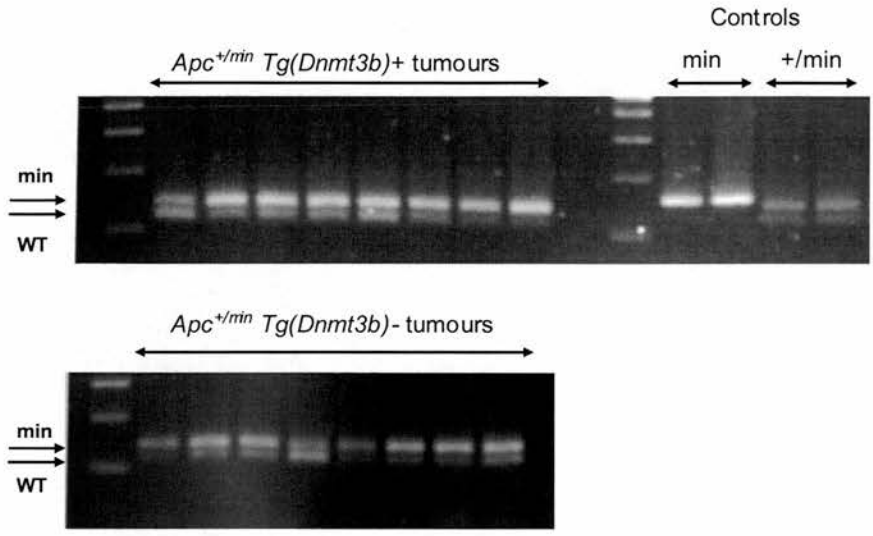


Figure 5.12 Loss of heterozygosity analysis from microdissected tumours. 3% agarose gel showing *HindIII* digested *Apc* gene PCR product. Results are shown for 16 individual microdissected tumours from *Apc* $^{+/min}$, *Tg(Dnmt3b)+* and *Apc* $^{+/min}$, *Tg(Dnmt3b)-* intestines. WT= wild type band (121 bp); min=mutant band (138 bp). Control samples: min= undigested PCR product;
+/min= *HindIII* digested PCR product from *Apc* $^{+/min}$, *Tg(Dnmt3b)-* tail tip genomic DNA.

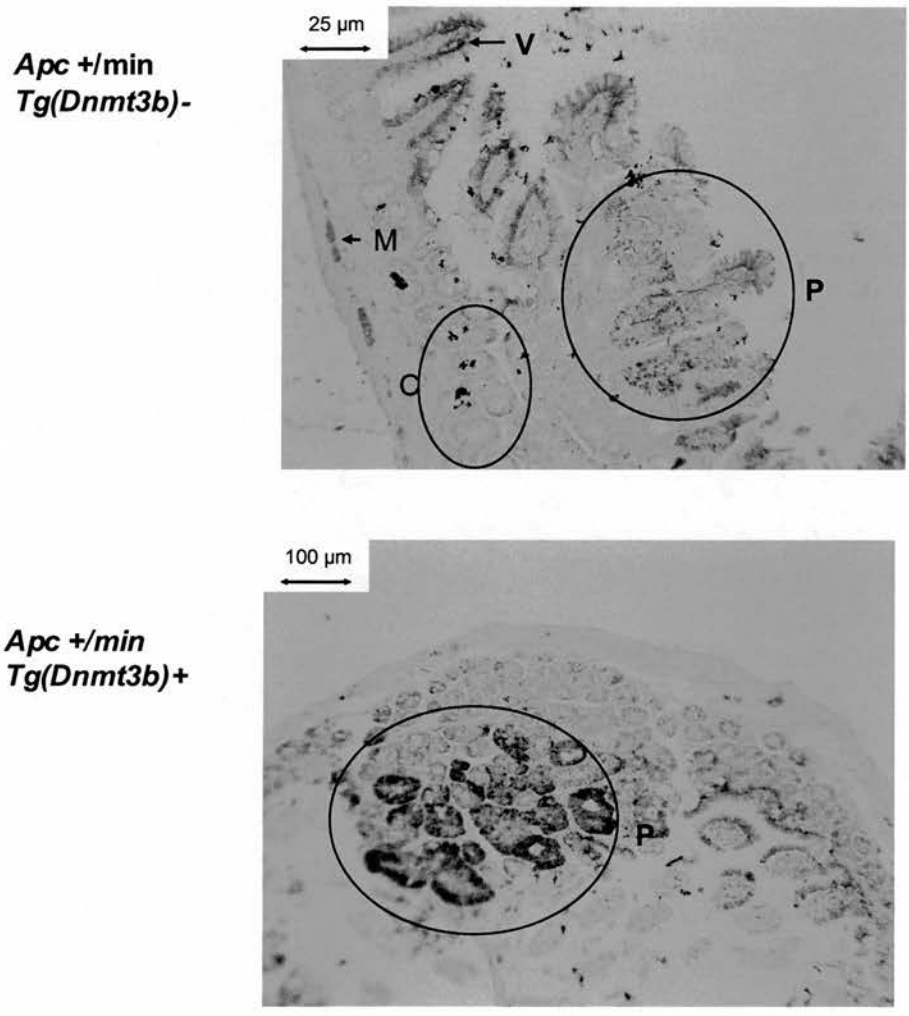


Figure 5.13 Beta catenin staining of distal small intestine from *Apc* $^{+/min}$ mice. The photomicrographs show nuclear beta catenin staining within polyps (P), membrane staining in villi (V), and largely absent staining in basal crypts (C). Meissner's plexus neurons (M) also express membrane beta catenin.

6. Concluding remarks

In this thesis, the production of what are to my knowledge are the first lines of transgenic mice which express a *de novo* methyltransferase in adult somatic cells. The production of mice which over-express Dnmt3b essentially completes the basic set of transgenic mouse experiments manipulating the DNA methyltransferases. The results of these different models are summarised in the table below.

Model (ref)	Devel	Phenotype	Global methylation	Methylation at specific loci	Apc +/min microadenomas	Apc +/min polyps	Other cancers
Dnmt1 Null (1)	Lethal E15.5	Absent forelimb development Reduced yolk sac haematopoiesis Open neural pore	Reduced 30%	Lost at Imprinted Genes	N/A	N/A	N/A
Dnmt1 Reduced (2-7)	Normal	Spontaneous lymphoma (2)	Reduced	?	Increased (3)	Reduced (3-5)	Spontaneous lymphoma.(2) Decreased Mlh1 intestinal tumours.(6) Decreased carcinogen lung tumours(7)
Dnmt1 increased (8)	Lethal E12.5	Multiple abnormalities	?	?	N/A	N/A	N/A
Dnmt3b Null (9)	Lethal E14.5	Ventricular septal defects Abnormal blood vessel formation	Reduced at centromeres	IgF2 Demethylation	Unchanged	Reduced	?
Dnmt3b reduced (10)	Survive to term	Facial/immune defects	Reduced minor satellite repeat	?	?	?	?
Dnmt3b increased (11)	Normal	Cardiac/ Kidney by 4 months	Increased 10%	Wnt3, Lrfn5 rmethylation	Unchanged	Increased dysplasia	DMBA carcinogen unaffected
Dnmt3a Null (9)	Survive to term	Runted, neural tube abnormalities	?	?	?	?	?

?= not done. N/A- impossible (e.g lethal phenotype prevents experiment). References in table: (1) Li, *et al.*,992. (2) Gaudet, *et al.* 2003 . (3) Yamada, *et al.* 2005. (4) Eads, *et al.* 2002. (5) Cormier, *et al.* 2000. (6) Trinh *et al.* 2002.(7) Belinsky, *et al.* 2003. (8) Biniszkiewicz, *et al.* 2002 . (9) Okano, *et al.* 1999 . (10) Ueda, *et al.* 2006. (11) This thesis.

Can we find any common features between these studies illustrating the role of DNA methylation *in vivo*? Reviewing the role in development first of all, it appears that at least some embryos are able to survive until E9.5 in both methyltransferase deficient and methyltransferase over-expressing mice. In normal mouse development rudimentary

organs are morphologically distinguishable by this stage. This suggests that the earliest stages of embryogenesis and organ specification are not critically regulated by DNA methylation. However DNA methylation does appear essential for later survival. Since DNA methylation patterns are largely established early in the normal embryo but lethality in the various transgenic models occurs later, it seems most likely that DNA methylation has a predominant role in maintaining, as opposed to initiating, differentiated expression patterns established in early embryogenesis. Since the effect of Dnmt3b deletion is severe whereas the effects of Dnmt3b over-expression are minor in development, it can be concluded that *de novo* methylation is necessary but that existing levels of Dnmt3b in the embryo are functionally saturated- in the sense that further increases do not cause further silencing. In contrast, both deletion and over-expression of Dnmt1 cause a severe phenotype. How do we reconcile this apparent difference? While it is of course difficult to draw general conclusions from specific phenotypes caused by specific over-expression models, one interpretation is that maintenance activity is the rate limiting step in establishing significant methylation under normal circumstances. Cooperation between Dnmt1 and the *de novo* methyltransferases (Kim, *et al* 2002) could lead to non-linear relationships between methyltransferase levels and steady state levels of methylation (Sontag, *et al* 2006), possibly creating increased sensitivity to disturbances in Dnmt1 levels . A second interpretation is of course that the toxic effects of Dnmt1 expression are not mediated through methylation, although the loss of methylation at certain imprinted loci in Dnmt1 over-expression (Biniszkiewicz, *et al* 2002) suggests that there is an effect on methylation.

The phenotypic changes between the different models seem to have little in common, other than the fact that CNS abnormalities appear to be common. Although *Tg(Dnmt3b)+* mice have no apparent structural, the abnormal maternal behaviour in these mice could suggest that DNA methylation changes have altered CNS functions. It is an attractive hypothesis, given the prevalence of DNA methylation in vertebrates that

a major function of DNA methylation is the maintenance of tissue specification following organ differentiation and has special importance in the central nervous system.

Moving on to the role of DNA methyltransferases in determining methylation patterns, the project has provided some insights. Considering firstly the question of whether CpG islands are unmethylated because of the lack of *de novo* methyltransferase activity, it has indeed been shown here that increases in *de novo* methyltransferase activity are sufficient to alter methylation patterns in somatic cells. However it is clear that upregulation of *de novo* methyltransferase activity is not a sufficient condition to cause methylation since the majority of CpG islands still remain unmethylated even when the gene is not transcribed. Moreover, loci at which aberrant methylation has been found are transcriptionally silent, demonstrating that aberrant methylation caused by increases in *de novo* methylation can occur at quiescent loci and would thus be predicted not to affect cellular function.

However, the development of phenotypic changes suggests that alterations in methylation are indeed able to alter cell biology. While, strictly speaking, it has not been proven that the phenotypic changes seen in *Tg(Dnmt3b)+* mice are caused by methylation, the existence of global methylation changes which precede the onset of the cardiac phenotype suggest that alterations in methylation are not simply secondary to the phenotype. The fact that the mice are developmentally normal and that phenotypic changes seem to be specific in nature and confined to defined organ systems suggests that the ability of methylation to disturb transcription in a manner that is important for the biology of the cell is quite restricted. In other words, although the number of loci prone to methylation as a result of *de novo* methyltransferase upregulation may not be insignificant, the number loci that can be actively silenced by methylation is very small. Moreover, phenotypic changes occur without exception in mice which show strong transgene expression. If abnormalities were caused by random inactivation of genes by methylation, one would expect sporadic occurrence of different abnormalities. The

results thus suggest that the susceptibility to silencing by methylation is tightly regulated and must be specified by coexisting factors.

The lack of spontaneous cancer formation is perhaps a special case of this hypothesis that active silencing by methylation is a rare event in the context of *de novo* methyltransferase over-expression, whereas methylation changes may be less so. It may also be a special case of the idea that silencing by *de novo* methylation is saturated, since many tumours are thought to arise from tissue stem cells which express Dnmt3b, and thus a further increase in Dnmt3b expression may not add much to gene silencing. Although there is detectable methylation of tumour suppressor genes, there has clearly not been widespread silencing. It could be argued that an increased tumour incidence could have been seen had the mice lived longer. However, there are genes which cause an aggressive cancer phenotype when inactivated (including the *Apc* gene). If inactivation by stochastic errors of methylation were a common cause of cancer it would be likely that at least some cancers would have been seen within the lifespan of the *Tg(Dnmt3b)+* mice. Given that methyltransferases are often increased in cancer, it follows from the preceding arguments that out of the many methylation events which undoubtedly occur in cancer, only a minority may have a causative role. It thus would seem likely that if methylation events do have a causative role in cancer, this may be restricted to a smaller number of loci than has often been proposed in the literature.

Nevertheless, the differences in tumour progression seen in $Apc^{+/min}$ mice suggest that abnormal methylation does have a role in cancer pathogenesis, but operates in this model mainly as a secondary phenomenon. In this hypothesis, changes in the transformed cell lead to an increased sensitivity to silencing by methylation. However the loci at which methylation silencing is important for tumour behaviour is still likely to be a subset of the loci which are abnormally methylated, and is likely to differ between different types of cancer. The hypothesis also emphasises the need for caution in interpreting the results of methylation studies performed in transformed cell lines since

the sensitivity to silencing by methylation in cell lines may not reflect the situation *in vivo*.

What implications do the findings have for the use of DNA methylation in clinical practice? To conclude this thesis, the clinical implications of this study are discussed in the form of a Socratic dialogue between the author, Dr. A (a clinician), and Prof. B (an advocate for the importance of DNA methylation in cancer). Any resemblance of these characters to individuals living or dead is purely coincidental.

Dr. A: Your mouse model has certainly shown some unexpected results. But it seems a long way from anything going on in the clinic.

TL: Indeed. But methylation inhibiting drugs such as 5'-azacytidine are currently being heavily promoted for clinical use, and so there is a current need for understanding of the biology of methylation. This project does raise questions about whether demethylating drugs will find a widespread role in cancer.

Prof B. But the importance of methylation in cancer is well established!

TL: I agree that there are specific examples such as microsatellite instability in colorectal cancer in which methylation changes appear to mimic the effects of genetic loss found in inherited forms of the cancer, and in these cases a role for methylation either in the initiation or maintenance of these changes seems very likely. Indeed the demonstration of the cardiac phenotype is very suggestive that in given circumstances, DNA methylation abnormalities can have important effects. However this study does offer proof that tissues can exist with increased methylation levels without apparent consequences. It also offers proof that abnormalities of methylation can occur at genes that are already silent and thus do not affect transcription. These findings raise doubts about the interest of the ever increasing number of papers published describing

methylation of more and more genes in different forms of cancer, since the silencing of many of these genes may not have functional importance...

Dr.A: And so reversing their methylation with 5'-azacytidine would not have any clinical effect. This is in keeping with the disappointing results of 5'-azacytidine therapy in almost all tumours. You have convinced me that many of the methylation changes in cancer may be secondary; this was previously suspected anyway. Is there any way in which we can sort out for which tumours which it might be useful?

TL: This is a difficult problem. The examples where DNA methylation is most strongly implicated rely on the existence of a known genetic abnormality which has a well defined molecular mechanism. The number of such well-defined abnormalities will no doubt increase rapidly with the results of ongoing genomic studies aimed at finding inherited polymorphisms which lead to differences in cancer incidence. However in the meanwhile we must rely on other criteria. Certainly the demonstration of downstream consequences of silencing and demonstration that the gene would normally be expressed in the normal counterpart cell are necessary, but this doesn't prove that the gene was not previously silenced by other means.

Dr A: It sounds rather labour intensive for routine clinical use. I think that what is likely to happen is that demethylating agents will be used empirically in clinical trials in conditions in which there are statistically high frequencies of important methylation abnormalities.

TL: I agree, and clinical efficacy is at the end of the day the most important evidence. It will be important nevertheless to try and distinguish whether any responses are due to cytotoxic or demethylating actions, emphasising the importance of having downstream effects to follow.

Prof B: Even if methylation abnormalities are mostly secondary, this does not rule out an important function. Your own data with the Apc mice shows this.

TL: Granted. But the distinction has implications for the aims of administering demethylating agents. It suggests that they may be more likely to have a role in preventing or delaying disease progression as removal of secondary methylation abnormalities would be unlikely to cure. Especially in the light of the suggestion from our decitabine treatment experiment suggesting that remethylation can occur rapidly in the context of active methyltransferase activity.

Prof B: Quite apart from the potential therapeutic implications, the patterns of methylation abnormalities surely can tell us about the nature of the disease. This would be useful for diagnosis and prognosis.

TL: Our demonstration that methylation patterns produced by Dnmt3b over expression seem to differ, independently of the strength of transgene expression or of the transcriptional state at a given locus. I think that this does indicate that other factors must regulate methylation patterns, and so I agree that methylation patterns may tell us about the nature of the disease. It is an attractive idea that methylation patterns may be established (in at least some cases) in tissue stem cells which express *de novo* methyltransferases, and one could speculate that perhaps methylation patterns may be a footprint of, for example, differences in chromatin structure or histone modifications in the cell of origin of the tumour.

Prof B: This is indeed just speculation. You need a more global technique to verify this. But I agree it might be interesting to use genomic array techniques to see whether methylation patterns show tissue-dependent clustering in the transgenic mice, and also whether tumour methylation patterns form distinct clusters.

