MBD Genes And Hedgehog Signalling In Cancer

Yanhua Zhu

MD

The University of Edinburgh

2003

Acknowledgements

I would like to thank my supervisors Prof. David Harrison and Dr. Scott Bader for their assistance during my study at Edinburgh. Firstly, my thanks should go to Prof. Harrison. With his support, I was able to have the opportunity to start study and research in the University of Edinburgh Department of Pathology, Sir. Alastair Currie Cancer Research UK Laboratories, Molecular Medicine Center, Western General Hospital, Edinburgh. I am most grateful to Dr. Bader for his daily basis guidance. Without his patience and consideration, my projects would not progress so smoothly. I also benefited from his extensive molecular biology techniques.

I am grateful to Ms Marion Walker and Ms Lindy Hutchison for their kindly help.

I am grateful to Dr. Marina Mackenzie for providing brand new powerful computer and other laboratory equipment. Also I would like to thank other members of CR UK Labs for their help.

Declaration

Except special statements, the research works in this thesis were author's own work. No part of the research has been accepted or is currently being submitted for the purpose of acquiring another degree.

Dr. Scott Bader designed all the primers in *MBD2* and *MBD3* genes. He also provided most of DNA samples.

Dr. Roberta James designed the primers in *SMO* and *PTC* genes.

The primers for bisulfite sequencing were designed by myself.

Yanhua Zhu

Dec.2003

Abstract

It is accepted that cancer progression is a multi-step process including multiple genetic and epigenetic events. In this study, two sets of candidate genes in colon and lung cancer tumourigenesis were studied. The first set comprised members of a family of genes whose proteins are important in the recognition of the methylation/epigenetic status of other genes. The second set were members of a pathway that normally regulate tissue development but whose abnormal, epigenetic loss of activity could lead to tissue dysregulation and tumourigenesis.

MBD3 and *MBD2* are two members of the MBD family of proteins with a methyl-CpG-binding domain (MBD) involved in transcriptional silencing of methylated genes. Both genes are located in chromosomal regions that suffer loss of heterozygosity in colon and lung cancers. By SSCP analysis and methylation sensitive restriction followed by PCR, 2 mutations were found in 28 cell lines and in no cases was there evidence of gene silencing by hypermethylation of putative promoter regions. RT-PCR and northern hybridisation showed expression of *MBD3* in all cancer cell lines examined. The results indicate that neither *MBD2* nor *MBD3* are major targets of genetic and epigenetic alteration in colon and lung cancers.

The Hedgehog (Hh) pathway is a highly conserved signaling cascade involved in many developmental processes. Mutations in elements of the pathway are associated with congenital diseases and a number of neoplasms. In this study, two genes of this pathway, *SMO* and *GLI3* were investigated for expression and epigenetic alterations in colon and lung cancers. In three cell lines expression of *SMO* was absent, the putative *SMO* promoter was fully methylated and *GLI3* was not expressed. Two other cell lines had a methylated wild-type *SMO* allele and expressed mutant *SMO*,

and also did not express *GLI3*. The results indicate that *SMO* is silenced by CpG island hypermethylation in colon and lung cancer cell lines, that *GLI3* is also silenced in colon and lung cancer cell lines by an as yet unrevealed mechanism and that *GLI3* is possibly regulated by *SMO* in a manner outside the normal sequence of steps currently thought to comprise the Hh pathway.

Contents

Chapter 1

Introduction

Contents

1. Overview of cancer genetics and epigenetics

1.1. Oncogenes

The concept that tumours derive from genetic alteration originated about a century ago, when Broca discovered that many members of a family suffered from breast or liver cancer, and he deduced that tumour development was caused by an inherited abnormality (Broca, 1866). Since then his hypothesis has been supported by accumulating evidence generated by familial, epidemiological and cytogenetic studies. In 1911, a breakthrough study in the campaign to target a genetic basis for cancer was reported by Rous (Rous, 1911). He found that cell-free filtrates of a sarcoma, which had previously arisen in a chicken, could cause a sarcoma reproducibly in chickens. Although Rous's observation strongly supported the idea that cancer could be induced virally, it also suggested that cancer could be caused by genetic materials. Since the oncogenic element of the Rous sarcoma virus was identified in 1976 (Stehlin *et al.*, 1976), it has been known that the oncogenicity of the virus relied on *v-src*, a transduced and mutated copy of the *c-src* cellular proto-oncogene. Also it has been confirmed that all of acutely transforming RNA tumour viruses harbour oncogenes that actually have their counterparts as transduced cellular genes (the proto-oncogenes). So far more than 50 different proto-oncogenes have been identified through various experimental strategies. In general, proto-oncogenes are involved in a variety of growth regulatory pathways, and their protein products are distributed throughout all subcellular compartments. The oncogenic mutant alleles present in cancers are activated by sustained gain-of-function alterations resulting from point mutation, chromosomal rearrangement, or gene amplification of the proto-oncogene sequences. In most cancers, mutations in proto-oncogenes are

somatic, although germline mutations exist in some cases. Oncogenes related to colon and lung cancers will be discussed in those sections.

1.2.Tumour suppressor gene (TSG)

The assay to detect chromosome deletions has played a major role in the process of identification and cloning of another class of cancer-associated genes, the tumour suppressor gene (TSG). Contrary to the oncogenes which are activated by dominant mutations and whose activity is to promote cell growth, tumour suppressor gene act in the normal cell as negative controllers of cell growth and lack normal function in tumour cells. In general, therefore, the mutations inactivating tumour suppressor genes are of the recessive type requiring mutation of both alleles to have an effect, which is proposed by Knudson's two-hit hypothsis (Knudson, 1971). A large number of tumour suppressor genes have been hypothesised to exist. Thus far, about 30 tumour suppressor genes have been identified and definitively implicated in cancer development. Like proto-oncogenes, the cellular functions of the tumour suppressor genes appear to be diverse. The well-known example of tumour suppressor genes is *p53*. The definition of *p53* as a tumour suppressor gene had experienced a dramatic turnaround. Initial findings suggested that *p53* functioned as an oncogene in some *in vitro* experiments (Lane and Benchimol, 1990; Eliyahu *et al.*, 1984; Jenkins *et al.*, 1984; Parada *et al.*, 1984). The first evidence to suggest that *p53* might frequently be inactivated in human cancers was obtained from studies demonstrating that chromosome 17p LOH was common in a number of different tumours types, including colorectal, bladder, breast and lung cancer (Fearon *et al.*, 1987; Vogelstein *et al.*, 1988; Baker *et al.*, 1989; Nigro *et at.,* 1989; Takahashi *et al.*, 1989). Analysis

of the sequence of the *p53* alleles retained in cancers with 17p LOH demonstrated that the remaining *p53* allele was mutated in the vast majority of such cases. Additional evidence that *p53* functions as a tumour suppressor gene in human cancer has been provided by gene transfer studies (Baker *et al.*, 1990). Based on the types of tumours in which *p53* mutations have been found and the prevalence of *p53* mutations in those tumour types, *p53* is believed to be among the most frequently mutated genes in human cancer. Although gross rearrangements of the *p53* gene are seen in some paediatric tumours like osteosarcoma and rhabdomyosarcoma and splice mutations are seen in some cancers, the vast majority of the somatic mutations in *p53* are missense mutations leading to amino acid substitutions in the central portion of the protein (Greenblatt *et al.*, 1994). The *p53* gene encodes a 53 kd nuclear protein that acts as a transcription factor, blocks the cell cycle at late G1 (El-Deiry *et al.*, 1993; Harper *et al.*, 1993; Hermeking *et al.*, 1997) and also can trigger apoptosis (Miyashita and Reed, 1995; Wu *et al.*, 1997; Polyak *et al.*, 1997; Moroni *et al.*, 2001; Robles *et al.*, 2001). p53 has a role in maintaining the stability of the genome during cellular stress from DNA damage, hypoxia and activated oncogenes(Wang *et al.*, 1995; Ford and Hanawalt, 1995). Subsequently, *RB, P16, APC, BRCAs,* etc. came into the category of tumour suppressor genes (Whyte *et al.*, 1988; Serrano *et al.*, 1993; Groden *et al.*, 1991; Hall *et al.*, 1990; Wooster *et al.*, 1995). Different experimental approaches demonstrated that they are important players in different subsets of tumours at different stages.

Generally speaking, there are two types of tumour suppressor gene: so called 'gatekeepers' and 'caretakers' (see figure 1). Gatekeeper genes act directly to

regulate cell proliferation and are rate limiting for tumourigenesis. Each cell type has only a few gatekeepers, examples being *APC*, *p53* and *RB*. Caretaker genes, by contrast, do not directly regulate proliferation but when their function is interrupted lead to accelerate conversion of a normal cell to a neoplastic cell. Caretaker genes are required for the maintenance of genome integrity. The existence of numerous so-called chromosomal instability disorders, in which germ-line mutations in a caretaker gene lead to both genome instability and a predisposition to cancer, attests to the importance of these genes in suppressing neoplastic transformation. Well-studied caretaker genes are *ATM*, *BRCA1 and BRCA2*.

A major group of caretaker genes, DNA mismatch repair genes, has been defined in tumours in which an inherited mutated predisposing gene plays a significant role. These tumours include cancers in patients suffering from hereditary nonpolyposis colorectal cancer (HNPCC) syndromes. The genes implicated in these tumours have been defined as mutator genes or genes involved in the DNA mismatch repair process. Mutational inactivation of both copies of a DNA mismatch repair gene results in a significant repair defect and progressive accumulation of mutations throughout the genome that eventually may cause the activation of oncogenes and the inactivation of other tumour suppressors. Five human DNA mismatch repair genes have been identified that are involved in HNPCC. Mutational analysis of MMR genes in HNPCC shows that defective *hMSH2* and *hMLH1* are the major cause of the disease accounting for more that 95% of the identified germline

Figure1. One possible relationship of caretaker and gatekeeper pathways.

Gatekeeper tumour suppressors are best defined by the fact that: first, their loss of function is rate-limiting for a particular step in multi-stage tumourigenesis; second, they act directly to prevent tumor growth, and third, restoring gatekeeper function to tumor cells suppresses neoplasia. By contrast, caretaker tumor suppressor genes act indirectly to suppress abnormal growth by ensuring the fidelity of the DNA code through effective repair of DNA damage or prevention of genomic instability (such as microsatellite or chromosome instability).

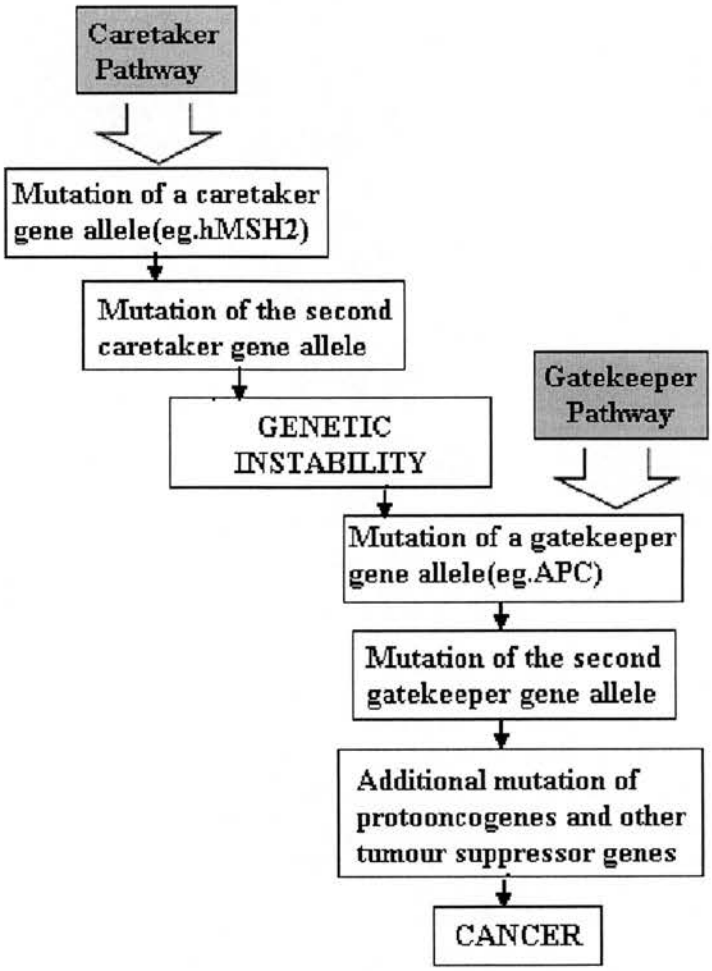


Figure 1. One possible relationship of caretaker and gatekeeper pathways

mutations, whereas mutation of *hPMS1* and *hPMS2* in HNPCC account for less than 5% of total germline mutations (Vasen *et al.*, 1991; Fishel *et al.*, 1993; Leach *et al.*, 1993; Lindblom *et al.*, 1993).

Mutations have been identified in tumour suppressor genes that are inherited and predispose human to cancer. Recent evidence indicates that genomic imprinting, an epigenetic form of gene regulation that results in uniparental gene expression, can also function as a cancer predisposing event. Thus, cancer susceptibility is increased by both inherited genetic and non-inherited epigenetic events. Consequently, carcinogenic agents can not only induce cancer through the formation of genetic mutations but also through epigenetic changes that result in the inappropriate expression of imprinted proto-oncogenes and tumour suppressor genes. Tumour suppressor genes related to colon and lung cancers will be discussed in those sections.

1.3. Haploinsufficiency effect of tumour suppressor gene

Although Knudson's two-hit model can explain the inactivation of some tumour suppressor genes, increasing evidence is changing our perspective of tumour suppressor gene function and regulation. Some evidence suggests that the haploinsufficiency effect of a tumour suppressor gene as well as LOH at a given tumour suppressor gene locus possess the same significance in the process of tumourigenesis induced by tumour suppressor gene inactivation. To date, there are three avenues by which the haploinsufficiency of certain tumour suppressor genes take their effect:

1) Haploid levels of a tumour suppressor gene product are insufficient to inhibit the activity of critical downstream target proteins involved in stimulating growth. For example, the *p27/KIP1* gene is haploinsufficient for tumourigenesis in the mouse and this appears to be due directly to the effects of haploid p27 protein levels because expression from the wild-type *p27* allele appeared normal (Fero *et al.*, 1998). p27 is one of the inhibitors of activity of cyclin/cdk complexes. Consequently, reduced levels of p27 would be expected to increase cyclin/cdk activity and promote tumourigenesis. *Pten* is another case of haploinsufficiency (Cristofano *et al.*, 1999).

2) The tumour suppressor mutation acts in a dominant-negative fashion to block the activity of wildtype tumour suppressor protein. Analysis of tumour formation in *p53* heterozygous mice showed that only ~50% of tumours show LOH at the *p53* locus (Venkatachalam *et al.*, 1998), suggesting that mutant *p53* allele may drive tumourigenesis by dominant-negative fashion.

3) Expression from the wild-type allele is shut off/reduced as a result of epigenetics events. Shoemaker *et al.* have shown that expression from the wild-type *APC* allele is reduced in heterozygous tumours, suggesting that LOH equivalent effect is achieved not by gene loss/mutation but by silencing *APC* gene expression from the wild-type allele in an alternative manner similar to genomic imprinting (Shoemaker *et al.*, 1998).

1.3.Epigenetic gene silencing in cancer

1.3.1. CpG island and DNA methylation

CpG islands are GC-rich regions of DNA, stretching for an average of about 1 kb, which are coincident with the promoters of ~60% of human RNA polymerase II-

transcribed genes, e.g. so called housekeeping genes (Bird, 1986). Methylation of the C5 position of 5′ -CpG-3′ dinucleotides of mammalian DNA is known to be a powerful mechanism for the suppression of gene activity. DNA methylation within some CpG islands is essential for human development, X-chromosome inactivation and genomic imprinting (Jaenisch, 1997). DNA methylation may also suppress transcriptional noise, defend the genome against retroviral elements, immobilise transposons and control tissue-specific gene expression.

CpG sites occur less frequently than expected in mammalian DNA (~1 per 100 bases) and are usually methylated, but are clustered at a higher frequency (~1per 10 bases) in CpG islands where they are typically found unmethylated (Bird *et al.*, 1985). The regulatory influences of CpG sequences may be most important when located in the 5′ promoter regions of genes, however some intron and 3′ methylation influences have also been reported. In contrast to the unmethylated CpG islands of all active housekeeping genes throughout the genome, the CpG islands of many genes on the inactive X chromosome are methylated, as are the CpG islands of many non-essential genes in long term cultured cells. The methylation status of a gene is usually inversely correlated with gene expression, such that hypermethylation of certain gene promoters yields gene inactivation and hypomethylation of these promoters activates or reactivates gene expression. Aberrant methylation of CpG islands in the promoter of many cancer-related genes results in silencing of their expression (discussed later).

1.3.2. The methylation machinery

It is thought that cellular DNA methylation patterns are established and maintained by a complex interplay of at least three independent DNA methyltransferases: DNMT1, DNMT3A and DNMT3B. DNMT1 was the first methyltransferase to be identified and characterized (Bestor *et.al.,* 1988). It is the most abundant methyltransferase in somatic cells, localises to replication foci and interacts with the proliferating cell nuclear antigen (Robertson *et al.*, 1999; Leonhardt *et al.*, 1992; Chuang *et al.*, 1997). It is often referred to as the maintenance methyltransferase because it is believed to be the enzyme responsible for copying methylation patterns after DNA replication. However, new evidences suggest that it can also act as a *de novo* methyltransferase. DNMT1 is required for proper embryonic development, imprinting and X-inactivation (Li *et al.*, 1992; Li *et al.*, 1993; Beard *et al.*, 1995). Subsequently the DNMT3 family of methyltransferases from mouse and human have been characterized. They are thought to be the main *de novo* methyltransferase activity in cells (Okano *et al.*, 1998). *Dnmt3a* knockout mice are born live but die at about four weeks of age. In contrast, *Dnmt3b* knockout mice are not viable (Li *et al.*, 1992; Okano *et al.*, 1999).

1.3.3. DNA methylation and cancer

In cancer, two different DNA methylation events happen. Hypomethylation of the wide genome occurs (Goelz *et al.*, 1985), while increased localised methylation of promoter-associated CpG islands is concurrently observed leading to silencing of genes with normal tumour-suppressive activity (Jones and Laird, 1999) (see figure 2).

Figure 2. CpG island hypermethylation is associated with gene silencing in cancer cell

Black circle represents methylated CpG sites while white circle represents unmethylated CpG sites. A. In normal cells, CpG islands around or upstream of the translation start are usually methylation-free while non-clustering CpG sites are usually methylated. B. In cancer cells, such CpG are often methylated while hypomethylation across the entire genome often occurs.

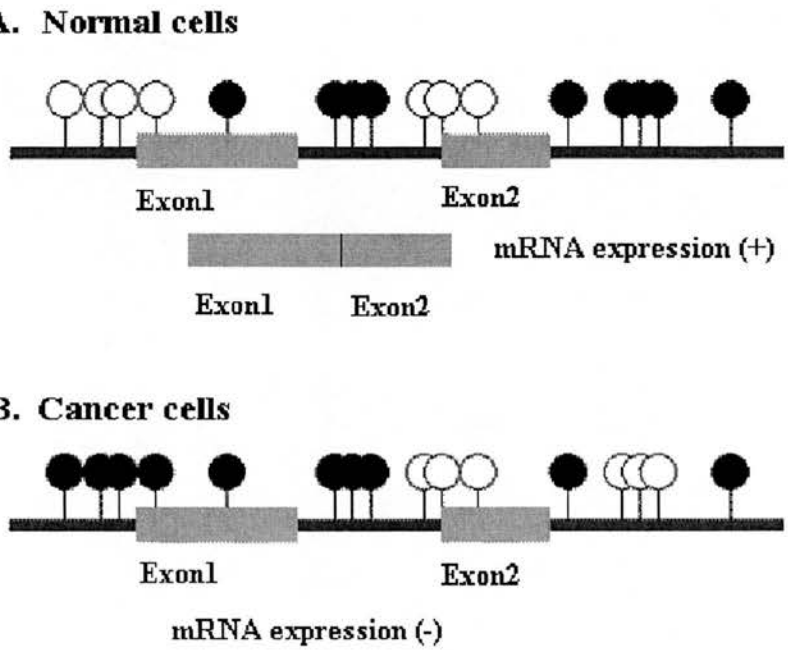


Figure 2. DNA hypermethylation is associated with gene silencing in cancer cell

A growing number of cancer genes are being recognised that suffer from hypermethylation in normally unmethylated promoter CpG islands. Whether DNA methylation initiates gene silencing or methylation is a secondary event following gene silencing has yet to be resolved. But DNA methylation both marks and plays a key role in an epigenetically mediated loss-of-gene function that is as critical and frequent in tumourigenesis as mutations in coding regions. An excellent example is the gene encoding the cyclin-dependent kinase inhibitor, *p16.* The loss of *p16* occurs through the different mechanisms of deletions, point mutation or promoter hypermethylation. The frequency of each mechanism varies with different tumour types. In certain tumours, like colon cancer, *p16* inactivation is seen only in association with promoter hypermethylation (Kamb *et al.*, 1994; Herman *et al.,* 1995). The incidence of functional disruption of *p16*, especially when promoter hypermethylation is included, indicates that this gene is as frequently inactivated as *p53* in cancer (Baylin, 1998). In addition to well-defined mutation of tumour-suppressor genes, promoter hypermethylation is being associated with an increasing list of other genes that have strongly been implicated in tumour development.

The demethylation happens mainly at transposable elements leading to aberrant expression (Cohen and Shapiro, 1980; Craig, 1990; Hartl, 1997; Moffat, 2000). This increased expression has the potential to cause mutational events through transposition. Demethylation may also contribute to genomic instability. For example, embryonic stem cells lacking DNMT1 display both a globally hypomethylated genome and a higher degree of genetic instability (Chen *et al.*, 1998). In addition, pericentromeric areas which are normally heavily methylated

exhibit increased levels of chromosomal translocations when demethylated by a demethylating substance (5-aza-dC)(Ji *et al.*, 1997).

1.4.4.MBD and cancer

1.4.4.1. Methyl-CpG-binding proteins (MBDs)

The mechanism by which DNA binding proteins transduce the signal encoded by a particular pattern of methylation to regulate gene expression has been studied for over 30 years. In the early 1990s, two complexes with ability to bind to methylated DNA were found, namely MeCP1 and MeCP2. MeCP2 was the first true member of the family of proteins that selectively recognise methylated CpG (Lewis *et al.*, 1992; Meehan *et al.*, 1989). Bird and colleagues characterised the minimal region of MeCP2 required for binding to methylated DNA, defining the so-called methyl-CpG-binding domain (MBD)(Hendrich and Bird, 1998). MBD is a short region of MeCP2 containing about 70 residues located within its N-terminal that has the ability to bind selectively methylated DNA. By database search for sequence homologous to the MBD, Bird's group identified a protein containing a MBD-like motif located at its N-terminus. This protein was named MBD1. MBD1 was shown to bind methylated DNA and to repress transcription from a methylated promoter *in vitro*. It was initially believed to be a component of the MeCP1 complex (Cross *et al.*, 1997). A further search of EST databases found three more genes in mammalian cells that encode proteins containing MBDs, namely MBD2, MBD3 and MBD4. Alignment of the MBD-like regions from the murine MBD1 to MBD4 and MeCP2 proteins showed that two subgroups could be established. The MBD of MBD4 is most similar to that of MeCP2 in primary sequence, while the MBDs of MBD1, MBD2 and MBD3 are

more similar to each other than to those of either MBD4 or MeCP2. The sequence similarity between these five proteins is largely limited to their MBD although MBD2 and MBD3 share greater homology along their whole length (~70%). Analysis of MBD genes expression in numerous murine tissues showed that they are expressed in all samples tested but embryonic stem cells, where DNA methylation is known to be dispensable, seem to have low levels of *mbd1* and *mbd2* transcripts (Hendrich and Bird, 1998).

1.4.4.2. MBD3

MBD3 shares about 70% of overall identity with MBD2 over most of their length. The greatest divergence occurs at the C-terminus, where MBD3 has 12 consecutive glutamic acid residues encoded by an imperfect trinucleotide repeat. This characteristic is retained in the human MBD3. MBD3 also has variants produced by alternative splicing. Recently, two *MBD3* like genes, *MBD3L1* and *MBD3L2* were cloned and identified. The *MBD3L1* is 42% identical to *MBD3* and 38% identical to *MBD2* but lacks the methyl-CpG binding domain. The *MBD3L1* gene is expressed specifically in testis, suggesting a role in the development of male germ cells. Interestingly, expression of *MBD3L2* was found in germ cell tumours and some somatic tissues (Jiang *et al.*, 2002).

The most abundant is a 32-kDa protein that shares high homology to MBD2b (80% similar, 72% identical). The second variant contains an insertion of a small exon (20 amino acids) in the MBD, with the rest of its sequence being identical with that of the short form of MBD3. These two MBD3 variants have been detected in human, mouse

and *Xenopus* systems. The recombinant MBD3 protein does not bind to methylated DNA in vitro. Considering its high similarity to MBD2b, demethylase activity (controversially reported for MBD2b) has also been tested for MBD3, however no demethylase activity was detected (Hendrich and Bird, 1998).

MBD3 is one member of the NuRD complex, which is a multisubunit complex containing nucleosome remodelling and histone deacetylase activities (Wade *et al.*, 1999, Zhang *et al.*, 1999; Wade *et al.*, 1998). One study indicates that Drosophila gene dMBD2/3 forms specialised nuclear compartments to keep certain genes epigenetically silenced during genome activation (Marhold *et al.*, 2002). One group hypothesised that the MBD2-MBD3 complex recognises hemi-methylated DNA concurrent with DNA replication and recruits histone deacetylase complexes, as well as DNMT1, to establish and/or maintain the transcriptionally repressed chromatin (Tatematsu *et al.*, 2000). Closely related proteins MBD2 and MBD3 play distinctive but interacting roles in mouse development (Hendrich *et al.*, 2001). MBD3, MeCP2 and MBD1 are also involving in imprinting selection (Fournier *et al.*, 2002).

MBD3 is located on chromosome 19p13.3, a region reported to suffer 20-50% LOH in sporadic colorectal carcinomas (Resta *et al.*, 1998; Trojan *et al.*, 2000; Dong *et al.*, 1998). According to data compiled by the Human Genome Mapping Project (available on the website http://www.ncbi.nlm.nih.gov/) *MBD3* is within about 500kb of the gene *LKB1/STK11* which is mutated or abnormally methylated in Peutz-Jeghers syndrome. Peutz-Jeghers patients have hamartomatous polyposis of the gastrointestinal tract and an increased risk of a range of cancers including colon.

LKB1 is rarely mutated or methylated (maximum about 20%) however in sporadic colorectal carcinomas (Resta *et al.*, 1998; Trojan *et al.*, 2000; Avizienyte *et al.*, 1998; Esteller *et al.*, 2000; Launonen *et al.*, 2000), raising the possibility that another gene in the vicinity is involved in these cancers. The short arm of chromosome 19 is also implicated in up to 86% of lung cancers (Lukeis *et al.*, 1990; Virmani *et al.*, 1998; Sanchez-Cespedes *et al.*, 2001). In the light of the location of *MBD3* in a region of chromosomal loss, its known functions in transcription suppression and data connecting other MBD family members with cancer it was considered as a candidate tumour suppressor gene.

1.4.4.3. MBD2

MBD2 is highly similar to MBD3 in a large region corresponding roughly to amino acids 140–400 (Hendrich and Bird, 1998). This region of MBD2 contains a repeat consisting of glycine and arginine residues (Hendrich and Bird, 1998). Compared to MBD3, MBD2 has a more restricted pattern of expression and an alternatively spliced mRNA is evident in testis (Hendrich and Bird, 1998). MBD2 binds methylated DNA in a way similar to MeCP2 (Hendrich and Bird, 1998; Wade *et al.*, 1999). Surprisingly, MBD2b (a version lacking the amino terminal 140 amino acids) has been reported to possess DNA demethylase activity (Bhattacharya *et al.*, 1999), although this finding has been questioned (Wade *et al.*, 1999; Ng *et al.*, 1999).

Immunoprecipitation studies demonstrate that MBD2 is physically associated with HDAC1 in mammalian cells and implicate MBD2 as the methyl CpG binding component of the MeCP1 complex (Ng *et al.*, 1999). While MBD2 is associated with

HDAC1 and with RbA p48/p46, coimmunoprecipitation analysis showed that it is not a component of the previously defined Sin3 and Mi-2/NURD complexes (Ng *et al.*, 2000). However, a direct interaction of Sin3A with MBD2b in the region sufficient to direct transcriptional repression has also been observed (Boeke *et al.*, 2000).

Whether MBD2 plays a role in cancer is attracting increasing attention. Our group screened MBD2 by SSCP but failed to find enough mutations to suggest that MBD2 suffers from genetic alterations as a tumour suppressor gene. Although two groups reported that the unregulated expression might be related to cancer progress (Billard *et al.*, 2002; Slack *et al.*, 2002), the finding still needs further study to confirm and elucidate.

1.4.4.4. MBD1, MBD4 and MeCP2 and their links to cancer

MBD1 is the largest member of the family and contains a sequence motif, the CXXC motif, shared with DNA methyltransferase I (Cross *et al.*, 1997). One group observed that, in BPH (benign prostatic hyperplasia) tissues and low-grade cancer tissues, MBD1 protein expression was very high and gradually decreased with increase of cancer grade (Patra *et al.*, 2003). However, our group screened colon and lung cancer cell lines and primary tumours by SSCP and only found a small number of mutations in the coding region of *MBD1* (Bader *et al.*, 2003). Thus, up to now, no strong evidence has suggested that MBD1 is crucial for tumour pathogenesis.