Dr. A: This might also resolve the question of the hypermethylator phenotype and its relationship to prognosis.

TL: As you know, I have problems with the definition of the hypermethylator phenotype- it does not seem clear to me whether there is a distinct phenotype or whether the numbers of CpG islands methylated is simply correlated with the aggressiveness of the tumour. I agree that array type techniques may help determine whether methylation patterns are closely determined, as I suggest, or simply occur at random. Paradoxically, taking a more abstract approach to determination of methylation patterns in this way may help to determine whether there are specific genes which have a particular role to play, although the statistical methodology would be critical. This could give new insight into the pathogenesis of tumours as well as giving a more manageable set of genes to examine for clinical predictive work. It is perhaps most likely that methylation abnormalities may fall in to two categories for clinical work- abnormalities which occur commonly in all cancers or all cancers of a given type, which could be used as a common marker for diagnosis or monitoring, and abnormalities at specific loci which may define specific subgroups of disease which may benefit from specific interventions.

Prof B: (aside to Dr.A) He does go on and on doesn't he. (To everyone) We can discuss these points for a long time until we have further data. I think we are in danger of missing closing time. Can I suggest we continue another day?

7. Methods

7.1 Production of Dnmt3b over-expressing mice

7.1.1 Preparation of Dnmt3b expression constructs

Plasmids with full length *Dnmt3b1* cDNA cloned in to PGEMT vector were obtained from Rudolph Jaenisch. The PCAGASIZ expression vector was obtained from Austin Smith (University of Edinburgh). Dnmt3b expression constructs were prepared by cutting the Dnmt3b plasmid with *Not1* and *XhoI* and ligating this insert between the *NotI* and *BamHI* sites in to the PCAGASIZ vector. The *XhoI* and *BamHI* sites had been partially filled in with Klenow polymerase to create compatible ends.

The preparation of the original Dnmt3b expression construct and verification of activity was performed by Bernard Ramsahoye, University of Edinburgh.

Fig 7.1 shows details of the expression vector and restriction sites.

Figure 7.2 gives a schematic overview of the strategy for producing Dnmt3b over-expressing mice.

7.1.2 Production of modified Dnmt3b expression construct.

A modified Dnmt3b expression construct with frt sequences flanking the resistance gene was produced as in Figure 7.5. Gel purified oligonucleotides containing the *frt* sequence or *frt* antisense sequence with added *RsrII* or *ClaI* restriction sequences at the 5' end were ordered from MWG and diluted to 100 μM:

ClaI oligos:

Sense:

CGATGAAGTTCCTATACTTTCTAGAGAATAGGAACTTCGGAATAGGAACAAC
AT

Antisense:

CGATGAAGTTCCTATTCGAAGTTCCTAGGCTCTAGAAAGTATAGGAACTTCA
T

RsrII oligos:

Sense:

GTCCGGAAGTTCCTATACTTTCTAGAGAATAGGAACTTCGGAATAGGAACTT
CC

Antisense:

GACCGGAAGTTCCTATTCCGAAGTTCCTATTCTCTAGAAAGTATAGGAACTT
CCG

Complementary oligos were annealed to form double stranded linkers with overhanging ends by mixing equal volumes of sense and antisense nucleotides, heating to 95°C for 1 hour, cooling to 72°C for 1 hour then allowing to cool to room temperature. Successful annealing was verified by running the product on a 3% agarose gel. The insert was then phosphorylated by incubating 5 µl of product with 1 µl T4 PN kinase and 50 nmolATP in a total volume of 50µl for 1 hour. The product was heat inactivated at 65°C for 20 mins, ethanol precipitated and resuspended in 10 µl.

Plasmid containing the original *Dnmt3b* construct was cut with *ClaI* by digesting 5µl of plasmid miniprep for 4 hours at 37°C then dephosphorylated by incubating with 1unit shrimp alkaline phosphatase (Boehringer) for 1 hour at 37°C before heat inactivating by heating to 65°C for 20 minutes. The linearised plasmid was run on a 1% agarose gel, the band was excised and DNA purified using the Qiagen minelute kit. *ClaI* linker was ligated into the vector by incubating 1 µl of dephosphorylated plasmid and 3 µl of the

ClaI linker ligated with T4 ligase (Promega) and the manufacturer's buffer for 2 hours before cloning into DH5alpha bacteria. Clones were expanded and plasmids extracted using the Qiagen Qiaprep kit according to the manufacturers instructions. Screening for integration and correct orientation was carried out by PCR (95°C melting 1 min, 60°C annealing 1 min, 72° extension 2.5 min x 35 cycles) using a forward primer within the *Dnmt3b* sequence AATTTGAGCCACCCAAGTTG and using either the sense and antisense *frt* oligonucleotides as reverse primer (Figure 7.5). Clones with the correct orientation were cut with *RsrII* and phosphorylated as above, and the *RsrII* linker ligated as above and cloned. Clones were screened for integration by *Xba1* restriction digestion (presence of integration giving a 1.1 kb fragment) and the correct orientation was confirmed by sequencing, using a sequencing primer within the zeocin resistance gene (GACGTGACCCTGTTCATCAGCGCG).

The resulting plasmid (clone FC2A) was expanded using the Qiagen maxiprep kit according to the manufacturer's instructions, linearised by digestion with *FspI*, ethanol precipitated and resuspended to a give a final concentration of 1 μg/μl under tissue culture conditions for incorporation into E14 embryonic stem cells by electroporation. E14 embryonic stem cells were originally obtained from the Centre for Genome Research (University of Edinburgh). Passage 13 ES cells were used for the original line; passage 22 cells were used for the second line of mice.

Results of cloning steps and a schematic view of the modified construct are shown in Figure 7.5.

7.1.3 Embryonic stem cell culture

Embryonic stem cells were maintained by growing in gelatinised 25 ml tissue culture flasks (Corning) in antibiotic free GMEM medium (Gibco) containing 0.25% sodium bicarbonate, 1% non-essential amino acids, 4 mM L-Glutamine, 2 mM sodium pyruvate 1, 0.1 mM 6-mercaptoethanol 10% fetal calf serum and 100 U/ml leukaemia inhibitory

factor (LIF). LIF was obtained as a culture supernatant from COS-7 cells transiently transfected with a murine LIF expression plasmid (pCAGGSLIF-418). LIF transfection and cell culture was performed by Helen Taylor and Julie Buchanan in the John Hughes Bennett Laboratory, University of Edinburgh. Cells were grown at 37°C in a humidified 5% CO_2 atmosphere and medium was changed on alternate days. Cells were passaged when almost confluent (approx twice weekly).

Cells were passaged by removing medium, washing once with 10 ml PBS, then incubating for 5 mins at 37°C with 2 ml trypsin solution (0.025% trypsin, 0.1% chicken serum, 1.3 mM EDTA in PBS). Cells were resuspended by gentle agitation of the flask and then trypsin activity was quenched by adding 8 ml of complete medium. The cell suspension was transferred to a 60 ml conical centrifuge tube and centrifuged at 100 RPM for 5 minutes. The supernatant was removed and the cells were resuspended in 10 ml of medium. 9 ml of medium and 1 ml cell suspension were then added to 25 ml flasks which had been gelatinised by adding 10 ml 1% gelatin in PBS for 20 mins and then removing the gelatine solution.

7.1.4 Cell culture

ES cell culture was performed in a dedicated ES cell tissue culture suite using sterile techniques in a laminar flow hood.

7.1.5 Cryopreservation of ES cells

ES cells were stored by resuspending cells from a sub-confluent flask as above in medium and adding DMSO to a final concentration of 10%. 1 ml aliquots were frozen overnight at -70°C and then stored at -140°C. Cells were thawed in a water bath at 37°C for the minimum amount of time, and the cell suspension was then added to centrifuge tubes containing 9 ml medium. The cell suspension was centrifuged at 1000 RPM for 5 minutes then the pellet resuspended in 10 ml fresh medium+LIF. The cell suspension

was transferred to a 25 ml gelatinised tissue culture flask and cultured as above. Medium was changed after approximately 8 hours to reduce any residual DMSO contamination.

7.1.6 Electroporation.

Cells for electroporation were cultured in a 75 ml flask until almost confluent. Cells were trypsinised as above, transferred to a centrifuge tube with four volumes of medium and centrifuged at 1000 RPM for 5 mins. The pellet was resuspended in 1 ml cold PBS and cells counted using a haemocytometer. Cell concentration was adjusted to 7 $x10^6$ cells/ml by adding further PBS. 40 µg linearised vector was added to 0.8 ml of cell suspension ($5x10^6$ cells) in an electroporation cuvette, and electroporation was carried out using a Biorad gene pulser at 240 V/ 500 µF. After incubating on ice for 20 minutes, the cell suspension was added to 10 ml medium+LIF.

10cm gelatinised cell culture dishes (Corning) were seeded with 1 ml of cell suspension($0.5x10^6$ cells) and 9 ml medium+LIF and cells were cultured as above. Two plates were seeded with non-electroporated cells as a control for adequacy of selection. Medium was replaced after 24 hours with 10 ml medium+LIF and 25 mg/l zeocin. Cells were cultured in selective medium, changing medium every 2 days until cell colonies were visible (approx 10 days).

Colonies were picked by removing the medium from the plates and replacing it with 10ml PBS. Colonies were aspirated using a 10 µl pipette and placed in to wells of a round bottomed 96 well plate containing trypsin. After picking colonies, plates were incubated for 5 minutes at 37°C. 150 µl medium+zeocin+LIF was then added to each well and cells were disaggregated by pipetting several times before transferring to a gelatinised 96 well flat bottomed plate. Cells were expanded, adding fresh zeocin containing medium every two days, until wells were subconfluent. Cells were trypsinised and transferred to gelatinised 12 well plates and expanded in selective medium until semi confluent. Finally, cells were trypsinised from the 12 well plates and

transferred to gelatinised 25 ml flasks and expanded in selective medium until subconfluent. Cells were washed and resuspended in fresh medium before cryopreserving cell suspensions until needed for electroporation. Transgene integration was verified by PCR (Figure 7.5d)

7.1.7 Blastocyst injection

Cryopreserved ES cells were thawed as above and passaged in culture with zeocin-free medium for 1 week before blastocyst injection. The protocol adopted for the generation of transgenic mice by injection of ES cells into blastocysts is described in detail in (Hogan 1986).

ES cell clones were cultured under normal conditions to ensure that they were approximately 30% confluent prior to the day of injection. Blastocysts from naturally mated C57BL/6 females were harvested at 3.5 days post-coitum. Injections were performed in M2 medium (Sigma) with 10 to 15 ES cells being injected into each blastocyst before transfer to pseudopregnant recipient females (12 blastocysts per recipient). All surgical procedures were performed in accordance with home office guidelines with appropriate analgesia and anaesthesia.

Blastocyst injections were performed by Dr. Jim Selfridge, Molecular Medicine Centre, University of Edinburgh.

7.1.8 Breeding of chimaeras

Male chimaeras were mated with C57BL6 females in trios, moving each chimaera to a new pair of females every week. Offspring derived from the ES cell contribution to the germline were selected by agouti coat colour. Agouti mice were genotyped by PCR and positive mice were bred with wild type C57BL6 mice.

7.2 Animal work

7.2.1 Mouse housing

Animals were maintained at the Biomedical Research Facility at University of Edinburgh. Animal work was carried out according to the provisions of the Animals (Scientific Procedures) Act (UK) 1986.

7.2.2 Genotyping *Tg(Dnmt3b)*+ mice

Routine *Tg(Dnmt3b)* genotyping was performed by western blot on ear clip samples taken at weaning. Ear clips were pulverized using liquid nitrogen and suspended in 50 µl SB solution (see materials Section 7.7). Suspensions were heated to 95°C for 15 min to remove protease activity and then centrifuged at 13000 RPM for 1 minute. 10 µl of supernatant was analysed by western blot for Dnmt3b (see below).

PCR genotyping, if required, was performed using primers:
L primer TGGCTCTGATATTCTAATGCCAAA
R primer TCATCCTGATACTCTGTGCTGTCTCC
PCR amplification was performed using Hot Star Taq mastermix (Qiagen) with conditions: Initial denaturation : 95°C 15 minutes; PCR cycle: Denaturation 95°C 30 sec, Annealing 60°C 30 sec, Extension 72°C 1 minute (45 cycles); Final extension: 72°C 15 minutes.

7.2.3 *Mbd2* $^{-/-}$ mice

A colony of *Mbd2* $^{-/-}$ mice was bred from founder animals obtained from Prof. Adrian Bird, University of Edinburgh. *Tg(Dnmt3b)+, Mbd2*$^{-/-}$ mice were produced by crossing male *Tg(Dnmt3b)*+ mice with female *Mbd2*$^{-/-}$ mice. The offspring were genotyped for *Mbd2* and Dnmt3b. Male F1 compound heterozygote offspring *(Tg(Dnmt3b)+Mbd2* $^{+/-}$)

were backcrossed on to *Mbd2*$^{-/-}$ females to produce offspring with the expected Mendelian proportions 25% *Tg(Dnmt3b)+ Mbd2*$^{-/-}$: 25% *Tg(Dnmt3b)+Mbd2*$^{+/-}$: 25% *Tg(Dnmt3b)- Mbd2*$^{-/-}$: 25% Tg*(Dnmt3b)-Mbd2*$^{+/-}$. Progeny were genotyped for Dnmt3b by western blot as above, and by PCR for *Mbd2* as below. Survival analysis was carried out using Kaplan-Meier estimation using XLStat.

7.2.4 Genotyping *Mbd2* mice

Mbd2 mice were genotyped by PCR. DNA was extracted from tail tips taken at weaning, using Qiagen DNAeasy kit according to the manufacturer's instructions. 2 µl DNA was amplified using Hotstar Taq Mastermix (Qiagen) with primer sets: Wild type-AAGAACAAGCAGAGATCCG and ACGCTGGCCTAGTGCCGTGC; Knockout TCCGCAAACTCCTATTTCTG and TTGTGGTTGTGCTCAGTT- with cycling conditions 95°C melting temperature 30sec, 55°C annealing for 30 sec, 72°C extension for 1 min for 35 cycles. PCR products were analysed on a 1% agarose gel. Typical results are shown in Figure 7.3a.

7.2.5 *Apc* $^{+/min}$ mice

Apc$^{+/min}$ mice were bred from founder animals obtained from Charles Patek, University of Edinburgh. The *Apc*$^{+/min}$ strain was maintained by crossing male mice positive for the mutation with wild type female littermates. *Tg(Dnmt3b)+, Apc*$^{+/min}$ compound mice were produced by crossing male *Tg(Dnmt3b)-, Apc*$^{+/min}$ mice with female *Tg(Dnmt3b)+* mice. Offspring were fostered at birth to CD1 females which had been mated at the same time as the transgenic mice in order to provide foster mothers. Fostering was necessary to prevent infanticide which occurred invariably within 24 hours of birth with *Tg(Dnmt3b)+* mothers.

7.2.6 Genotyping *Apc*$^{+/min}$ mice

The Apc$^{+/min}$ mice were genotyped by PCR and restriction digestion. DNA was extracted from tail tips taken at weaning, using Qiagen DNAeasy kit according to the manufacturer's instructions. 2 µl of the extracted DNA was amplified using Hotstar Taq Mastermix (Qiagen), left primer TCTCGTTCTGAGAAAGACAGAAGCT and right primer TGATACTTCTTCCAAAGCTTTGGCTAT, with cycling conditions: 95°C melting temperature 30 sec, 60°C annealing for 30 sec, 72°C extension for 1 min for 35 cycles. 5 µl of PCR product were digested overnight with 10 u *HindIII*. Digests were run on a 3% agarose gel, with wild type DNA giving a band at 121 bp and DNA containing the min mutation giving a band of 138 bp. Typical results are shown in Figure 7.3b.

7.2.7 Analysis of tumours from *Apc*$^{+/min}$ mice

Mice for analysis of intestinal tumours were sacrificed between 120 and 130 days of age by cervical dislocation. Intestines were dissected out and perfusion fixed by injecting with formalin using a 23 gauge needle, until the intestine was just inflated, then stored in formalin for at least 24 hours before processing. The entire intestine was unravelled and transverse sections were taken at 5 cm intervals along the length of the small intestine for paraffin embedding. Multiple transverse sections were cut at different levels in the block and stained with H+E.

Scoring of tumours was performed by Dr David Brownstein, Dept of Veterinary Pathology. Scoring was performed blinded to the genotype of the animal. Criteria used for scoring were as follows: Polyp: elevation >1 mm above the surrounding gut lumen. Invasiveness: Presence or absence of invasion into the muscularis mucosa. Dysplasia: presence or absence of abnormal nuclear morphology.

For molecular biology analysis, individual tumours were microdissected from the paraffin blocks. 3 μm sections were cut from blocks and stained with H+E. The H+E image was projected on to the paraffin block to use as a guide to identify the location of polyps, which were then cut from the block. 20 tumours were isolated from *Apc*$^{+/min}$, *Tg(Dnmt3b)+* animals and 20 tumours from *Apc*$^{+/min}$, *Tg(Dnmt3b)-* animals. Microdissection was performed by Dr David Brownstein, Dept of Veterinary Pathology.