A link between *MBD4* and the MMR system was provided by the observation that between 20% and 43% of primary human colorectal carcinomas that displayed

microsatellite instability (MSI) also harbored inactivating mutations in *MBD4* (Riccio *et al.*, 1999, Bader *et al.*, 1999; Miyaki *et al.*, 2001). In addition, *MBD4* mutations were also frequently observed in other microsatellite unstable cancers such as gastric, endometrial, and pancreatic carcinomas (Riccio *et al.*, 1999; Yamada *et al.*, 2002). However, the failure to detect *MBD4* mutations in microsatellite stable tumours together with the lack of mutations occurring outside this mononucleotide repeat track suggested that *MBD4* mutations were likely the result, rather than the cause, of MMR deficiency (Bader *et al.*, 1999).

The methyl CpG binding protein 2 gene, *MECP2*, encodes a global transcriptional silencer and was identified as the gene defective in RTT (Rett syndrome, a neurodevelopmental disorder that mostly affects females) (Amir *et al.*, 1999). Mutations in *MECP2* so far identified in association with RTT are thought to cause loss of function of the protein and therefore a generalised derepression of transcription. Many other groups have undertaken mutation analysis, using a variety of molecular techniques. Mutation detection in classically affected females varies from 46% to 80% (Cheadle *et al.*, 2000; Amano *et al.*, 2000; Bienvenu *et al.*, 2000). So far, no substantial evidence has been found that MeCP2 is linked to tumourigenesis.

1.5. Developmental gene regulation-hedgehog pathway

1.5.1. General introduction

The *Hh* gene was first identified as one of the many players that are required for segmentation of the *Drosophila* embryo (Nüslein-Volhard and Wieschaus, 1994).

We now know that it is also involved in numerous other aspects of embryonic, larval and adult development in the fly and that it has homologs that play key roles in human, mouse, frog, fish and chick development. The Hh proteins are secreted and are thought to function as morphogens, signals that elicit concentration-dependent responses from target cells. Increasing evidence show that Hedgehog (Hh) signalling is a universal pathway that specifies and regulates the growth and differentiation of organ systems in eukaryotes. It is crucial in patterning a diverse range of vertebrate structures and organs during embryo development (Ingham and McMahon, 2001).

In contrast with *Drosophila,* which has only one hedgehog gene, three vertebrate homologues have been identified, *Sonic (Shh), Desert (Dhh)*, and *Indian hedgehog (Ihh)*. *In vitro* studies show that each of these proteins can act via the same signal transduction pathway and that the different hedgehog genes regulate patterning of different organ systems by their unique expression pattern. The most extensively studied vertebrate hedgehog gene is *Shh* that is expressed widely in the developing CNS, limb, lung, gut, teeth and hair-follicle. *Dhh* and *Ihh* are found mainly involved in development of the germline and skeletal system respectively. (Ingham, 1998; Goodrich and Scott, 1998; Muenke and Beachy, 2001; McMahon, 2000).

1.5.2. The mechanism of hedgehog signalling

Studies in vertebrate systems have built up a model for hedgehog signalling (see figure 3). Hedgehog signals are received at the cell surface by a complex consisting of the patched and smoothened proteins. According to this model, patched, a 12-pass transmembrane protein is the ligand-binding component of the receptor complex.

Smoothened, a protein with homology to a serpentine G-protein coupled receptor, is responsible for transducing the hedgehog signal. In the absence of hedgehog binding, patched is thought to hold smoothened in an inactive state and thus inhibit signalling to downstream genes. With the binding of hedgehog, patched inhibition of smoothened is released and SMO then transduces a signal. A downstream complex composed of fused, suppressor of fused, costal 2 and ci dissociates, and an active form of ci translocates to the nucleus where it switches on transcription of the target genes, *wingless* (homologous to the vertebrate Wnt genes), *decapentaplegic* (a member of the superfamily most homologous to the vertebrate bone morphogenetic proteins (Bmps), as well as *patched* itself (Saldanha, 2001).

So far, it is known that the Hh signal regulates target genes by two ways. One is to activate Gli proteins to induce gene transcription and the other is to inhibit the formation of Gli repressors (mostly those of Gli3) to derepress targets. Specifically, Hh signaling induces *Gli1* transcription and thus Gli1 acts as an activator to amplify the Hh response. However, Gli2 and Gli3 functions are more complex than the role of Gli1, and possibly Gli3 can be variable in relation to Hh signaling in different situations. Hh signaling represses both the transcription of *Gli3* and the proteolytic formation of Gli3 repressors. Hh signaling turns full-length Gli2 into a potent activator. Besides, both *Gli2* and *Gli3* could be involved in other signaling pathways, as they are often expressed independent of Hh signalling (Matise and Joyner, 1999).

Figure 3. Elements of hedgehog pathway conserved from Drosophila to vertebrates

Hedgehog binds to patched releasing smoothened to transduce a signal. A downstream complex composed of fused, suppressor of fused, costal 2 and ci dissociates, and an active form of ci translocates to the nucleus where it switches on transcription of the target genes, *wingless*, *decapentaplegic* and *patched*. PKA, probably regulated by a parallel pathway, can inhibit activation of ci.

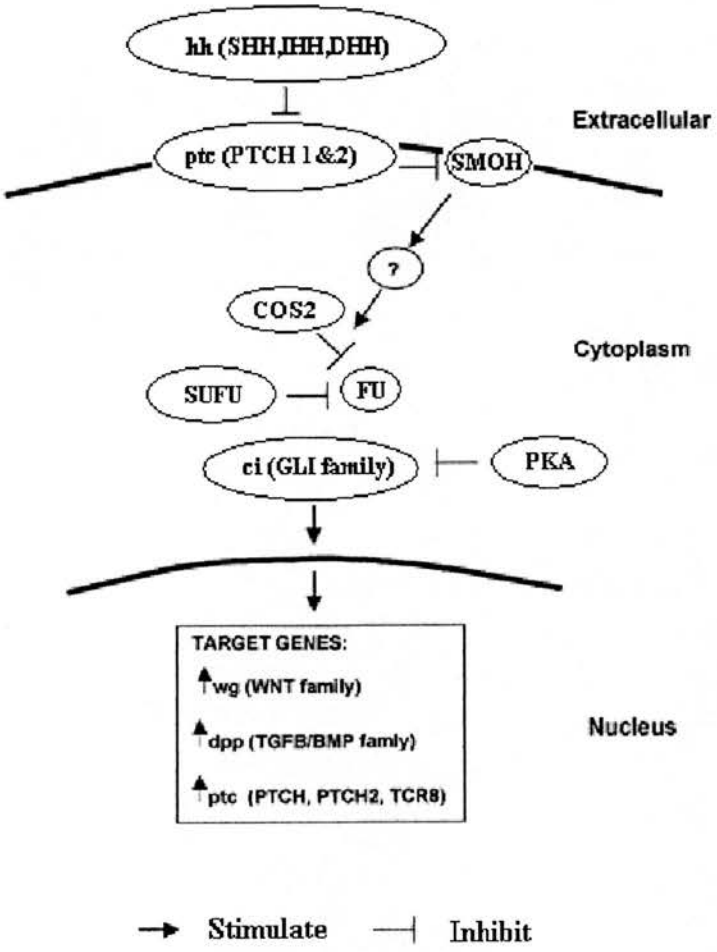


Figure 3. Elements of hedgehog pathway conserved from *Drosophila* to vertebrates (human gene families in parentheses).

1.5.3. The role of Hh in animal development

Different roles have been described for Shh during development, acting as a morphogen, mitogen or differentiation factor (Altaba *et al.*, 2003). Hedgehog is also important for development of the gut in *Drosophila* larvae. *Shh* is required for the development of small intestine (Zhang *et al.*, 2001) and for larval to adult intestine remodelling in *Xenopus* (Ishizuya-Oka, 2001). In the chick, *Shh* is involved in establishing the anterior/posterior and radial axis of the gut (Roberts *et al.*, 1998 and Sukegawa *et al.*, 2000). Shh knockout mice have been reported to show foregut, trachea and lung abnormalities (Litingtung *et al.*, 1998), reduction of intestinal smooth muscle, gut malrotation and annular pancreas (Ramahlo-Santos *et al.*, 2000) and persistent cloaca, where the distal intestinal and genitourinary tracts remain in a common channel.

1.5.4. Hedgehog pathway and cancer

Since *Gli1* was first identified as an amplified gene in a human glioblastoma, it has also been found that several key components of the hedgehog pathway have been implicated in tumour formation.

1.5.4.1 PTCH and cancer

It was the study that showed *patched* is mutated in both familial (NBCC, nevoid basal cell carcinoma syndrome or Gorlin's syndrome) and sporadic forms (BCC) of basal cell carcinoma that firmly consolidated this pathway as a crucial player of tumourigenesis. The evidence of *PTCH* mutations in sporadic BCCs, in many cases with both alleles inactivated by either mutation or loss of heterozygosity supports

that *PTCH* acts as a putative tumour suppressor gene (Wicking *et al.*, 1997; Gillies *et al.*, 1997; Gailani *et al.*, 1996). $Ptc^{+/-}$ mice also displayed a phenotype partly similar to that of NBCC patients(Wetmore *et al.*, 2000).

1.5.4.2. SMO and cancer

Mutations in *SMO* have also been detected in 10–20% of BCCs and primitive neuroectodermal tumours. In these cases, the mutations appear to activate *SMO*, and so activate the HH pathway (Lam *et al.*, 1999; Reifenberger *et al.*, 1998).

1.5.4.3. Glis and cancer

Dysregulation of GLI1 has been implicated in HH pathway diseases. Since the original isolation of *GLI1* from a glioma line (Kinzler *et al.*, 1987), the evidence of GLI1 involvement in brain tumours has increased more recently. Gli1 misregulation was subsequently shown to lead to the development of BCC-like tumours in frog embryos and mice (Dahmane *et al.*, 1997; Oro *et al.*, 1997), Because the Hh–Gli pathway is active in precursor populations in a variety of tissues and organs, deregulation of Gli1 function is likely to be involved in tumour development in tissues that utilise the Hh pathway for their development or maintenance. In contrast to the gain-of-function phenotype of GLI1, there are no human diseases associated with loss of GLI1 function to date, and in mice Gli1 appears to be redundant. *GLI1* has been classified as an oncogene on the basis of its ability to transform cells in cooperation with adenovirus E1A (Ruppert *et al.*, 1991).

So far, although no defects in *GLI2* have been found associated with human diseases, studies in mice suggest that involvement in basal cell carcinomas, skeletal defects,

and other disorders are likely (Sasaki *et al.,* 1999; Grachtchouk *et al.*, 2000; Park *et al.*, 2000).

GLI3 frameshift and nonsense mutations have been linked to Pallister-Hall syndrome (PHS), an autosomal dominant disease involving hypothalamic hamartoma (a kind of benign midline tumours of the ventral forebrain), central or postaxial polydactyly, syndactyly, imperforate anus, anteverted nares and other facial abnormalities, and associated HPE and malformations of the axial skeleton (Jones 1997; Kang *et al.*. 1997). The mutation in PHS maps to the C-terminus of the *GLI3* gene downstream of the zinc-finger coding region, and would be predicted to give rise to a C-terminally truncated protein with constitutive repressor activity. This hypothesis is supported by the fact that in GCPS (Greig cephalopolysyndactyly syndrome) patients do not develop hypothalamic hamartomas. Mutations in human *GLI3* have been implicated in several types of birth defects. Translocations, deletions, and point mutations throughout the *GLI3* gene cause GCPS, which is characterized by syndactyly, predominantly preaxial polydactyly, broad thumbs and first toes, and facial anomalies such as hypertelorism and frontal bossing (Vortkamp *et al.*, 1991; Jones, 1997; Wild *et al.*, 1997; Kalff-Suske *et al.*, 1999).

2. Colon cancer

2.1. General introduction

After lung and breast cancer, colorectal cancer is the most common cause of death from malignant disease in the Western world. But unlike for the commonest cause of lung cancer, the basis of the initiation of this disease is currently not understood.

Fortunately the morphological observation that most colorectal cancers develop from normal epithelium through sequentially worsening degrees of adenomatous dysplasia provide a good model for us to study the genetic alterations that underlie the visible progression (see figure 4).

Colon cancer is usually observed in one of two specific patterns: sporadic, inherited (or familial). Sporadic disease, with no familial or inherited predisposition, accounts for approximately 70% of colorectal cancer in the population. Sporadic colon cancer is common in persons older than 50 years of age, probably as a result of dietary and environmental factors as well as normal aging. While the majority of cases of colorectal cancer are sporadic, significant minorities occur as a result of an inherited genetic mutation. Familial adenomatous polyposis (FAP), the polyposis syndromes and hereditary non-polyposis colorectal cancer (HNPCC) account for about 5% of all colorectal cancers (CRCs).

In most colorectal cancers the causative mutations in tumour suppressor genes and oncogenes occur at different times during tumourigenesis, specifically, adenomatous polyposis coli (*APC*) gene mutations, global hypomethylation, K-ras mutations, deleted in colon cancer (*DCC*) gene mutations, and finally mutations in the p53 gene.

Figure 4. Molecular basis of colon adenoma-carcinoma progression

APC mutations initiate the neoplastic process, and tumour progression results from mutations in the other genes indicated. Patients with FAP inherit *APC* mutations and develop numerous dysplastic aberrant crypt foci (ACF), some of which progress as they acquire the other mutations indicated in the figure. The tumours from patients with HNPCC go through a similar, though not identical, series of mutations; MMR deficiency speeds up this process. *K-RAS* is an oncogene that requires only one genetic event for its activation. The other specific genes indicated are tumour suppressor genes that require two genetic events (one in each allele) for their inactivation. Chromosome 18q21 may contain several different tumour suppressor genes involved in colorectal neoplasia, with *DCC* and other genes proposed as candidates. A variety of other genetic alterations have each been described in a small fraction of advanced colorectal cancers. These may be responsible for the heterogeneity of biologic and clinical properties observed among different cases.

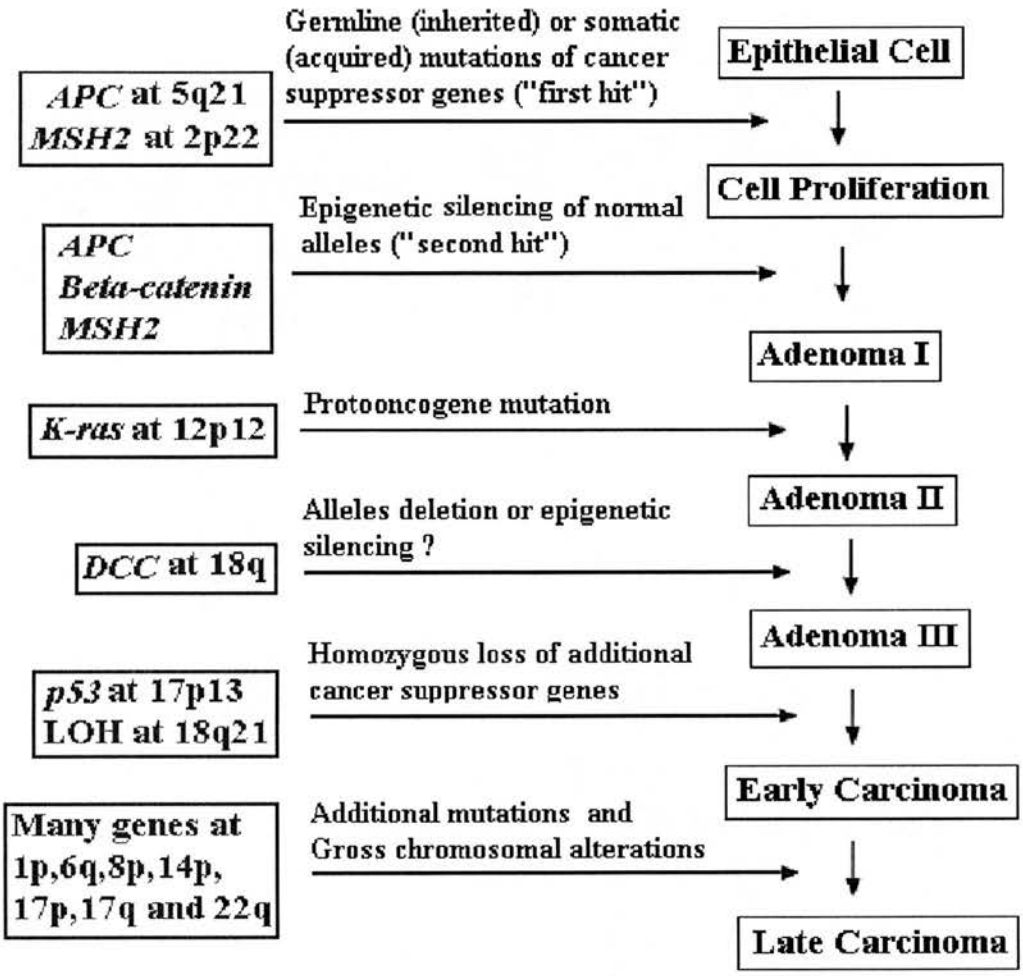


Figure 4. Molecular basis of colon adenoma-carcinoma progression

Figure 5. Mechanism of tumour suppressor gene inactivation

In the early 1970s, Alfred Knudson suggested that two mutations or 'hits' were sufficient for the development of a retinoblastoma and that the inheritance of one of these mutations could account for the earlier onset and frequent bilateral occurrence of the hereditary form of this tumour. Subsequent molecular studies supported Knudson's hypothesis by demonstrating mutations of both alleles of a tumour suppressor gene-'*RB1*' in both hereditary and sporadic retinoblastomas. In hereditary retinoblastoma, individuals begin life with a constitutional mutation that inactivates one allele of the *RB1* gene. The 'second hit' occurs somatically and usually involves loss of all or part of the chromosome containing the normal *RB1* allele. Because of the nature of this second mutation, other genes and genetic markers in the region of the normal *RB1* allele are often lost within the tumour cells as well. If some of these genetic markers happen to be heterozygous in the individual, loss of one allele on the same chromosome as the normal *RB1* allele produces loss of heterozygosity (LOH), a cell-specific phenotype that is relatively easy to find in tumour tissue. Genes like *RB1* that permit tumour development when both alleles are inactivated or lost are known as 'tumour suppressor genes', and dozens of such genes have now been identified.

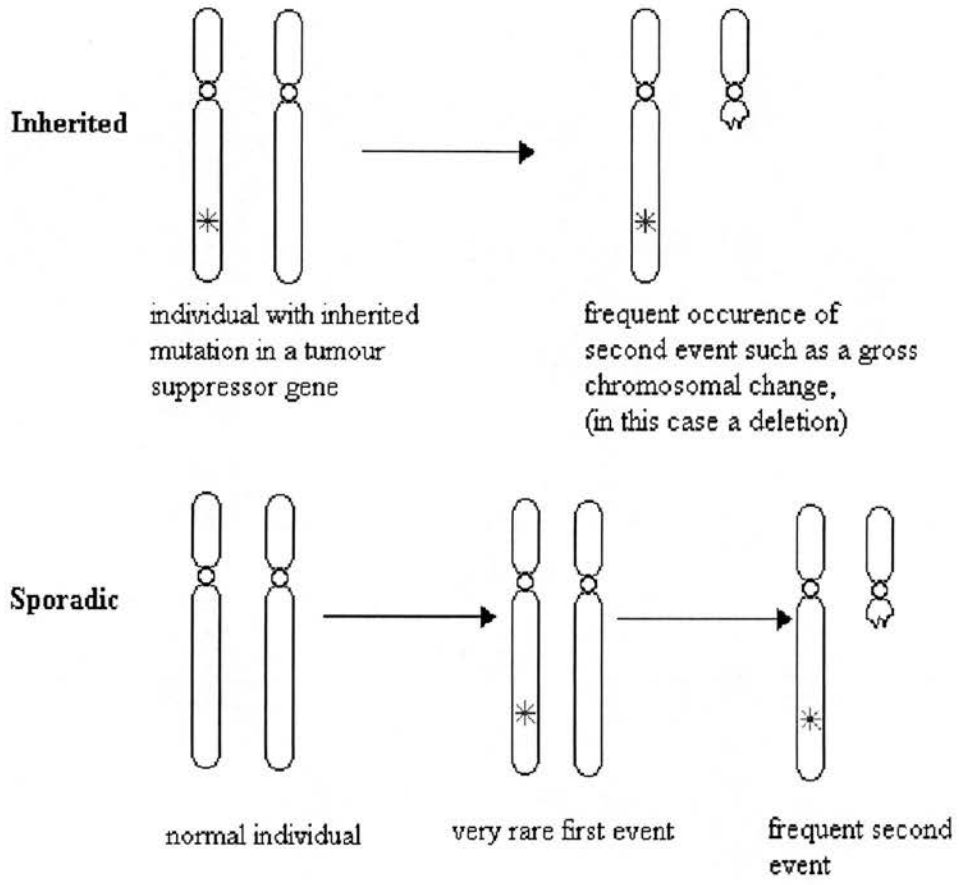


Figure 5. Mechanism of tumour suppressor gene inactivation

(Knudsen's two hit hypothesis for *RB*)

2.2. *p53*

There is no doubt about the role of *p53* mutations in the progression of colorectal tumours. Genetic alternation in the *p53* gene occur in around 80% of colorectal cancers, and *p53* mutations are rare in adenomas, suggesting that *p53* mutations occur more frequently in high-grade dysplastic polyps and are thought to mark the malignant transition from adenoma to carcinoma (Baker *et al.*, 1990).

2.3. *APC* and the Wnt signalling pathway

The genetic basis for FAP mainly lies in germline (inherited) mutation of the adenomatous polyposis coli (*APC*) gene. About one-quarter of all cases are caused by mutations that maintain the incidence of FAP (Bisgaard, 1994). The first clue to the localisation of of the *APC* gene came from identification of a patient with colorectal polyposis and mental retardation who had a deletion of the chromosomal fragment 5q21(Herrera,1986). Linkage analysis of families with FAP led to the mapping of the *APC* gene to 5q21 in 1987(Bodmer *et al.*, 1987). The *APC* gene was then cloned, identified and characterized in 1991(Groden *et al.*, 1991; Kinzler *et al.*, 1991). Exon 15 comprises more than 75% of the coding sequence of *APC* and is the most common target for both germline and somatic mutations (Beroud *et al.*, 1996). Germline mutations in the *APC* gene have been reported in most FAP patients (Cottrell *et al.*, 1992). The majority (95%) of *APC* mutations are nonsense or frameshift mutations that result in a truncated protein product with abnormal function. As expected from Knudson's two-hit hypothesis (see figure 5), colon tumours from FAP patients nearly all harbour either additional somatic *APC* mutations or loss of heterozygosity at the *APC* locus in addition to the original

germline mutation. The type of germline *APC* mutation in FAP appears to be correlated with the nature of the second somatic hit to *APC*. If the germline mutation occurs between codons 1194 and 1392, then there is strong preference for allelic loss of *APC* as the second hit in the development of a colorectal adenoma. If the germline mutation lies outside this region, the second hit in tumourigenesis is most likely to produce a truncating mutation in the somatic mutation cluster region (MCR) between codons 1286 and 1513 (Lamlum *et al.*, 1999).

In addition to the role of the APC gene in the aetiology of familial adenomatous polyposis coli (FAP), mutations in this gene cause sporadic colorectal cancer. *APC* mutations, which generally bring about a truncated APC protein (Miyoshi *et al.*, 1992), or take the form of allele loss (Solomon *et al.*, 1987), are found in about 75% of sporadic colorectal cancers (Miyaki *et al.*, 1994) and are observed in the earliest adenomas (Powell *et al.*, 1992). Accumulating evidence shows that the APC may function as a gatekeeper tumour suppressor in a wide variety of cellular processes including cytoskeletal organization, migration, adhesion, proliferation, even perhaps aspects of chromosome stability and the control of cellular proliferation, possibly by affecting the rate of cell division or apoptosis (Baeg *et al.*, 1995; Su *et al.*, 1993; Rubinfeld *et al.*, 1993; Browne *et al.*, 1994; Burchill *et al.*, 1994; Munemitsu *et al.*, 1994).

Figure 6. The canonical Wnt pathway

Activators of the pathway are white; negative regulators are gray. Left, in the absence of Wnt stimulation, the Axin complex actively earmarks β-catenin/Armadillo (white circles) for degradation by the proteasome. The levels of cytoplasmic β-catenin/Armadillo are low, and TCF is repressed. Right, after Wnt stimulation of the Frizzled receptor (arbitrarily drawn to be in the apical membrane), Dsh is recruited to the membrane where it binds to Axin to inhibit the Axin complex. β-catenin/Armadillo accumulates and, after translocation into the nucleus, binds to TCF to coactivate Wnt target genes. Inhibition of the Axin complex by GBP/Frat appears to be an alternative to Wnt-mediated inhibition. CBP switches from being a negative regulator to being a coactivator, apparently depending on the stimulation status of the cell. Note also the apicolateral adherens junctions to which the Axin complex appears to be anchored. These junctions are formed by the transmembrane protein E-cadherin (black bars) which is linked to the actin cytoskeleton (thin lines) by β-catenin/Armadillo and α-catenin (black dots).

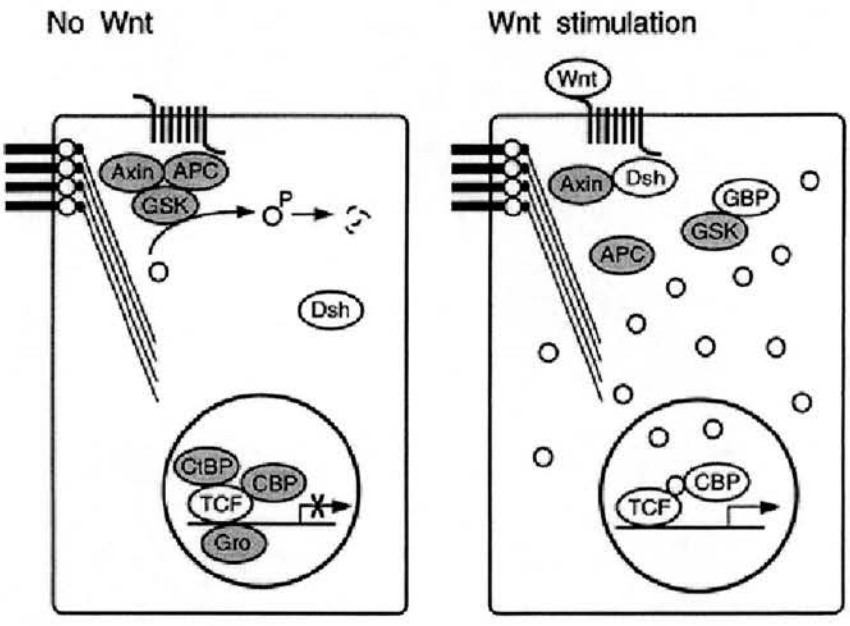


Figure 6. The canonical Wnt pathway
(Cell, Vol. 103, 311-320)

APC plays a major role in the Wnt signalling pathway where, in the absence of Wnt, its main partners are axin and glycogen synthase kinase 3β (GSK3β)(Behrens *et al.*, 1998). In these conditions, the complex actively targets β–catenin for degradation by the proteasome. In the absence of this regulation, β–catenin escapes degradation and translocates to the nucleus where it complexes with one of the TCF/LEF transcription factors and initiates transcription of a wide variety of genes (Polakis, 1999). Mutations of *APC* inhibit its ability to signal degradation of β–catenin leading to constitutive activation of downstream genes. In many of those relatively few colon cases where *APC* is not mutated, there is a mutation of *β–catenin* (Iwao *et al.*, 1998; Sparks *et al.*, 1998; Mirabelli-Primdahl *et al.*, 1999; Miyaki *et al.*, 1999) or one of the other members of the pathway (e.g. axin)(Liu *et al.*, 2000 and Webster *et al.*, 2000).

2.4. *K-ras*

Although mutations in the *APC* gene are initiating events in colorectal tumourigenesis, these mutations are not enough for adenomas to progress malignancy. Several other genes are involved in the progression of early adenomas to early carcinomas. More than 50% of colorectal cancers display specific mutations in the *K-ras* gene, with an increasing frequency in larger and more advanced lesions (Fearon and Vogelstein, 1990). The consequence of *K-ras* mutations during tumour development may be a growth advantage of those cells with both *APC* and *K-ras* mutations over cells with *APC* mutations alone.

2.5. *DCC* and chromosome 18q

During the past 10 years, other mutations that occur in colorectal cancer have been identified. The *DCC* gene was identified as a result of the frequent allele loss close to its location on 18q21.3 in colorectal cancers (Vogelstein *et al.*, 1988). DCC is a neural cell adhesion molecule homologue and *DCC* mutations may therefore have a role in colorectal tumour progression, invasion and metastasis. Although genetic mutations occur relatively infrequently (Cho *et al.*, 1994), a high incidence of reduced expression has been reported (Shibata *et al.*, 1996; Saito *et al.*, 1999). This may be due to the allelic imbalance (seen as LOH), and the role of DCC in colorectal carcinogenesis remains equivocal. Apart from *DCC*, there is also evidence showing that *SMAD4* is also a target of allele loss on chromosome 18q in some cancers. Smad4 acts as a cofactor that binds transforming growth factor-β (TGF-β) receptor-activated Smad2 and Smad3 generating transcriptional complexes (Lagna, 1996). Inactivating mutations in *Smad4/DPC4* can cause the post-receptor defects of TGF-β and so promote cell proliferation. Following the initial identification of *Smad4* mutations in half of all pancreatic carcinomas (Hahn *et al.*, 1996), *Smad4* mutations were reported in colon cancer (Takagi *et al.*, 1996; Thiagalingam *et al.*, 1996; Howe *et al.*, 1998) and other gastrointestinal cancers (Powell *et al.*, 1997). In colon cancer, the incidence of Smad4 mutations appears to increase with tumour progression (Howe *et al.*, 1998). Although it is rare as an initiating event and relatively infrequent (approximately 10% of cases) in colon adenomas and nonmetastatic carcinomas (Riggins *et al.*, 1997), *Smad4* mutations have been reported in more than 30% of invasive metastatic carcinomas and in colon cancer metastases (Miyaki M *et al.*, 1999).