Individual tumours were de-waxed by heating in 1 ml xylene for 5 minutes at 50°C, centrifuging for 1 minute at 13000RPM, then washing in 1 ml 100% ethanol twice. DNA was extracted using the Qiagen DNAAmp micro kit according to the manufacturer's instructions but with an extended proteinase K digestion time of 24 hours. DNA from each tumour was eluted in a final volume of 20 μl of 10 mM Tris.

Loss of heterozygosity analysis was performed by PCR amplification and *HindIII* digestion (see Section 7.2.6) using 2 μl of eluted DNA from each tumour as a template.

Bisulphite modification was performed on pooled DNA in order to have sufficient DNA for analysis following bisulphite treatment. 5 μl of DNA eluted from each of the 20 tumour samples from *Apc*$^{+/min}$, *Tg(Dnmt3b)-* and *Apc*$^{+/min}$, *Tg(Dnmt3b)+* tumours was pooled, ethanol precipitated and resuspended in 20 μl. The pooled DNA was bisulphite treated using the Qiagen Epitect bisulphite kit using the manufacturer's protocol for bisulphite modification of formalin-fixed-paraffin-embedded samples.

The promoter regions of both promoters of the *Apc* gene were amplified using 2 rounds of fully nested PCR, then cloned and sequenced (see Bisulphite sequencing Section 7.3.3).

7.2.8 Multistage skin carcinogenesis

Skin carcinogenesis was performed using a modification of the protocol in Owens et al. 5-6 week old mice were shaved on the right flank and 200 nmol DMBA (Sigma) in 200 µl acetone was applied topically using a pipette. DMBA treatment was repeated on weeks 2 and 4. Tumour promotion was performed with the mitogen TPA (Sigma). 5 nmol TPA in 200 µl acetone was applied topically to shaved skin. TPA was administered twice weekly for 10 weeks. Experiments were performed with mice housed in IVC filtered cages separate from non-experimental mice.

7.2.9 5'aza 2-deoxycytidine treatment of mice

5' Aza 2-deoxycytidine (AdC) treatment was performed using two separate protocols. For the first experiment AdC was administered at a dose of 1 mg/kg in a 100 µg/ml solution in PBS by subcutaneous injection. The protocol is based on that in (Laird, *et al* 1995) which was sufficient to cause demethylation in the gut in $Apc^{+/min}$ mice. Animals were injected weekly for 11 weeks starting from 5-6 weeks of age. Animals were weighed weekly during administration to monitor for toxicity.

In the second protocol, AdC was administered twice weekly at a dose of 2.5 mg/kg by intraperitoneal injection. This protocol was based on that used in (Hu, *et al* 1998) to cause demethylation of the *Igf2r* imprinted gene. Animals were treated either with a course of 5 injections or continued treatment. Animals treated for 5 injections were sacrificed for analysis, either 24 hours after the final injection or 14 days after the final injection.

Hearts were dissected out from sacrificed animals. Half was stored in formalin for histology and half was stored at -70°C in RNAlater (Qiagen) for later measurement of methylation. Demethylation was measured by extracting DNA using the Qiagen

DNAeasy kit according to the manufacturer's instructions then measuring methylation of the *Igf2r* differentially methylated region by M-QPR.
The AdC was stored at -70°C in a 2.5 mg/ml solution in PBS and diluted in PBS to working concentration prior to administration.

7.2.10 Analysis of cardiac histology

Animals were sacrificed by cervical dislocation and weighed. Hearts were dissected out and weighed. Half of the heart was frozen in RNAlater for molecular biology analysis and half was fixed in formalin for at least 24 hours before paraffin embedding according to standard protocols. 3 μm sections were cut and stained with H+E and Sirius Red (collagen) stains according to standard techniques. Histological preparation was performed by the technical staff of the Dept of Veterinary Pathology, University of Edinburgh. Microscopy was performed by Dr David Brownstein, Dept of Veterinary Pathology, University of Edinburgh.

7.2.11 Immunohistochemistry

3μm sections were cut from paraffin embedded tissue blocks, mounted on slides, de-waxed and rehydrated according to standard procedures. Antigen retrieval was performed by immersing slides in 100 mM citrate pH 6.0 and heating in a microwave for 20 minutes in a plastic slide holder, before being allowed to cool to room temperature. Tissue sections were incubated in 3% hydrogen peroxide for 15 minutes to block endogenous peroxide activity then washed in TBS-T for 5 minutes. Primary antibodies were applied to the tissue section as follows: Beta catenin: Mouse monoclonal anti-beta catenin (BD Biosciences catalogue 610153) diluted 1:50 in 100 mM Tris pH 7.4; PCNA: Rabbit polyclonal anti-PCNA (Abcam catalogue AB2426-1) diluted 1:50 in 100 mM Tris pH7.4. Primary antibodies were incubated at room temperature for 1 hour before washing in TBS-T for 15 mins. HRP-conjugated secondary antibodies (Dako Envision system) were applied undiluted and incubated for 1 hour at room temperature before

washing in TBS-T for 15 minutes. Visualisation was by incubating in diaminobenzidine (Dako) according to the manufacturers instructions for 3-5 minutes or until brown staining was visible. Slides were counterstained in Mayer's haematoxylin for 5 seconds and rinsed in tap water before mounting with aqueous mounting medium (Crystal Mount, Sigma). The majority of the immunohistochemical staining has been performed by Shin Teoh, University of Edinburgh undergraduate student, under my supervision. Independent review of the β-catenin stained sections was performed by Prof. David Harrison, University of Edinburgh.

For *in vivo* BrdU labelling, mice of 6 weeks and 4 months of age were injected intraperitoneally with 0.1 μmol/kg BrdU (Roche). Mice were sacrificed 6 hours after injection and tissues immediately dissected out and fixed in formalin. After paraffin embedding according to standard procedures, 3μm sections were cut and stained for BrdU using the BD Pharmingen BrdU In Situ Detection Kit (BD Biosciences) according to the manufacturer's instructions.

7.3 DNA methylation techniques

7.3.1 Methylation sensitive restriction fingerprinting

MSRF was performed according to the methods in (Davies 2002). A schematic summary is shown in Figure 7.4. Briefly, genomic DNA was extracted from the heart, lung and gut of normal and transgenic mice by phenol/chloroform extraction. Positive control methylated DNA was produced by incubating 1 μg genomic DNA with 10 u/μg DNA of *SssI* methylase and S-adenosyl methionine for 4 hours. 1μg DNA was digested with 10u of the methylation inhibited enzyme *BstUI* for 4 hours or underwent control digestion with buffer overnight, then both *BstUI* treated and control digests were digested overnight with 10 u *MseI*.

The digested DNA was PCR amplified at low specificity using short primers A1= AGCGGCCCGCG and either A2: ACCCCCAGCCG or A3: TGGGGTCGGCGC using HotStarTaq mastermix (Qiagen) with Q-solution (Qiagen) incorporating 0.2 μCi ^{32}P dCTP (Amersham Redivue) per reaction. Cycling conditions were (melting temperature 94°C 2 mins, annealing temperature 40°C 1 min , extension temperature 72°C 2min) x35 cycles with final extension time 10 min at 72°C. PCR products were separated by loading 5 μl on a 6% TBE/acrylamide gel and running at 1000V/60W until the bromophenol blue front from the loading dye reached the end of the plate. The gel was then taken up on to filter paper and dried for 1 hour at 70°C in a gel dryer before developing for 12-48 hrs by autoradiography.

Fragments showing abnormal methylation were excised from the gel. DNA was eluted by incubating excised gel fragments in 100 μl water for 1 hour, heating to 95°C for 10 minutes and centrifuging at 800 g for 5 mins. DNA was precipitated from the DNA by ethanol precipitation and resuspended in 10 μl water. 4 μl was re- amplified using the same PCR protocol and primers as the initial reaction (without ^{32}P dCTP incorporation). PCR products were separated on a 1% agarose gel and bands were excised, DNA extracted from the gel using Qiagen Minelute kit, and then cloned and sequenced as in Section 7.5.2. Products were identified by BLAST search on the NCBI website (http://www.ncbi.nlm.nih.gov/BLAST/) against the mouse genome. The methylation status of regions identified was confirmed by bisulphite sequencing.

7.3.2 Bisulphite modification of DNA

Bisulphite modification for confirmation of the MSRF experiments was performed according to the methods in (Hajkova, *et al* 2002). 600 μg genomic DNA was digested for 4 hours with *SacI*. Digested DNA was melted at 100°C for 5 mins then incubated at 50°C in a final concentration of NaOH of 0.3 M for 15 mins. NaOH treated DNA was embedded in 2 volumes of low melting point agarose at 50°C. 2.5M sodium bisulphite solution was prepared as described. 1 ml samples of bisulphite solution were overlaid

with mineral oil and chilled at -20°C before adding the DNA. The DNA/agarose solution was added to the cold bisulphite solution in 10µl aliquots by forming beads in the mineral oil layer. DNA was placed on ice for 20 mins then incubated at 50°C for 4 hours. The bisulphite solution was then removed and the agarose pellet washed 4x in TE buffer pH 8, 2x in NaOH 0.2M then TE buffer x3. Agarose pellets containing the bisulphite modified DNA were stored under 100µl TE buffer at 4°C.

Bisulphite modification for analysis of the *Apc* gene was performed using the EZ-DNA methylation kit (Cambridge Biosciences) according to the manufacturer's instructions.

7.3.3 Bisulphite sequencing

Bisulphite treated DNA was amplified using PCR primers surrounding the region of interest designed against the predicted bisulphite modified genomic sequence. Primers were designed so as not to include C residues immediately 5' to G anywhere in the primer sequence, in order to amplify methylated and unmethylated sequences with equal efficiency. Primer design was assisted by the web based design programs Methprimer: (http://www.urogene.org/methprimer/index1.html)
and Primer3 :
(http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi). Primer sequences are given in Table 7.1. Amplification was performed using HotStarTaq (Qiagen), for 45 cycles with cycling conditions: Denaturation 1min x 95°C, annealing 1 min (temperature in Table 1), extension 1 min x 72°C. A second round of hemi-nested PCR was necessary to obtain specific products in the case of *Lrfn5* and the CpG island *Apc* promoter. A fully nested second round of PCR was used for the *Apc* internal promoter when used for DNA extracted from paraffin embedded tissue. PCR products were separated on a 1% agarose gel, and bands of the correct size were cut from the gel and DNA extracted using Minelute columns (Qiagen). Purified DNA was ligated in to PGEMT-Easy vector, and cloned as in Section 7.5.2. Colonies were expanded in 96 well plates and

automated plasmid extraction and sequencing reaction with T7 primer was carried out using a Biomek 2000 Automated Workspace robot.

7.3.4 Quantitative PCR for methylation (M-QPR)

Approximately 1 μg of genomic DNA of each was digested overnight with methylation sensitive enzymes (10 u *HpaII* and 10 u *HhaI*) followed by heat inactivation at 65°C for 15 minutes.

Quantitative PCR for genomic DNA was performed in 384 well plates on an Applied Biosciences 7900 analyser using sybr green detection. Template DNA was amplified using Quantitect sybr green mastermix (Qiagen) with gene specific primers at a concentration of 1 μM with addition of Q solution (Qiagen) if this was found to improve efficiency in the validation experiments.

Primers were designed using Primer3 software (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) with optimal conditions specified – product length 100-200 bp, primer Tm 59-61°C, maximum primer Tm difference 0.5°C). Amplicons were chosen so as to include the maximum number of restriction sites for *HpaII* and *HhaI*, in order that the assay should be specific for regions that are densely methylated since a PCR product would only be produced if all sites included in the region are methylated.

Assays were validated by amplifying a standard curve of 4-fold dilutions of undigested genomic DNA from mouse tail-tips, using cycling conditions- initial activation 95°C for 15 mins, denaturation 95°C 15 sec, annealing 55°C 30 sec, extension 72°C 30 sec. Following 50 cycles of amplification, a DNA melting curve was performed to check that the product Tm was as expected and that there was little non-specific amplification (as defined by presence of products with multiple Tm values). Assays were chosen for experimental analysis if there was near 100% amplification efficiency in the standard

curve with a minimum of non-specific amplification. Primer sets used are detailed in Table 7.2. Examples of standard curves and dissociation curves are shown in Figure 7.6

DNA concentrations were standardised for the experimental reaction by performing an initial quantification of the digested DNA using the NOR control primer set (which has no *HpaII* restriction sites) and a check for complete digestion using primers in the unmethylated CpG island extending from the clathrin light chain (*Ctla*) promoter region. Samples were diluted on the basis of this initial assay in *HpaII* buffer to the same concentration. Samples giving a product in at least 2/3 replicates in the clathrin assay were excluded from the methylation analysis.

For the experimental determination of methylation, digested and standardized DNA as above was amplified using gene specific primers using the same conditions as in the validation step with quantification by measurement of sybr green fluorescence at 1°C below the product Tm at the end of each amplification cycle. Absolute quantities of DNA surviving digestion was measured by reference to a standard curve of undigested DNA run in the same plate using linear regression. All reactions were performed in triplicate, and reactions where a product was seen in at least 2/3 replicates were considered as positive. The quantity of DNA was recorded as the arithmetic mean of all successful replicates.

A "methylation index" was calculated as a measure of methylation by dividing the absolute quantity of DNA surviving digestion by the starting quantity of DNA as determined by amplification with the NOR control primer set. Theoretically this should give the fraction of DNA which is methylated at all restriction sites between the two primer sites. A measure of the confidence of the calculated methylation index can be given by a transformed standard error $10^{[\text{Log(std error of methylated copies)+Log(std error of initial copies)}]}$, which estimates the variation in the sample mean methylation index based on the standard error of the difference in PCR detection thresholds.

In order to validate the assay, the methylation status of the *Wnt3* and *Lrfn5* promoter CpG islands in the heart were measured, since the methylation status of these regions had been shown by bisulphite sequencing (Section 3.3).

The *Wnt3* assay appears to be particularly robust at a qualitative level, with clear differences between the normal and transgenic animals which is consistent with the bisulphite data (Figure 7.7a). The methylation index does appear to be of a similar order of magnitude to the percentage of densely methylated clones in the bisulphite sequencing data, but there is insufficient data at this stage to confirm the accuracy of quantification of methylation levels

The *Lrfn5* assay is less robust since there is amplification of non specific products which will alter the efficiency of amplification of the intended product. Nevertheless, there is a clearly significant difference in methylation index between the wild type and transgenic animals (Figure 7.7 b). In addition, there is a qualitative difference in the dissociation curves, presumably as the increased quantity of methylated template in the transgenic gives a selective advantage to amplification of the correct product (Figure 7.7c). The assay is therefore adequate as a comparative assay to look at differences in methylation levels between paired sample groups, and the lack of a detectable PCR product probably does indicate that the sample studied is unmethylated, but positive results may need confirmation by bisulphite sequencing if accurate quantification is needed.

As an overview of the comparison with the bisulphite sequencing data, therefore, the methylation index appears to underestimate the proportion of methylated chromosomes in an a sample in which the region studied is genuinely methylated in a proportion of tissues. This is unsurprising as the test is fairly stringent, requiring all restriction sites to be methylated to achieve a PCR product, thus quantifying only the most densely methylated alleles. Conversely, there is likely to be a false positive rate of detection, either as a result of non-specific PCR products or as a result of inadequate digestion which is below the sensitivity of the clathrin screening test. In addition, since the

numbers of restriction sites differ between PCR amplicons, it is not possible to use the technique to compare methylation levels between different loci in a quantitative way, and the technique is most valid for comparison of differences in methylation at a particular locus between an experimental and control sample.

7.3.5 Nearest neighbour analysis

Quantification of global CpG methylation levels by nearest neighbour analysis was performed by Dr Bernard Ramsahoye, John Hughes Bennett Laboratory, Edinburgh according to the methods in (Ramsahoye, *et al* 2000): Briefly, genomic DNA digested with *MseI* (TTAA) and separated on a 1% agarose gel. The 2-4 kb fraction was cut from the gel in order to select for CpG island- rich DNA regions. DNA was extracted from the gel using Qiagen Minelute kit according to the manufacturer's instructions. The extracted DNA was then digested with the restriction enzyme *MboI* (N||GATC). Following heat inactivation, ethanol precipitation and resuspension, the digested DNA was filled in with Klenow polymerase and α-^{32}P dGTP. The reaction was terminated by addition of EDTA and unincorporated nucleotides removed using a Sephadex spin column. DNA was digested to single nucleotides using micrococcal nuclease and phosphodiesterase. The digested nucleotides were separated by thin layer chromatography. Quantification of nucleotides was performed by autoradiography and densitometry analysis of the autoradiographs.

7.4 Microarray Analysis

Total RNA was prepared from the heart of 1 wild type and 1 transgenic mouse, both aged 130 days, by phenol extraction.

Microarray hybridisation was performed by Dr Kevin Robertson University of Edinburgh. RNA quality was initially assessed using formaldehyde gels and by measurement of absorbance at 260/280 nm following standard procedures. A further

assessment of RNA quality was then undertaken using an Agilent Bioanalyser (Agilent Technologies, GmBH).