2.6. *LKB1* and chromosome 19

Chromosome 19 abnormality is also a feature in some cancers. The evidence that chromosome 19 might harbour tumour suppressor genes was found in lung cancer (Virmani *et al.*, 1998; Sanchez-Cespedes *et al.*, 2001) and prostate cancer (Gao *et al.*, 1999). In these studies, the deletions of chromosome 19 were observed in cancer cell lines and primary tumours.

LKB1, also *STK11* or *LKB1/STK11* is located on chromosome 19p13.3 (Hemminki *et al.*, 1997) and encodes a serine-threonine kinase, a human homologue of *Xenopus* early embryonic kinase 1. (Su *et al.*, 1996). LKB1 is suggested to act as a tumour suppressor gene in PJS (Peutz-Jeghers syndrome) as hamartoma formation in PJS patients with inactivating *LKB1* germline mutations is associated with somatic loss of the wild-type *LKB1* allele. (Hemminki *et al.*, 1997, Gruber *et al.*, 1998; Trojan *et al.*, 1999)

The development of cancer in patients with PJS does not only arise in association with hamartomas (Giardiello *et al.*, 1987) but dysplasia with consecutive neoplastic transformation within hamartomatous polyps accounts for at least some malignancies in this syndrome (Spigelman *et al.*, 1989; Boardman *et al.*, 1998). In contrast with the tumourigenesis of sporadic colorectal cancer, which is frequently involving *APC, K-ras*, *DCC* and *p53*,(Kinzler and Vogelstein, 1996; Lynch,1998; Vogelstein *et al.*, 1988) the molecular mechanisms of cancer in PJS patients remain unclear. Since patients with PJS are prone to develop colorectal cancer, *LKB1* may also be a vulnerable target during the pathogenesis of sporadic colorectal cancer. Although

some reports showed a low frequency of somatic mutations of the *LKB1* gene in colorectal tumour, (Aviziente, 1998; Wang, 1998; Resta, 1998) conflicting data were reported by Dong and colleagues (Dong, 1998). This group identified somatic *LKB1* mutations in one third of left sided colorectal cancers and in two colonic adenomas. Also, silencing of *LKB1* associated with promoter hypermethylation was found in a few cancer cell lines and primary tumours (Esteller, 2000).

2.7. DNA mismatch repair genes

Mutations in three MMR genes primarily cause the dominantly inherited syndrome hereditary non-polyposis colorectal cancer (HNPCC): MSH2 on chromosome 2p, MLH1 on chromosome 3p, and MSH6 on chromosome 2p16(Leach *et al.*, 1993; Fishel *et al.*, 1993; Aaltonen *et al.*, 1993; Kolodner *et al.*, 1999). Colorectal cancers from patients with mutations in these MMR genes consistently show microsatellite instability (MSI), a form of replication error (RER). It is generally accepted that mutations in the MMR genes in HNPCC families act to not only increase the mutation rate of the whole genome, but promote tumourigenesis when a gatekeeper gene (e.g. APC) is affected. MSI also occur in ~15% sporadic Colorectal cancers (CRCs), but is mainly due to hypermethylaion of MLH1 (see next section), and mutation of MSH2 to a lesser extent (Liu *et al.*, 1995). Like p53 mutations, MMR abnormalities often occur in late colonic adenomas, and are therefore involved in the progression of tumours rather than initiation (Tomlinson *et al.*, 1996).

2.8. Other genes in colon carcinogenesis

Other genes such as *PTEN* (Zhou *et al.*, 2002), *FHIT* (Mady and Melhem, 2002), *p16* (Herman *et al.*, 1995) and *NM23* (Hsu *et al.*, 1995) have been found to be involved in colorectal tumourigenesis in a small proportion of cancer. PTEN and FHIT may play a role as serine/threonine kinase and protein kinase B (PKB) inhibitors. *p16* is an important member of the cell cycle control system. NM23 is associated with cell adhesion and metastasis of several cancers.

2.9. Epigenetic events in colon cancer

In addition to common deletions and inactivating mutations, epigenetic events have recently been recognised as an alternative mechanism of gene silencing. The hypermethylation of CpG islands in the promoter region, an important epigenetic mark, has been found important for colon cancer by reports of homozygous promoter hypermethylation of the *hMLH1* gene in about 80% of sporadic RER+ tumours (Cunningham *et al.*, 1998). In addition, promoter hypermethylation has been shown associated with the silencing of *p16* in some colorectal cancer cell lines and patients.

One group found that *Drg-1* was silenced by promoter hypermethylation (Guan *et al.*, 2000). It is a differentiation-related and putative metastatic suppressor gene. It may induce colon cancer cell differentiation and partially reverse the metastatic phenotype. CDX1, a homeobox protein that inhibits proliferation of intestinal epithelial cells and regulates intestine-specific genes involved in differentiation, was reported to be down regulated by aberrant methylation of the CpG island in the *CDX1* promoter colorectal cancer cell lines (Suh *et al.*, 2002). Recently, one study

showed *SLC5A8,* a sodium transporter, was silenced by methylation in human colon aberrant crypt foci and cancers and suggested that it may be a tumour suppressor gene (Li *et al.*, 2003).

3. Lung cancer

3.1. General introduction

Lung cancer has become the leading cause of cancer death in many industrialised countries. A better understanding of the molecular pathogenesis of this fatal disease is thus urgently needed in order to provide effective preventive (other than abstinence) or therapeutic routes for clinical management. It is widely accepted that carcinogens in the cigarette smoke are major triggers of lung cancer by the formation of DNA adducts that induce multiple genetic alterations. Well functioning cell cycle regulation and checkpoints are crucial for maintaining genomic integrity by allowing time for repair of damage, and their abnormality are thought to contribute to genomic instability, thus playing an important role even in the early steps of cancer development. Many of the tumour suppressor genes and oncogenes altered in lung cancer are believed to play a role in the regulation of cell cycle progression in either a direct or an indirect manner, and a large proportion of the lung cancer-related genes are a component of the checkpoint mechanisms (see table 1).

3.2. *p53*

Loss of Heterozygosity at 17p13, which contains p53, is very frequent in lung cancer, associated with mutation of the other p53 allele. Such inactivation of p53

occurs in ~90% of SCLC (Small cell lung cancer) and 40~70% of NSCLC (Non-small cell lung cancer)(Harris, 1996).

3.3. *RB*

The *RB* gene is located on chromosome 13q14, and its protein product is a nuclear phosphoprotein initially identified as a tumour suppressor gene in childhood retinoblastomas. The phosphorylation status of the RB protein and its interaction with transcription factor E2F is one of the most important determinants in the regulation of G0/G1 transition. Abnormalities of the *RB* gene in lung cancer include deletions, nonsense mutations and pathogenic splicing variations. More than 90% of the SCLCs, and 15-30% of the NSCLCs have abnormal or no *RB* expression (Reissmann *et al.*, 1993, Cagle *et al.*, 1997 and Dosaka-Akita *et al.*, 1997). In addition germline carriers of a *RB* mutation are 15 times more likely to die from lung cancer than unaffected individuals (Sanders *et al.*, 1989).

3.4. *p16*

p16/INK4 is an inhibitor of RB phosphorylation and thus it is also a tumour suppressor gene. The *p16/INK4* gene is most commonly altered in NSCLCs by aberrant promoter methylation (25%)(Zöchbauer-Müller *et al.*, 2001; Merlo *et al.*, 1995) and homozygous deletions or point mutations (10%-40%)(Rusin *et al.*, 1996; Marchetti *et al.*, 1997; Shapiro *et al.*, 1995; Okamoto *et al.*, 1994). In lung cancer, apart from mutational inactivation, promoter region methylation has been found to be a major alternative mechanism silencing the expression of *p16.*

3.5. *p19*

p19/ARF is a p16 splice variant leading to an altered reading frame from *p16/INK4*. p19/ARF was shown to play an important role in tumour suppression with binding to the MDM2-p53 complex and thus preventing p53 degradation. p19/ARF was found more frequently lost in lung tumours with neuroendocrine features(Gazzeri *et al.*, 1998). Thus the p16/INK4 p19/ARF locus products interact with both the Rb and p53 pathways.

3.6. Other candidate tumour suppressor genes for lung cancer

The most significant chromosome event is the abnormality of chromosome 3p in SCLC and NSCLC samples. This region harbours many candidate genes in 4 localised regions (3p25-p26, 3p21-p22, 3p14, and 3p12) (Hibi *et al.*, 1992; Latif *et al.*, 1992), which are undergoing active investigation. These genes include retinoic acid receptor genes, phosphatase genes, members of the ubiquitin activation family, *DUTT1*, *BAP1*, *RASSF1A* and many others. The most prominent candidate chromosome 3p gene is the *FHIT* gene at 3p14.2 which undergoes genomic alterations with absent protein expression in a large subset of human lung tumours (Zochbauer-Muller *et al.*, 2001). The function of its protein product and its role in lung tumourigenesis, however, is still unknown.

3.7. *Ras*

The dominant oncogene *RAS* plays a key role in signal transduction and cell proliferation. *RAS* mutations are rare or non-existent in SCLC, but are present in 15-

20% of NSCLC. Up to 30% of the adenocarcinomas carry *RAS* mutations usually affecting codon 12 for *KRAS* (85% of cases), and uncommonly codon 13 of *HRAS* and codon 61 of *NRAS* (Richardson and Johnson, 1993).

3.8. *MYC*

The *MYC* proto-oncogenes (*MYC, MYCN, and MYCL*) encode nuclear phosphoproteins that have a role in transcriptional regulation by heterodimerizing with proteins such as MAX, MAD or MX11. The MYC-MAX complex represses transcriptional activation. MAX can bind MAD and MX11, thereby MYC is released from the complex and functions as a transcriptional activator. MYC can cooperate with a mutant RAS gene to transform primary rat embryo fibroblasts to malignancy. The activation of the *MYC* genes by amplification or loss of transcriptional control resulting in protein overexpression is a major molecular mechanism in the pathogenesis of human lung cancers. *MYC* gene activation has been observed in both NSCLC and SCLC whereas *NMYC* and *LMYC* abnormalities mainly occur in SCLC. MYC amplification occurs in 15-30% of SCLC and 5-10% of NSCLC (Richardson and Johnson, 1993).

3.9. Epigenetic events in lung cancer

Gene expression can be turned off by aberrant promoter methylation in cancer as mentioned before. Alteration in 5′-CpG island methylation has been reported for several genes in lung cancer (Esteller *et al.*, 1999). One recent study shows that 8 genes are aberrantly methylated in NSCLC. Specifically, *RARB* in 40%, *metalloproteinase-3 inhibitor (TIMP-3)* in 26%, *p16INK4a* in 25%, *O6-*

methylguanine-DNA-methyltransferase (MGMT) in 21%, *death-associated protein kinase (DAPK)* in 19%, *E-cadherin (ECAD)* in 18%, *p14ARF* in 8% and *glutathione-S-transferase P1* in 7% samples tested were hypermethylated (Zöchbauer-Müller *et al.*, 2001).

Table 1. Oncogenes and tumor-suppressor genes altered in lung cancer

Oncogenes	SCLC	NSCLC
	c-myc*	K-ras*
	L-myc	N-ras
	N-myc	H-ras
	c-raf	c-myc
	c-myb	c-raf
	c-erbB-1(EGF-R)	c-fur*
	c-fms	c-fes
	c-rlf	c-erbB-1(EGF-R)
		c-erbB-2(Her-2/neu)
		c-sis
		Bcl-1
Tumour suppressor enes	p53*	p53*
	p16	
	p19	
	RB*	RB
EGF-R=epidermal growth-factor receptor; NSCLC=nonsmall cell lung cancer; SCLC=small cell lung cancer. *Most frequently altered genes in tumours or cell lines evaluated		

4. Summary and aims

It is well known that cancer progression is a multi-step process and that multiple genetic and epigenetic events will occur in tumours. Given that there are about 3 x 10^9 nucleotides encoding more than 30000 to 40000 genes in the human genome, identifying these genetic and epigenetic abnormalities is a difficult task. This is further complicated by the fact that since tumour cells are genetically and

epigenetically unstable, a range of abnormalities may additionally occur in genes that are not involved in the initiation and/or development of the tumour.

Despite this complexity, many genes that are involved in tumourigenicity have been identified, some of which are common to many tumour types and some of which are more restricted in their tissue distribution. The techniques used to discover these genes have included loss of heterozygosity analysis, positional cloning, cloning of chromosomal breakpoints, candidate gene analysis and combinations of these. In both types of cancer many genes have been identified, some common to both (e.g. *p53*) and some specific to one (e.g. *APC* in colon). Damage to these genes in lung cancer seems to have a relatively random pattern, perhaps due to the massive chemical assault caused by smoking, whereas a particular pattern of genetic and epigenetic events is generally thought to occur in sporadic colon cancers as cells progress from normal to malignant stages.

However, there are still many gaps to be filled. For example, there are chromosomal regions that are known to suffer loss of heterozygosity that harbour tumour suppressor genes, but for which there is a relatively limited incidence of abnormality in the cases studied. This disparity of LOH and gene abnormality suggests a lack of complete analysis of the known tumour suppressor gene by techniques to screen for all forms of potential inactivation (e.g. mutations are known already but epigenetic screening has not been done to look for silencing of the gene), or it suggests the association of another gene from the region. So, for example in the case of colon cancer, LOH occurs at chromosome 18q21 in about 65% of cases, but mutations of

the known tumour suppressor genes of the region (*DCC, SMAD4*) occur at lower frequencies (5% and 10~30%). This implies either another neighbouring gene of greater importance or some combination of known and unknown tumour suppressor genes from the area is necessary for tumourigenesis. The same is true for lung cancer in 18q21, and for both cancers in chromosome 19p13, where this thesis will address the role of candidate genes as outlined below.

It is also possible that other pathways are disrupted or disregulated that are not part of the current paradigm of tumourigenesis, especially the much favoured genetic steps in the adenoma-carcinoma progression of colon cancer. Reasons for the need to keep an open mind on this include several observations. For example, the paradigm involves mutations in all of the three main genes studied (*APC*, *K-ras* and *p53*), but in a study of all of these genes in a set of colon cancers by Smith *et al.* (Smith *et al.*, 2002) they found only 6% had all three mutations and a significant percentage had a mutation in only any one of them. They concluded that the mutations lay on alternate pathways of colorectal development. Similarly, although many of the genes found to be abnormal in colon cancers derived from ulcerative colitis are the same as those in sporadic cancers, the frequencies of defects are significantly different so as to indicate an alternative molecular pathway for aetiology (Wong and Harrison, 2001; Benhattar and Saraga, 1995).

The aim of this study was to take the candidate gene approach to test two examples of genes or intracellular pathways that could control or suffer epigenetic changes in colon and lung cancer:

1) **to continue the investigation of members of the MBD family whose proteins are important in recognition of methylation patterns and epigenetic regulation of other genes.**

MBD2: This gene lies within chromosome region 18q21, a major zone of LOH in colon and lung cancers, near to the *DCC* and *SMAD4* tumour suppressor genes. A survey for mutations in *MBD2* was published by our group (Bader *et al.*, 2003) in which it was reported that the gene suffered few changes. However, mRNA levels of *MBD2* are reported to be reduced in colon and lung tumours (Kanai *et al.*, 1999, Müller-Tidow *et al.*, 2001) suggesting inactivation by gene silencing is more important for this gene. The aim of this study was therefore to screen our own set of cell lines for expression and the presence of methylated sequences in the putative promoter region of *MBD2*.

MBD3: This gene is located within chromosome region 19p13, a zone of LOH, and within 500kbp of the *LKB1/STK11* tumour suppressor gene. The latter gene is involved in cancers from patients with Peutz-Jeghers syndrome but is relatively rarely affected in sporadic colon cancers or lung cancers. *MBD3* is thus a suitable alternative candidate and this study aimed to screen for mutations, abnormalities of expression or methylation of the promoter of this gene.

2) **to investigate the Hedgehog pathway whose activity is involved in normal tissue development**.

As stated above, the Hedgehog pathway is both important in colon and lung development and has been implicated in tumourigenesis of certain other tissues. Preliminary data from our laboratory has shown that proteins of the hedgehog pathway are still expressed in normal adult (as opposed to foetal) colon and also in

hyperplastic, adenomatous and carcinomatous colon tissues. It was therefore of interest to assay genes of the pathway for mutations and abnormalities of methylation or expression at the RNA level, to see if Hh signalling is an alternative pathway of tumourigenesis in these tissues.

Chapter 2

Materials and Methods

Contents

1.Cell culture and primary tumour samples

1.1.Cell culture

Aseptic techniques were strictly applied and cells were inspected frequently to ensure they were in good condition without microbiological contamination. All cell lines were cultured in RPMI-1640 medium supplemented with 5% FBS and incubated at 37°C with 5% CO_2.

Cell feeding

Prewarmed media were used and cells were taken out of the incubator for as little time as possible, 10-15 ml for T-25's, 25-35 ml for T-75's and 50-60 ml for T-150's. For adherent cells, about every 2-3 days, old media was aspirated from culture flasks and replaced with fresh media.

Subculturing adherent cells

When adherent cells became 80% confluent, they were subcultured using trypsin/EDTA. Medium was removed from culture dish and cells were washed in PBS. The wash solution was removed. An appropriate amount of trypsin-EDTA solution was added to cover the bottom of the culture vessel. The culture was placed in the 37°C incubator for 2 minutes. Cells were monitored under microscope. Cells were beginning to detach when they appear rounded. As soon as cells were in suspension, culture medium containing serum was immediately added. Cells were washed once with serum containing medium and diluted as appropriate (generally 4-20 fold).

Harvesting

Cells were harvested (for RNA) when they were still growing in log phase, before they had reached a population density that suppresses growth. Medium was removed

from culture flask and cells were washed in PBS. An appropriate amount of trypsin-EDTA solution was added to cover the cell layer. The culture was incubated 37°C incubator for an appropriate time. Then cells were collected as normal.

1.2. Primary tumour samples

Primary colon tumour samples were part of an unselected, anonymised collection from patients at the Royal Infirmary Edinburgh. DNAs had been extracted from small portions of frozen samples, prior to the start of this study.

2.DNA manipulation and techniques

2.1.DNA preparation from culture cells

Suspension cultures were pelleted out of its serum-containing medium. Adherent cells were trypsinized and collected from the flask. Cells suspension was centrifuged 5 min at 500 × *g*, and supernatant was discarded. Cells were resuspended with 1 to 10 ml ice-cold PBS. Cells suspension was centrifuged 5 min at 500 × *g* and supernatant was discarded. This resuspension and centrifugation step were repeated once. Cells were resuspended in 1 volume digestion buffer (100ug/ml protease K, 10mmol Tris-Cl-ph8.0, 15mmolNaCl, 10mmol EDTA-pH8.0, 0.4%SDS). For <3 $\times 10^7$ cells, 0.3 ml digestion buffer was used. For larger numbers of cells 1 ml digestion buffer/10^8 cells was used. The sample was incubated with shaking at 50°C for 12 to 18 hr in tightly capped tubes. The sample was thoroughly extracted with an equal volume of phenol/chloroform/isoamyl alcohol. The sample was centrifuged 10 min at 1700 x *g*. The aqueous (top) layer was transfered to a new tube and 1/10 vol

of 3 M sodium acetate and 2 vol (of original amount of top layer) of 100% ethanol were added. The DNA formed a stringy precipitate. DNA was recovered by centrifugation at >8000 × *g* for 15 min, 4°C. The pellet was rinsed with 70% ethanol. Ethanol was decanted and the pellet was air-dried. DNA was resuspended at ~1 mg/ml in TE buffer until dissolved. The DNA solution was shaked gently at room temperature or at 65°C for several hours to facilitate solubilization. DNA was stored indefinitely at 4°C.

2.2.DNA quantification

Spectrophotometric method. For pure solutions of DNA, the simplest method of quantification was acquired by reading the absorbance at 260 nm where an OD of 1 in a 1 cm path length = 50 ug/ml for double-stranded DNA, 40 ug/ml for single-stranded DNA and RNA and 20-33 ug/ml for oligonucleotides. An absorbance ratio of 260 nm and 280 nm gave an estimate of the purity of the solution. Pure DNA and RNA solutions have OD260/OD280 values of 1.8 and 2.0, respectively.

2.3.Extraction of DNA from low-melting temperature agarose

Gel slice containing DNA was cut out, as smallest size as possible. The volume was estimated and doubled with TE (10 mM Tris-HCl, pH 8.0/1 mM EDTA). And then the solution was melt at 65 °C, 5-10 min. One volume Tris-buffered phenol was added and mixed by inversion at room temperature. The sample was spun 15 min at 10-12k rpm, 4°C and aqueous phase was transfered to new tube. Phenol extraction was repeated once. The sample was transfered to new tube. One ul carrier

(glycogen), 1/10 vol of 3 M sodium acetate and 2.5 volumes cold ethanol were added and mixed, left at -70 C 5-10 min, and spun as above, 10 min. Pllet was washed with 1 ml 70% ethanol, dried under vacuum, and resuspended in 10-20 ul water or TE.

2.4. Agarose gel electrophoresis of DNA

One gram agarose was dissolved in 100 ml of 1× TAE or TBE buffer (gives a 1% gel). 10 µl ethidium bromide solution per 100ml of buffer was added. The gel was cast with the comb in place. 6× gel loading buffer was added to sample and the samples were loaded into wells. The gel was run in 1× TAE or TBE (30-60 min. at 100-150V). The photo was taken on a long wave UV transilluminator.

2.5. Polymerase chain reaction (PCR)

Following reagents were added to a microfuge tube: 10 ul reaction buffer, 5 ul 20 uM forward primer, 5 ul 20 uM reverse primer, 1 ul template DNA, 16 ul 1.25 uM dNTP, 3 ul 50 mM MgCl2 or MgSO4 (volume variable) and water (to make up to 100 ul). The tube was placed in a thermocycler. Sample was heated to 95 °C, then 0.5 -1 ul of enzyme (*Taq, etc.*) was added. A few drops of mineral oil were added to cover the reaction solution. The PCR cycles were started according the following schemes: a) denaturing - 94 ° C, 30-90 sec. b) annealing - 55 °C (or Tm -5°), 0.5-2 min. c) extending - 72 °C, 1 min. (time depends on length of PCR product and enzyme used). Cycles were repeated 30 times. A final extension step of 5 min was added to fill in any uncompleted polymerisation. Then the reaction was cooled down to 4- 25 °C.

When high CG content DNA needed to be amplified, *Pfx* polymerase with appropriate amount of enhancer solution was used to provide high fidelity PCR product.

2.6. Restriction digest of DNA

For a typical digestion of 10 µg of DNA, water was added to a volume of 17 µl. Two µl of the appropriate 10× reaction buffer and 1 µl of enzyme were added. The sample was incubated at 37 °C for 2~16 hours. DNA was run on gel. For double digestion with 2 different restriction enzymes, both enzymes were added at the same time if they use the same buffer. If not, digested with one enzyme, precipitated the DNA. Then the DNA was resuspended and digested with the second enzyme in its appropriate buffer.

2.7.DNA sequencing

The following protocol used USB Sequenase DNA sequencing kit

Pre-treatment of PCR product

Ten ul PCR product, 2ul Exo, 2ul SAP were added in 0.5ml tube. The sample was heated at 37°C 15min, 80°C 15min and then put on ice.

Preparation of reaction mix

The following reagents were added: 2ul reaction buffer (5×), DNA (7ul), 1ul primer (20uM), 8ul double distilled water and 2ul sequenase

Preparation of termination mix

Four tubes were prepared for each labelling reaction of ddGTP,ddATP, ddTTP and ddCTP. The following reagents were added to each tube: 2ul master mix and 0.5ul α-33p(ddNTP).

Cycle sequencing

4.5ul reaction mix was added to each termination mix. Normal PCR cycles were applied, PCR condition varied with different template. Stop solution (4ul) was added, mixed and kept on ice.

Preparation of sequencing gel

Plates were cleaned with ethanol and one side of the shorter plate was siliconised. Spacers were put between plates. Plates' sides were clamped with clips to form mould. The combs were checked to make sure that they fit between plates. Plates were laid on horizontal surface. Gel mixture was prepared (6% sequence gel, 555ml) as following: 240g Urea, 75ml 40% acrylamide solution, 50ml TBE (10×), and 190ml H2O. To start polymerisation reaction, following reagents were added: 60ml 6% sequence gel solution, 400ul 10% APS, and TEMED 30ul. Gel mixture was mixed and cast quickly. Combs were put in place with the straight side down. Gel was left to polymerise. When gel had set, tape and comb were removed. The wells were rinsed. Plates were placed in electrophoresis apparatus. Combs were placed with the serrated edge down into the well, with the teeth just slightly penetrating the gel. Top and bottom tanks were filled with 1× TBE. Wells were rinsed to remove any bubbles and particles. A pre-run (80W, 30 min) was performed before samples were loaded to preheat the gel.

Sample loading

Samples were heated to 95 °C for >3 min, then put on ice immediately. Wells were rinsed to remove any bubbles. 4ul of sample was loaded in the sequence GATC or ACGT.

Electrophoresis of DNA

Gel was run at constant power (80W) for ~ 2 - 4 hours. If second loading of sample was needed, the second sample was loaded after the xylene cyanol light blue dye had reached the bottom.

Gel drying

After the electrophoresis had finished, spacers and plates were removed from apparatus. The plates were taken apart carefully to make sure that the gel only sticks to the longer plate. A large piece of Whatman paper was placed on gel and gel was transferred onto the paper from plate. Cling film was placed on gel. Gel was dried in gel dryer (80 °C under vacuum).

2.8. PCR cloning

2.8.1. Set up ligation (pGEM-T Easy Vector System)

The pGEM®-T Easy Vector tube was briefly centrifuged to collect contents at the bottom of the tube. The 2× Rapid Ligation Buffer was vigorously vortexed before each use. Ligation reaction was set up as described below. 2×Rapid Ligation Buffer 5µl, PCR product 4µl, pGEM®-T Easy Vector 0.5µl and T4 DNA Ligase 0.5µl were added to a tube. The reaction was incubated in 16°C for overnight.

2.8.2.Transformation (One Shot. TOP10 Competent Cells)

The vial containing the ligation reaction was centrifuged briefly and placed on ice. One 50ul vial of One Shot® cells was thawed on ice for each ligation/transformation. 1 to 5ul of each ligation reaction was pipetted directly into the vial of competent cells and mixed by tapping gently. The vial was incubated on ice for 30 minutes. Then the vial was incubated for exactly 30 seconds in the 42°C water bath. The vial was removed from the 42°C bath and placed on ice. 250ul of pre-warmed S.O.C medium was added to the vial. The vial was placed in a microcentrifuge rack on its side and secured with tape to avoid loss of the vial. The vial was shaked at 37°C for exactly 1 hour at 225 rpm in a shaking incubator. 20ul to 200ul from transformation vial was spreaded on separate, labeled LB agar plates with 50ug/ml ampicillin and X-gal. The remaining transformation mix was stored at +4°C and plated out the next day, if desired. The plate was inverted and incubated at 37°C overnight. White colonies were selected with tooth sticks and analyzed by lysates PCR of boiled lysates.

2.9.Plasmid isolation (Qiagen Plasmid Preparation Kits)

All steps were carried out at room temperature.

Things done before starting:

The provided RNase A solution was added to Buffer P1 before use. One vial of RNase A (spin down briefly before use) was used for one bottle of Buffer P1, to give a final concentration of 100μg/ml. Buffer P2 was checked for SDS precipitation due to low storage temperatures. If necessary, the SDS was dissolved by warming to 37°C. Buffer P3 was pre-chilled to 4°C.

2.9.1. Plasmid mini-preparation

To purify up to 20μg of high-copy plasmid (pGEM-T) DNA from1-5ml overnight cultures of *E. coli* in LB medium. A single colony was picked from a freshly streaked selective plate and inoculated a culture of 5ml LB medium containing the appropriate selective antibiotic. The culture was incubated for 16~18h at 37°C with vigorous shaking (~300rpm). The bacterial cells were harvested by centrifugation at 6000 × *g* for 15min at 4°C. Pelleted bacterial cells were resuspended in 250μl Buffer P1 (containing RNase A) and transfered to a microcentrifuge tube. 250μl Buffer P2 was added and the tube was gently inverted 4–6 times to mix. 350μl Buffer N3 was added and the tube was inverted immediately but gently 4–6 times. The tube was centrifuged for 10min at maximum speed in a tabletop microcentrifuge. A compact white pellet would form. The supernatant was applied to the QIAprep column by decanting or pipetting. The column was centrifuged for 30–60s. The flow-through was discarded. QIAprep spin column was washed by adding 0.75ml Buffer PE and centrifuging for 30–60s. The flow-through was discarded. The column was centrifuged for an additional 1 min to remove residual wash buffer. The QIAprep column was placed in a clean 1.5ml microcentrifuge tube. To elute DNA, 50μl Buffer EB (10mM Tris·Cl, pH8.5) or water was added to the center of each QIAprep column. The column was let stand for 1min, and then centrifuged for 1min.

2.9.2. Plasmic midi/maxi preparation

This procedure was to prepare up to 100μg /2.5mg of plasmid from 25ml/500ml overnight cultures of *E. coli* in LB medium. A single colony was picked from a freshly streaked selective plate and inoculated a starter culture of 2–5ml LB medium

containing the appropriate selective antibiotic. The culture was incubated for ~8 h at 37°C with vigorous shaking (~300 rpm). A tube or flask with a volume of at least 4 times the volume of the culture was used. The starter culture was diluted 1/500 to 1/1000 into selective LB medium. For high-copy plasmids (pGEM-T), a culture of 25ml/500ml medium was inoculated and growed at 37°C for 12–16h with vigorous shaking (~300rpm). A flask or vessel with a volume of at least 4 times the volume of the culture was used. The culture should reach a cell density of approximately 3–4×10^9 cells per ml, which typically corresponds to a pellet wet weight of approximately 3g/liter medium. The bacterial cells were harvested by centrifugation at 6000 ×*g* for 15min at 4°C. The bacterial pellet was resuspended in 4ml/50ml of Buffer P1. 4ml/50ml of Buffer P2 was added, mixd gently but thoroughly by inverting 4–6 times, and incubated at room temperature for 5min. 4ml /50ml of chilled Buffer P3 was added, mixed immediately but gently by inverting 4–6 times, and incubated on ice for 15min. The tube was centrifuged at 20,000 ×*g* for 30 min at 4°C. Supernatant containing plasmid DNA was removed promptly. The supernatant was centrifuged again at 20,000×*g* for 15 min at 4°C. Supernatant containing plasmid DNA was removed promptly. A QIAGEN-tip 100 /2500column was equilibrated by applying 4ml QBT, and the column was allowed to empty by gravity flow. The supernatant was applied to the QIAGEN-tip 100/2500 column and allowed to enter the resin by gravity flow. The QIAGEN-tip 100/2500 was washed with 2×10ml/2×100ml Buffer QC.

Buffer QC was allowed to move through the QIAGEN-tip 100/2500 by gravity flow. DNA was eluted with 5ml/35ml Buffer QF. DNA was precipitated by adding

3.5ml/24.4ml (0.7 volumes) room-temperature isopropanol to the eluted DNA. The DNA was mixed and centrifuged immediately at 15,000×*g* for 30 min at 4°C. The supernatant was carefully decanted. DNA pellet was washed with 2ml/7ml of room temperature 70% ethanol, and centrifuged at .15,000×*g* for 10 min. The supernatant was carefully decanted without disturbing the pellet. The pellet was air-dried for 5–10min, and redissolved the DNA in a suitable volume of buffer (e.g., TE buffer, pH8.0, or 10mM Tris·Cl, pH8.5).

2.10.DNA methylation analysis

2.10.1. Methylation sensitive restriction enzyme analysis

By using the isoschizomer pair (e.g. *HpaII/MspI* and *SmaI/XmaI*) which display differential sensitivity to cytosine methylation, a simple PCR was used to assay the DNA methylation status of restricted DNAs. The selection of enzymes varied according to different CpG sequence with different recognised sites, to ensure enough sites are assayed. The presence of the PCR product from *Hpa II* or *SmaI* and other methylation sensitive enzymes digest indicates that the DNA is protected by methylation from being cut, while no PCR product results if the DNA is not methylated. No PCR product will be obtained from *MspI* and *XmaI* digested DNA because they cut irrespective of its methylation status.

Restriction:

1ug genomic DNA was used for restriction, 20units enzyme was added in 50ul reaction volume. The digestion was incubated overnight at 37°C (or appropriate temp.) in water bath.

Precipitation:

Normal ethanol precipitation was used to precipitate DNA. The DNA pellet was washed with 70% ethanol once.

PCR detection

The samples cut by methylation non-sensitive restriction enzyme *MspI* but not cut by methylation sensitive restriction enzyme *HpaII* are methylated.

2.10.2. Bisulfite modification of DNA

20mM hydroquinone (Sigma; #H 9003) and 4.8M sodium bisulfite (Sigma; #S 8890) were prepared immediately before use. Above solutions were dissolved by gently inverting with a minimum amount of mixing and kept cold and in the dark as much as possible. 2.5μg DNA was denatured by adding freshly prepared NaOH (3M) to a final concentration of 0.3M, and then incubated at 42°C for 30min. 500μl 4.8M sodium bisulfite and 30μl 20mM hydroquinone were added to denatured DNA tube. The tube was gently mixed and overlaid with mineral oil. The tube was wrapped with aluminium foil to shield from the light. The tube was incubated at 55°C for 16−18h. DNA was precipitated by 0.6 volume isopropanol, 0.3M NaOAc and 2.5ul glycogen. The solution was mixed and centrifuged at max speed for 30mins at room temperature. The pellet was washed carefully with 70% ethanol and air-dried for about 1 hour. TE was added to a final volume of 100μl. The sample was desulfited with freshly prepared NaOH (as above) and incubated at 37°C for 15min. The DNA was precipitated with three volumes of ethanol, centrifuged for 10min (14,000rev/min) at room temperature, washed twice with 70% ethanol and dried under a vacuum. Resuspended in 100μl TE, and stored at −20°C wrapped in foil.

Treated DNA was used within two weeks as degradation still occurs in the cleaned and frozen sample.

Care was taken to make sure that the DNA was completely denatured prior to and in the presence of the bisulfite solution or the modification would not be complete. To ensure complete denaturation, no more than 5μg of starting material was used, the DNA was digested with restriction enzymes and the initial alkaline denaturation was at 42°C for 30min.

2.10.3. Primer design for bisulfite converted DNA sequence

When primers were designed for bisulfite converted DNA sequence, particular attention was paid to remember that after bisulfite modification, all the unmethylated C except methylated C had changed to U (recognised by Taq as T). Briefly, sequence was copied and pasted into a text editor software. All CG was first converted to XG. Then all C was converted to T. And then all X was converted to C. A restriction map of this converted sequence (methylated map) was made. All remaining C was converted to T. A restriction map of this converted sequence (unmethylated map) was made. Restriction enzyme sites that are unique to the methylated map (not in the unconverted or unmethylated map) were selected. These were the best to use. If none was available, restriction sites that are present in the methylated map but absent in the unmethylated map were selected. Primers were designed for PCR of a region that containes usable enzymes and that is close to the transcription start. The methylated/converted sequence was used to design primers, but having C in the sense primer or G in the anti-sense primer was avoided. If no suitable primers was found, up to one C was included in each primer, but they are in the 5' end of the primer and

they were synthesized with a mix of C and T (sense strand) or a mix of G and A (antisense strand) instead of simply C or G. PCR was amplified with the calculated annealing temperature and optimised PCR condition.

2.10.4. Experiment procedure for sequencing bisulfite converted DNA

Bisulfite-PCR products were subcloned into pGEM^{T}and multiple clones (usually 10) were sequenced to get more methylation detail than other methods. Any C in the sequence reflects methylation. Nested primers were used to improve PCR proficiency.

Outline steps:

1) bisulfite treatment described as before
2) nested PCR
3) cloning described as before
4) cycle sequencing described as before

2.11. Single-strand conformation polymorphism analysis (SSCP)

PCR reaction was set up according to standard recipe (see table2). Optimised PCR conditions was applied to each PCR reaction

MDE Buffer

590ml H_2O, 60ml 10×TBE, and 100ml Glycerol were added to make 750ml solution.

SSCP Gel

To make 50ml gel, 12.5ml MDE gel, 37.5ml MDE buffer, 400µl 10% APS, and 30µl TEMED were added. Gel was cast with appropriate sharkstooth comb. Gel would polymerize in about 1-2 hour.

Loading Buffer

The recipe for loading buffer: 95% formamide, 10mM NaOH, 0.025% Bromophenol Blue, and 0.025% Xylene Cyanol. Gel was run in 0.6X TEB buffer. Samples were heat-denatured at 94°C for 5 minutes and then placed on ice for 3-5 minutes. 4µl of each sample was loaded.

Electrophoresis conditions

a). Fragment Size: 150-200bp

Run under 6 Watts for 10-12 hours at room temperature

b). Fragment Size: > 200bp

Run under 8 Watts for 10-12 hours at room temperature

Exposure

Gel was dried and exposed to autoradiographic film at room temperature for 16-18 hours.

Table 2. "Hot" PCR recipe

10X PCR Buffer	1.0µl
$MgCl_2$ (50mM)	0.3µl
dNTPS (1.25µM)mix*	1.6µl
Primer-F (20µM)	0.5µl
Primer-R (20µM)	0.5µl
Taq (5U)	0.1µl
^{33}P dATP (10µci/µl)	0.1µl
ddH_2O	5.4µl
SUBTOTAL	9.5µl
Template DNA	0.5µl

3. RNA manipulation and techniques

RNA is more susceptible to degradation than DNA, due to the ability of the 2′ hydroxyl groups adjacent to the phosphodiester linkages in RNA to act as intramolecular nucleophiles in both base- and enzyme-catalyzed hydrolysis. Whereas deoxyribonucleases (DNases) require metal ions for activity and can therefore be inactivated with chelating agents (e.g. EDTA), many ribonucleases (RNases) bypass the need for metal ions by taking advantage of the 2′ hydroxyl group as a reactive species. Therefore special care was taken to protect RNA from Rnases: Gloves were always worn when working with RNA. Sterile, disposable plasticware was used. DEPC was used to treat solutions if necessary.

3.1.Total RNA-isolation (TRIZOL method)

Isolation of RNA from cultured cells was performed with TRIZOL reagent (Gibco-BRL). Culture medium was changed 2 hours before harvesting cells. Cells were collected by trypsinizing and washing. Cells were spun down at 1200rpm in a microfuge for 3 minute. Excess medium was removed and 1 ml Trizol reagent was added. The tube was vortexed, inverted and left at room temperature for 10minutes (stringy-like material should be seen). The tube was spun in a microfuge at 13000rpm for 10minutes at 4° C. Supernatant was transferred into fresh RNAse free eppendorf tube and 200µl of chloroform was added. The tube was vortexed for 15 seconds and left at room temperature for 3 minutes. The tube was spun in a microfuge at 13000rpm for 15 minutes at 4° C. The top layer (clear) was carefully transferred into a new RNAse free eppendorf tube and 500µl of isopropanol was added. The tube was inverted to mix and left for 10 minutes at room temperature. The tube was spun in a microfuge at 13000 for 10 minutes at 4°C. (The RNA formed a white pellet). Pellet was washed with 100µl of 75% ethanol (made by diluting into DEPC-treated water). The tube was spun in a microfuge at 7500×g for 5 minutes at 4°C. Supernatant was removed. Pellet was air-dried for 10 minutes. Pellet was dissolved in 25µl of DEPC-treated water. The pellet was heated for 10 minutes at 60°C to help dissolve RNA.

3.2. mRNA-isolation (Qiagen Oligotex method)

Isolation of Poly A+ mRNA (from 5×10^5 -5×10^7 culture cells) was performed with Oligotex Direct mRNA kit.

Preparation before starting

Oligotex suspension was heated to 37°C in a water bath or heating block, mixed by vortexing, and then placed at room temperature. A water bath or heating block was heated to 70°C, and buffer OEB was heated to 70°C. 30μl β-Mercaptoethanol was added to 1ml buffer OL1 in a fume hood. buffer OW1 and buffer OL1 were redissolved by warming at 37°C, and then placed at room temperature. Cell pellets was stored at –70°C for later use or used directly in the procedure. Unless otherwise indicated, all protocol steps, including centrifugation, were performed at 20 to 30°C. Unless otherwise indicated, all centrifugation steps were performed in a microcentrifuge at maximum speed 13000rpm.

Procedure

Sample preparation for cells grown in suspension. Used less than 5×10^7. The desired numbers of cells were spun down for 5min at 1000rpm in an RNase-free polypropylene centrifuge tube. All supernatant was carefully removed by aspiration. Cells were disrupted by addition of room temperature 0.6ml buffer OL1. For pelleted cells, the cell pellet was loosened by flicking the tube. 0.6ml buffer OL1 was added. The tube was vortexed for 5–10s or pipetted up and down to mixed thoroughly, and proceeded at once with next step. The sample was homogenized. The lysate was pipetted directly onto a QIAshredder spin column placed in 2ml collection tube, and centrifuged for 2min at maximum speed. Action was taken to make sure that homogenization is complete. 1.2ml buffer ODB was added to the lysate, and mixed thoroughly by pipetting. The tube was centrifuged in a microcentrifuge for 3min at maximum speed. The supernatant was transferred to a new RNase-free tube. 70ul oligotex suspension was added to the sample, mixed thoroughly by pipetting or

vortexing, and placed at room temperature for 10min. The Oligotex:mRNA complex was pelletted by centrifuging in a microcentrifuge for 5min at maximum speed (13000rpm). The supernatant was carefully removed by pipetting. The Oligotex:mRNA pellet was resuspended thoroughly in 100 µl buffer OL1 by vortexing or pipetting. 400µl buffer ODB was added, incubated at 70°C for 3 min and then placed at room temperature for 10 min. The Oligotex:mRNA complex was pelletted by centrifugation in a microcentrifuge for 5 min at 13000rpm, and the supernatant was carefully removed by pipetting. The pellet was resuspended in 350µl buffer OW1 by vortexing or pipetting. The sample was pipetted onto a small spin column in a 1.5ml microifuge tube and centrifuged for 1 min at 13000rpm. The flow-through was discarded. The spin column was transfered to a new RNase-free 1.5 ml microfuge tube. 350µl buffer OW2 was pipetted onto the column. The column was centrifuged for 1 min at 13000rpm, and the flow-through was discarded. Last step was repeated once, using the same microfuge tube. The spin column was transferred to a new RNase-free 1.5ml microfuge tube. 100µl hot (70°C) buffer OEB was pipetted onto the column. The resin was pipetted up and down 3 or 4 times to resuspend, and centrifuged for 1 min at 13000rpm. To ensure maximal yield, another 100µl hot (70°C) buffer OEB was pipetted onto the column. The resin was pipetted up and down 3 or 4 times to resuspend and then centrifuged for 1 min at 13000rpm.

3.3. RT-PCR

First-strand cDNA synthesis using M-MLV RT

A 50μl reaction volume were used for 25μg total RNA or 2.5-1.25μg mRNA. The following components were added to a nuclease-free microcentrifuge tube: 1.5μl random hexamer (10mg/ml), or 20μM gene-specific primer, 25μg total RNA or 1.25μg mRNA, 2.5μl 10mM dNTP Mix (10mM each dATP, dGTP, dCTP and dTTP at neutral pH) and sterile, distilled water to the volume of 33.5μl. The mixture was heated to 65°C for 5min and quick chilled on ice. The contents of the tube was collected by brief centrifugation and added: 10μl 5×First-Strand Buffer and 5μl 0.1M DTT. Contents of the tube was mixed gently and incubated at 37°C for 2min. 1.5μl (300units) of M-MLV RT was added, and mixed by pipetting gently up and down. If random primers were used, tube was incubated at 25°C for 10min. The tube was incubated 1hour at 37°C. The reaction was inactivated by heating at 95°C for 15min. The cDNA was stored at 4°C as a template for amplification in PCR.

PCR Reaction

10-20% of the first-strand reaction (5-10μl of the reaction from the previous page) was used for PCR. PCR was performed under normal condition. If necessary, the PCR condition was optimised.

3.4. Northern blot analysis

The procedure was divided into three sections:

a). Electrophoresis of an RNA sample under denaturing conditions in an agarose-formaldehyde gel.

b). Transfer of the RNA from the gel to a nitrocellulose membrane by upward capillary transfer.

c). Hybridisation analysis of the RNA sequences of interest using a labeled cDNA probe.

3.4.1.Agarose/formaldehyde gel electrophoresis

Cast gel(150ml)

Following reagents were added: 1.5g agarose, 15ml 10×MOPS and 109.5ml ddH2O. The mixture was boiled to dissolve agarose, and cooled to 60°C. Then 25.5ml 38% formaldehyde was added. The gel was mixed and cast.

Prepare running buffer (1000ml)

100ml 10×MOPS, 52.6ml 38% formaldehyde, and 847.4ml H2O were added and mixed.

Prepare RNA sample buffer (1000μl)

100μl 10×MOPS, 500μl formamide, 180μl 38% formaldehyde, 216μl ddH2O, and 4μl ethidium bromide were added and mixed. The solution was stored at 4°C. 12μl RNA sample buffer was used for each RNA sample.

Prepare dye (1ml)

500μl glycerol, 2.5mg xylene cyanol, 2.5mg bromophenol blue,

2 μl 0.5M EDTA and ddH2O to the volume of 1ml were added and mixed. Loading buffer was stored at 4°C. 3μl loading buffer was used for each RNA sample.

Run gel

The gel was run in 1× running buffer at 150 volt for 3 hours, at 4°C.

2.4.2.Transfer of RNA from gel to membrane

Prepare gel for transfer

The gel was placed in a clean dish, rinsed and soaked with ddH_2O for 2 15 min.

Transfer RNA from gel to membrane

A glass dish was filled with enough 10×SSC. 3 pieces of Whatman 1MM paper was cut, placed on the glass plate and wetted with 10×SSC. The gel was placed on the filter paper and air bubble was squeezed out by rolling a glass pipe. A used film was cut into suitable shape and placed over the edges of the gel to prevent the solution bypassing the gel. A piece of nitrocellulose membrane was cut just large enough to cover the gel and wetted in water. The wetted membrane was placed on the surface of the gel. Action was taken to avoid getting air bubbles under the membrane. The surface of the membrane was flooded with 10×SSC. 5 sheets of whatman 3MM paper were cut to the same size as membrane and placed on top of the membrane. Paper towels were put on top of the whatman 3MM paper to a height of ~6cm and a weight was added to hold everything in place. The transferring set was left overnight.

Prepare membrane for hybridisation

paper towels and filter papers were removed and the membrane and flattened gel were recovered. A mark was made by pencil to label the position of the wells on the membrane and ensure that the up-down and back-front orientation were

recognizable. The membrane was rinsed with 5×SSC, then placed on a sheet of Whatman 3MM paper and allowed to air dry. The membrane was placed RNA-side-up on a UV transilluminator (254nm wavelength), and irradiated for appropriate length of time.

3.4.3.Hybridisation analysis

Prepare cDNA probe

The probe was labeled with Radprimer DNA labelling system (Invitrogen). 25ng DNA dissolved in 5-20μl of sterile distilled water or TE was denatured in a microcentrifuge tube by heating for 5min in a boiling water bath; then immediately cooled on ice. The following additions were performed on ice: 1μl 500μM dATP, 1μl 500μM dGTP, 1μl 500μM dTTP, 20μl 2.5× Random Primers Solution 5μl (approximately 50μCi) [α-32P]dCTP (3000Ci/mmol, 10mCi/ml) and dd Water to a total volume of 49μl. The tube was mixed briefly. Then 1μl Klenow fragment was added. The tube was mixed gently but thoroughly and centrifuged 15-30s. The tube was incubated at 37°C for 10-60min. And then 5μl Stop buffer was added to finish the reaction.

Hybridisation

Hybridisation was performed with ExpressHyb™ Hybridisation Solution (Clontech). ExpressHyb Solution was heated to 68°C, and stirred well to completely dissolve any precipitate. Membrane was prehybridized in a minimum of 5 ml of ExpressHyb Solution, with continuous rotation for 30 min at 68°C in a hybridisation oven. The marked side of the membrane was flush against the side of the bottle. Bubbles

between the membrane and the bottle were ruled out as they could prevent hybridisation to those areas. Radioactively labeled probe was denatured at 95–100°C for 2–5 min, then chilled quickly on ice. Radiolabeled probe was added to 5 ml of fresh ExpressHyb, and mixed thoroughly. The ExpressHyb Solution was replaced with the fresh solution containing the radiolabeled probe. All air bubbles were removed from the container, and ExpressHyb Solution was made evenly distributed over the membrane. The hybridisation was incubated with continuous shaking for 1 hr at 68°C. The membrane was rinsed in wash solution 1 (2× SSC, 0.05% SDS) several times at room temperature and washed for 30–40 min with continuous agitation; the wash solution was replaced several times. The membrane was washed two times in wash solution 2 (0.1× SSC, 0.1% SDS) with continuous shaking for 40 min at 50°C. The membrane was removed with forceps and shook off excess wash solution. Even partially dry was prevented as allowing the membrane to dry could cause high background and would make subsequent probe removal difficult. The membrane was immediately covered with plastic wrap, mounted on Whatman 3 MM chromatography paper and wrapped again with plastic wrap.

Autoradiography

Blot was exposed at -80 °C using Kodak XAR film and x-ray intensifying screens.

Strip probe from the membrane

100ml sterile H2O containing 0.5% SDS was heated to 90–100°C. Plastic wrap was removed from membrane and immediately placed in the heated solution. The exposure to air was reduced as less as possible.

The membrane was incubated for 10 min, shaking frequently. The hot solution was left to cool for 10 min. The membrane was removed and air-dried until it was dry

enough to be slipped into a plastic bag. The membrane was stored at –20°C until next hybridisation.

4. Materials

4.1. General laboratory reagents and suppliers

Amersham Life Science UK Ltd.: Hybond-N nylon membrane

BDH Ltd.: dimethyldichlorosilane, hydrochloric, and xylene cyanol

Boehringer Mannheim UK Ltd.: proteinase K.

Difco Laboratories: agar, bacto-tryptone

Fisher Scientific UK Ltd.: 3M blotting paper, EDTA, isopropanol, sodium acetate, sodium chloride, sodium hydroxide

Gibco BRL Life Technologies Ltd.: 100bp DNA marker ladder, agarose,

Formamide, phenol, Tris, low melting point agarose, urea.

Millipore (UK) Ltd.: disposable sterile filters.

New England Biolabs (UK) Ltd.: restrict enzyme

Promega (UK) Ltd.: oligo dT

Qiagen Ltd.: plasmid preparation kit series

Scotlab Ltd.: acrylamide/bis-acrylamide solutions

Sigma-Aldrich Company Ltd.: bromophenol blue, DTT, ethidium bromide, formaldehyde, glycerol, IPTG, mineral oil, MOPS, parafilm M, X-gal and Fully methylated DNA (methylation status was confirmed by bisulfite sequencing)

Strategene. Ltd.: Foetal and adult colon RNA, Foetal and adult lung RNA

4.2. Bacterial strains used in this study

Invitrogen Life technologies Ltd.: One Shot Top 10 competent cells

4.3. Plasmid

Promega Ltd.: PGEM-T

4.4. DNA/RNA modifying enzymes

Boehringer Mannheim UK Ltd.: Klenow polymerase, restriction endonucleases

Gibco BRL Life Technologies Ltd.: restriction endonucleases,

T4 DNA kinase, *Taq* DNA polymerase

New England Biolabs (UK) Ltd.: restriction endonucleases, T4 DNA ligase

4.5. Radioactive reagents

Amersham International plc. : Redivue [α-^{32}P]-dCTP (~3000Ci/mmole, 10mCi/ml),

Redivue [α-^{33}P]-ddNTPs (4×25µl, 450µCi/ml), dATP [α-^{33}P](3000 Ci/mmol).

4.6. Mammalian cell culture reagents

Becton Dickinson Labware: plasticware.

Difco Laboratories: trypsin

Gibco BRL Life Technologies Ltd.: RPMI1640 medium

Sigma-Aldrich Company Ltd.: β-mercaptoethanol.

4.7. Cancer cell lines

Cancer cell lines (see table 3) were obtained from ECACC (the European Collection of Cell Cultures/ATCC).

Table 3 Cell lines used in this study

tissue	description	name
colon	microsatellite stable	HT29, SW480, COLO320
	microsatellite unstable	HCT116, DLD1/HCT15, LOVO, LS180
lung	small cell (SCLC)	NCI-H69, -H524, -H740, -H1672, -H1092, -H1184, -H1838, COR-L24, -L47, -L51, -L88, -L279, -L311
	non-small cell (NSCLC)	NCI-H358, -H835, -H920, -H1648, -H2122, COR-L23, -L105

4.8. General solutions and buffers

PBS (1× in 500 mls): added 4.09 g NaCl, 0.093g KCl,

2.028 g Na_2HPO_4 (7 H_2O) and 0.109g KH_2PO_4 pH 7.2

1×TE : 10mM Tris-HCl, 1mM EDTA, pH 8.0

1 M Tris-Cl (500 ml): added Tris base 60.5 g, adjusted pH with concentrated HCl and then add double distilled water to 400 ml.

20× SSC (2L pH 7.0): added NaCl 350.6 g and Na3Citrate 176.6g, then Mix and adjusted pH to 7.0 with HCl.

10× TBE Buffer (4L, pH 8): added Tris-Base 432 g, Boric acid 220g,and Na2EDTA 37.2 g, Then mix and store at room temperature.

50× TAE Buffer (50×, pH 8): added Tris Base 242 g, acetic acid 57 ml and Na2EDTA 37g. Then mix and store at room temperature.

4.9. Primers used in this study (see table 4)

Table 4. PCR primers used in this study (1)

name	sequence	target DNA	product size	specific condition
MBD3/1 MBD3/2a	5'-GAAGAAGTTCCGCAGCAA 5'-GGTCGCTCTTGACCTTGT	MBD3 cDNA	260 bp	95 3'(95 30", 52 30", 72 30")x35,72 3'
MBD3/3 MBD3/30	5'-ACATGCTGGGGGACGTGGA 5'-GCTGCACAGTGGGTGATGTGA	MBD3 cDNA	587bp	95 3'(95 30", 65 30", 72 1')x2 (95 30", 62 30", 72 1')x2 (95 30", 60 30", 72 1')x2 (95 30", 58 30", 72 1')x2 (95 30", 55 30", 72 1')x35,72 3' Taq polymerase+1M betaine
MBD3/17 MBD3/18	5'-ACTGGCAGCTCGCAAGGCACA 5'-CGCTGGGAGGAGCCCGTTGAG	MBD3 gDNA, CpG island	531bp	95 3'(95 30", 55 30", 68 2')x35,68 3' Pfx polymerase + 5X Enhancer
MBD3/17 MBD3/15	5'-ACTGGCAGCTCGCAAGGCACA 5'-GCACGCACGCACGACGCA	exon1	564bp	95 3', (95 30", 55 30", 68 2') x 35, 68 3', Pfx polymerase + 5X Enhancer
MBD3/5 MBD3/6	5'-TGGGTTTGGGGTCTTGGGGT 5'-GTCACCTGCGTGACGCCA	exon2	240bp	94 3', (94 30", 55 30") x 35, 72 3' Taq polymerase + 10% DMSO
MBD3/7 MBD3/8	5'-CAGGCCGGACTGCATATC 5'-TTGGGGTCGCTGTGCGTT	exon3	195bp	94 3', (94 30", 58 30") x 35, 72 3' Taq polymerase
MBD3/9 MBD3/10	5'-GGGCCACTCTTGAGGTTCACA 5'-TCCGCCTCCCCTCAGGGA	exon4	156bp	94 3', (94 30", 58 30") x 35, 72 3' Taq polymerase
MBD3/11 MBD3/12	5'-CTTGCGGCTGTTTGTCCA 5'-TGGAGCAGCAGGGGACCA	exon5	239bp	94 3', (94 30", 58 30") x 35, 72 3' Taq polymerase + 5% DMSO
MBD3/13 MBD3/4	5'-TGGTGTAACGCAAGGTCCA 5'-CACTGCCAGGACCCGACT	exon6	280bp	94 3', (94 30", 58 30") x 35, 72 3' Taq polymerase + 5% DMSO

Table 4. PCR primers used in this study (2)

name	sequence	target DNA	product size	PCR condition
MBD2/15 MBD2/16	5′-GGATTCCAAGGGCTCGGTTACGG 5′- TGCCCAGGCCCGCTCTTGACC	MBD2 CpG island	670bp	95 3', (95 30", 55 30", 68 2') x 35, 68 3', Pfx polymerase + 6X Enhancer
2b NcoMet 2b TerBam	5′-ACGTCCATGGATTGCCCGGC 5′-ACTGGGATCCTTAGGCTTCATCT	MBD2 cDNA	750bp	94 3', (94 30", 58 30", 72 1') x 35, 72 3', Pfx polymerase
PTC1 PTC2	5′-GCGAGCCAATCGCGTCCGCA 5′-GGAGGGCAAGCGCACAGC	PTC CpG island	519bp	95 3', (95 30", 55 30", 68 1') x 35, 68 3', Pfx polymerase + 4X Enhancer
PTC3 PTC4	5′-GCTCTGCGCTTGCCCTCC 5′-CGCTGCTGCTGCTCACACG	PTC CpG island	357bp	95 3', (95 1', 66 1') x 35, 68 3', Pfx polymerase + 4X Enhancer
PTC9 PTC6	5′-CAGAAGCGTCCTCGCAAG 5′-GAAATCTGCTCCAGAGCGAAGGC	PTC CpG island	316bp	95 3', (95 1',55 30", 68 1') x 35, 68 3', Pfx polymerase + 3X Enhancer
PTC9 PTC8	5′- CAGAAGCGTCCTCGCAAG 5′-TCTCGAGGTTCGCTGCTT	PTC cDNA	486bp	95 3', (95 1',55 30", 68 1') x 35, 68 3', Pfx polymerase + 2X Enhancer
SMO1 SMO2	5′-CCGCCGAGGTCGTGCGTGTG 5′-GTCGCGTTCCCGCTCGAGGC	SMO cDNA	346 bp	95 3'(95 30", 62 30", 72 45")x35,72 3'
SMO13 SMO43	5′-ACGAGGACGTGGAGGGCTG 5′-CGCACGGTATCGGTAGTTCT	SMO cDNA	590bp	95 3'(95 30", 62 30", 72 45")x35,72 3' Taq polymerase+1M betaine
SMO21 SMO52	5′-TGTGGTCCTCACCTATGCCT 5′-ATGAGCACCAGGCCGATT	SMO cDNA	365bp	95 3'(95 30", 66 30", 72 1')x3 (95 30", 62 30", 72 1')x3 (95 30", 58 30", 72 1')x35,72 3' Taq polymerase

Table 4. PCR primers used in this study (3)

name	sequence	target DNA	product size	PCR condition
SMObis1 SMObis4	5′- GAYGATTTTAGATTAAGTAAGGT GTT 5′-TCCCACCATTAAAACCACCT	bisulfite treated DNA	469bp	94 3'(94 30", 60 30", 72 1')x2 (94 30", 58 30", 72 1')x2 (94 30", 56 30", 72 1')x3 (94 30", 53 30", 72 1')x4 (94 30", 50 30", 72 1')x25,72 3' Taq polymerase
SMObis3 SMObis4	5′-TTYGTGTATTTTAGAGAGTTTAG 5′-TCCCACCATTAAAACCACCT	1st round PCR product	212bp	94 3'(94 30", 45 30", 72 1")x35,72 3' Taq polymerase
Gli3/5 Gli3/6	5'-AGGGGACATCCATCCTCG 5'-GTCCCTTCTGTTGAGCAT	Gli cDNA	615bp	95 3'(95 30", 58 30", 72 1')x35,72 3' Taq polymerase