Total RNA was prepared for microarray analysis as follows. In brief, 10 μg of total RNA from each sample was used to generate double stranded cDNA using an oligodT primer containing a T7 promoter. This template was used to generate amplified, biotinylated cRNA using a Bioarray High Yield RNA Transcription Labeling Kit (Enzo Diagnostics, Farmingdale, NY, USA). After confirming that all cRNA samples were of good quality using the Agilent Bioanalyser, 10 μg of labelled cRNA was hybridized to Affymetrix MG-U74Av2 GeneChip arrays for 16 hours at 45^0C. Arrays were then washed in an Affymetrix Fluidics Station 400 (Affymetrix, USA) stained with streptavidin-R phycoerythrin (Molecular Probes, USA) and biotinylated goat anti-streptavidin antibody (Sigma Chemical, St. Louis, MO, USA) and scanned according to standard protocols.

All analyses were undertaken in Affymetrix Microarray Suite 5.1 software as follows: arrays were scaled (normalised) to an overall target intensity of 100 prior to comparison and notable differences in gene expression were identified by a 'comparative analysis' of the two samples (detailed in http://www.affymetrix.com/support/technical/whitepapers/sadd_whitepaper.pdf). Filtering was undertaken on the basis of the detection call. The detection call algorithm is used to provide an indication of whether the expression level of a transcript is below the threshold of detection for the system. As an output of this analysis step, a value for the probability that a transcript expression level is provably different from zero is generated. If a transcript has a p value of less than 0.04 then it is 'called' present, between 0.04 and 0.06 it is called marginal and greater than 0.06 it is called absent.

7.5 Basic molecular biology techniques

7.5.1 Western blot for Dnmt3b

Protein extracts were loaded in 10% acrylamide precast gels (Biorad) and were separated at 75 V by electrophoresis in SDS buffer. Proteins were transferred on to nitrocellulose membrane at 14 V using semi-dry transfer cell with semi-wet buffer. Membranes were probed by blocking in 5% low fat milk (Marvel) in TBS-T for 1 hour then incubating with polyclonal sheep anti-Dnmt3b serum (raised against the N-terminal 200 amino acids) diluted 1:1500 in 5% milk/TBS-T at room temperature on a rotary shaker. Membranes were washed in TBS-T for 2x15 minutes then incubated at room temperature for 1 hour with HRP-conjugated donkey anti-sheep secondary antibody (Sigma) at 1:5000 concentration. Following 2x15 min washes in TBS-T, membranes were developed by covering the membranes with approx 1ml ECL solution and exposing X-ray film (Kodak).

If a loading control was required, membranes were re-probed using rabbit polyclonal antiserum raised against the C-terminus of histone H3 as primary antibody (produced by Alain Verrault, University of Montreal) at 1:30000 dilution, and HRP-conjugated sheep anti-rabbit (Sigma) in 1:5000 concentration as secondary antibody. Some of the western blots in Figure 2.3 were performed by my lab technician Jayne Culley.

7.5.2 Cloning PCR products

PCR products for cloning were ligated into PGEMT-easy vector using the PGEMT-easy vector system I kit (Promega) according to the manufacturers instructions. Ligated vector was added to 20 µl JM109 competent cells (Promega) and incubated for 15 minutes on ice before subjecting to heat shock for 45 seconds at 42°C. Mixtures were incubated on ice for 2 minutes before adding 80µl LB broth and incubating at 37°C for

1hour with shaking. Following incubation the mixture was spread on to LB-agar plates containing 100 mg/ml ampicillin with 35 µl of 20 mg/ml X-gal and 12ul of 0.2 g/ml IPTG per plate. Plates were incubated at 37°C overnight. White coloured colonies were picked with a sterile toothpick and incubated in 3ml LB broth with 100mg/ml ampicillin overnight. Bacteria were pelleted by centrifugation and plasmid extracted using the Qiagen Qiaprep Spin kit according to the manufacturer's instructions.

7.5.3 Sequencing

Sequencing of plasmids was performed by adding 5 µl of plasmid to 0.8 µl of BigDye 3.1 terminator mix (ABI), 4 µl Better Buffer (Web Scientific), 1.2 µl distilled water and 1 µl 4 nM T7 sequencing primer (GTAATACGACTCACTATAGGGC) for PCR products or sequence specific primers as detailed in the relevant sections. Sequencing reactions were performed using the following conditions: Initial denaturation- 96°C 1 minute, (melting 96°C 20 sec, annealing 50°C 10 sec extension 60°C 4 min)x 25 cycles, final extension time 60°C 4 mins. Products were ethanol precipitated, washed once in 70% ethanol and air dried for 1 minute at 90°C before storing at -20°C while awaiting reading. Sequence was read using an Applied Biosystems 3100 genetic analyzer.

7.5.4 Quantitative RT-PCR

RNA was extracted from tissues of *Tg(Dnmt3b)+* mice and age matched wild-type C57Bl6 mice by phenol/chloroform extraction. Approximately 1µg RNA was used for cDNA production using the Promega RT-PCR kit according to the manufacturer's instructions using oligo-dT as template.

Gene expression of *Wnt3* and *Lrfn5* was measured by quantitative real time RT-PCR using Taqman sequence detection. 10 ng of cDNA was amplified using Taqman mastermix (Applied Biosciences) and standard off-the-shelf probe/primer sets from

Applied Biosciences (reference numbers: *Wnt3*= Mm00437336_m1; *Lrfn5*=Mm00724566_m1). Amplification was performed for 50 cycles at a melting temperature of 95°C for 15 seconds and a combined annealing/extension temperature of 60°C for 1 minute. Expression levels were calculated relative to beta-actin expression measured by quantitative RT-PCR using a standardised control probe/primer set (Applied Biosciences ref: 4352933E). All reactions were performed in triplicate, and results are expressed as the mean of all successful replicates. Gene expression was considered as detectable if a PCR product was detected with at least 2/3 replicates.

7.5.5 DNA extraction

Tissue samples were disrupted by freezing in liquid nitrogen then pulverising with a hammer. Samples were digested in DNA extraction buffer containing 1mg/ml proteinase K at 50°C until all solid material had dissolved. DNA was extracted by adding an equal volume of pH 8.0 Tris saturated phenol. The aqueous phase was taken off and phenol extraction was repeated once. The excess phenol was removed from the aqueous phase of the second extraction by adding an equal volume of 49:1 chloroform/IAA. The aqueous phase was taken off and DNA was recovered from this by ethanol precipitation. The pellet was resuspended in 10 mM Tris pH 8.0.

DNA extraction from mouse tail tips was performed using the Qiagen DNA Easy kit according to the manufacturer's instructions.

7.5.6 RNA extraction

Tissue samples were disrupted by freezing in liquid nitrogen then pulverising with a hammer. 2 ml of Solution D (see materials) was added to the tissue sample and then the mixture was homogenised using a rotor homogeniser or by multiple passage through a 19 gauge needle and syringe. 500 µl aliquots were either stored at -70°C or RNA was extracted immediately by phenol extraction. 50 µl of 2 M sodium acetate pH 4.0, 500 µl

water saturated phenol and 100 µl 49:1 chloroform:IAA was added to the homogenate and mixed by vortexing. The mixture was centrifuged at 13000 RPM for 20 mins at 4°C. The aqueous layer was removed by pipetting and extracted a second time with phenol/chloroform as above. The aqueous phase from the second extraction was added to 500 µl chloroform/IAA, mixed by vortexing, and separated by centrifugation for 1 minute. The aqueous phase was ethanol precipitated before resuspending in water.

7.5.7 Ethanol precipitation of nucleic acids

2.5 volumes of 100% ethanol and 0.1 volumes of 3M sodium acetate (pH 5.4 for DNA, pH 4 for RNA) were added. The mixture was allowed to precipitate for 1 hour at -70°C before centrifuging at 13000 RPM for 20 mins. The pellet was washed in 100 µl 70% ethanol and centrifuged for 5 minutes, before resuspending in water or Tris buffer as required.

7.6 Bioinformatics and Statistics

Sequence analysis was performed using the program Bioedit (available as freeware from http://www.mbio.ncsu.edu/BioEdit/bioedit.html). Multiple sequence alignments were performed using the ClustalW algorithm (http://www.ebi.ac.uk/clustalw/). Bisulphite sequencing data was analysed using a PERL script (source in Appendix). Statistical analysis was performed using the Analyse-It plug-in for Microsoft Excel (http://www.analyse-it.com/) or XLStat (Addinsoft). Numerical simulations for bootstrap analysis were performed using Mathematica (Wolfram Research). Source code for the bootstrap analysis is available from http://library.wolfram.com/infocenter/MathSource/4272/. All statistical tests are two tailed.

7.7 Materials

All reagents, unless otherwise specified were obtained from Sigma. Oligonucleotides were ordered from MWG and restriction enzymes were obtained from New England Biolabs. Radionucleotides were obtained from Amersham.

DNA extraction buffer: 10 mM Tris pH 7.4, 100 mM EDTA, 0.5% w/v SDS.

ECL solution was prepared as 2 stock solutions: Solution 1: 100 mM Tris pH 8.5 to which was added 1ml 250 mM Luminol in DMSO and 440 µl coumaric acid per 100 ml. Solution 2: 100 mM Tris pH 8.5 to which was added 64 µl 30% Hydrogen peroxide per 100 ml. The working solution was prepared by mixing equal volumes of solutions 1 and 2 immediately before use.

SB solution was made up as a 4x stock solution (40 ml glycerol, 12 g SDS, 25 ml 1M Tris-HCl pH 6.8, 20 ml 2-mercaptoethanol made up to 100 ml with distilled water). Stock solution was diluted with distilled water and bromophenol blue added as required.

SDS buffer: 12 g/l Tris, 57.6 g/l Glycine, 4g/l SDS.

Semi wet transfer buffer: 5.8 g/l Tris, 2.9 g/l Glycine, 0.38 g/l SDS, 20% v/v Methanol.

Solution D: 4 M Guanidium thiocyanate, 25 mM sodium thiocyanate, 0.5% Sarkosyl, 0.1 M 2-mercaptoethanol. Stock solutions without mercaptoethanol were prepared and stored at -20°C. Mercaptoethanol was added to stock solution before use.

TBS-T: 8 g/l NaCl, 0.2g/l KCl, 6g/l Tris, 0.1% v/v Tween20. pH adjusted to 7.4.with HCl.

7.8 Figures and tables

Gene	Primers	Product (bp)	Region amplified	Anneal Temp(°C)
Hop(CpG island1)	ttttagataggaattttattaataaaggg ttattcaaatttaatacttaaataataact	421	T-216 to+205	50
Hop(CpG island 2)	aaagagtaaatatttattggtagtattgtt ctaactacaaaatttaaaaaaaaaccctcct	370	See text	50
Ces3	gtaggtatttggttgttgttgtttgttt ctctataaaatctctaaactatctactac	351	T-322 to +29	48
Viaat	tagaaagttttggtttattat aaaaaaaacttctccaacacttc	512	Exon2 328 to 841	48
Wnt3	gaaggggtgttttaggttagtag caaataaaactccattaaaaacaa	407	T-357 to+50	48
Lrfn5 (1st round	gtgattaaagggttatgtgagaa ctacccttaccataactcaacac	578	T-442 to +136	48
Lrfn5 (heminested)	gtgattaaagggttatgtgagaa aaataactcctacccctacctttc	332	T-442 to -110	48
Apc (internal promoter) 1st ROUND	tgttgtatttgaattatgggaatg ccaaaaaaacaccaaacctcttat	183	T-385 to -202	50
Apc (internal promoter) (2ND ROUND nested)	tgggaatgaaatttatttattgttttt aaacaccaaacctcttatacacctt	160	T-369 to -198	46
Apc (CpG island promoter) (1ST ROUND)	gtagaggtagggtataggttgttg ctaaaaaaaccactcctcactcc	295	T-421 to -126	50
Apc (CpG island promoter) (2ND ROUND heminested)	gtagaggtagggtataggttgttg aaactcccccaaaataac	196	(T-421 to -225)	46

Table 7.1 Bisulphite amplification primers. Details of primers used for PCR amplification of bisulphite modified DNA, regions amplified and annealing temperatures used. T– indicates distance from transcriptional start site.

Gene	Primers	Amplicon start	Amplicon length	HpaII sites in product	HhaI sites in product	Product Tm (C)
NOR control	aacgagccccataaaatagaaatac tctatgttgctgtatcagtgctagg	*	114	0	0	76
Clathrin	gttgagttgggctgtcatgg cccgcaatctcactctcttg	T+97	149	4	4	85*
Wnt3	cttcccagagccggagtt gactttcgggcttgtcagag	T-244	193	4	4	84*
Lrfn5	ggatcttgagaaggaccagaact gctcacacacactcacacattct	T-426	110	3	4	80*
Hic1	cctcttggcccaggtgtg cccggtcggtctgtcttg	T-188	197	3	3	91*
Cdkn1a (p21)	caagagaatagcccaggtgtg tcgagctgcctccttatagc	T-205	192	2	5	87
cdkn2a (p19)	ggcgtcaggcatagacgaaagt cttcgctcgccacccatc	T-420	171	3	4	91
Cdkn1b (p27)	gtttgttggcagtcgtacacc aggaggagctccattggtct	T-219	175	2	9	89*
Cdkn1c (p57)	acgttaccgcccgcagag ggcgattggagggtgcag	T-302	210	2	5	89*
Mlh1	cctattggctggagatttcg ctgattgggcagcatgaat	T-258	187	0	3	84
Igf2r	ttgtcaggcctcgagtaggta gtcctcgttcaggtgctctc	T-212	115	2	3	76*
Apc CpG promoter	ccggagcaaagaccagaagag atcttgttggaggcgggaag	T-574	156	3	3	86*

Table 7.2 M-QPR primers Details of primers used for real time PCR amplification for M-QPR, regions amplified, number of restriction sites in amplicon and melting temperature (Tm) of PCR product. Amplification data for each cycle was recorded at 2°C below the melting temperature. * indicates PCR was performed with addition of Q solution, Tm quoted is Tm including Q solution.

Figure 7.1 Preparation of Dnmt3b expression constructs

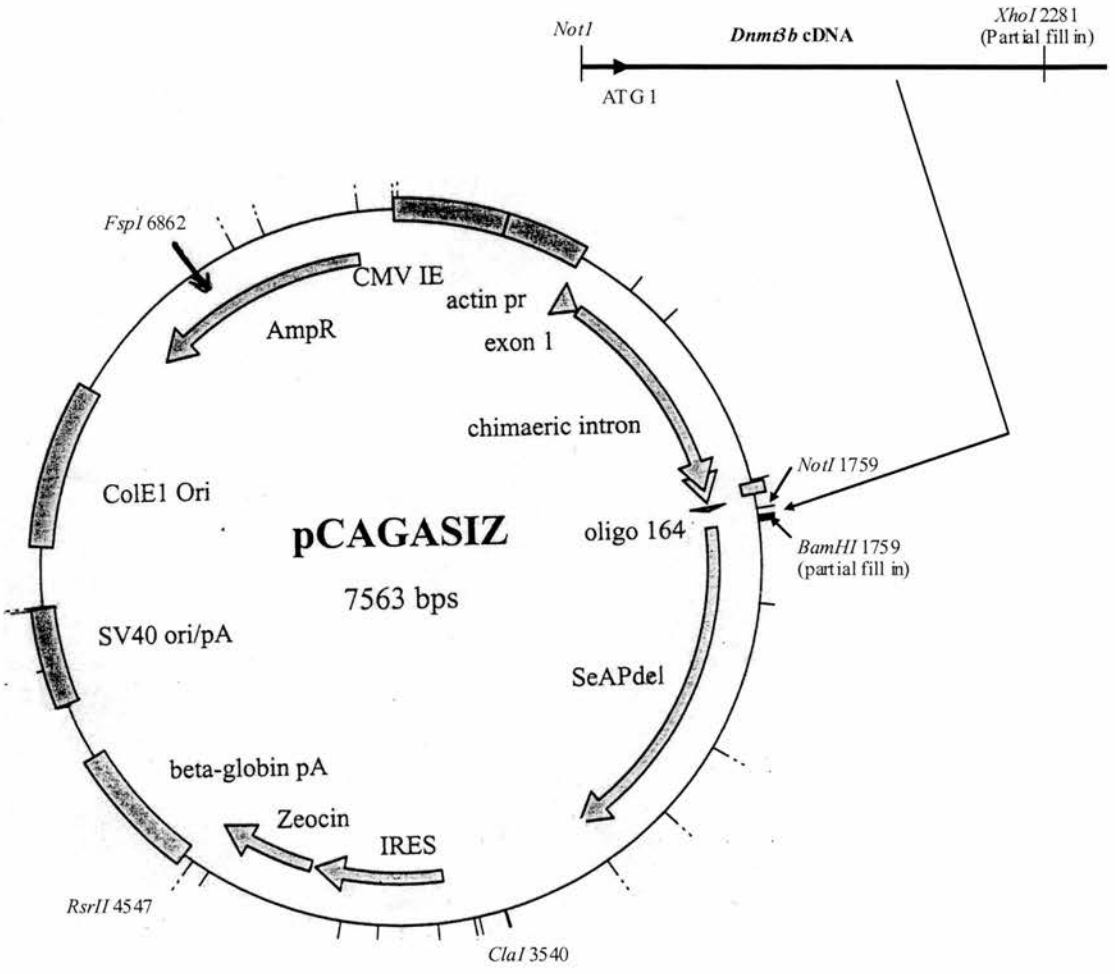


Figure 7.1 Preparation of Dnmt3b expression constructs. Strategy for production of Dnmt3b overexpression construct. CMV IE– CMV enhancer. IRES interribosomal entry site. pA– polyA. Co1E1– *E.Coli* replication origin.
AmpR -Ampicillin resistance gene. Zeocin– Zeocin resistance gene.