Chapter 3

Genetic and Epigenetic Analysis of

MBD3

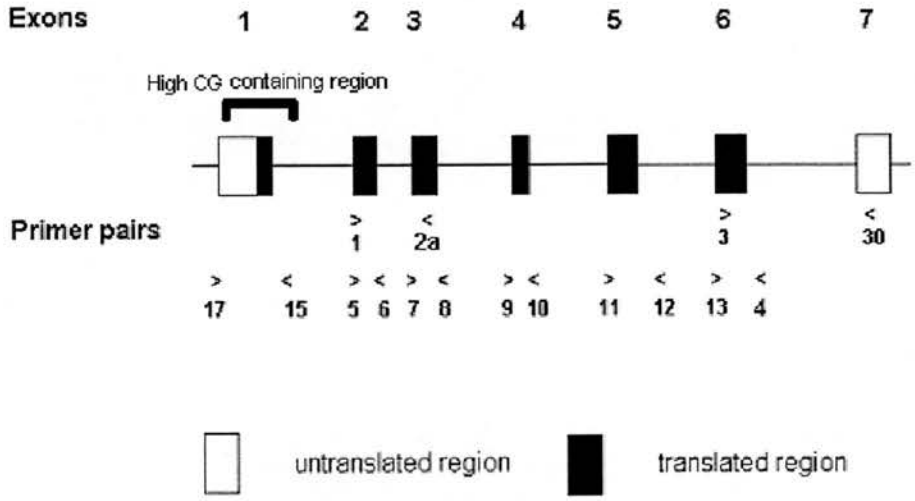
Contents

Page

1. Introduction

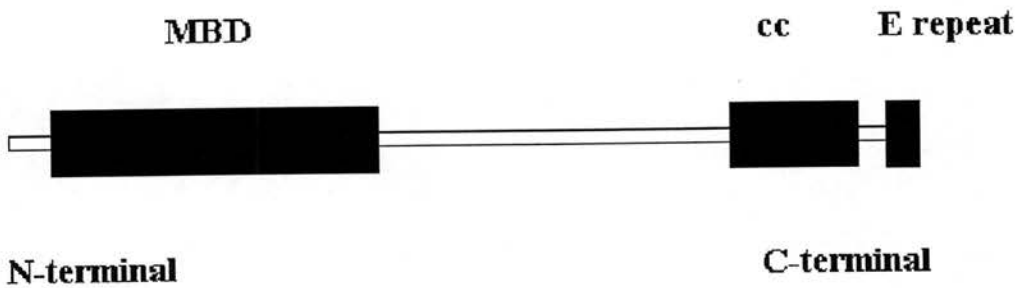
MBD3 is a member of the methyl-CpG-binding domain (MBD) protein family. It is an important subunit of the NuRD complex, a multisubunit complex containing nucleosome remodeling and histone deacetylase activities(Wade *et al.*, 1998; Wade *et al.*, 1999; Zhang Y *et al.*, 1999). *MBD3* encodes a 32 kD protein. The MBD domain is encoded by exons 1 and 2 (see Figure 7). Like MBD2, MBD3 also contains one of C-terminal coiled-coil (cc) domains (see Figure 8), which are ubiquitous and highly versatile assembly motifs found in a wide range of structural and regulatory proteins (Hendrich and Bird, 1998). Coiled-coil domains are involved in protein-protein interactions and exhibit a broad range of different functions related to their specific folding of coiled-coil domains. *MBD3* is located on chromosome 19p13.3, a region of Loss of Heterozygosity in both colon and lung cancers (Resta *et al.,* 1998; Dong *et al.,* 1998; Trojan *et al.,* 2000). Inactivating *LKB1/STK11* germline mutations in combination with loss of the wildtype allele by chromosomal loss or methylation are responsible for the development of hamartomatous polyps and adenocarcinomas in Peutz-Jeghers syndrome patients. *LKB1/STK11* however is rarely involved in sporadic colon cancer cases and at most 33% of NSCLC cases (Resta *et al.*, 1998; Avizienyte *et al.,* 1998; Trojan *et al.,* 2000; Esteller *et al.,* 2000; Launonen *et al.,* 2000), leading us to consider the role of *MBD3* as an alternative tumour suppressor gene on this location. This gene is tightly linked to *LKB1/STK11*, being only 500kb away on chromosome 19p13.3(data available on the website http://www.ncbi.nlm.nih.gov). To investigate its possible role as a tumour suppressor gene in colon and lung carcinogenesis, we screened a set of colon and lung cancer samples for genetic and epigenetic abnormalities in *MBD3*.

Figure 7 *MBD3* gene structure



There are seven exons in MBD3 gene. Exon 2,3,4,5and 6 are translated. Exon 1 is partially translated. Exon 7 is an untranslated exon. Exon1 is a high CG (78%)region in which a CpG island is located.

Figure 8 MBD3 protein structure



MBD domain is located at N-terminal. A glutamic acid repeat (E) and coil-coiled domain (cc) are at C-terminal

2.Mutational analysis by SSCP and sequencing

PCR of genomic DNA was carried out using 6 pairs of intron primers as listed in Table4. The primer sets covered the entire coding region (exons 1 to 6) of *MBD3* and included splice acceptor and donor sites (see Figure 7). Seven colon cancer, 11 SCLC, 8 NSCLC, one immortalised non-tumourigenic human bronchial epithelial cell, and 51 MSS primary CRC tumours were subjected to mutation screening.

Two aberrant SSCP bandshifts in these cell lines and primary tumour samples were found, that were then sequenced. Both of the changes were heterozygous as seen by the retention of normal bands in SSCP or sequencing gels and all were in cell lines. Specifically, in exon 3 of DLD1/HCT15 (cell lines derived from the same tumour), a G to A transition leads to a silent change (T104)(see Figure 9,10). In exon 6 of DLD1/HCT15, a C to A transversion leads to a novel leucine to methionine substitution at residue 248 (L248M) (see Figure11, 12). Since matching normal tissues were not available for the cell lines to confirm somatic mutation events in these cases, independent normal DNAs were then screened to see if these differences exist in cells of non-cancerous individuals. The exon 3 variant was found in 1 of 47 normals indicating that it is probably a rare polymorphism, but the exon 6 substitution was not seen in 54 normals, nor the silent exon3 variant in 47 normals. There is a possibility that the L248M change is a true cancer-related mutation although it is a conservative substitution and so may have little functional effect.

No mutations were found in the primary colon tumours. Interestingly, the same bandshift was found using primer pair MBD3/7 and MBD3/8 for exon3 in two

patients, which was postulated to be a rare single-nucleotide polymorphism (SNP) (see Figure 13) as it was also found in normal matched tissue from the same patients.

Considering the low frequency of changes in the lung cancer cell lines, it was decided not to screen a panel of NSCLC primary samples that would have been available.

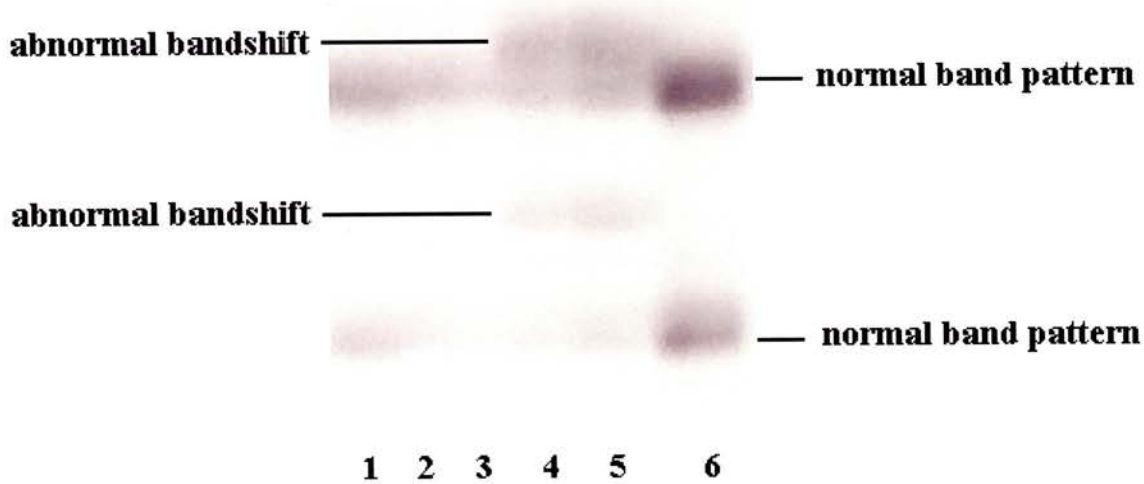
3.Analysis of methylation of the promoter CpG island

The method used to screen for methylation concentrated on the nine *HpaII/MspI* sites in the CpG island (see Figure 14 and 15), all of which must be methylated to allow PCR to give a positive result after *HpaII* digestion, and therefore gives a qualitative (total vs. partial/no methylation) rather than quantitative assessment of methylation. Results from the cell lines indicated an absence of total methylation (see Figure 16).

4.Expression of *MBD3* in colon and lung cancer cell lines

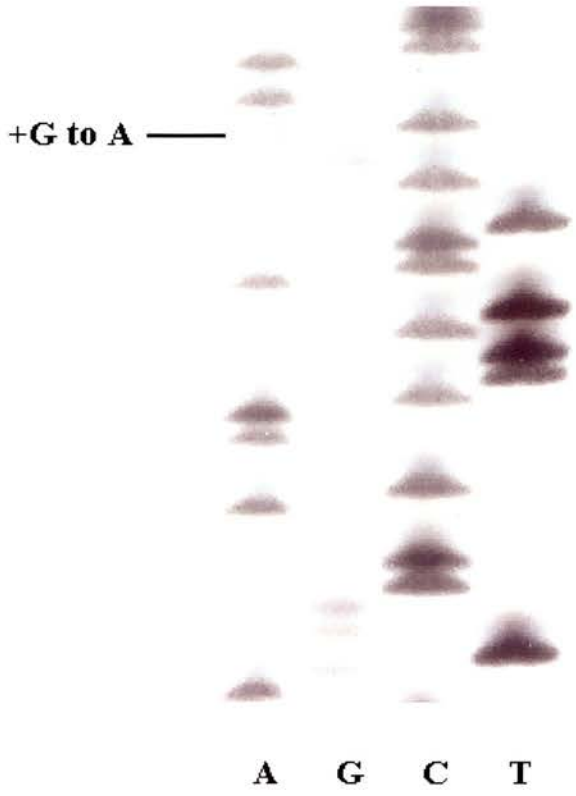
RT-PCR showed that *MBD3* expression is detectable in all the cell lines tested (see Figure17 and 18). To further confirm the expression, northern hybridisation analysis was performed. The result is consistent with that of RT-PCR (see Figure 19). Since 25ug total normal adult colon RNA was used as positive control, the band is weaker than other bands generated by 10ug polyA RNA.

Figure 9 Exon 3 abnormal bandshift on sscp gel



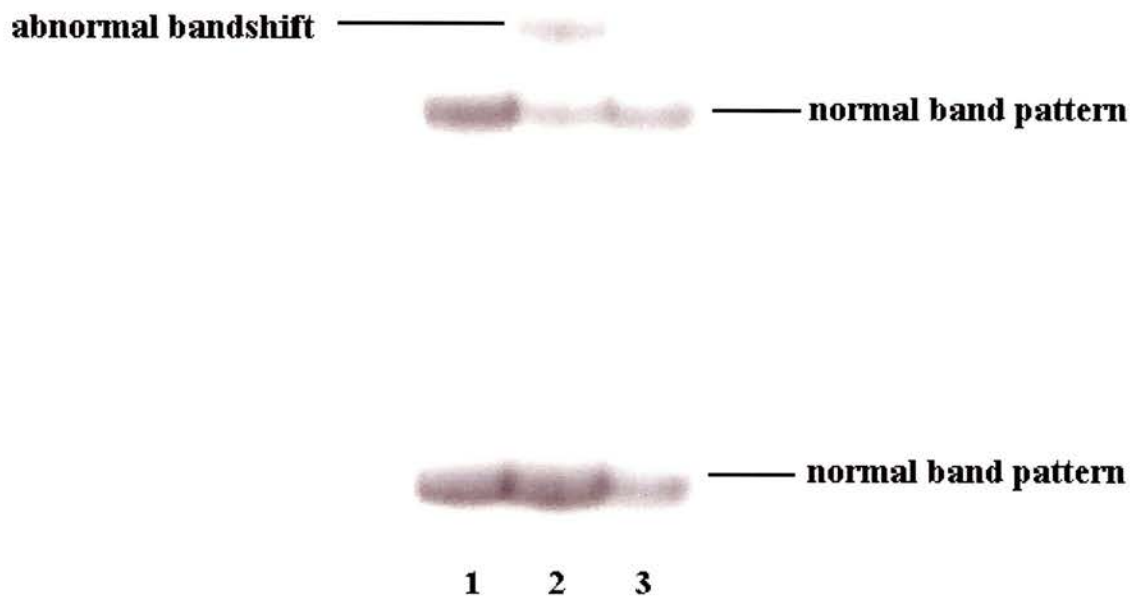
Lanes are: 1.HT29, 2.SW480, 3.COLO320, 4.DLD1, 5.HCT15 and 6.LOVO.

Figure 10 Heterozygous sequence change in exon 3 of DLD1



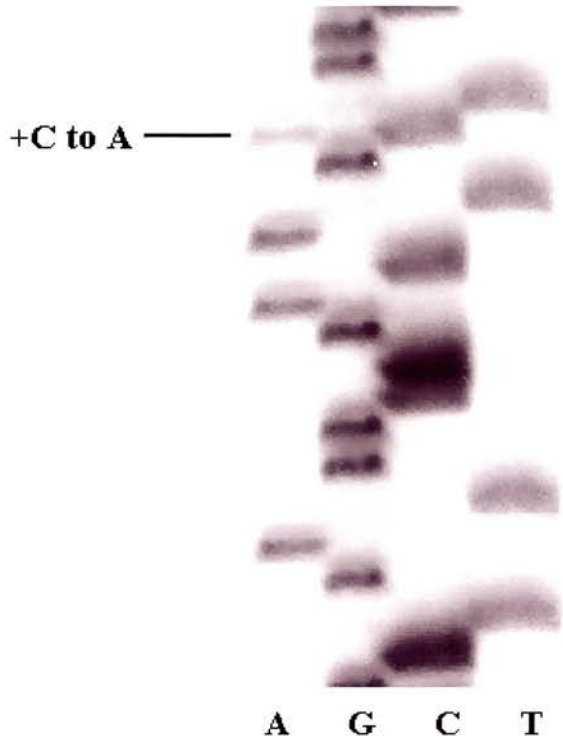
A G to A transition was found by DNA sequencing

Figure 11 Exon 6 abnormal bandshift on sscp gel



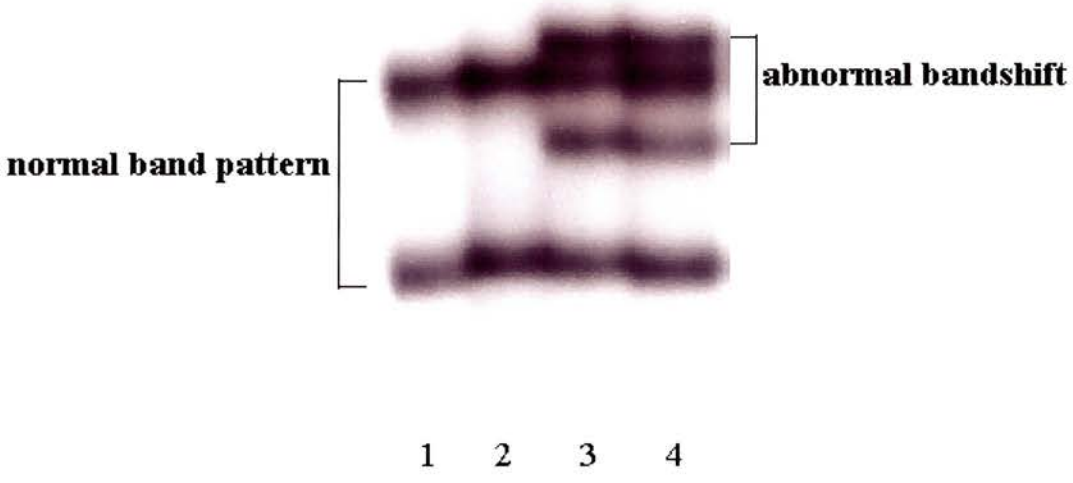
Lanes are: 1. HCT116, 2. DLD1 and 3.LOVO

Figure 12 Heterozygous sequence change in exon 6 of DLD1



A C to A transversion was found by DNA sequencing

Figure13 Putative SNP in primary tumour



Lanes are: 1. patient A tumour, 2. patient A normal tissue, 3. patient B tumour and 4. patient B normal tissue.

Figure 14 *MBD3* CpG island analysed by CpGPlot

The original definition suggested by Gardiner-Garden and Frommer is a region greater than 200 base pairs (bp) with a high-GC content and an observed/expected ratio for the occurrence of CpG > 0.6.

For *MBD3*, the input sequence is AC005943 nt8000-10104.

Window size = 100

Observed/Expected ratio > 0.60

Percent C + Percent G > 50.00

Length=912 (8540-9451) > 200

Figure 14 ***MBD3*** **CpG island CpGPLOT analysis**

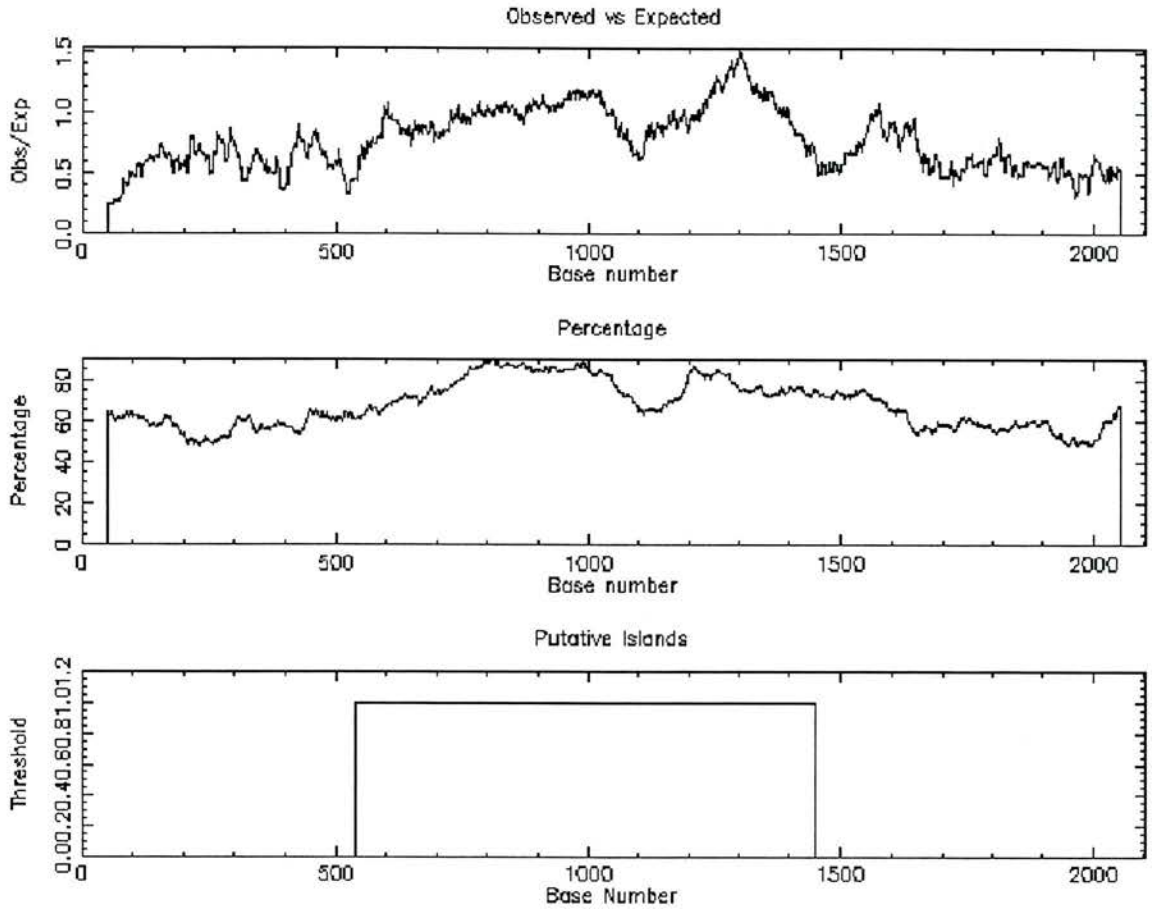


Figure 15 *MBD3 HapII/MspI* recognised CpG sites in the PCR region

GCATGTCCCCGCGCACGCGCACCAGGCGGCCGGCAGAGACAGGGCACAG

TAAAAGGAGCCGCGCGGACTCCGGGCACAGCTGCTGGGC***ACTGGCAGCT***

Primer MBD3.17

CGCAAGGCACAGCGCGCAGGCGCGCACGCCACTTAAGGCCCCGCCCCCTC

CCAGCGCGCTGATTGGCGGGGCCGTAGAGGGGCGGGGGCATGGCGCCGC

GGCTGCGCAGCGCGGGGCCCCGCGCGGGCGGGTGGGCGGAGCGGGCCCC

CCGAGCGGGGGCTGGCGAGCGCGGGGAGGGGGCCGAAAGGGCGGCAAA

G**CCGG**CGCGCGC**CCGG**CTGGGGGCGGAGGGCGGGGGCCCGCGGG**CCG**

GAACAGCCGCGGCAAGTGGCGGCGGCGGCAGCGGCAGCGGAGGCAGCT

GAGGCGGCGGCGGCGAGTGGGGGT**CCGG**GCGGCGGCGGCGGGC**CCGG**C

GGCGGGCCGAGGAG**CCGG**GCGCAATGGAGCGGAAGAGGTGGGAGTGC**C**

CGGCGCTCCCGCAGGGCTGGGAGAGGGAAGAAGTGCCCAGAAGGTCGG

GGCTGTCGG**CCGG**CCACAGGGATGTCTTTTACTATAGGTGAATGAGCGC

G**CCGG**CCGCGCAGGGCC***TCAACGGGCTCCTCCCAGCG***GCCTCTGCGGCG

Primer MBD3.18

GCCTCGGCGCCCCTGCCCCGGGCCGCGGCCTGGGCCCCGCGCGTGCGTCG

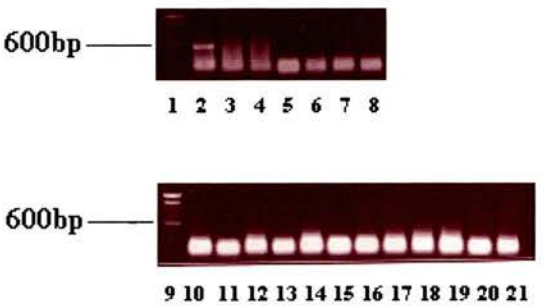
TGCGTGCGTGCGTGCGTCTTGCGTGCGTGTGCGCCCGGCGCGCGGAGGCC

CAGTGCGCGTGCGCCGCTTGTTGGGCCGCATGGCTTTCGGCCGGCCGTTC

AGGGACTGGCTGCCGGACG

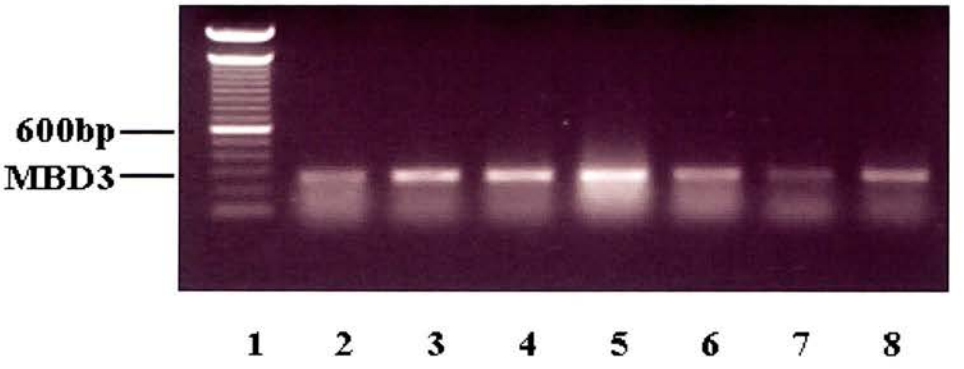
The sequence Genebank accession number: AC005943 reverse-complement nt8500-9300. There are 9 *HapII/MspI* recognised CpG sites in the PCR region generated by primer pair MBD3/17 and MBD3/18. The translation start ATG is located between sixth and seventh *HapII/MspI* recognised CpG sites.

Figure 16 Methylation sensitive *HpaII* restriction followed by PCR



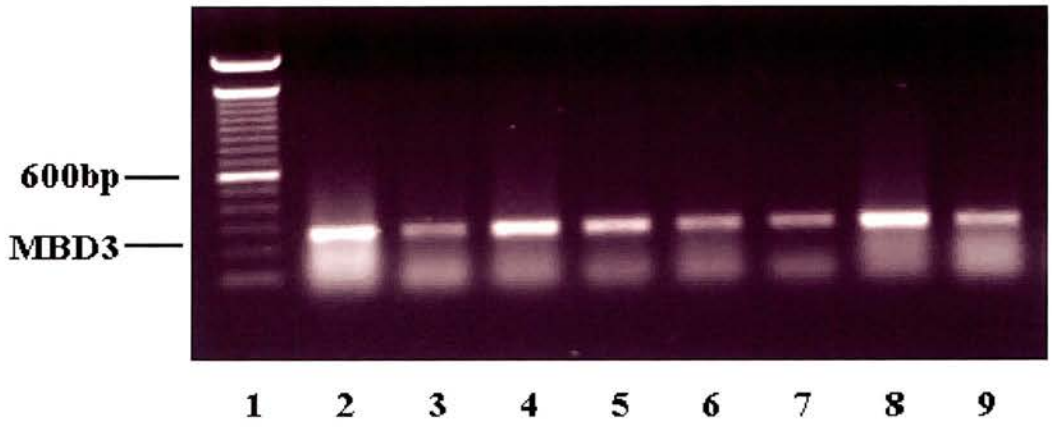
Lanes are: 1. DNA marker, 2. totally methylated DNA, 3. HT29, 4. SW480, 5. COLO320, 6. HCT116. 7. DLD1, 8. LOVO, 9. DNA marker, 10. H69, 11. H524, 12. H740, 13. H1672, 14. COR-L24, 15. COR-L47, 16. COR-L51, 17. H358, 18. H835, 19. H920, 20. H1648 and 21. H2122.

Figure17 ***MBD3*** **RT-PCR of colon cancer cell lines**



Lanes are: 1. DNA marker, 2. HT29, 3. SW480, 4. COLO320, 5. HCT116, 6. DLD1, 7. LOVO and 8. normal adult colon.

Figure 18 ***MBD3*** **RT-PCR of lung cancer cell lines**

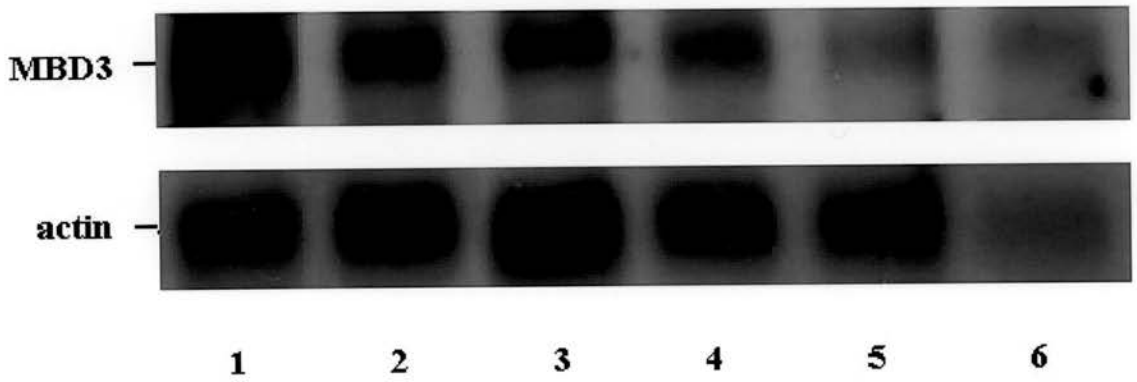


Lanes are: 1. DNA marker, 2. H69, 3. H524, 4. H740, 5. H1672, 6. COR-L24, 7. H358, 8. H835 and 9. H920

Figure 19 *MBD3* polyA+ Northern hybridisation

Lane 1-6 represents HT29, SW480, COLO320, DLD1, LOVO and normal adult colon total RNA respectively. *β-actin* was included as positive control. The transcript size of *MBD3* and *β-actin* are 2.4kb and 1.8kb respectively.

Figure 19 ***MBD3*** **polyA+ Northern hybridisation**



5.Discussion

To investigate the role of *MBD3* in colon and lung tumourigenesis this gene was assayed for mutations, promoter methylation and lack of expression. The *MBD3* gene spans more than 16kb of genomic DNA but contains 7 relatively small exons(six of which are coding). PCR followed by SSCP is a simple, sensitive method to screen for nucleotide substitutions in DNA. Exon1 (containing the start methionine codon) is located within a typical CpG island which stretches about 1kb with observed vs expected ratio >0.6 and 78% CG content (see Figure19). As high CG content DNA sequence is usually refractory to normal Taq PCR, it was necessary to use high fidelity DNA polymerase *Pfx* with high concentration Enhancer solution (5×) and high denaturation temperature to facilitate PCR amplification of this region. The remaining 5 exons(2-6) could be amplified using standard Taq polymerase.