Figure 7.2 Production of *Tg(Dnmt3b)*+ mice

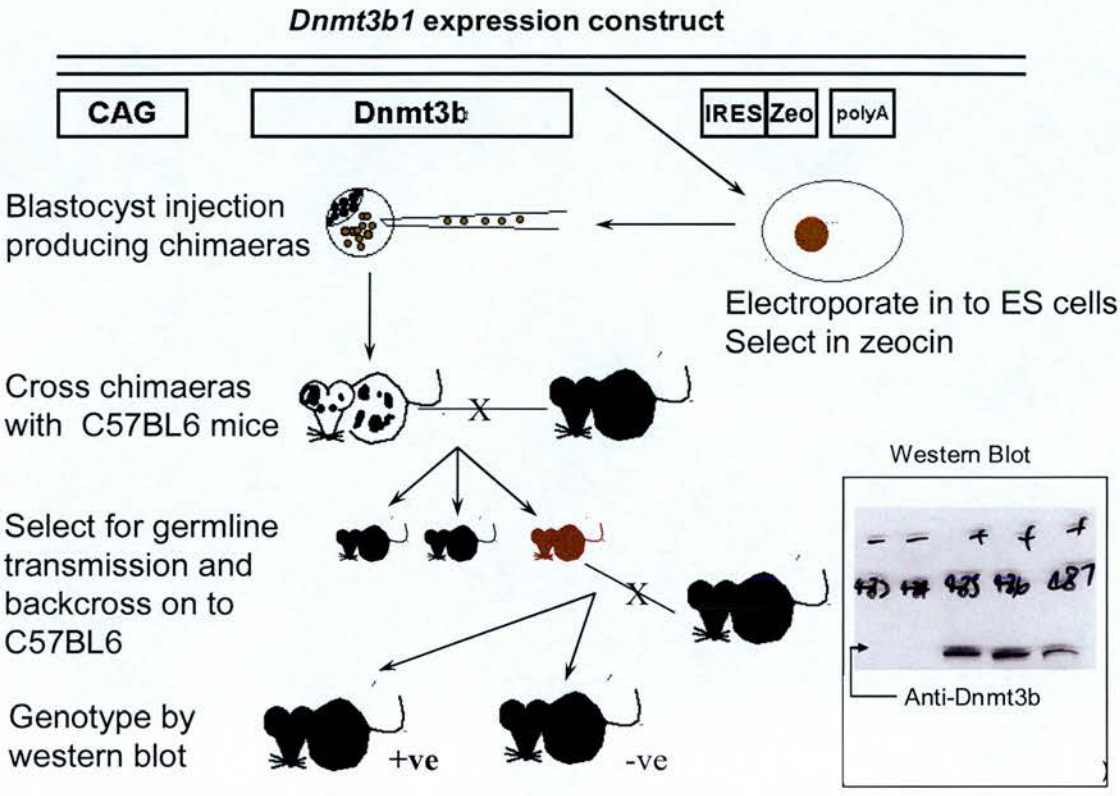
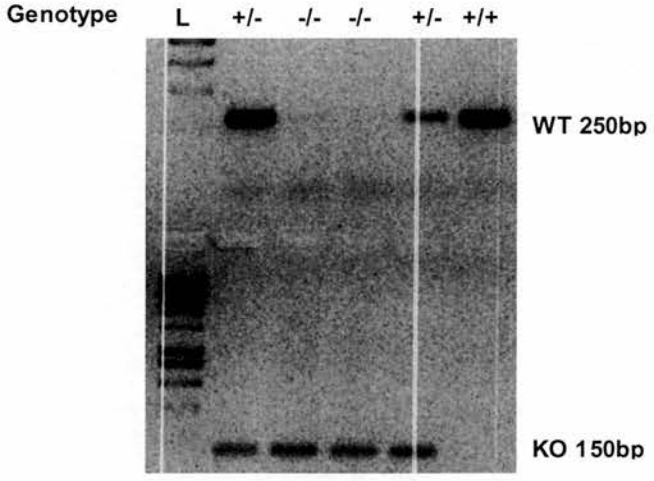


Fig 7.2 Production of *Tg(Dnmt3b)*+ mice. CAG-CAG promoter IRES interribosomal entry site
Zeo– zeocin resistance gene. Dnmt3b– full length Dnmt3b1 cDNA

A) *Mbd2* mice



B) *Apc*^{min} mice

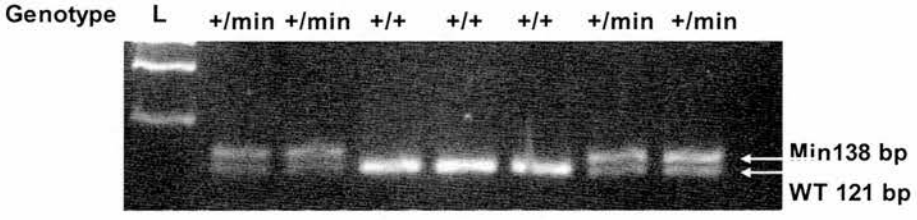


Figure 7.3 *Apc* and *Mbd2* genotyping. Ethidium bromide stained agarose gel electrophoresis showing typical results from genotyping mice for A) *Mbd2* mice B) *Apc*^{min}. L=size marker

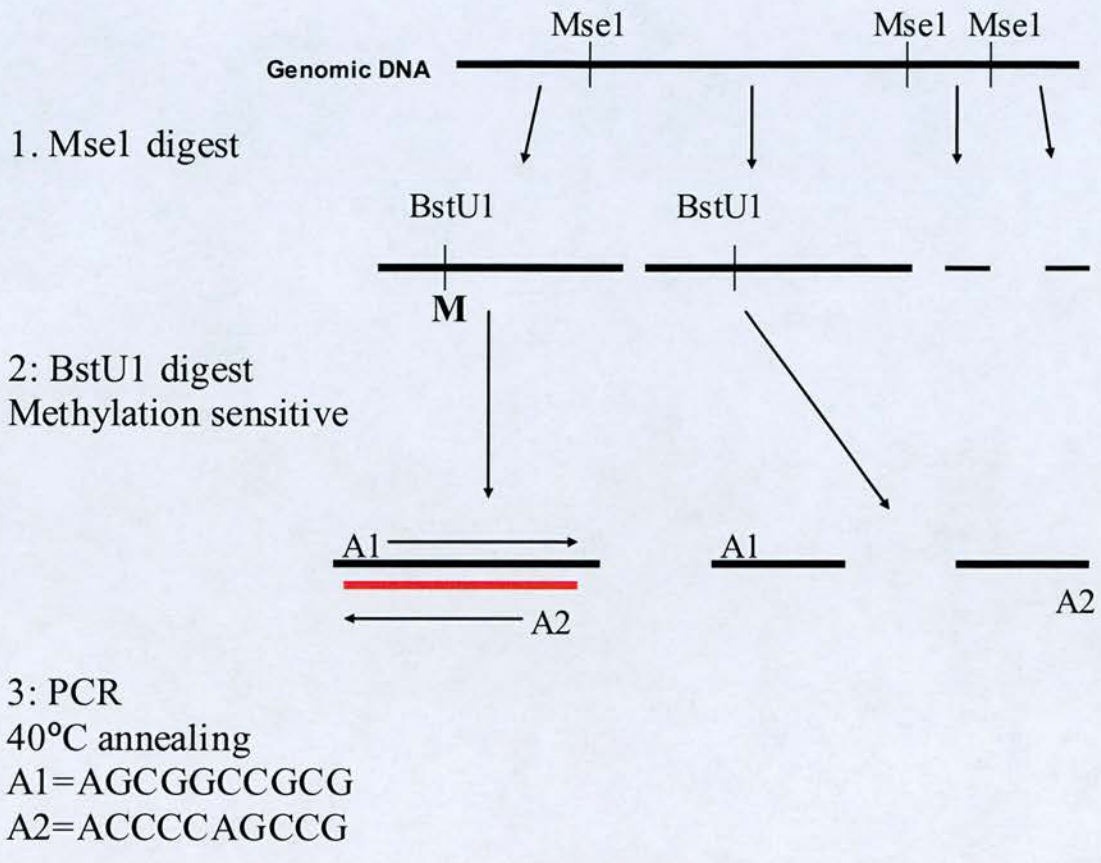


Figure 7.4 Schematic summary of methylation sensitive restriction fingerprinting (MSRF)

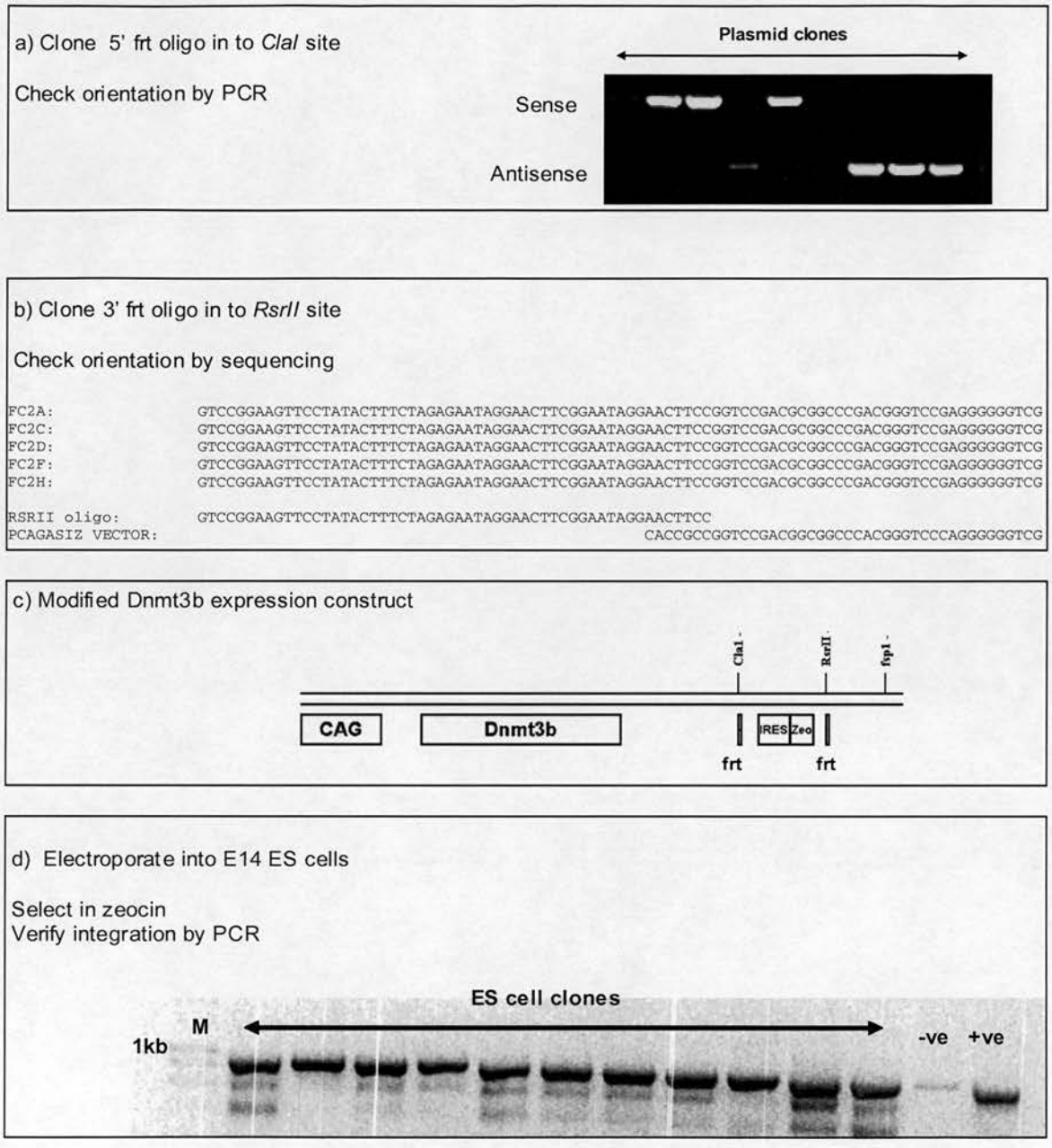


Figure 7.5 Production of second line of *Tg(Dnmt3b)* mice.

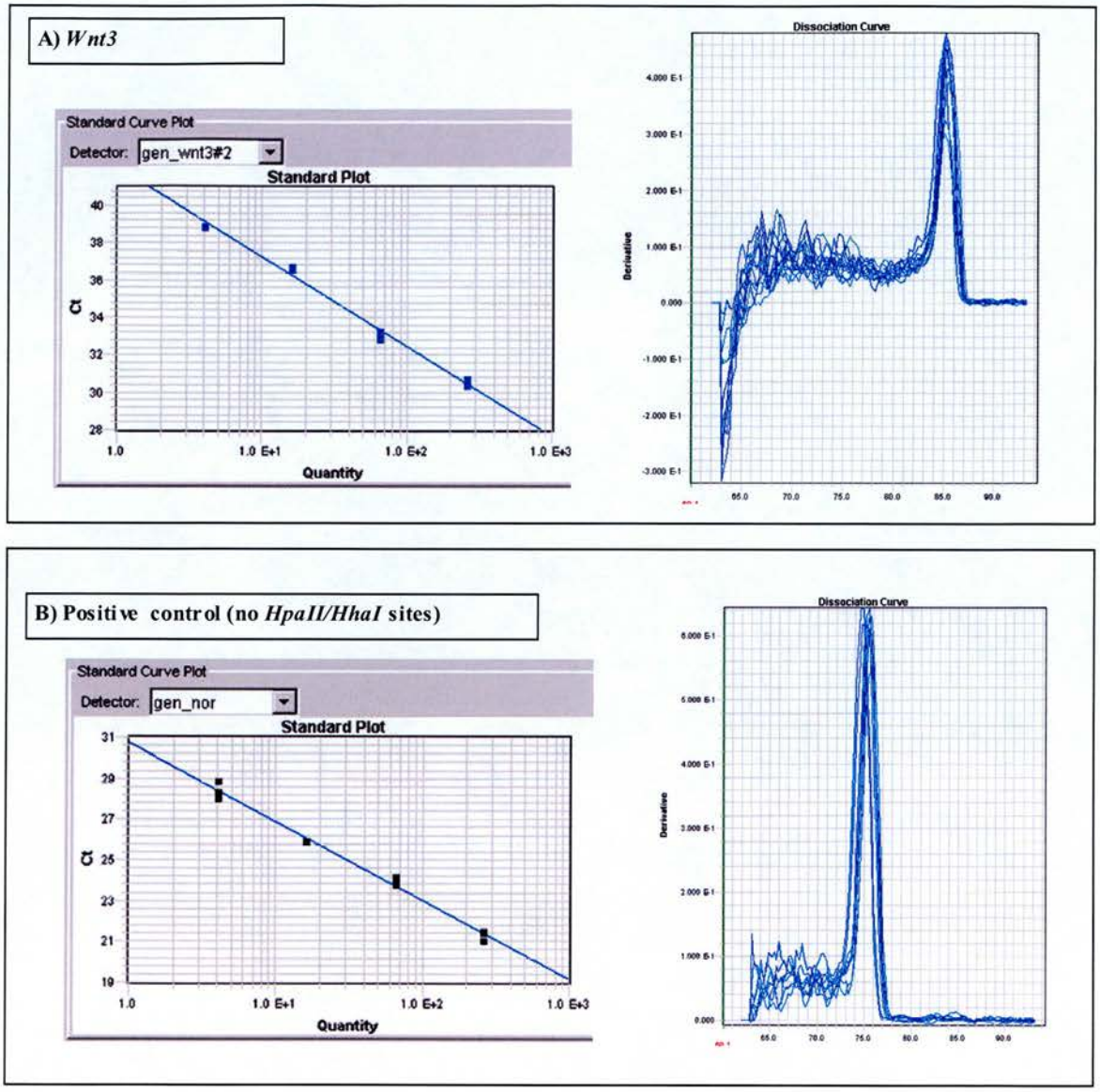


Figure 7.6 Quantitative PCR standard curves/ dissociation curves. Examples of validation experiments for sybr green real time PCR experiments. Left hand figures show standard curves for 4-fold dilutions of undigested genomic DNA. Ct=PCR cyle where detection threshold is reached. Quantity-arbitrary units (initial dilution=256 arbitary units). Right hand figures show dissociation curves for PCR product following amplification- 1st differential of sybr green fluorescence against temperature as temperature is raised. Melting temperature is estimated from the maximum of the curve.

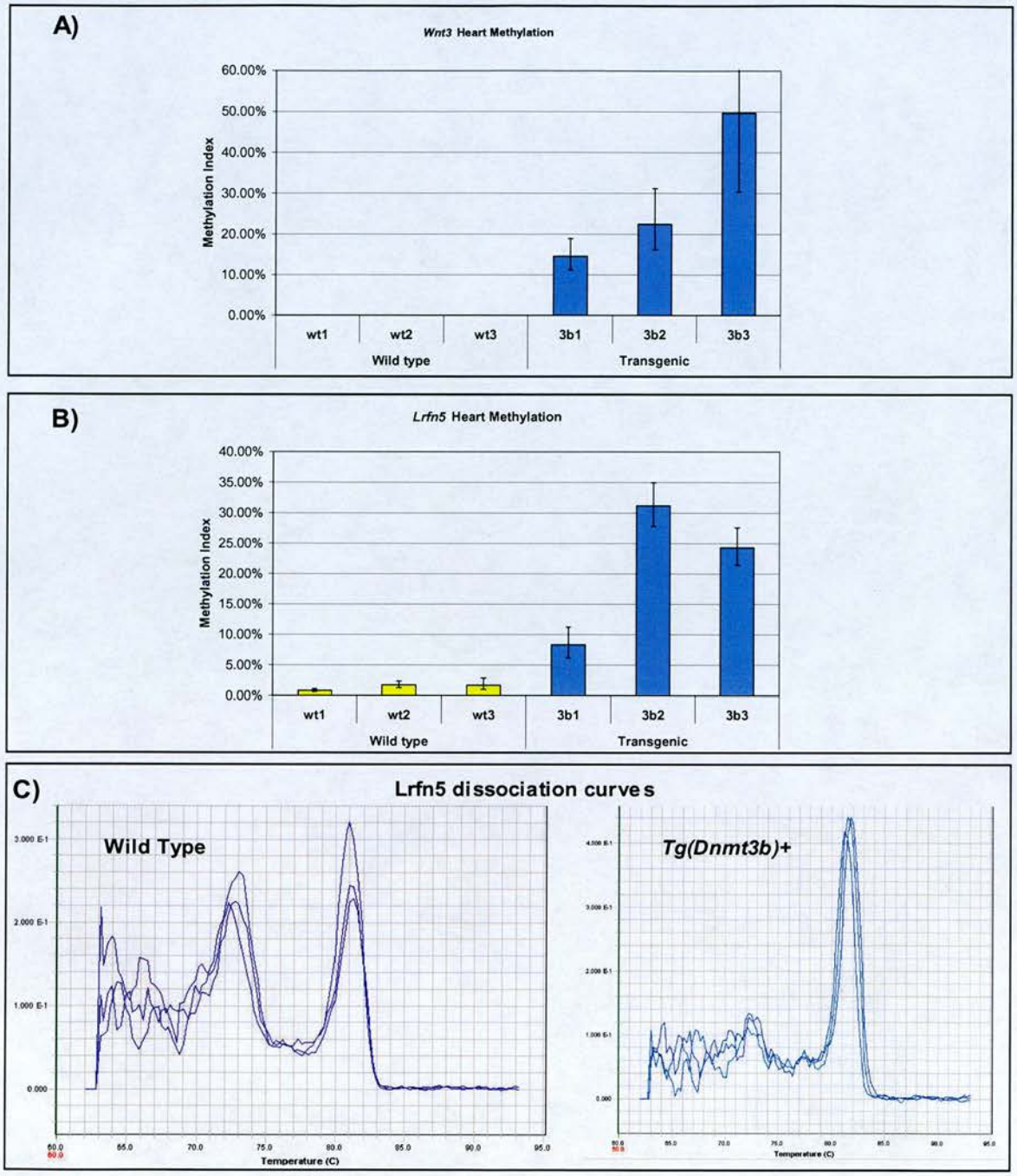


Figure 7.7 Validation of M-QPR. Measurement of methylation in heart of known methylated genes by M-QPR. A) *Wnt3* B)*Lrfn5*. Error bars represent transformed standard errors. C) shows dissociation curve for *Lrfn5* PCR products, (differential of sybr green fluorescence with respect to temperature), showing qualitative difference between transgenic (methylated template) and wild type samples.

8. References

Aapola, U., Lyle, R., Krohn, K., Antonarakis, S.E. & Peterson, P. (2001) Isolation and initial characterization of the mouse Dnmt3l gene. *Cytogenet Cell Genet,* **92,** 122-126.

Aronow, B.J., Toyokawa, T., Canning, A., Haghighi, K., Delling, U., Kranias, E., Molkentin, J.D. & Dorn, G.W., 2nd (2001) Divergent transcriptional responses to independent genetic causes of cardiac hypertrophy. *Physiol Genomics,* **6,** 19-28.

Bachman, K.E., Park, B.H., Rhee, I., Rajagopalan, H., Herman, J.G., Baylin, S.B., Kinzler, K.W. & Vogelstein, B. (2003) Histone modifications and silencing prior to DNA methylation of a tumor suppressor gene. *Cancer Cell,* **3,** 89-95.

Belinsky, S.A., Klinge, D.M., Stidley, C.A., Issa, J.P., Herman, J.G., March, T.H. & Baylin, S.B. (2003) Inhibition of DNA methylation and histone deacetylation prevents murine lung cancer. *Cancer Res,* **63,** 7089-7093.

Bell, A.C. & Felsenfeld, G. (2000) Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. *Nature,* **405,** 482-485.

Bestor, T., Laudano, A., Mattaliano, R. & 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,** 971-983.

Biniszkiewicz, D., Gribnau, J., Ramsahoye, B., Gaudet, F., Eggan, K., Humpherys, D., Mastrangelo, M.A., Jun, Z., Walter, J. & Jaenisch, R. (2002) Dnmt1 overexpression causes genomic hypermethylation, loss of imprinting, and embryonic lethality. *Mol Cell Biol,* **22,** 2124-2135.

Bird, A. (2002) DNA methylation patterns and epigenetic memory. *Genes Dev,* **16,** 6-21.

Birger, Y., Shemer, R., Perk, J. & Razin, A. (1999) The imprinting box of the mouse Igf2r gene. *Nature,* **397,** 84-88.

Blaxall, B.C., Spang, R., Rockman, H.A. & Koch, W.J. (2003) Differential myocardial gene expression in the development and rescue of murine heart failure. *Physiol Genomics,* **15,** 105-114.

Bourc'his, D., Xu, G.L., Lin, C.S., Bollman, B. & Bestor, T.H. (2001) Dnmt3L and the establishment of maternal genomic imprints. *Science,* **294,** 2536-2539.

Brembeck, F.H., Rosario, M. & Birchmeier, W. (2006) Balancing cell adhesion and Wnt signaling, the key role of beta-catenin. *Curr Opin Genet Dev,* **16,** 51-59.

Brenner, C., Deplus, R., Didelot, C., Loriot, A., Vire, E., De Smet, C., Gutierrez, A., Danovi, D., Bernard, D., Boon, T., Pelicci, P.G., Amati, B., Kouzarides, T., de Launoit, Y., Di Croce, L. & Fuks, F. (2005) Myc represses transcription through recruitment of DNA methyltransferase corepressor. *Embo J,* **24,** 336-346.

Busslinger, M., Hurst, J. & Flavell, R.A. (1983) DNA methylation and the regulation of globin gene expression. *Cell,* **34,** 197-206.

Castanotto, D., Tommasi, S., Li, M., Li, H., Yanow, S., Pfeifer, G.P. & Rossi, J.J. (2005) Short hairpin RNA-directed cytosine (CpG) methylation of the RASSF1A gene promoter in HeLa cells. *Mol Ther,* **12,** 179-183.

Chen, R.Z., Pettersson, U., Beard, C., Jackson-Grusby, L. & Jaenisch, R. (1998) DNA hypomethylation leads to elevated mutation rates. *Nature,* **395,** 89-93.

Chen, T. & Li, E. (2004) Structure and function of eukaryotic DNA methyltransferases. *Curr Top Dev Biol,* **60,** 55-89.

Chen, T., Ueda, Y., Dodge, J.E., Wang, Z. & Li, E. (2003a) Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b. *Mol Cell Biol,* **23,** 5594-5605.

Chen, T., Ueda, Y., Xie, S. & Li, E. (2002) A novel Dnmt3a isoform produced from an alternative promoter localizes to euchromatin and its expression correlates with active de novo methylation. *J Biol Chem,* **277,** 38746-38754.

Chen, W.Y., Zeng, X., Carter, M.G., Morrell, C.N., Chiu Yen, R.W., Esteller, M., Watkins, D.N., Herman, J.G., Mankowski, J.L. & Baylin, S.B. (2003b) Heterozygous disruption of Hic1 predisposes mice to a gender-dependent spectrum of malignant tumors. *Nat Genet,* **33,** 197-202.

Clark, S.J. & Melki, J. (2002) DNA methylation and gene silencing in cancer: which is the guilty party? *Oncogene,* **21,** 5380-5387.

Cormier, R.T. & Dove, W.F. (2000) Dnmt1N/+ reduces the net growth rate and multiplicity of intestinal adenomas in C57BL/6-multiple intestinal neoplasia (Min)/+ mice independently of p53 but demonstrates strong synergy with the modifier of Min 1(AKR) resistance allele. *Cancer Res,* **60,** 3965-3970.

Cross, S.H., Meehan, R.R., Nan, X. & Bird, A. (1997) A component of the transcriptional repressor MeCP1 shares a motif with DNA methyltransferase and HRX proteins. *Nat Genet,* **16,** 256-259.

Dai, Z., Lakshmanan, R.R., Zhu, W.G., Smiraglia, D.J., Rush, L.J., Fruhwald, M.C., Brena, R.M., Li, B., Wright, F.A., Ross, P., Otterson, G.A. & Plass, C. (2001) Global methylation profiling of lung cancer identifies novel methylated genes. *Neoplasia,* **3,** 314-323.

Davies, C.S. (2002) Methylation-sensitive restriction fingerprinting. *Methods Mol Biol,* **200,** 43-51.

de Ruijter, A.J., van Gennip, A.H., Caron, H.N., Kemp, S. & van Kuilenburg, A.B. (2003) Histone deacetylases (HDACs): characterization of the classical HDAC family. *Biochem J,* **370,** 737-749.

De Smet, C., Lurquin, C., Lethe, B., Martelange, V. & Boon, T. (1999) DNA methylation is the primary silencing mechanism for a set of germ line- and tumor-specific genes with a CpG-rich promoter. *Mol Cell Biol,* **19,** 7327-7335.

Deng, C., Zhang, P., Harper, J.W., Elledge, S.J. & Leder, P. (1995) Mice lacking p21CIP1/WAF1 undergo normal development, but are defective in G1 checkpoint control. *Cell,* **82,** 675-684.

Di Croce, L., Raker, V.A., Corsaro, M., Fazi, F., Fanelli, M., Faretta, M., Fuks, F., Lo Coco, F., Kouzarides, T., Nervi, C., Minucci, S. & Pelicci, P.G. (2002) Methyltransferase recruitment and DNA hypermethylation of target promoters by an oncogenic transcription factor. *Science,* **295,** 1079-1082.

Dodge, J.E., Ramsahoye, B.H., Wo, Z.G., Okano, M. & Li, E. (2002) De novo methylation of MMLV provirus in embryonic stem cells: CpG versus non-CpG methylation. *Gene,* **289,** 41-48.

Douglas, D.N., Dolinsky, V.W., Lehner, R. & Vance, D.E. (2001) A role for Sp1 in the transcriptional regulation of hepatic triacylglycerol hydrolase in the mouse. *J Biol Chem,* **276,** 25621-25630.

Eads, C.A., Nickel, A.E. & Laird, P.W. (2002) Complete genetic suppression of polyp formation and reduction of CpG-island hypermethylation in Apc(Min/+) Dnmt1-hypomorphic Mice. *Cancer Res,* **62,** 1296-1299.

Eggan, K., Baldwin, K., Tackett, M., Osborne, J., Gogos, J., Chess, A., Axel, R. & Jaenisch, R. (2004) Mice cloned from olfactory sensory neurons. *Nature,* **428,** 44-49.

Ehrlich, M., Norris, K.F., Wang, R.Y., Kuo, K.C. & Gehrke, C.W. (1986) DNA cytosine methylation and heat-induced deamination. *Biosci Rep,* **6,** 387-393.

Esteller, M. (2002) CpG island hypermethylation and tumor suppressor genes: a booming present, a brighter future. *Oncogene,* **21,** 5427-5440.

Esteller, M., Fraga, M.F., Guo, M., Garcia-Foncillas, J., Hedenfalk, I., Godwin, A.K., Trojan, J., Vaurs-Barriere, C., Bignon, Y.J., Ramus, S., Benitez, J., Caldes, T., Akiyama, Y., Yuasa, Y., Launonen, V., Canal, M.J., Rodriguez, R., Capella, G., Peinado, M.A., Borg, A., Aaltonen, L.A., Ponder, B.A., Baylin, S.B. & Herman, J.G. (2001) DNA methylation patterns in hereditary human cancers mimic sporadic tumorigenesis. *Hum Mol Genet,* **10,** 3001-3007.

Fatkin, D. & Graham, R.M. (2002) Molecular mechanisms of inherited cardiomyopathies. *Physiol Rev,* **82,** 945-980.

Feltus, F.A., Lee, E.K., Costello, J.F., Plass, C. & Vertino, P.M. (2003) Predicting aberrant CpG island methylation. *Proc Natl Acad Sci U S A,* **100,** 12253-12258.

Festing, M.F.W. (1979) *Inbred strains in biomedical research.* MacMillan.

Fraga, M.F., Herranz, M., Espada, J., Ballestar, E., Paz, M.F., Ropero, S., Erkek, E., Bozdogan, O., Peinado, H., Niveleau, A., Mao, J.H., Balmain, A., Cano, A. & Esteller, M. (2004) A mouse skin multistage carcinogenesis model reflects the aberrant DNA methylation patterns of human tumors. *Cancer Res,* **64,** 5527-5534.

Fujita, N., Shimotake, N., Ohki, I., Chiba, T., Saya, H., Shirakawa, M. & Nakao, M. (2000) Mechanism of transcriptional regulation by methyl-CpG binding protein MBD1. *Mol Cell Biol,* **20,** 5107-5118.

Fujita, N., Takebayashi, S., Okumura, K., Kudo, S., Chiba, T., Saya, H. & Nakao, M. (1999) Methylation-mediated transcriptional silencing in euchromatin by methyl-CpG binding protein MBD1 isoforms. *Mol Cell Biol,* **19,** 6415-6426.

Futscher, B.W., Oshiro, M.M., Wozniak, R.J., Holtan, N., Hanigan, C.L., Duan, H. & Domann, F.E. (2002) Role for DNA methylation in the control of cell type specific maspin expression. *Nat Genet,* **31,** 175-179.

Gandelman, R., Zarrow, M.X., Denenberg, V.H. & Myers, M. (1971) Olfactory bulb removal eliminates maternal behavior in the mouse. *Science,* **171,** 210-211.

Gardiner-Garden, M. & Frommer, M. (1987) CpG islands in vertebrate genomes. *J Mol Biol,* **196,** 261-282.

Gaudet, F., Hodgson, J.G., Eden, A., Jackson-Grusby, L., Dausman, J., Gray, J.W., Leonhardt, H. & Jaenisch, R. (2003) Induction of tumors in mice by genomic hypomethylation. *Science,* **300,** 489-492.

Goffin, J. & Eisenhauer, E. (2002) DNA methyltransferase inhibitors-state of the art. *Ann Oncol,* **13,** 1699-1716.

Goll, M.G., Kirpekar, F., Maggert, K.A., Yoder, J.A., Hsieh, C.L., Zhang, X., Golic, K.G., Jacobsen, S.E. & Bestor, T.H. (2006) Methylation of tRNAAsp by the DNA methyltransferase homolog Dnmt2. *Science,* **311,** 395-398.

Gould, K.A., Dietrich, W.F., Borenstein, N., Lander, E.S. & Dove, W.F. (1996) Mom1 is a semi-dominant modifier of intestinal adenoma size and multiplicity in Min/+ mice. *Genetics,* **144,** 1769-1776.

Guy, J., Hendrich, B., Holmes, M., Martin, J.E. & Bird, A. (2001) A mouse Mecp2-null mutation causes neurological symptoms that mimic Rett syndrome. *Nat Genet,* **27,** 322-326.

Hajkova, P., el-Maarri, O., Engemann, S., Oswald, J., Olek, A. & Walter, J. (2002) DNA-methylation analysis by the bisulfite-assisted genomic sequencing method. *Methods Mol Biol,* **200,** 143-154.

Hamra, F.K., Schultz, N., Chapman, K.M., Grellhesl, D.M., Cronkhite, J.T., Hammer, R.E. & Garbers, D.L. (2004) Defining the spermatogonial stem cell. *Dev Biol,* **269,** 393-410.

Harbers, K., Schnieke, A., Stuhlmann, H., Jahner, D. & Jaenisch, R. (1981) DNA methylation and gene expression: endogenous retroviral genome becomes infectious after molecular cloning. *Proc Natl Acad Sci U S A,* **78,** 7609-7613.