Two missense mutations were found (1/7 colon and 1/20 lung) cancer cell lines, and none in colon primaries. Both mutations were located outside the MBD (methyl-CpG-binding domain), and one appears to be a naturally occurring rare polymorphism. The coincidence of the missense and silent mutations in DLD1/HCT15 may simply reflect the mismatch repair defect of these cell lines due to *MSH6* mutation. The L248M change may be a somatic event, but since it is a conservetive substitution it may not be significant for tumourigenesis. Considering the expressions of other genes involved in tumourigenesis are silenced or reduced by epigenetic alterations, with which hypermethylation is the most common associated events. Hypermethylation usually shows methylation across the bulk of the associated CpG island. We expected to find such abnormalities with *MBD3*.

Hypermethylation is an alternative mechanism for inactivation of tumour suppressor gene and tumour-related genes in many human cancers (Momparler, 2003) even in some tumour types where mutations rarely occur (e.g. GSTP1 in prostate cancers, Millar *et al.,* 1999). In sporadic Colorectal cancer, methylation of *MLH1* is the predominant mechanism of inactivation rather than mutation (Cunningham *et al.,* 1998), while the reverse is true for inactivation in HNPCC cases (Leach *et al.,* 1993; Fishel *et al.*, 1993; Aaltonen *et al.*, 1993; Kolodner *et al.*, 1999).

Methylation analysis using restriction enzyme digestion followed by PCR is a simple and rapid method of methylation analysis based on the fact that some restriction enzymes are unable to cut methylated DNA but their isoschizomers can. *HpaII* - *MspI* were selected as they have 9 recognition sites in the promoter region we were looking at. Both enzymes recognize CCGG sequence; however *HpaII* is unable to cut DNA if cytosine is methylated. In this study, there are nine CpG sites of the *MBD3* recognised by the isoschizomers pair. If all of the nine sites are methylated, the *HpaII* digested DNA will give a clear PCR band but *MspI* will not. Since hypermethylation usually occurs across the bulk of the associated CpG island, the nine CpG sites of the *MBD3* CpG island (about 50% of the putative CpG island) were expected to be methylated in a significant proportion of cells to give a detectable PCR result if this phenomenon had occurred. The data showed that none of the cell lines were fully methylated across the region tested.

In addition to epigenetic abnormality due to hypermethylation, gene silencing can also occur in the absence of methylation (perhaps due to inactivation of transfactors). Thus RT-PCR and confirming northern hybridisation analyses were performed to see if the expression of *MBD3* is absent or reduced. RT-PCR and northern hybridisation analyses results showed that *MBD3* is expressed in all the cell lines assayed at significant levels. These results are consistent with a lack of aberrant, tumour-associated silencing of the gene. In conclusion, *MBD3* is not a target of genetic or epigenetic alteration in colon and lung cancer.

Chapter 4

Analysis of Expression and Methylation Status of MBD2 Putative Promoter Region

Contents

1.Introduction

MBD2 is a transcriptional co-repressor (Ng *et al.*, 1999) that binds specifically to mCpG residues in DNA and exhibits a preference for densely methylated regions. It is associated with HDAC1 and HDAC2. MBD2 may direct Mi-2 remodeling complexes to methylated regions in chromatin, mediating transcriptional silencing (Zhang *et al.*, 1999 and Wade *et al.*, 1999). *MBD2*, mapped on the human chromosome 18q21 (a region of LOH in colon and lung cancer), comprises six coding exons and one non-coding exon. The MBD domain is encoded within exons 1 and 2.

Prior to this study, the colon and lung cancer cell line panel had been screened for mutations in exons 2-6, but very few changes had been found (Bader *et al.*, 2003). Exon 1 containing the N-terminal portion of the MBD was not studied because its sequence is very G/C rich and so difficult to PCR amplify. Although a method was then developed to PCR amplify across the region using Pfx polymerase and its enhancer solution, this region was still too large for simple SSCP analysis, and no convenient enzymes were found to cut the PCR fragments into smaller pieces before SSCP. No other methods were developed or attempted to screen this region. Reports have been made of a decrease in expression of MBD2 in some tumours (Muller-Tidow *et al.*, 2001) raising the possibility that MBD2 is inactivated not by mutation but by downregulation. The aim of this project was therefore to look at expression of the gene in the cell lines and to screen for hypermethylation of the putative promoter region.

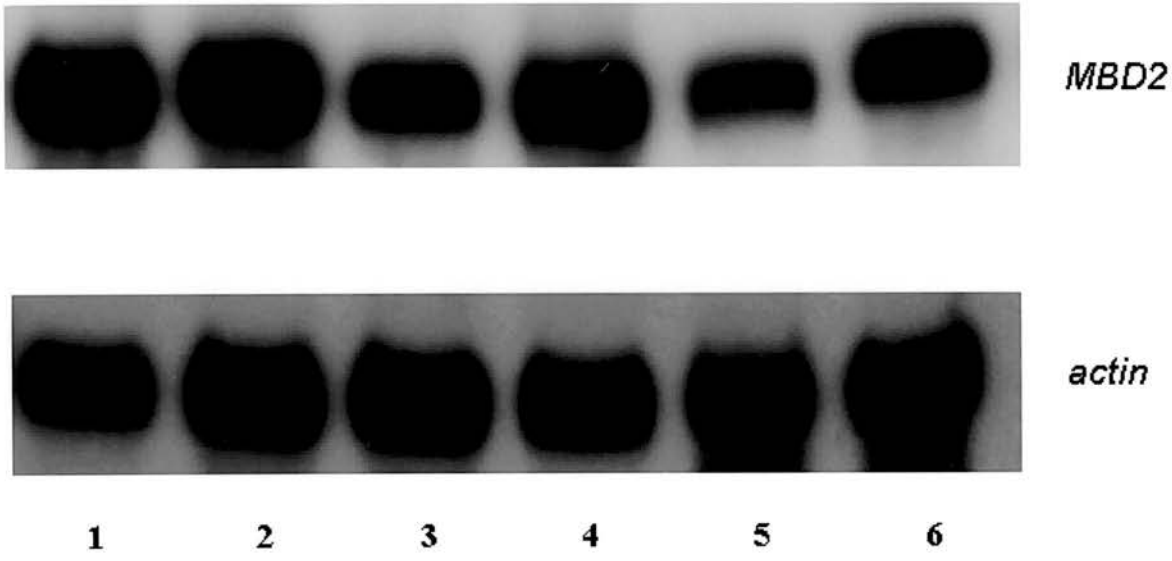
2. *MBD2* is expressed in colon and lung cancer cell lines

Northern hybridisation analysis showed that *MBD2* expression is detectable in the selection of cell lines tested (see Figure 20 and 21). By eye, the intensity of *MBD2* bands indicated that expression was abundant and roughly equivalent. β-actin was used as a loading control to assist with comparisions. No attempt was made to generate quantitative measurements of expression levels by densitometry or real time PCR (as in references Billard *et al.*, 2002 and Muller-Tidow *et al.*, 2001). At best, such data would simply have replicated that already published and would not address the probable mechanism(s) involved in downregulation of the gene, e.g. hypermethylation.

3. No dense methylation found in the promoter CpG island

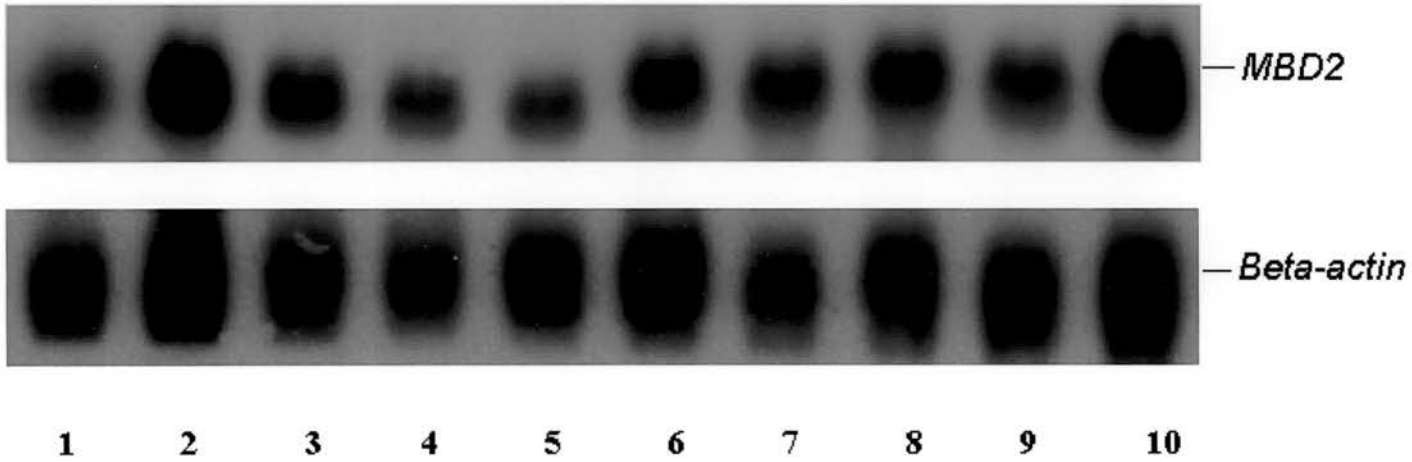
The method used to screen for methylation concentrated on the two *SmaI/XmaI* sites within a 670bp region of the predicted CpG island (nt425-1095 of sequence AF120988)(see figure 22 and 23). Both *SmaI* and *XmaI* recognize CCCGGG sequence. Two recognition sites of both restriction enzymes contain CpG pair(s). When the CpG sites are methylated, *SmaI* will not cut, but *XmaI* is unaffedted. Both sites must be methylated to allow PCR and to give a positive result after *SmaI* digestion. The assay therefore gives a qualitative (total vs. partial/no methylation) rather than a quantitative assessment of methylation. *XmaI* was used for a standard control as it can cut methylated CpG sites. Results from the cell lines indicated an absence of total methylation as judged by the lack of PCR. A totally methylated genomic DNA control (Sigma) was used in parallel as positive control for the methylation (see Figures 24 and 25).

Figure 20 ***MBD2*** **Northern hybridisation analysis of colon cancer cell lines**



Lanes are: 1. HT29, 2. SW480, 3. HCT116, 4. DLD1, 5. LOVO and 6. normal adult colon total RNA. The MBD2 transcript is 2.6kb.

Figure 21 ***MBD2*** **Northern hybridisation analysis of lung cancer cell lines**



Lanes are: 1. H69, 2. H524, 3. H740, 4. H1672, 5. COR-L24, 6. H358, 7. H835, 8. H920, 9. H1648 and 10.normal adult lung total RNA.

Figure 22 *MBD2* CpG island analysed by CpGPlot

The input *MBD2* exon1 sequence is AF120988 nt1-1993

Window size = 100

Observed/Expected ratio > 0.60

Percent C + Percent G > 50.00

Length 799 (169-967), Length > 200

Figure 22 ***MBD2*** **CpG island CpG analysis**

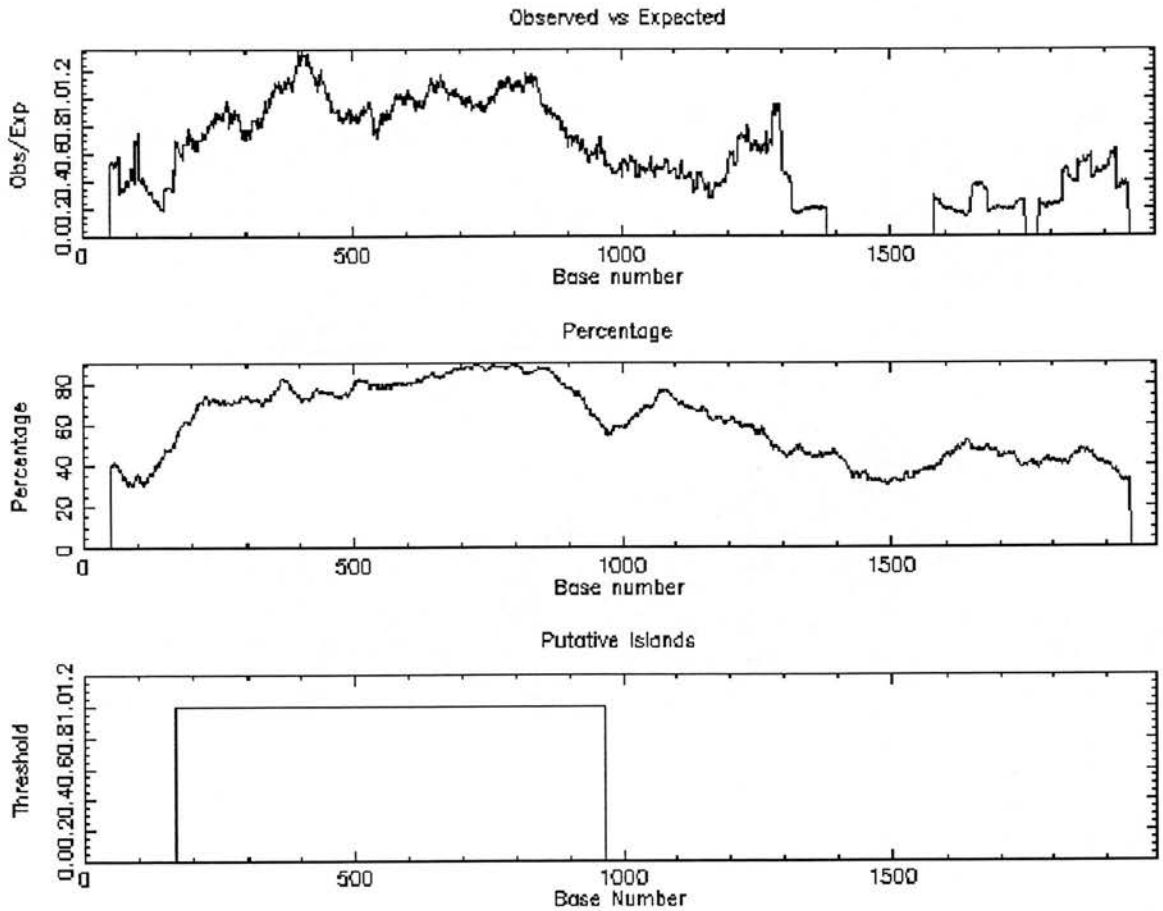


Figure23 *SmaI/XmaI* restriction CpG sites in the PCR region

TTGGAGGACCTAAGAGGCGGTGGCCGGGGCCACGCCCCGGGCAGGAGGGCCGCTCTGT

GCGCGCCCGCTCTATGATGCTTGCGCGCGTCCCCCGCGCGCCGCGCTGCGGGCGGGGCG

GGTCTCCG**GGATTCCAAGGGCTCGGTTACGGA**AGAAGCGCAGCGCCGGCTGGGGAGG

Primer MBD2.15

GGGCTGG**ATG**CGCGCGCA**CCCGGG**GGGAGGCCGCTGCTGCCCGGAGCAGGAGGAGGGG

MBD2a *SmaI/XmaI*

GAGAGTGCGGCGGGCGGCAGCGGCGCTGGCGGCGACTCCGCCATAGAGCAGGGGGGCC

AGGGCAGCGCGCTCGCCCCGTCCCCGGTGAGCGGCGTGCGCAGGGAAGGCGCTCGGGGC

GGCGGCCGTGGCCGGGGGCGGTGGAAGCAGGCGGGCCGGGGCGGCGGCGTCTGTGGCCG

TGGCCGGGGCCGGGGCCGTGGCCGGGGACGGGGACGGGGCCGGGGCCGGGGCCGCGGCC

GTCCCCCGAGTGGCGGCAGCGGCCTTGGCGGCGACGGCGGCGGCTGCGGCGGCGGCGGC

AGCGGTGGCGGCGGCGCCCCCCGGCGGGAGCCGGTCCCTTTCCCGTCGGGGAGCGCGGG

GCCGGGGCCCAGGGGAC**CCCGGG**CCACGGAGAGCGGGAAGAGG**ATG**GATTGCCCGGCC

SmaI/XmaI MBD2b

CTCCCCCCCCGGATGGAAGAAGGAGGAAGTGATCCGAAAATCTGGGCTAAGTGCTGGCAA

GAGCGATGTCTACTACTTCAGGTACCTCCCTGGGGGCGGGGAGGGGGTGGCGGGGTCAG

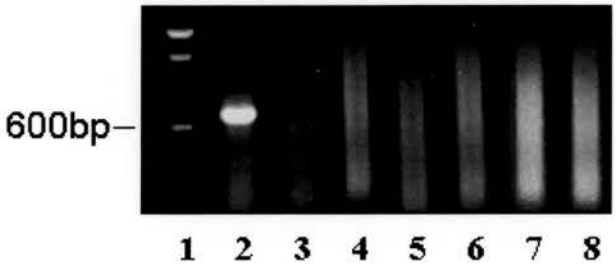
GCCGGGGTCAGG**GGTCAAGAGCGGGCCTGGGCA**GAGGATGAGCGCTGGGGCCCGGGG

Primer MBD2.16

GGCATGTGGCAGGGACAGGCTG

The sequence accession number at Genebank is AF120988 (nt301-1141). There are two SmaI/XmaI restriction sites in the PCR region amplified by primer pair MBD2.15 and MBD2.16. The translation start ATGs for both MBD2a and MBD2b locate in this region.

Figure 24 *MBD2* methylation analysis of colon cancer cell lines



Lanes are: 1. DNA maker, 2. total methylated DNA. 3. HT29, 4. SW480, 5. COLO320, 6. HCT116, 7. DLD1 and 8. LOVO.

Figure 25 *MBD2* methylation analysis of lung cancer cell lines

Lanes are: 1. DNA marker, 2. total methylated DNA, 3. H69, 4. H524, 5. H740, 6. H1672, 7. COR-L24, 8. H358, 9. H835 and 10. H920.

4. Discussion

The maintenance of CpG-methylation is essential for normal embryonic development and deletional mutants that disturb this process can be lethal (Li *et al.*, 1992). Promoter hypermethylation has been shown to be associated with the silencing of tumour suppressor genes. As an important element in the process of methylation-associated transcriptional repression, MBD2 was considered to be an attractive candidate as tumour suppressor gene, the loss of whose activity could lead to further, genomewide changes in gene regulation, thus contributing to tumourigenic progression.

Northern hybridization was used to look for the absence of *MBD2* expression in a panel of human tumour cell lines and showed that it is expressed in all the cell lines assayed. To a first approximation therefore, *MBD2* is expressed at significant levels in these colon and lung lines, but other reports describe a decrease in expression of between 20 and 80% of some primary cancers compared with normal tissue (Müller-Tidow C *et al.,* 2001). One possible mechanism for such a decrease in expression involves hypermethylation of the promoter region. Since no matching normal tissue was available for comparison with the cell lines studied, assays were done to look for such methylation in the tumour lines.

MBD2 contains a typical CpG island starting upstream of exon1 and crossing the intron-exon conjunction, stretching about 800bp. Restriction enzyme digestion followed by PCR was used to assay *MBD2* promoter methylation status. *SmaI/XmaI were* selected to assay two CpG sites within the promoter region. Reports of

hypermethylation of other genes involved in tumourigenesis usually shows methylation across the bulk of the associated CpG island. We would therefore have expected the two CpG sites of the *MBD2* CpG island to have been methylated in a significant proportion of cells to give a detectable PCR result after digestion by these enzymes that have at most 2 sites within the PCR fragment. Our result shows that none of the cell lines were methylated at the sites tested. In combination therefore, our expression analysis and methylation results are consistent with a lack of aberrant, tumour-associated silencing of the gene.

The primary tumour-associated decreased in expression of *MBD2* that is reported in the literature may still be correct, despite our lack of supporting evidence in the cell lines. Another mechanism by which expression could be decreased may be simply due to a lack of compensatory upregulated expression from the remaining wild type allele after LOH of the region. There are no reports yet of methylation of the *DCC* gene, suggesting that its decreased expression in colon tumours may also simply be due to the 18q21 LOH.

On the other hand, this data is consistent with reports showing that *MBD2* overexpression can occur in breast cancer (Billard *et al.*, 2002), and that *MBD2* downregulation inhibits the growth of cultured transformed cells (Slack *et al.*, 2002). Recently it was reported that deficiency of Mbd2 suppresses intestinal tumourigenesis (Owen *et al.*, 2003). In that study, $APC^{Min/+}Mbd2^{-/-}$ mice survived significantly longer than control mice ($APC^{Min/+}Mbd2^{+/+}$), whereas $Mbd2^{+/-}$ mice survived for an intermediate length of time. At death, $Mbd^{-/-}$ mice had 10 times fewer

adenomas than $APC^{Min/+}Mbd2^{+/+}$controls. Taken together, it is more likely that MBD2 acts as an oncogenic factor in tumour progression.

Chapter 5

Analysis of Methylation Status of PTC Putative Promoter and CGG repeat Mutation

Contents **Page**

1. Introduction

As mentioned in Chapter 1, the hedgehog pathway is important for colon and lung development. *Shh* knockout mice have been reported to show foregut, trachea and lung abnormalities, reduction of intestinal smooth muscle, gut malrotation and annular pancreas and persistent cloaca, where the distal intestinal and genitourinary tracts remain in a common channel. PTC normally maintains the hedgehog pathway in an inactive state by inhibiting the signalling effector SMO (reviewed in Ingham and McMahon, 2001). Shh activates signalling in target cells by inhibiting PTC, resulting in derepression of SMO and activation of Hedgehog target genes such as *PTC* and *Gli1*. Whereas this pathway normally is regulated by the spatially and temporally restricted expression of Shh, loss-of-function *PTC* mutations are associated with constitutive hedgehog signalling in human cancers, particularly BCCs (reviewed in Mullor *et al.*,2002). The findings of mutations in the *PTC* gene in both Gorlin's syndrome and sporadic basal cell carcinomas are thus precedents for abnormalities of this gene in human disease.

The *PTC* gene is located on chromosome 9q22.3 where it contains 23 exons spanning approximately 34 kb. There are two CpG islands, the first one covering the putative promoter region (5'-untranslated region and exon1)(see figure 26) and the second one covering exon 2 and stretching into both intron1 and intron2. In its 5'-untranslated region, there is a seven triplet CGG repeat region (see figure 27), a sequence that can be vulnerable to expansion and deletion in some diseases and microsatellite unstable cancers. As a part of a study of *PTC* in colon and lung cancers, a panel of cancer cell lines and primary tumours were assayed for *PTC*

expression, hypermethylation of the first CpG island and instability of the CGG repeat.

2. No cancerous mutation found in the triplet repeat sequence

4/8 colon cancer cell lines and 6/21 lung cancer cell lines presented identically shifted bands when assayed by SSCP across the repeat region. In the colon cancer set, all of these bandshifts occurred in MSI lines. In the lung cancer set, 4 bandshifts occurred in SCLC cell lines, 1 in an NSCLC cell line and 1 in the immortalised non-tumourigenic human bronchial epithelial cell. Further, MSS primary tumours with matched normal tissue and MSI primary tumours with matched normal tissue were also screened. Although 42/97 (42%) bandshifts were found in the MSS primary tumours and 10/18 (56%) bandshifts in the MSI primary tumours, all of these bandshifts also occurred in matched normal tissues. Considering the incidence in both normal and tumour samples, and the apparent identical nature of the bandshift, it was concluded that this is a naturally occurring polymorphism in the gene.

3. *PTC* is expressed in all cancer cell lines tested.

Northern hybridisation analysis (see figures 28 and 29) showed that *PTC* is expressed in colon and lung cancer cell lines. The intensity of bands is almost equal to those of normal foetal and adult colon and lung.

4. No dense methylation found in the promoter CpG island

The method used to screen for methylation concentrated on the 5 *XmaI/SmaI* sites, 4 *NotI* sites, all of which must be methylated to allow PCR and give a positive result

after *SmaI* digestion, and therefore gives a qualitative assessment of methylation(as described in Chapter 2). Results from the cell lines indicated an absence of total methylation (see Figures 30, 31 and 32).

Figure 26 *PTC* CpG island analysed by CpGPlot

The input sequence is NT008470 (reverse complemented, nt96957701-96959386).

Window size = 100

Observed/Expected ratio > 0.60

Percent C + Percent G > 50.00

Length 1068 (154-1221) Length > 200

Figure 26 ***PTC*** **CpG island CpGPlot analysis**

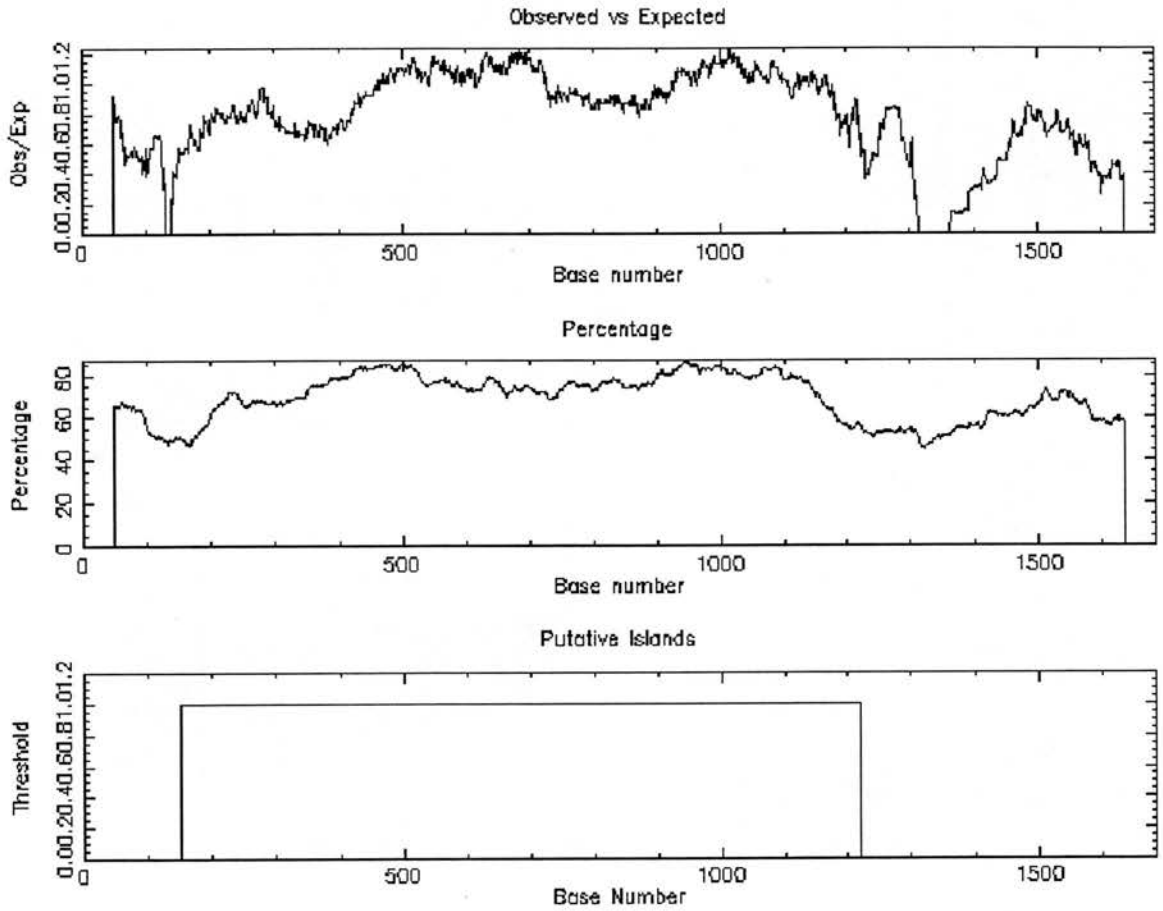


Figure 27 *SmaI/XmaI* restriction CpG sites in the PCR regions

C***GCGAGCCAATCGCGTCGCA***GAAACCCAAGCCATCTGACAGC**CCCGGG**GGCGGGAGCT
Primer PTC 1 *SmaI/XmaI*
CCAGCCTTTTTCTCTTTCGCTCGCTCTCTCCCCCTCCACCCCCTCCCCTCCCCTCCTCCTC

TTCTTTC TCCAAATGGAGAGAGTTCTTTTTTTTCTTCTTCCATCTCATCTATTGAATCAAG

GAGCTGCTGCGGCCGCTGCCCGCTGCACACACACAGAGCGGAGTCCCCAGGT**CCCGGG**A
SmaI/XmaI
GCGAGAGAGGCG CGCTGCACGGGGACAGAATGGTCCAGCGGGTTGCCGAGGAGCACAA

GAAAGCAGAGTCCGGGACCGAGCAGCCACCGCGAACCCAGCAGCCAGAGCCCGAGCAG

CCCGAGCAGCAGCTCCTGG GCCGCCACCGCCAGCAGCAGCCACCGCGGGAGCAGCGGC

AGCTGCGGCTG**CCCGGG**CCCGGCAGCGCTGAGACCCGCCGGGCACCCTCGGACCCCGCG
SmaI/XmaI
GCGGCGGCG***GCTCTGCGCTTGCCCTCC***GCGGCCGCTCGGGCGA**CCCGGG**AGGCGCCGA
Primer PTC2 and 3 *SmaI/XmaI*
GAGAGCCAGCGCGGCCGGCGGGAGCAGCGGGGATTCGCTGGCTCTTTCTGCAGTGAAGG

GGTCGCGGCGCGGGGCGGGGGGCGGGTAGGGGGAAGTTGGGGGACCGCAAGGAGTGCC

GCGGAAGCGCCCGAAGGACAGGCTCGCTCGGCGCGCCGGCTCTCGCTCTTCCGCGAACT

GGATGTGGGCAGCGGCGGCCGCAGAGACCTCGGGACCCCCGCGCAATGTGGCAATGGAA

GGCGCAGGGTCTGACTCCCCGGCAGCGGCCGCGGCCGCAGCGGCAGCAGCGCCCGC***CGT***

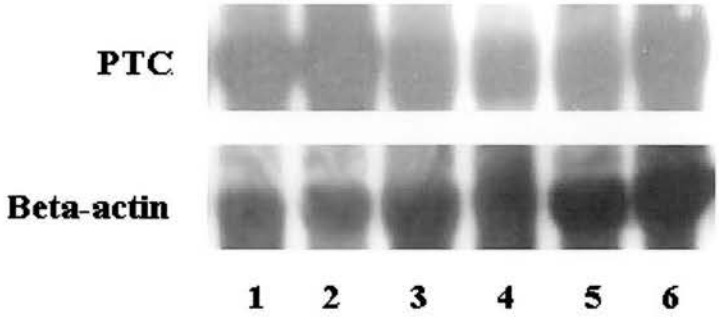
GTGAGCAGCAGCAGCGGCTGGTCTGTCAACCGGAGCCCGAGCCCGAGCAGCCTGCGGC***C***
Primer PTC4
AG CAGCGTCCTC GCAAGCCGAGCGCCCAGGCGCGCCAGGAGCCCGCAGCAGCGGCAGC
Primer PTC9
AGCGCGCCGGGCCG**CCCGGG**AAGCCTCCGTCCCCG**CGGCGGCGGCGGCGGCGGCG**GC
SmaI/XmaI Seven CGG triplets repeat
AAC**ATG**GCCTCGGCTGGTAACGCCGCCGAGCCCCAGGACCGCGGCGGCGGCGGCAGCGG

CTGTATCGGTGC**CCCGGG**ACGGCCGGCTGGAGGCGGGAGGCGCAGACGGACGGGGGGG
SmaI/XmaI
CTGCGCCGTGCTGCCGCGCC GGACCGGGACTATCTGCACCGGCCCAGCTACTGCGACGC

C***GCCTTCGCTCTGGAGCAGATTTC***CAAGGTGCATTTCAGACTCT CT
Primer PTC6

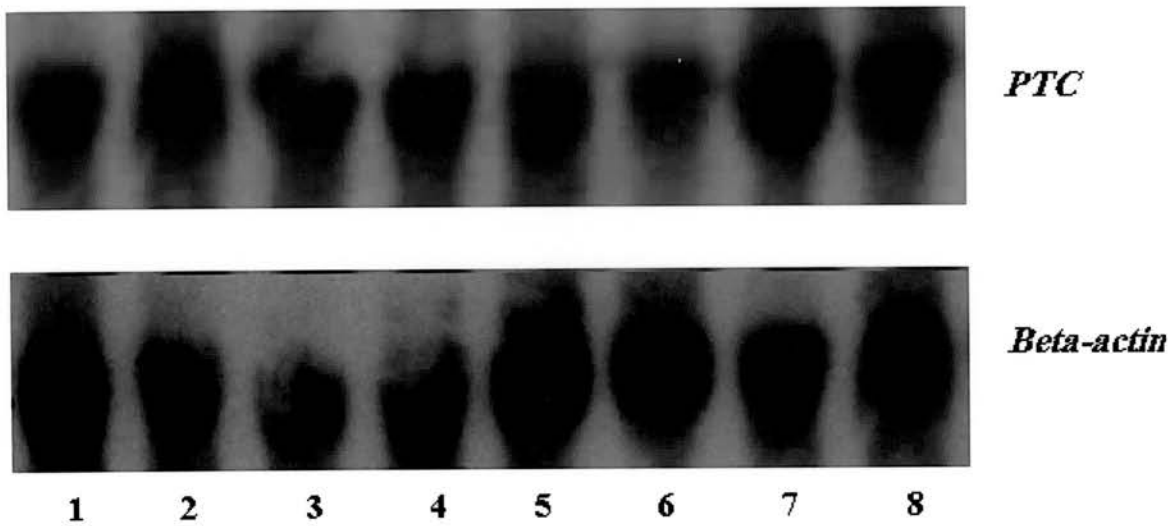
The sequence accession number at Genebank is NT008470 (reverse complemented, nt96957701-96958922). There are 3 *SmaI/XmaI* sites in the PCR region amplified by PTC1/2, 1 site in the PCR region amplified by PTC3/4 and 2 sites in the PCR region amplified by PTC9/6. One of the translation start ATG locate in the region between PTC9 and PTC6. The seven CGG triplets repeat locates upstream of the translation start ATG.