Hata, K., Okano, M., Lei, H. & Li, E. (2002) Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. *Development,* **129,** 1983-1993.

Hendrich, B. & Bird, A. (1998) Identification and characterization of a family of mammalian methyl-CpG binding proteins. *Mol Cell Biol,* **18,** 6538-6547.

Hendrich, B., Guy, J., Ramsahoye, B., Wilson, V.A. & Bird, A. (2001) Closely related proteins MBD2 and MBD3 play distinctive but interacting roles in mouse development. *Genes Dev,* **15,** 710-723.

Hendrich, B. & Tweedie, S. (2003) The methyl-CpG binding domain and the evolving role of DNA methylation in animals. *Trends Genet,* **19,** 269-277.

Herman, J.G. & Baylin, S.B. (2003) Gene silencing in cancer in association with promoter hypermethylation. *N Engl J Med,* **349,** 2042-2054.

Herman, J.G., Umar, A., Polyak, K., Graff, J.R., Ahuja, N., Issa, J.P., Markowitz, S., Willson, J.K., Hamilton, S.R., Kinzler, K.W., Kane, M.F., Kolodner, R.D., Vogelstein, B., Kunkel, T.A. & Baylin, S.B. (1998) Incidence and functional consequences of hMLH1 promoter hypermethylation in colorectal carcinoma. *Proc Natl Acad Sci U S A,* **95,** 6870-6875.

Hogan, B. (1986) *Manipulating the Mouse Embryo: A Laboratory Manual.* Cold Spring Harbor Laboratory.

Hsieh, C.J., Klump, B., Holzmann, K., Borchard, F., Gregor, M. & Porschen, R. (1998) Hypermethylation of the p16INK4a promoter in colectomy specimens of patients with long-standing and extensive ulcerative colitis. *Cancer Res,* **58,** 3942-3945.

Hu, J.F., Oruganti, H., Vu, T.H. & Hoffman, A.R. (1998) Tissue-specific imprinting of the mouse insulin-like growth factor II receptor gene correlates with differential allele-specific DNA methylation. *Mol Endocrinol,* **12,** 220-232.

Huang, W.Y. & Liew, C.C. (1998) Chromatin remodelling of the cardiac beta-myosin heavy chain gene. *Biochem J,* **330 (Pt 2),** 871-876.

Ikeguchi, Y., Wang, X., McCloskey, D.E., Coleman, C.S., Nelson, P., Hu, G., Shantz, L.M. & Pegg, A.E. (2004) Characterization of transgenic mice with widespread overexpression of spermine synthase. *Biochem J,* **381,** 701-707.

Isoda, K., Nishikawa, K., Kamezawa, Y., Yoshida, M., Kusuhara, M., Moroi, M., Tada, N. & Ohsuzu, F. (2002) Osteopontin plays an important role in the development of medial thickening and neointimal formation. *Circ Res,* **91,** 77-82.

Issa, J.P., Ahuja, N., Toyota, M., Bronner, M.P. & Brentnall, T.A. (2001) Accelerated age-related CpG island methylation in ulcerative colitis. *Cancer Res,* **61,** 3573-3577.

Issa, J.P., Ottaviano, Y.L., Celano, P., Hamilton, S.R., Davidson, N.E. & Baylin, S.B. (1994) Methylation of the oestrogen receptor CpG island links ageing and neoplasia in human colon. *Nat Genet,* **7,** 536-540.

Jackson, M., Krassowska, A., Gilbert, N., Chevassut, T., Forrester, L., Ansell, J. & Ramsahoye, B. (2004) Severe global DNA hypomethylation blocks differentiation and induces histone hyperacetylation in embryonic stem cells. *Mol Cell Biol,* **24,** 8862-8871.

Jahner, D., Stuhlmann, H., Stewart, C.L., Harbers, K., Lohler, J., Simon, I. & Jaenisch, R. (1982) De novo methylation and expression of retroviral genomes during mouse embryogenesis. *Nature,* **298,** 623-628.

Jair, K.W., Bachman, K.E., Suzuki, H., Ting, A.H., Rhee, I., Yen, R.W., Baylin, S.B. & Schuebel, K.E. (2006) De novo CpG island methylation in human cancer cells. *Cancer Res,* **66,** 682-692.

Kanellopoulou, C., Muljo, S.A., Kung, A.L., Ganesan, S., Drapkin, R., Jenuwein, T., Livingston, D.M. & Rajewsky, K. (2005) Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. *Genes Dev,* **19,** 489-501.

Karagianni, N., Ly, M.C., Psarras, S., Chlichlia, K., Schirrmacher, V., Gounari, F. & Khazaie, K. (2005) Novel adenomatous polyposis coli gene promoter is located 40 kb upstream of the initiating methionine. *Genomics,* **85,** 231-237.

Kee, H.J., Sohn, I.S., Nam, K.I., Park, J.E., Qian, Y.R., Yin, Z., Ahn, Y., Jeong, M.H., Bang, Y.J., Kim, N., Kim, J.K., Kim, K.K., Epstein, J.A. & Kook, H. (2006) Inhibition of histone deacetylation blocks cardiac hypertrophy induced by angiotensin II infusion and aortic banding. *Circulation,* **113,** 51-59.

Kim, G.D., Ni, J., Kelesoglu, N., Roberts, R.J. & Pradhan, S. (2002) Co-operation and communication between the human maintenance and de novo DNA (cytosine-5) methyltransferases. *Embo J,* **21,** 4183-4195.

Klose, R.J. & Bird, A.P. (2006) Genomic DNA methylation: the mark and its mediators. *Trends Biochem Sci,* **31,** 89-97.

Kong, Y., Tannous, P., Lu, G., Berenji, K., Rothermel, B.A., Olson, E.N. & Hill, J.A. (2006) Suppression of class I and II histone deacetylases blunts pressure-overload cardiac hypertrophy. *Circulation,* **113,** 2579-2588.

Kook, H., Lepore, J.J., Gitler, A.D., Lu, M.M., Wing-Man Yung, W., Mackay, J., Zhou, R., Ferrari, V., Gruber, P. & Epstein, J.A. (2003) Cardiac hypertrophy and

histone deacetylase-dependent transcriptional repression mediated by the atypical homeodomain protein Hop. *J Clin Invest,* **112,** 863-871.

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

Lee, Y.W., Klein, C.B., Kargacin, B., Salnikow, K., Kitahara, J., Dowjat, K., Zhitkovich, A., Christie, N.T. & Costa, M. (1995) Carcinogenic nickel silences gene expression by chromatin condensation and DNA methylation: a new model for epigenetic carcinogens. *Mol Cell Biol,* **15,** 2547-2557.

Lehnertz, B., Ueda, Y., Derijck, A.A., Braunschweig, U., Perez-Burgos, L., Kubicek, S., Chen, T., Li, E., Jenuwein, T. & Peters, A.H. (2003) Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr Biol,* **13,** 1192-1200.

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

Leonhardt, H., Page, A.W., Weier, H.U. & Bestor, T.H. (1992) A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. *Cell,* **71,** 865-873.

Leu, Y.W., Rahmatpanah, F., Shi, H., Wei, S.H., Liu, J.C., Yan, P.S. & Huang, T.H. (2003) Double RNA interference of DNMT3b and DNMT1 enhances DNA demethylation and gene reactivation. *Cancer Res,* **63,** 6110-6115.

Lewis, A., Mitsuya, K., Umlauf, D., Smith, P., Dean, W., Walter, J., Higgins, M., Feil, R. & Reik, W. (2004) Imprinting on distal chromosome 7 in the placenta involves repressive histone methylation independent of DNA methylation. *Nat Genet,* **36,** 1291-1295.

Lewis, J.D., Meehan, R.R., Henzel, W.J., Maurer-Fogy, I., Jeppesen, P., Klein, F. & Bird, A. (1992) Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. *Cell,* **69,** 905-914.

Li, E., Bestor, T.H. & Jaenisch, R. (1992) Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell,* **69,** 915-926.

Li, L., Keverne, E.B., Aparicio, S.A., Ishino, F., Barton, S.C. & Surani, M.A. (1999) Regulation of maternal behavior and offspring growth by paternally expressed Peg3. *Science,* **284,** 330-333.

Liang, G., Gonzales, F.A., Jones, P.A., Orntoft, T.F. & Thykjaer, T. (2002) Analysis of gene induction in human fibroblasts and bladder cancer cells exposed to the methylation inhibitor 5-aza-2'-deoxycytidine. *Cancer Res,* **62,** 961-966.

Lin, H., Yamada, Y., Nguyen, S., Linhart, H., Jackson-Grusby, L., Meissner, A., Meletis, K., Lo, G. & Jaenisch, R. (2006) Suppression of intestinal neoplasia by deletion of Dnmt3b. *Mol Cell Biol,* **26,** 2976-2983.

Lindahl, T. & Nyberg, B. (1974) Heat-induced deamination of cytosine residues in deoxyribonucleic acid. *Biochemistry,* **13,** 3405-3410.

Lindsay, S. & Bird, A.P. (1987) Use of restriction enzymes to detect potential gene sequences in mammalian DNA. *Nature,* **327,** 336-338.

Liu, P., Wakamiya, M., Shea, M.J., Albrecht, U., Behringer, R.R. & Bradley, A. (1999) Requirement for Wnt3 in vertebrate axis formation. *Nat Genet,* **22,** 361-365.

Lyko, F., Ramsahoye, B.H., Kashevsky, H., Tudor, M., Mastrangelo, M.A., Orr-Weaver, T.L. & Jaenisch, R. (1999) Mammalian (cytosine-5) methyltransferases cause genomic DNA methylation and lethality in Drosophila. *Nat Genet,* **23,** 363-366.

Mager, J., Montgomery, N.D., de Villena, F.P. & Magnuson, T. (2003) Genome imprinting regulated by the mouse Polycomb group protein Eed. *Nat Genet,* **33,** 502-507.

Mass, M.J. & Wang, L. (1997) Arsenic alters cytosine methylation patterns of the promoter of the tumor suppressor gene p53 in human lung cells: a model for a mechanism of carcinogenesis. *Mutat Res,* **386,** 263-277.

McKinsey, T.A., Zhang, C.L., Lu, J. & Olson, E.N. (2000) Signal-dependent nuclear export of a histone deacetylase regulates muscle differentiation. *Nature,* **408,** 106-111.

Meehan, R.R., Lewis, J.D. & Bird, A.P. (1992) Characterization of MeCP2, a vertebrate DNA binding protein with affinity for methylated DNA. *Nucleic Acids Res,* **20,** 5085-5092.

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

Millar, C.B., Guy, J., Sansom, O.J., Selfridge, J., MacDougall, E., Hendrich, B., Keightley, P.D., Bishop, S.M., Clarke, A.R. & Bird, A. (2002) Enhanced CpG mutability and tumorigenesis in MBD4-deficient mice. *Science,* **297,** 403-405.

Miyamoto, S., Kawamura, T., Morimoto, T., Ono, K., Wada, H., Kawase, Y., Matsumori, A., Nishio, R., Kita, T. & Hasegawa, K. (2006) Histone acetyltransferase activity of p300 is required for the promotion of left ventricular remodeling after myocardial infarction in adult mice in vivo. *Circulation,* **113,** 679-690.

Morimura, N., Inoue, T., Katayama, K.I. & Aruga, J. (2006) Comparative analysis of structure, expression and PSD95-binding capacity of Lrfn, a novel family of neuronal transmembrane proteins. *Gene.*

Morris, K.V., Chan, S.W., Jacobsen, S.E. & Looney, D.J. (2004a) Small interfering RNA-induced transcriptional gene silencing in human cells. *Science,* **305,** 1289-1292.

Morris, R.J., Liu, Y., Marles, L., Yang, Z., Trempus, C., Li, S., Lin, J.S., Sawicki, J.A. & Cotsarelis, G. (2004b) Capturing and profiling adult hair follicle stem cells. *Nat Biotechnol,* **22,** 411-417.

Moser, A.R., Pitot, H.C. & Dove, W.F. (1990) A dominant mutation that predisposes to multiple intestinal neoplasia in the mouse. *Science,* **247,** 322-324.

Murchison, E.P., Partridge, J.F., Tam, O.H., Cheloufi, S. & Hannon, G.J. (2005) Characterization of Dicer-deficient murine embryonic stem cells. *Proc Natl Acad Sci U S A,* **102,** 12135-12140.

Mutskov, V. & Felsenfeld, G. (2004) Silencing of transgene transcription precedes methylation of promoter DNA and histone H3 lysine 9. *Embo J,* **23,** 138-149.

Ng, H.H., Zhang, Y., Hendrich, B., Johnson, C.A., Turner, B.M., Erdjument-Bromage, H., Tempst, P., Reinberg, D. & Bird, A. (1999) MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex. *Nat Genet,* **23,** 58-61.

Niemann, S., Zhao, C., Pascu, F., Stahl, U., Aulepp, U., Niswander, L., Weber, J.L. & Muller, U. (2004) Homozygous WNT3 mutation causes tetra-amelia in a large consanguineous family. *Am J Hum Genet,* **74,** 558-563.

Ohtani-Fujita, N., Dryja, T.P., Rapaport, J.M., Fujita, T., Matsumura, S., Ozasa, K., Watanabe, Y., Hayashi, K., Maeda, K., Kinoshita, S., Matsumura, T., Ohnishi, Y., Hotta, Y., Takahashi, R., Kato, M.V., Ishizaki, K., Sasaki, M.S., Horsthemke, B., Minoda, K. & Sakai, T. (1997) Hypermethylation in the retinoblastoma gene is associated with unilateral, sporadic retinoblastoma. *Cancer Genet Cytogenet,* **98,** 43-49.

Okabe, M., Ikawa, M., Kominami, K., Nakanishi, T. & Nishimune, Y. (1997) 'Green mice' as a source of ubiquitous green cells. *FEBS Lett,* **407,** 313-319.

Okano, M., Xie, S. & Li, E. (1998a) Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. *Nat Genet,* **19,** 219-220.

Okano, M., Xie, S. & Li, E. (1998b) Dnmt2 is not required for de novo and maintenance methylation of viral DNA in embryonic stem cells. *Nucleic Acids Res,* **26,** 2536-2540.

Owens, D.M., Wei, S. & Smart, R.C. (1999) A multihit, multistage model of chemical carcinogenesis. *Carcinogenesis,* **20,** 1837-1844.

Powell, S.M., Zilz, N., Beazer-Barclay, Y., Bryan, T.M., Hamilton, S.R., Thibodeau, S.N., Vogelstein, B. & Kinzler, K.W. (1992) APC mutations occur early during colorectal tumorigenesis. *Nature,* **359,** 235-237.

Pradhan, S., Bacolla, A., Wells, R.D. & Roberts, R.J. (1999) Recombinant human DNA (cytosine-5) methyltransferase. I. Expression, purification, and comparison of de novo and maintenance methylation. *J Biol Chem,* **274,** 33002-33010.

Prokhortchouk, A., Hendrich, B., Jorgensen, H., Ruzov, A., Wilm, M., Georgiev, G., Bird, A. & Prokhortchouk, E. (2001) The p120 catenin partner Kaiso is a DNA methylation-dependent transcriptional repressor. *Genes Dev,* **15,** 1613-1618.

Ramsahoye, B.H., Biniszkiewicz, D., Lyko, F., Clark, V., Bird, A.P. & Jaenisch, R. (2000) Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proc Natl Acad Sci U S A,* **97,** 5237-5242.

Reiners, J.J., Jr., Nesnow, S. & Slaga, T.J. (1984) Murine susceptibility to two-stage skin carcinogenesis is influenced by the agent used for promotion. *Carcinogenesis,* **5,** 301-307.

Rhee, I., Bachman, K.E., Park, B.H., Jair, K.W., Yen, R.W., Schuebel, K.E., Cui, H., Feinberg, A.P., Lengauer, C., Kinzler, K.W., Baylin, S.B. & Vogelstein, B. (2002) DNMT1 and DNMT3b cooperate to silence genes in human cancer cells. *Nature,* **416,** 552-556.

Robertson, K.D., Uzvolgyi, E., Liang, G., Talmadge, C., Sumegi, J., Gonzales, F.A. & Jones, P.A. (1999) The human DNA methyltransferases (DNMTs) 1, 3a and 3b:

coordinate mRNA expression in normal tissues and overexpression in tumors. *Nucleic Acids Res,* **27,** 2291-2298.

Roelink, H., Wagenaar, E., Lopes da Silva, S. & Nusse, R. (1990) Wnt-3, a gene activated by proviral insertion in mouse mammary tumors, is homologous to int-1/Wnt-1 and is normally expressed in mouse embryos and adult brain. *Proc Natl Acad Sci U S A,* **87,** 4519-4523.

Ruchirawat, M., Noshari, J. & Lapeyre, J.N. (1987) Kinetic mechanisms and interaction of rat liver DNA methyltransferase with defined DNA substrates. *Mol Cell Biochem,* **76,** 45-54.