Figure 28 ***PTC*** **Northern hybridisation analysis of colon cancer cell lines**



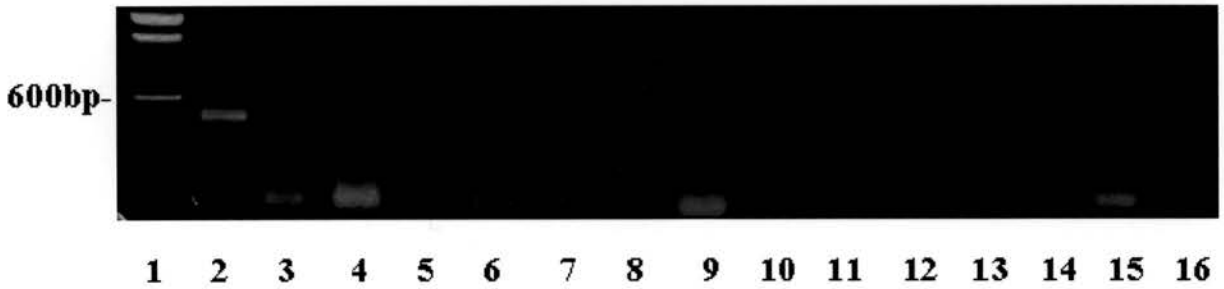
Lanes are: 1. HT29, 2. SW480, 3. COLO320, 4. HCT116, 5. DLD1 and 6. normal adult colon total RNA.

Figure 29 ***PTC*** **Northern hybridisation of lung cancer cell lines**



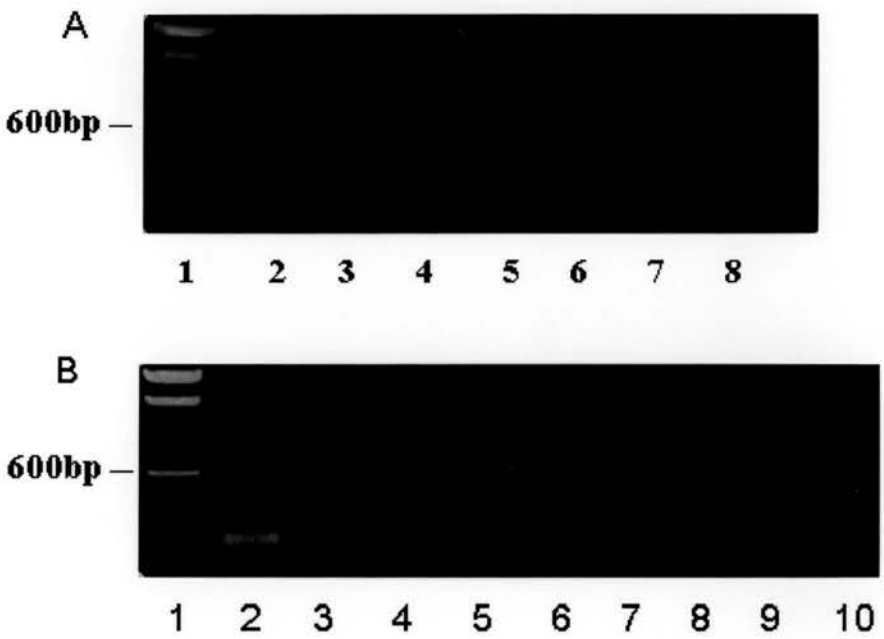
Lanes are: 1. H69, 2. H524, 3. H740, 4. H1672, 5. COR-L24, 6. H358, 7. H835 and 8. normal adult lung .

Figure 30 ***PTC*** **methylation analysis of colon and lung cancer cell lines (PTC1/2)**



Lanes are: 1. DNA marker, 2. totally methylated DNA, 3. HT29, 4. SW480, 5. COLO320, 6. HCT116, 7. DLD1, 8. LOVO, 9. H69, 10. H524, 11. H740, 12. H1672, 13. COR-L24, 14. COR-L51, 15. H358 and16. H835 .

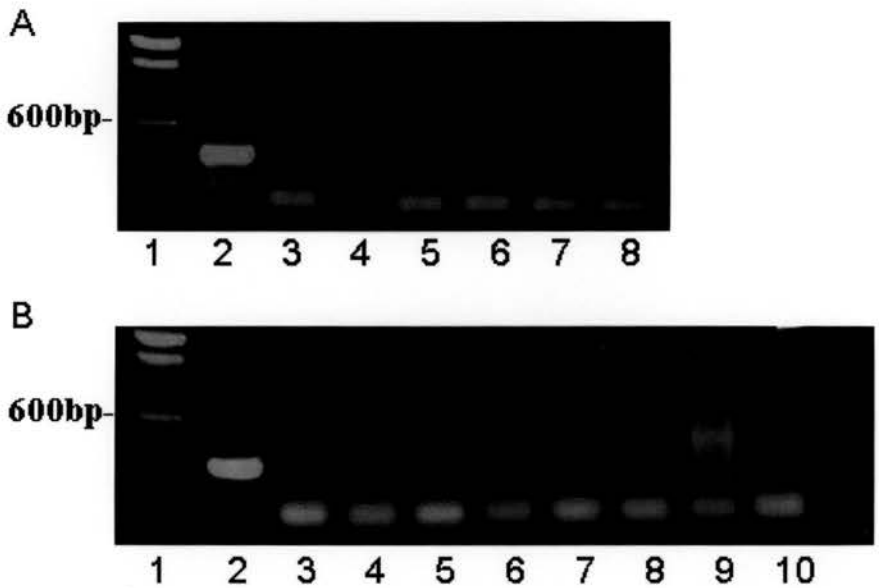
Figure 31 ***PTC*** **methylation analysis of colon and lung cancer cell lines (PTC3/4)**



A. Lanes are: 1. DNA marker, 2. totally methylated DNA, 3. HT29, 4. SW480, 5. COLO320, 6. HCT116, 7. DLD1 and 8. LOVO

B. Lanes are: 1. DNA marker, 2. totally methylated DNA, 3. H69, 4. H524, 5. H740, 6. H1672, 7. COR-L24, 8. H358, 9. H835 and 10. H920.

Figure 32 *PTC* methylation analysis of colon and lung cancer cell lines(PTC9/6)



A. Lanes are: 1. DNA marker, 2. totally methylated DNA, 3. HT29, 4. SW480, 5. COLO320, 6. HCT116, 7. DLD1 and 8. LOVO.
B. Lanes are: 1. DNA marker, 2. totally methylated DNA, 3. H69, 4. H524, 5. H740, 6. H1672, 7. COR-L24, 8. H358, 9. H835 and 10. H920.

5. Discussion

Mutation in triplet repeat sequences can cause human diseases by forming a variety of DNA conformations (see review by Reddy PS *et al.,* 1997). Although the mechanism by which trinucleotide expansion occurs is not completely understood, it is currently thought that DNA secondary structure plays an important role in the process. Recent studies have shown that CAG, CTG, CCG and CGG repeats form stable secondary structures *in vivo* that defeat DNA repair enzymes (Morrone *et al.,* 1997). During replication, a structure-specific nuclease (FEN-1) is required to cleave displaced Okazaki fragments (flaps). The presence of a hairpin formed by CNG (where N=C or G) expansion results in stimulation of flap formation and inhibition of FEN-1 cleavage, thus increasing the likelihood of occurrence of uncleaved flaps. Hairpins formed by the (GCC).(CGG) expansions responsible for Fragile X syndrome were also found to be 10–15 times more efficient substrates for methyltransferase (the enzyme that methylates the CpG sites in DNA) than the corresponding Watson-Crick duplexes (Chen *et al.,* 1998). Expansion of this repeat, which is located in the 5'-untranslated region of the Fragile X mental retardation gene (*FMR1*), results in hypermethylation, formation of a CpG island within the *FMR1* promoter and reduced *FMR1* transcription (reviewed in Cummings *et al.,* 2000). Similarly, expanded (CGG).(GCC) repeats in the promoter region of the *FMR2* gene are hypermethylated and lead to transcriptional silencing of *FMR2* in FRAXE patients (reviewed in Cummings *et al.,* 2000). One group showed that alteration of triplet repeat sequences may cause improper expression of disease related genes, through their effects on chromatin structure (Tomita *et al.,* 2003).

They found that $(CGG)_{12}$ disrupts an array of positioned nucleosomes and insertion of $(CGG)_{12}$ increases gene expression about 10-fold of a UAS-less promoter.

In the case of *PTC*, it was hypothesised that the alteration of $(CGG)_8$ in *PTC* promoter region may be associated with tumourigenesis and so mutation analysis was performed to look for CGG expansion and deletion. The results showed that there was one commonly occurring alteration that was not related to cancer as it also occurred in matched normal tissues. The change was thus concluded to be a polymorphism.

Tumour suppressor genes are often down regulated in cancer, especially by promoter hypermethylation. *PTC* expression and methylation status were also studied therefore. The results showed expression of *PTC* in both colon and lung cancer cell lines and no methylation, at least not at the two *SmaI* sites, in the promoter region, which is consistent with the northernblot data. In addition to SSCP data screening for mutation of *PTC* (from our laboratory, unpublished), there is thus no evidence to support the hypothesis of abnormalities of *PTC* in association with either colon or lung cancers.

Chapter 6

SMO and *GLI3* Expression Is Down Regulated in Cancer

Conetnts **Page**

1.Introduction

As mentioned in Chapter 1, the Hedgehog pathway plays an important role in development of several tissues, including colon and lung. Studies have shown that several members are implicated in cancers. As a key member of the hedgehog pathway, SMO is essential for pattern formation and morphogenesis in multicellular organisms. In the absence of Hh, PTC binds to SMO and inhibits its activity, but when present Hh binds to PTC and stops PTC inhibition of SMO, that then transmits a signal to a cytoplasmic complex that ultimately leads to GLI activation and downstream gene expression. In mouse knock-out models of *SMO*, the phenotype essentially mimics knock-out of *Shh* and includes developmental abnormalities of the gut. Mutations have been published for *Hh* and *PTC* in medulloblastomas and basal cell carcinomas presumably due to consitutive activation of the Hh pathway. Oncogenic mutations have also been reported for *SMO* in sporadic basal cell carcinoma. Colon and lung cancer cell lines and primary tumours have been screened for abnormalities in *SHH* and *PTC* with very few changes identified (other unpublished work from our laboratory and Chapter 5 for PTC). *SMO* is also a reasonable candidate in the Hh pathway to screen for mutations and epigenetic changes in the same samples.

2. *SMO* is down regulated in colon cancer and lung cancer cell lines.

Colon cancer and lung cancer cell lines were analysed by RT-PCR for expression (see Figures 33 and 34). Colon cancer cell lines were further assayed by Northern Blot analysis of polyA RNA, with total RNA from foetal and adult colon as normal controls (see Figure 35). A PCR product amplified by primers in exon3 and

6(SMO13 and SMO43 see Table4) was used as the probe. Foetal and adult colon expressed *SMO*; however, 3/7 colon cancer cell lines and 6/9 lung cancer cell lines lacked expression. Colon cancer cell lines COLO320, HCT116 and LOVO clearly express *SMO*, while HT29, DLD1 and LS180 were negative. The results of Northern hybridisation analysis were consistent with those of RT-PCR except that the SW480 expression is very low. In this case, there seems to be a very little expression, below the level of detection by Northern blot hybridisation but detectable by the much more sensitive method of PCR.In lung cancer cell lines, COR-L47, COR-L51, H358, H2122 and COR-L23 lack expression of *SMO* as detected by PCR. In contrast, COR-L24, COR-L279, H1648, COR-L105 and the immortalised non-tumourigenic human bronchial epithelial cell line BW1799 (a "normal" control) expressed *SMO*, although expression COR-L279 appears low. Expression and lack of expression did not correlate with lung cancer type (SCLC vs NSCLC).

3. *SMO* is methylated in non-expressing cancer cell lines.

Since some of the cell lines were negative for *SMO* expression, it was important to assay for methylation of the promoter region. Although there are as yet no reports of the identity of the promoter region, the gene has a typical CpG island (see Figure 36), stretching for 1kb and including part of exon 1. The CpG rich region upstream of exon1 is likely to contain part or all of the promoter. With the expectation of methylation in some or all of the SMO-negative cell lines, it was decided to use a more sensitive and informative assay to gauge the incidence of methyl-CpG than the methylation-sensitive restriction enzyme/PCR assay used for *MBD3* and *PTC*. In the latter assay, only a limited number of CpG sites is tested. Bisulfite modification of

DNA followed by sequencing is the gold standard with highest resolution to identify single methylated C nucleotides. Direct sequencing or cloning of PCR products followed by sequencing individual clones can be done (see Figure 37). Direct sequencing is faster, but cloning can resolve issues of heterogeneity of methylation on separate alleles. It was decided to use cloning in these experiments.

After bisulfite modification of DNA from the colon cancer cell lines, a 200bp stretch of the putative promoter region containing 22 CpG sites (see Figure 38), was PCR amplified and subcloned into pGEMT for sequencing. For each cancer cell line, 10 PCR clones were sequenced. The four cell lines that did not express *SMO* were all fully methylated in the amplified region (see Figures 39 and 40). The cell lines HCT116 and LOVO appeared to have one methylated and one un-methylated allele. In the case of LOVO, 50% of sequenced clones were totally unmethylated CpG (suggestive of one unmethylated allele)(see Figure 39), and 50% were completely or partially methylated CpG (suggestive of a second variably methylated allele)(see Figure 42).

Since LOVO does express *SMO* as seen on the Northern blot, one might predict that the unmethylated allele is responsible for the mRNA that is produced. Alternatively, the methylation seen on the second allele may not be extensive enough or may not affect the key CpG sites to inhibit expression, and so the gene continues to be expressed from both alleles. One cannot be certain of either situation without knowledge of independent sequence information for the two genomic alleles. Such information however was gathered in subsequent experiments (see next section). Not

only did the data imply that one allele was largely silent (predicted to be the partially methylated one), but the particular distribution of methylated CpG sites implicates what sites are the key players, and perhaps also the affected transcription factors, that are involved in silencing of the gene. In the case of HCT116, two patterns were seen (allowing for sampling error in the number of representative clones picked for sequencing), one fully methylated one fully unmethylated, again implying opposing states of the two alleles. Contrary to LOVO, one predicts for HCT116 that the mRNA seen by Northern hybridisation certainly comes only from the unmethylated allele.

The cell lines SW480 and COLO320 showed strange results. SW480 showed very little expression detectable by Northern blot hybridisation, but was apparently almost entirely unmethylated. The interpretation is that although *SMO* of SW480 is not methylated, the gene is kept almost silent by the activity of inhibitory transcriptional regulators. Silencing of genes in cancers by other mechanisms beside hypermethylation has been seen in several cases. For example, heterogeneous methylation of some genes has been reported for genes like *Rb* in retinoblastoma tumours (Stirzaker *et al.,* 1997), *p15* in leukaemia (Melki *et al.,* 1999), where regardless of the extent of methylation, the gene appears to be silenced in the tumour. Similarly, *BRCA1* expression is low or absent in most ductal carcinomas and cell lines (Thompson *et al.,* 1995), while methylation occurs in only about 20% of patients and much less frequently in cell lines (Dobrovic and Simpfendorfer 1997; Bianco *et al.,* 2000; Catteau *et al.,* 1999; Esteller *et al.,* 2000; Rice *et al.,* 1998).

Likewise, some leukaemias have little or no methylation of the *E-cadherin* promoter yet lack expression of that gene (Melki *et al.*, 2000).

In contrast, COLO320 showed comparatively high expression of *SMO* by Northern blot hybridisation yet had total methylation across the region screened in 9 out of 10 clones sequenced. The almost total methylation could simply be due to incomplete bisulfite modification of DNA before PCR amplification, but this is not likely because a) the experiment was repeated and produced the same results, and b) the PCR primers do not work on unmodified DNA as they span non-CpG cytosines that will always be changed following bisulfite treatment thus inhibiting primer annealing (data not shown). Allowing for sampling error of clones picked for sequencing, the interpretation of results is that one allele of COLO320 is methylated (and presumably silenced) but the other allele remains unmethylated from which a high level of expression is generated. The lung cancer cell line COR-L51 appeared fully methylated, however the immortalised non-tumourigenic human bronchial epithelial cell BW1799 was completely unmethylated. These results are consistent with the observations of expression by cDNA PCR.

4. *SMO* mutation occurred in the expressed allele of hemi-methylated cell lines.

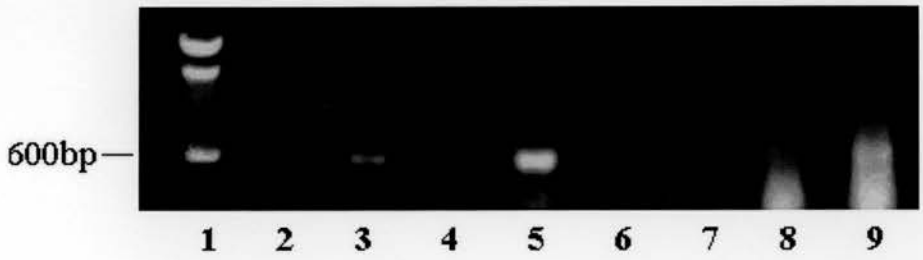
A heterozygous base substitution in exon 6 of *SMO* in HCT116 (G to A) was identified during other work of the laboratory (unpublished data) and that was not seen in 40 independent normal blood samples. This substitution was thus considered to be a candidate mutation (in the absence of normal tissue to confirm somatic aetiology). This base change was used to study the expressed message of *SMO* in

HCT116 to identify the allelic expression of the gene in the light of the hemi-methylated data. Primers SMO21 and SMO52 were used to amplify cDNA from this cell line and others normal for this exon and run on an SSCP gel. The results showed a bandshift for HCT116 without the normal pattern suggesting that only the mutant was expressed (see Figure 43).

Similarly, it was interesting to clarify the expression of alleles in LOVO (unpublished group data). Again, a heterozygous candidate mutation was found in exon 1 for this cell line: an insertion of two CTG trinucleotides in a stretch of seven CTGs, presumably due to the microsatellite instability of LOVO. Primers SMO1 and SMO2 were used to amplify cDNA from this cell line as for the exon 6 change in HCT116 and run on an SSCP gel. The results showed again a bandshift for LOVO, with a very low intensity of the normal pattern, suggesting that the mutant allele was the one expressed predominantly.

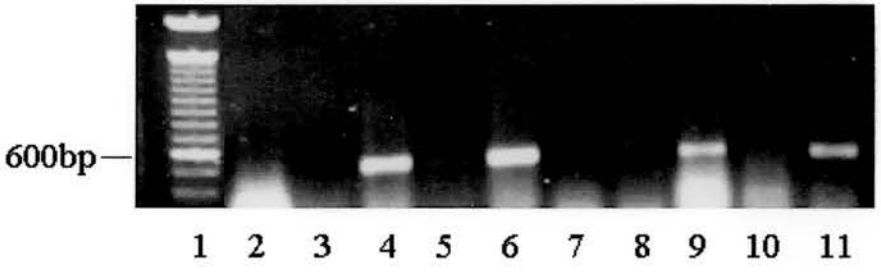
To confirm these findings, cDNA from HCT116 and LOVO was amplified again and the products were cloned and sequenced. 10/10 clones in HCT116 had the mutant allele. 2/8 clones in LOVO (unpublished group data)had the wild-type allele (coding for 7 leucines in the signal peptide), while the remaining 6 clones had 9 leucines, indicative of the mutant allele. These results suggested that *SMO* is preferentially expressed from the unmethylated mutant HCT116 allele, and that LOVO has one allele that is similarly unmethylated allowing normal expression carrying a mutation, while the other allele that is wild type but partially methylated is almost completely suppressed.

Figure 33 The RT-PCR analysis of *SMO* in colon cancer cell lines



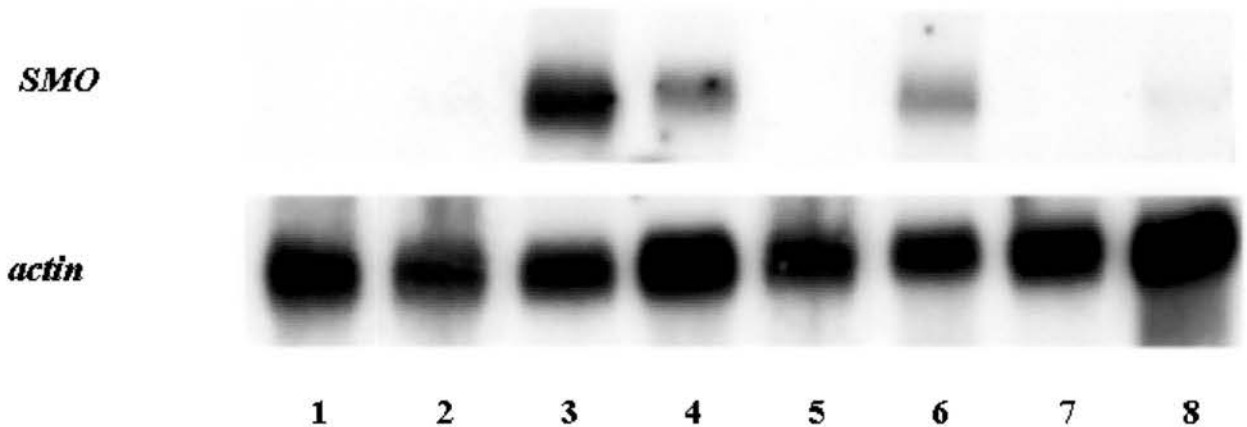
Lanes are: 1. DNA marker, 2. HT29, 3. SW480, 4. COLO320, 5. HCT116, 6. DLD1, 7. LOVO, 8. LS180 and 9. normal adult colon.

Figure 34 The RT-PCR of *SMO* in lung cancer cell lines



Lanes are: 1. DNA marker, 2. COR-L47, 3. COR-L51, 4. COR-L24, 5. COR-L279, 6. H1648, 7. H358, 8. H2122, 9. COR-L105, 10. COR-L23 and 11. BW1799

Figure 35 *SMO* Northern hybridisation analysis



Lanes are: 1. HT29, 2. SW480, 3. COLO320, 4. HCT116, 5. DLD1, 6. LOVO, 7. LS180 polyA+ RNA and 8. adult colon total RNA.

Figure 36 *SMO* CpG island CpGPlot analysis

The input sequence is NT007659 (nt1-1685).