Rush, L.J., Dai, Z., Smiraglia, D.J., Gao, X., Wright, F.A., Fruhwald, M., Costello, J.F., Held, W.A., Yu, L., Krahe, R., Kolitz, J.E., Bloomfield, C.D., Caligiuri, M.A. & Plass, C. (2001) Novel methylation targets in de novo acute myeloid leukemia with prevalence of chromosome 11 loci. *Blood,* **97,** 3226-3233.

Rush, L.J., Raval, A., Funchain, P., Johnson, A.J., Smith, L., Lucas, D.M., Bembea, M., Liu, T.H., Heerema, N.A., Rassenti, L., Liyanarachchi, S., Davuluri, R., Byrd, J.C. & Plass, C. (2004) Epigenetic profiling in chronic lymphocytic leukemia reveals novel methylation targets. *Cancer Res,* **64,** 2424-2433.

Sado, T., Okano, M., Li, E. & Sasaki, H. (2004) De novo DNA methylation is dispensable for the initiation and propagation of X chromosome inactivation. *Development,* **131,** 975-982.

Samowitz, W.S., Albertsen, H., Herrick, J., Levin, T.R., Sweeney, C., Murtaugh, M.A., Wolff, R.K. & Slattery, M.L. (2005) Evaluation of a large, population-based sample supports a CpG island methylator phenotype in colon cancer. *Gastroenterology,* **129,** 837-845.

Sanoudou, D., Vafiadaki, E., Arvanitis, D.A., Kranias, E. & Kontrogianni-Konstantopoulos, A. (2005) Array lessons from the heart: focus on the genome and transcriptome of cardiomyopathies. *Physiol Genomics,* **21,** 131-143.

Sansom, O.J., Berger, J., Bishop, S.M., Hendrich, B., Bird, A. & Clarke, A.R. (2003) Deficiency of Mbd2 suppresses intestinal tumorigenesis. *Nat Genet,* **34,** 145-147.

Sansom, O.J., Reed, K.R., Hayes, A.J., Ireland, H., Brinkmann, H., Newton, I.P., Batlle, E., Simon-Assmann, P., Clevers, H., Nathke, I.S., Clarke, A.R. & Winton, D.J. (2004) Loss of Apc in vivo immediately perturbs Wnt signaling, differentiation, and migration. *Genes Dev,* **18,** 1385-1390.

Schmelz, K., Sattler, N., Wagner, M., Lubbert, M., Dorken, B. & Tamm, I. (2005) Induction of gene expression by 5-Aza-2'-deoxycytidine in acute myeloid leukemia (AML) and myelodysplastic syndrome (MDS) but not epithelial cells by DNA-methylation-dependent and -independent mechanisms. *Leukemia,* **19,** 103-111.

Schoenherr, C.J., Levorse, J.M. & Tilghman, S.M. (2003) CTCF maintains differential methylation at the Igf2/H19 locus. *Nat Genet,* **33,** 66-69.

Sharpless, N.E., Ramsey, M.R., Balasubramanian, P., Castrillon, D.H. & DePinho, R.A. (2004) The differential impact of p16(INK4a) or p19(ARF) deficiency on cell growth and tumorigenesis. *Oncogene,* **23,** 379-385.

Silverman, K.A., Koratkar, R., Siracusa, L.D. & Buchberg, A.M. (2002a) Identification of the modifier of Min 2 (Mom2) locus, a new mutation that influences Apc-induced intestinal neoplasia. *Genome Res,* **12,** 88-97.

Silverman, L.R., Demakos, E.P., Peterson, B.L., Kornblith, A.B., Holland, J.C., Odchimar-Reissig, R., Stone, R.M., Nelson, D., Powell, B.L., DeCastro, C.M., Ellerton, J., Larson, R.A., Schiffer, C.A. & Holland, J.F. (2002b) Randomized controlled trial of azacitidine in patients with the myelodysplastic syndrome: a study of the cancer and leukemia group B. *J Clin Oncol,* **20,** 2429-2440.

Singer, J., Roberts-Ems, J. & Riggs, A.D. (1979) Methylation of mouse liver DNA studied by means of the restriction enzymes msp I and hpa II. *Science,* **203,** 1019-1021.

Soejima, K., Fang, W. & Rollins, B.J. (2003) DNA methyltransferase 3b contributes to oncogenic transformation induced by SV40T antigen and activated Ras. *Oncogene,* **22,** 4723-4733.

Song, F., Smith, J.F., Kimura, M.T., Morrow, A.D., Matsuyama, T., Nagase, H. & Held, W.A. (2005) Association of tissue-specific differentially methylated regions (TDMs) with differential gene expression. *Proc Natl Acad Sci U S A,* **102,** 3336-3341.

Sontag, L.B., Lorincz, M.C. & Georg Luebeck, E. (2006) Dynamics, stability and inheritance of somatic DNA methylation imprints. *J Theor Biol.*

Stirzaker, C., Song, J.Z., Davidson, B. & Clark, S.J. (2004) Transcriptional gene silencing promotes DNA hypermethylation through a sequential change in chromatin modifications in cancer cells. *Cancer Res,* **64,** 3871-3877.

Su, L.K., Kinzler, K.W., Vogelstein, B., Preisinger, A.C., Moser, A.R., Luongo, C., Gould, K.A. & Dove, W.F. (1992) Multiple intestinal neoplasia caused by a mutation in the murine homolog of the APC gene. *Science,* **256,** 668-670.

Swartz, M.N., Trautner, T.A. & Kornberg, A. (1962) Enzymatic synthesis of deoxyribonucleic acid. XI. Further studies on nearest neighbor base sequences in deoxyribonucleic acids. *J Biol Chem,* **237,** 1961-1967.

Takai, D. & Jones, P.A. (2002) Comprehensive analysis of CpG islands in human chromosomes 21 and 22. *Proc Natl Acad Sci U S A,* **99,** 3740-3745.

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

Thorvaldsen, J.L., Duran, K.L. & Bartolomei, M.S. (1998) Deletion of the H19 differentially methylated domain results in loss of imprinted expression of H19 and Igf2. *Genes Dev,* **12,** 3693-3702.

Trinh, B.N., Long, T.I., Nickel, A.E., Shibata, D. & Laird, P.W. (2002) DNA methyltransferase deficiency modifies cancer susceptibility in mice lacking DNA mismatch repair. *Mol Cell Biol,* **22,** 2906-2917.

Tufarelli, C., Stanley, J.A., Garrick, D., Sharpe, J.A., Ayyub, H., Wood, W.G. & Higgs, D.R. (2003) Transcription of antisense RNA leading to gene silencing and methylation as a novel cause of human genetic disease. *Nat Genet,* **34,** 157-165.

Turek-Plewa, J. & Jagodzinski, P.P. (2005) The role of mammalian DNA methyltransferases in the regulation of gene expression. *Cell Mol Biol Lett,* **10,** 631-647.

Tweedie, S., Ng, H.H., Barlow, A.L., Turner, B.M., Hendrich, B. & Bird, A. (1999) Vestiges of a DNA methylation system in Drosophila melanogaster? *Nat Genet,* **23,** 389-390.

Ueda, Y., Okano, M., Williams, C., Chen, T., Georgopoulos, K. & Li, E. (2006) Roles for Dnmt3b in mammalian development: a mouse model for the ICF syndrome. *Development,* **133,** 1183-1192.

Van den Wyngaert, I., Sprengel, J., Kass, S.U. & Luyten, W.H. (1998) Cloning and analysis of a novel human putative DNA methyltransferase. *FEBS Lett,* **426,** 283-289.

Vertino, P.M., Spillare, E.A., Harris, C.C. & Baylin, S.B. (1993) Altered chromosomal methylation patterns accompany oncogene-induced transformation of human bronchial epithelial cells. *Cancer Res,* **53,** 1684-1689.

Vire, E., Brenner, C., Deplus, R., Blanchon, L., Fraga, M., Didelot, C., Morey, L., Van Eynde, A., Bernard, D., Vanderwinden, J.M., Bollen, M., Esteller, M., Di Croce, L., de Launoit, Y. & Fuks, F. (2006) The Polycomb group protein EZH2 directly controls DNA methylation. *Nature,* **439,** 871-874.

Waki, T., Tamura, G., Sato, M. & Motoyama, T. (2003) Age-related methylation of tumor suppressor and tumor-related genes: an analysis of autopsy samples. *Oncogene,* **22,** 4128-4133.

Weaver, I.C., Cervoni, N., Champagne, F.A., D'Alessio, A.C., Sharma, S., Seckl, J.R., Dymov, S., Szyf, M. & Meaney, M.J. (2004) Epigenetic programming by maternal behavior. *Nat Neurosci,* **7,** 847-854.

Weisenberger, D.J., Siegmund, K.D., Campan, M., Young, J., Long, T.I., Faasse, M.A., Kang, G.H., Widschwendter, M., Weener, D., Buchanan, D., Koh, H., Simms, L., Barker, M., Leggett, B., Levine, J., Kim, M., French, A.J., Thibodeau, S.N., Jass, J., Haile, R. & Laird, P.W. (2006) CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. *Nat Genet,* **38,** 787-793.

Wheeler, J.M., Loukola, A., Aaltonen, L.A., Mortensen, N.J. & Bodmer, W.F. (2000) The role of hypermethylation of the hMLH1 promoter region in HNPCC versus MSI+ sporadic colorectal cancers. *J Med Genet,* **37,** 588-592.

Wiekowski, M.T., Chen, S.C., Zalamea, P., Wilburn, B.P., Kinsley, D.J., Sharif, W.W., Jensen, K.K., Hedrick, J.A., Manfra, D. & Lira, S.A. (2001) Disruption of neutrophil migration in a conditional transgenic model: evidence for CXCR2 desensitization in vivo. *J Immunol,* **167,** 7102-7110.

Wilkinson, C.R., Bartlett, R., Nurse, P. & Bird, A.P. (1995) The fission yeast gene pmt1+ encodes a DNA methyltransferase homologue. *Nucleic Acids Res,* **23,** 203-210.

Wong, E., Yang, K., Kuraguchi, M., Werling, U., Avdievich, E., Fan, K., Fazzari, M., Jin, B., Brown, A.M., Lipkin, M. & Edelmann, W. (2002) Mbd4 inactivation increases Cright-arrowT transition mutations and promotes gastrointestinal tumor formation. *Proc Natl Acad Sci U S A,* **99,** 14937-14942.

Xie, S., Wang, Z., Okano, M., Nogami, M., Li, Y., He, W.W., Okumura, K. & Li, E. (1999) Cloning, expression and chromosome locations of the human DNMT3 gene family. *Gene,* **236,** 87-95.

Yamada, Y., Jackson-Grusby, L., Linhart, H., Meissner, A., Eden, A., Lin, H. & Jaenisch, R. (2005) Opposing effects of DNA hypomethylation on intestinal and liver carcinogenesis. *Proc Natl Acad Sci U S A,* **102,** 13580-13585.

Yamashita, S., Tsujino, Y., Moriguchi, K., Tatematsu, M. & Ushijima, T. (2006) Chemical genomic screening for methylation-silenced genes in gastric cancer cell lines using 5-aza-2'-deoxycytidine treatment and oligonucleotide microarray. *Cancer Sci,* **97,** 64-71.

Yisraeli, J., Frank, D., Razin, A. & Cedar, H. (1988) Effect of in vitro DNA methylation on beta-globin gene expression. *Proc Natl Acad Sci U S A,* **85,** 4638-4642.

Yoder, J.A. & Bestor, T.H. (1998) A candidate mammalian DNA methyltransferase related to pmt1p of fission yeast. *Hum Mol Genet,* **7,** 279-284.

Yoder, J.A., Soman, N.S., Verdine, G.L. & Bestor, T.H. (1997) DNA (cytosine-5)-methyltransferases in mouse cells and tissues. Studies with a mechanism-based probe. *J Mol Biol,* **270,** 385-395.

Zhang, C.L., McKinsey, T.A., Chang, S., Antos, C.L., Hill, J.A. & Olson, E.N. (2002) Class II histone deacetylases act as signal-responsive repressors of cardiac hypertrophy. *Cell,* **110,** 479-488.

Zhang, Y., Ng, H.H., Erdjument-Bromage, H., Tempst, P., Bird, A. & Reinberg, D. (1999) Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation. *Genes Dev,* **13,** 1924-1935.

Zoumpourlis, V., Solakidi, S., Papathoma, A. & Papaevangeliou, D. (2003) Alterations in signal transduction pathways implicated in tumour progression during multistage mouse skin carcinogenesis. *Carcinogenesis,* **24,** 1159-1165.

9. Appendix

9.1 Perl source code for bisulphite sequencing analysis

```
#! C;\perl\bin\perl

#bisulphite reporter
#Input format- copy and paste pre-aligned sequences into txt file with path specified in $infile
#-the first sequence is the reference genomic sequence and the 5'end must extend
# beyond the 5' end of the other sequences  or be filled up with -.

#The 5'end is filled up with "-" as in ClustalW output
#file structure for each sequence is ">title" <cr> "sequence"

#report is written to file contained in $outfile.  Output is space delimited text
#so it can be imported in to excel.

#GETTING SEQUENCE FILE (put path in to $infile)
#open sequence file- contained in array @sequence

$infile = "C:/Documents and Settings/Tom Latham/My Documents/Sequences/BS8(Apc
        tumours)/external/external(2nd fit)infile.txt";

$outfile = "C:/Documents and Settings/Tom Latham/My Documents/Sequences/BS8(Apc
        tumours)/external/(2nd fit)outfile.txt";
$runtime = localtime(time);

open (TXT, "<$infile");
@alignment = <TXT>;
close (TXT);

#Process reference sequence

open (TXT, ">$outfile");

print TXT "$runtime \n $infile\n $alignment[0]cpg positions \n";
close TXT;

$str=lc($alignment[1]);
@seq=split(//,$str);
$len=@seq;

#loop to find positions of cpgs and cp(not g);

for ($i=0; $i<$len;$i++)
   {
  if (@seq[$i] eq "c" && $seq[$i+1] eq "g") {$cpg .= $i." "};

  if (@seq[$i] eq "c" && $seq[$i+1] ne "g") {$cpn .= $i." "};
   }
open (TXT, ">>$outfile");
print  TXT "$cpg \n \n";
close TXT ;
```

```
@cpg=split(/ /,$cpg);
$ncpg=@cpg;
$nseq=@alignment/2;


#@cpg is position of CpGs.  $ncpg is number of cpg in ref sequence.
#nseq is number of sequences in alignment.

open (TXT, ">>$outfile");
for ($i=1; $i<$nseq;$i++)
        {

        @clone=split(//,lc($alignment[(2*$i)+1]));
        for ($j=0; $j<$ncpg;$j++)
                {
                print TXT cpgmatch(@clone[@cpg[$j]],@clone[@cpg[$j]+1]);
         print  TXT " ";
                }
    print TXT " @alignment[2*$i]";

        }


#--------Report on CpH  frequency as measure of conversion.----------
print  TXT "\n Report on Bisulphite conversion \n\n";
  @cpn=split(/ /,$cpn);
  $ncpn=@cpn;


 for ($i=1; $i<$nseq;$i++)
        {
        @clone=split(//,lc($alignment[(2*$i)+1]));
        $conv=0;
        $uconv=0;
        $noseq=0;
        $poorseq=0;
        $unmatch=0;
        for ($j=0; $j<$ncpn;$j++)
                {
         $match=cpnmatch(@cpn[$j]);
                if ($match eq "C") {$conv++};
         if ($match eq "U") {$uconv++};
         if ($match eq "-") {$noseq++};
         if ($match eq "n") {$poorseq++};
         if ($match eq "?") {$unmatch++};
        }
    $efficiency = (($uconv+$conv) ne 0) ? int((100*$conv)/($uconv+$conv)) : "n/a" ;
    $name=@alignment[(2*$i)];
    chomp($name);
    print  TXT "$name  $efficiency % \n" ;

    }
    close TXT;
```

```
    print "0/0 $runtime";
#--------subroutines-------------------------------------
sub cpgmatch
#cpgmatch(a,b) returns methylation status of dinucleotide passed as arguments a,b
        {
        my $ret;
        $ret="?";
        if (@_[0] eq "c" && @_[1]  eq "g") {$ret= "M"};
        if (@_[0] eq "t" && @_[1]  eq "g") {$ret= "U"};
    if (@_[0] eq "-" && @_[1]  eq "-") {$ret= "-"};
    if (@_[0] eq "n" | @_[1]  eq "n") {$ret= "n"};
        return $ret;
        }


sub cpnmatch
#cpnmatch(n) returns conversion status of dinucleotide at position n
# nb unmethylated ccg will be returned as "?"
    {
    my $ret;

    $ret="?";
    if (@clone[@_[0]] eq "t" && @clone[@_[0]+1]  eq @seq[@_[0]+1] && @clone[@_[0]+2]  eq
        @seq[@_[0]+2]) {$ret= "C"};
        if (@clone[@_[0]] eq "c" && @clone[@_[0]+1]  eq @seq[@_[0]+1] && @clone[@_[0]+2]  eq
        @seq[@_[0]+2]) {$ret= "U"};
    if (@clone[@_[0]] eq "-" | @clone[@_[0]+1]  eq "-") {$ret= "-"};
    if (@clone[@_[0]] eq "n" | @clone[@_[0]+1] eq "n") {$ret= "n"};
        return $ret;
        }
```