Window size = 100

Observed/Expected ratio > 0.60

Percent C + Percent G > 50.00

Length 1068 (154-1221), Length > 200

Figure 36 ***SMO*** **CpG island CpGPlot analysis**

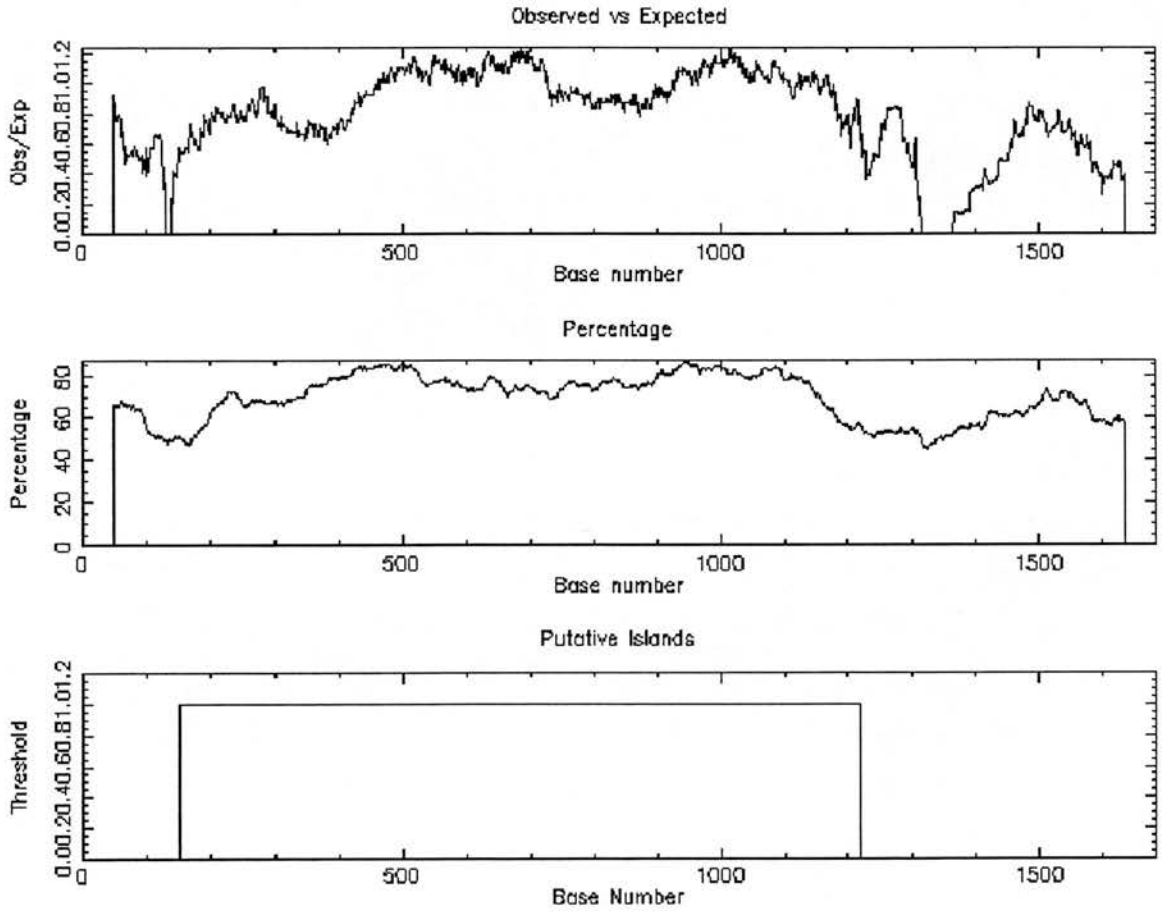
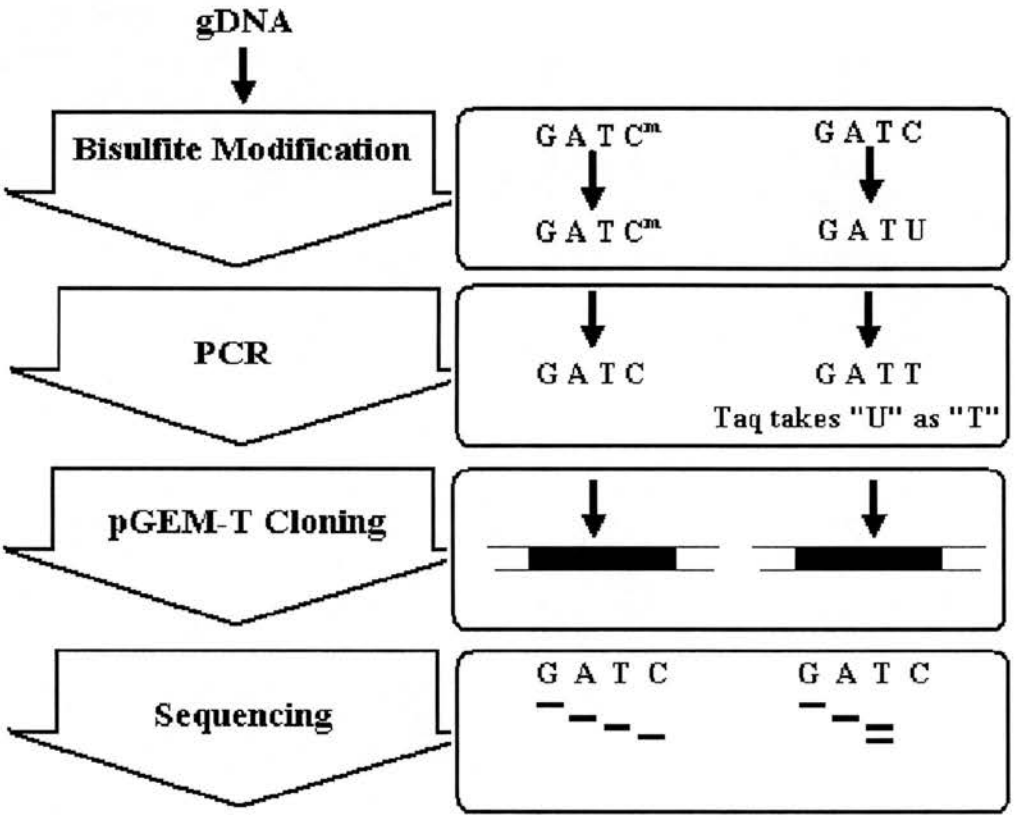


Figure 37 Bisulfite sequencing procedure



After bisulfite modification, all of the unmethylated C are changed to U. When being PCR amplified, those Us was taken as T by Taq. Thus, after being sequenced, unmethylated Cs are replaced by Ts while methylated Cs remain C.

Figure 38 *SMO* CpG sites in the bisulfite sequenced region

AAAAGTCCGCTTCCTGCA***GACGACCTCAGACCAAGCAAGGTGCC***CGCCGA
Primer SMObis1
GTCTC TCCTTGCAGGTCCGGCCCACGATTTCCACTCATCTCTTTCCCCCGG

GCGCGGGGCGGCGGCAGGCGGGGTCACCAGATCCCCCTAGCCCGGGCCC

CTCCAGGCGCCAGGGACGCTGACGCTCGCGCTCTTCCTCTCTCGCCTCCC

CTCCCCACCTCTCCGCTCCTTCGTCCAGTCCCTCCCCCAGCCTCGGCGCAG

GGGGGCCGGGCTTGGCTCCGCGAGG***CCCGTGCATTCCAGAGAGCCCAG***C
Primer SMObis3
GAGCTAGAGCAACAAAGGAGC**CG**GGT**CG**C**CGGCG**GGGAGAGTT**CG**GG

GGGCTG**CGGCGCG**CTGGGG**CG**AAGGTGGCTGCTGGGC**CGCG**GGCTGG**C**

GCGGGGG**CG**GAGC**CG**GAGCTGCACT**CG**CACCCC**CG**GCC**CGCG**TCTGGC

CTCCCT**CGCG**GGG**CG**GGG***AGGTGGCTTTAATGGTGGGAGAGGGA***ATGGG
Primer SMObis4
GCTGGGGATTGGGGGCCCAGGGGTCTCCTAGGGCTGAAGACAACTTGGA

TTGCGAGGCTAGGGCTTGGGGAGTCGTGCATCCCGTTCCGGGCCTCCGCA

GCCCAACATGGGCCCCGGGTTCCAAAGTTTGCGAAGTTGGGCGCCGAGG

GGCCGGGGCGCGCGGAGCGTCCGGGGGGGCCCGGGCCCGGATTCTCTGG

GCGCACAGGTCGCCTGAGCCGCCTCCGCGGCCGCCGAGGTCGTGCGTGT

GGCCGGGGGGCTCCGAGGAGCAGGCGGGGGCGCCGGGGCTTTTGCTGAG

TTGGCGGGGTTGGCC**ATG**GCCGCTGCCCGCCCAGCGCGGGGGCCGGAGC

TCCCGCTCCTGGGGCTGCTGCTGCTGCTGCT

There are 23 CpG sites in the bisulfite sequenced region by SMObis3 and SMObis4. CpG sites and translation start site ATG are in bold. Primer sequences are in bold and underlined.

Figure 39 Methylation map of individual clone (1)

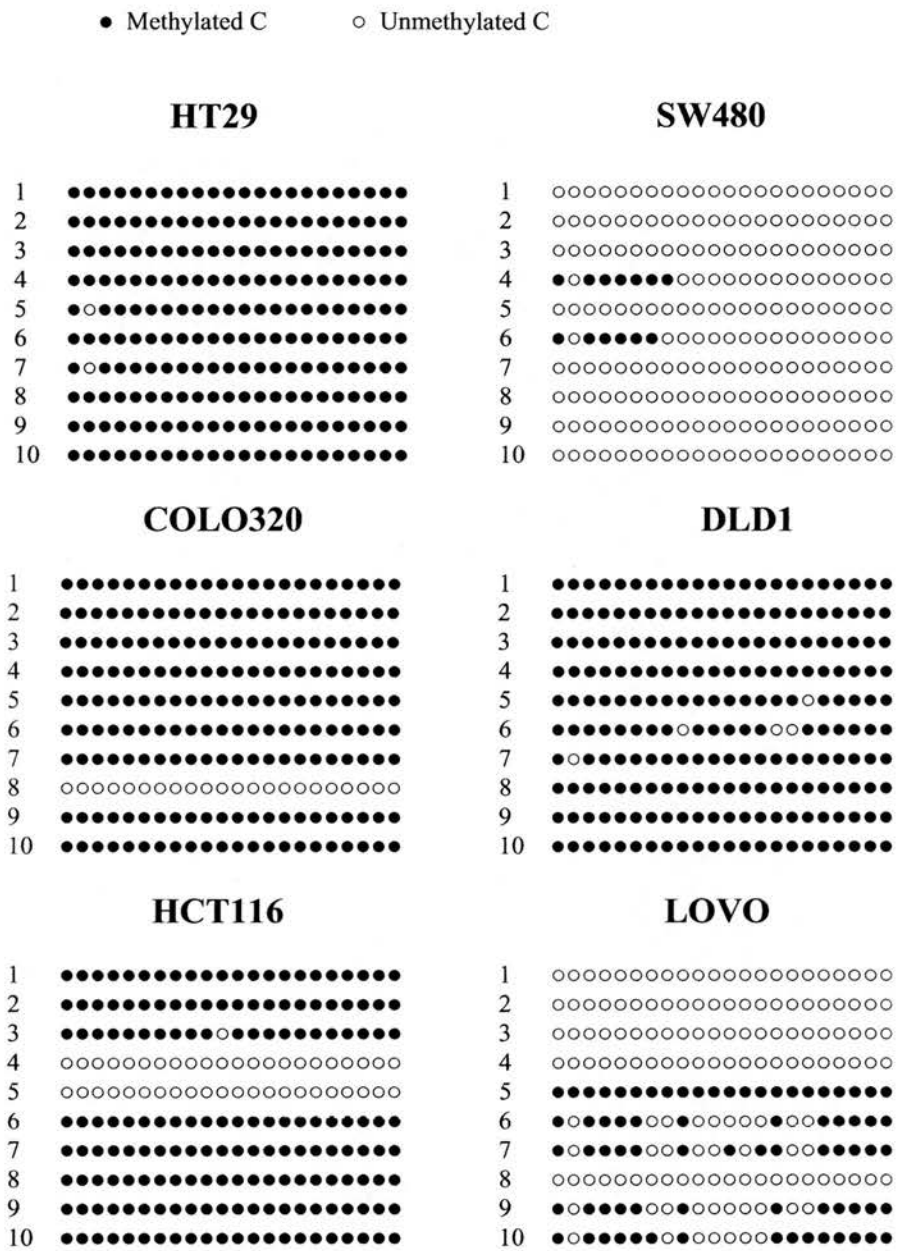
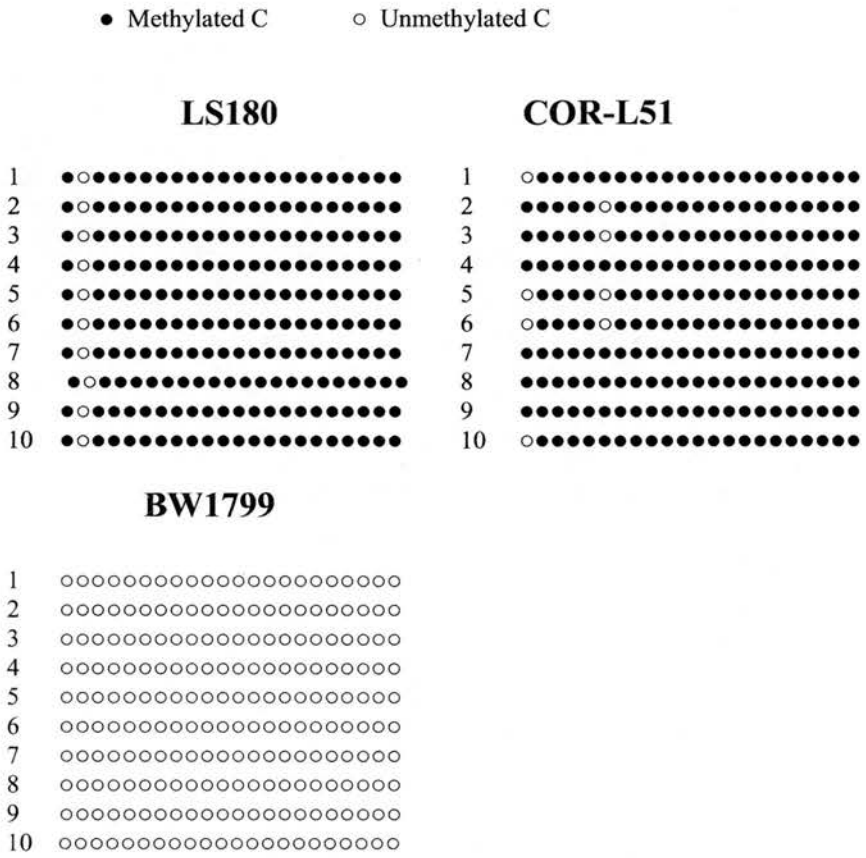


Figure 39 Methylation map of individual clone (2)



For every cell line, PCR product after bisulfite modification was cloned. Ten clones were subjected to sequencing. Twenty-two CpG sites were examined to observe their methylation status.

Figure 40 **Fully methylated sequence in the *SMO* CpG island of DLD1**

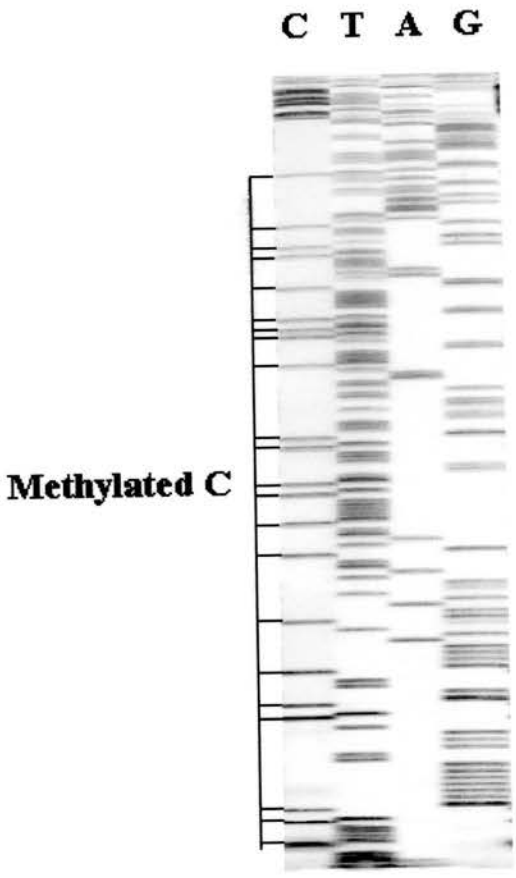


Figure 41 No methylated C found in the *SMO* CpG island of some SW480 cells.

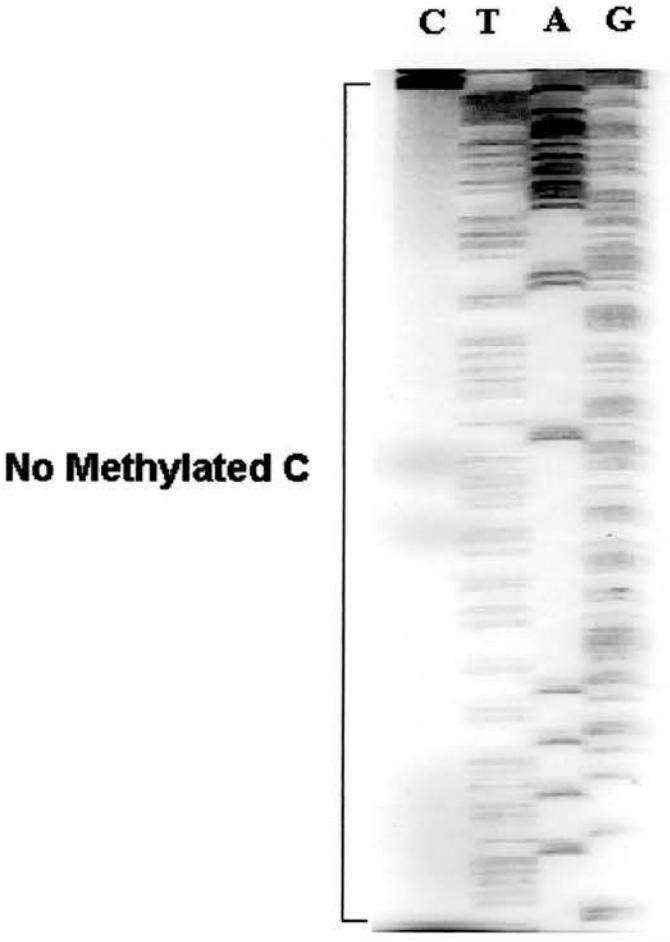


Figure 42 Partially methylated sequence in the promoter of *SMO* in some SW480 cells

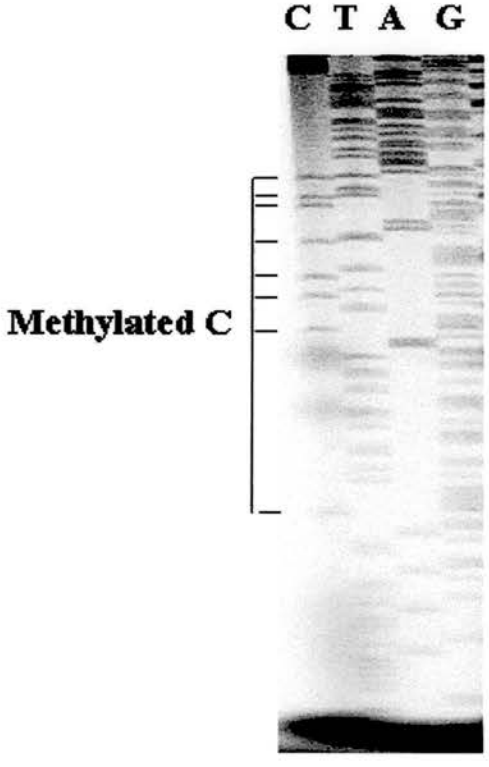
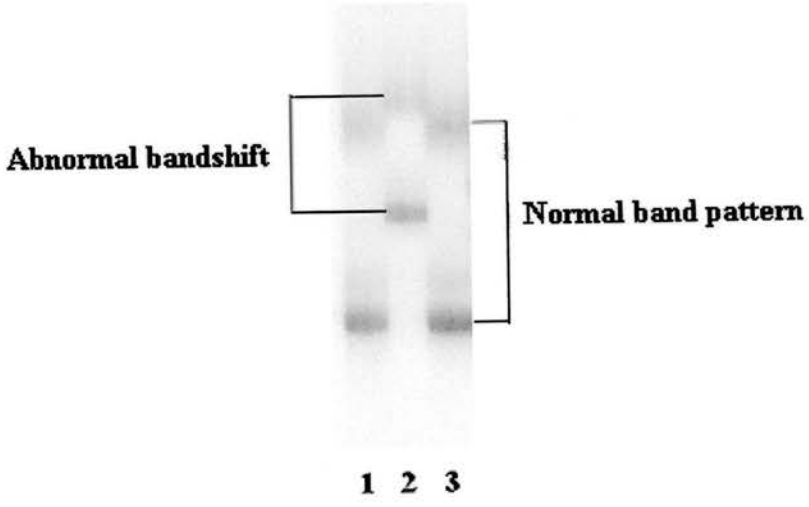


Figure 43 *SMO* SSCP analysis of HCT116 cDNA between exon5 and 6

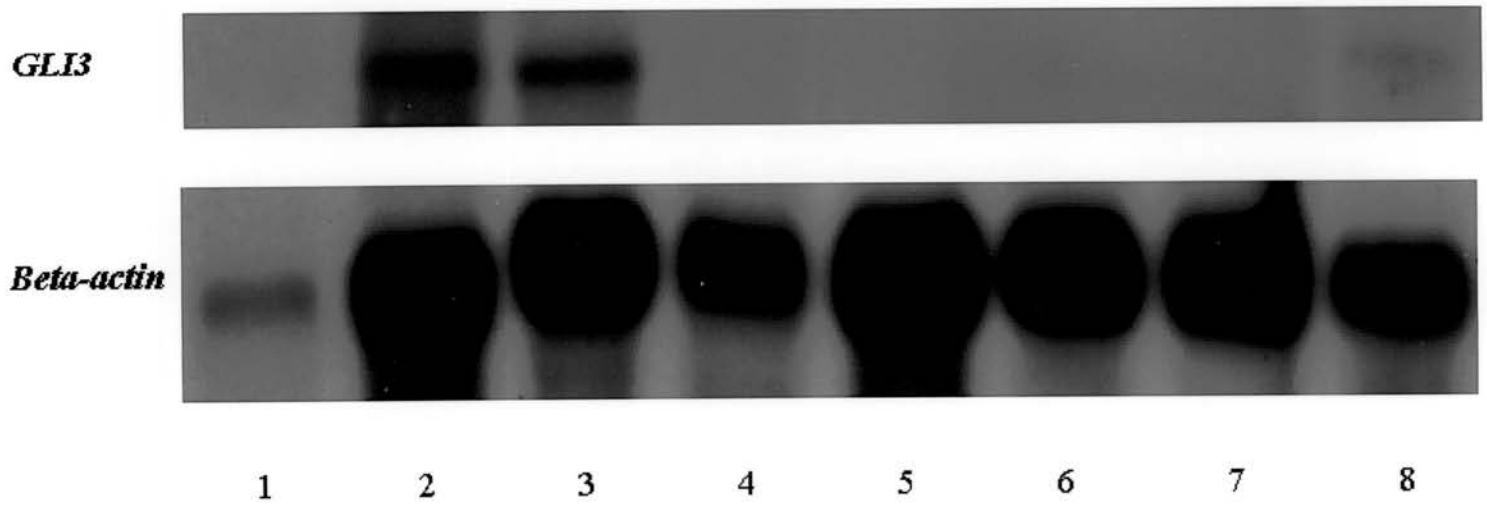


Lanes are: 1. SW480, 2. HCT116 and 3. LOVO

5. *SMO* and *GLI3* show co-ordinated expression in colon cancer cell lines.

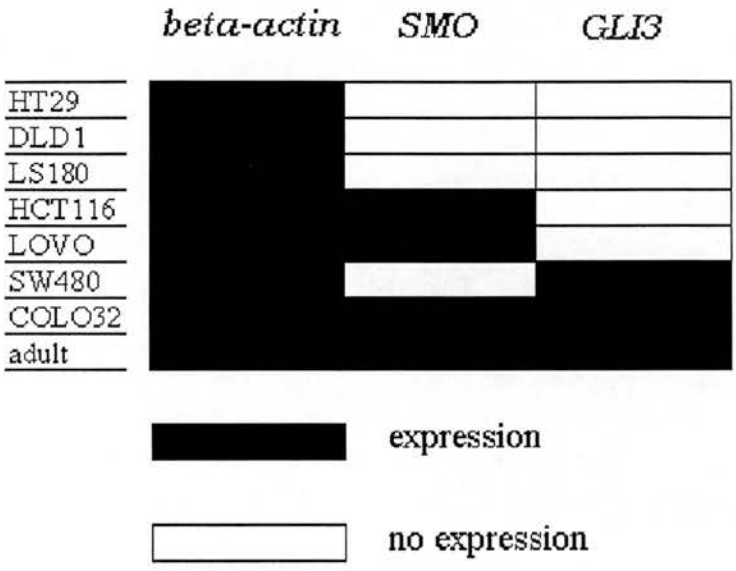
As part of a study of the expression of many of the genes in the canonical Hh pathway, colon cancer cell lines were analysed by Northern hybridisation analysis for expression of *GLI3* (see Figure 44). A 615bp PCR product amplified by primers (Gli5 and Gli6 see Table 4) in exons 12 and 14 was used as the probe. Expression of this gene was not found to be uniform for the set of cell lines used. Taking the results of *SMO* and *GLI3* expression analyses together (see Figure 45), it was found that cell lines HT29, DLD1 and LS180 lacked expression of both *SMO* and *GLI3*. HCT116 and LOVO expressed *SMO* although this was mutated in each case, but not *GLI3*. Expression of both *SMO* and *GLI3* in normal adult colon was low but this RNA was derived from whole tissue, a mix of cell types that may include some/many that do not express these genes at all thus reducing the relative signal strength. Uniquely, expression of *SMO* in cell line SW480 was very low, although *GLI3* expression was high.

Figure 44 *GLI3* Northern hybridisation analysis



Lanes are: 1. HT29, 2. SW480, 3. COLO320, 4. HCT116, 5. DLD1, 6. LOVO, 7. LS180 polyA+ RNA and 8. normal adult colon total RNA

Figure 45 Diagram of the expression status of *SMO* and *GLI3* in the cell lines



6. Discussion

Screening of *SMO* in colon cancer cell lines has revealed a low frequency of mutations (unpublished group data). However, analysis of expression in these cell lines and in lung cancer cell lines as described here, has found that downregulation of *SMO* is a common phenomenon, and that in the two colon cell lines with heterozygous mutations, it is the mutant allele that is solely or predominantly expressed. The mechanism of silencing of the normal gene (whether mono- or bi-allelic) appears to be via epigenetic methylation of the promoter region. These data are the first demonstration of loss of *SMO* expression associated with tumours.

A simple model generally accepted is that PTC inhibits SMO by a direct, stable, physical interaction, and Hh binding to PTC causes a conformational change in this PTC-SMO receptor complex which releases SMO from inhibition. Whereas this pathway normally is regulated by the spatially and temporally restricted expression of Shh, loss of function *PTC* mutations are associated with human cancers, particularly BCCs (Wicking *et al.,* 1997; Gillies S *et al.,* 1997; Gailani MR *et al.,* 1996). In addition, activating mutations in *SMO*, making SMO protein resistant to the inhibition of PTC, have been identified in BCC tumours that do not contain detectable alterations in *PTC* (Lam *et al.,* 1999; Reifenberger *et al.,* 1998). But the model of PTC-SMO receptor complex has been facing increasing arguments as a number of studies are presenting different explanations for the mechanism of the hedgehog pathway. For example, some studies show that SMO protein is not present in a complex bound to PTC and suggest that in the absence of Hh, PTC indirectly modifies SMO protein to generate an inactive moiety, in a non-stochiometric

reaction (reviewed in Kalderon, 2000). Also SMO does not appear to play its role in tumourigenesis as it was described in BCCs before. One group reported that *PTC* and *SMO* mRNA levels have an inverse correlation with histological malignancy in astrocytic tumours and suggest that these gene products are implicated in the suppression of astrocytic tumours (Katayam *et al.,* 2002). Another study showed that the expression of an activated *SMO* mutant in keratinocytes is not sufficient for the development and maintenance of BCCs (Grachtchouk *et al.,* 2003).

The results reported here suggested that the down regulation of *SMO* is a common phenomenon in colon and lung cancer cell lines and that *SMO* may also be able to act as a tumour suppressor gene. Consistent with this idea is the observation that in cell lines HCT116 and LOVO, the one allele that remains ummethylated and is expressed is mutant while the wild-type allele is methylated and totally or almost totally silenced. The dual role of *SMO* as an oncogene and as a tumour suppressor gene has precedent. For example, *p53* suffers mutations and deletions in many tumours, but some these mutations lead to a loss of function typical for a tumour suppressor gene while others lead to a gain of function according to an oncogene (Zalcenstein *et al.,* 2003). The data reported here is however the first example of phenomena specifically relating to *SMO* as a tumour suppressor gene, and in a tissue type different from earlier associations with the gene.

In recent months reports have been published of an activation of the Hh pathway in SCLC cell lines and those of upper parts of the digestive tract (Watkins *et al.,* 2003; Berman *et al.,* 2003). In each of these reports, activity of the pathway was indicated

by expression of the ligand Shh and its downstream gene transcription activator GLI1, and the implication was that tumourigenesis was promoted due to the stimulatory activity of the pathway initiated by *SHH* expression. This was supported by data showing that the SMO-antagonist cyclopamine inhibited growth of several SCLC, oesophageal, gastric, biliary and pancreatic cell lines *in vitro* and in vivo.

Colon cancer cell lines were presented as exceptions to this trend however, especially in the paper by Berman *et al.* There they showed by cDNA PCR that although the cell lines expressed Hh ligands, many did not express *GLI1* and none expressed *PTC*. Curiously, data presented here (Chapter 5) shows exactly the opposite regarding *PTC*. Furthermore, cDNA PCR data from our group also shows expression of *GLI1* (data not shown), and for both genes several of the cell lines tested were used by Berman *et al.* (DLD1, HCT116, HT29, LOVO). A simple explanation for the difference in results would be merely technical problems of the other group since the positive results here were reproducible and appropriate negative PCR blank controls were used to rule out contamination artifacts. Further equivocal results were also observed when the supplementary table from Watkins *et al.* was compared with the data from Berman *et al.* The cDNA PCR data (the latter paper) frequently showed positive results for *SHH* and *GLI1* that were scored as negative cell lines in the former paper where the assay done was Western analysis. It could be explained that the two assays may apparently conflict in this way if the more sensitive cDNA PCR detects gene expression that the Western cannot show. Yet Watkins *et al.* were able adequately to detect the proteins in other cell lines. Most interesting was their observation that cyclopamine could not inhibit growth or reporter gene activity in HCT116 (Watkins and Berman papers) and DLD1 (Berman) colon cancer cell lines.

It was shown here by cDNA PCR and Northern analyses that DLD1 is refractory to cyclopamine because it does not express *SMO* whose protein product can therefore not be inhibited. In the case of HCT116, although it expresses *SMO* it was found to carry a mis-sense mutation of a conserved residue of mouse and human SMO and related WNT receptor Frizzled. Sequence and single stranded conformation polymorphism analyses of cDNA from HCT116 also indicated that the wild type allele was not expressed implying that only protein carrying the putative mutation was present. Since Berman *et al.* showed that HCT116, like DLD1, was refractile to cyclopamine this mutation may alter SMO conformation enough to escape inhibition by this drug and may also cause the expressed SMO protein to lose normal function. Thus, colon is similar to other parts of the digestive tract as shown by Berman *et al.*, in that several genes along the Hh pathway are expressed suggesting the requirement for Hh ligand stimulation for tumour growth. However, as it is described above a significant proportion of tumour cell lines from this tissue do not express wild type *SMO*. Thus, despite expression of Hh ligand, the Hh pathway may actually be dysfunctional or underactive during tumourigenesis of colon due to absence of SMO, akin to tumour suppressor gene function for this gene. Alternatively, a combination of Hh ligand expression coupled with loss of SMO may drive the unregulated Hh signal down another pathway whose effects are tumourigenic. The fact that five of the lung cancer cell lines also lacked *SMO* expression also raises the possibility that similar changes may be occurring in lung cancer, and that the hypothesis of canonical Hh pathway activation in relationship to lung tumourigenesis may need to be re-examined.

Cancer cell lines are commonly used in cancer research, including studies designed to assess methylation abnormality. Studies have revealed that methylation is very common in colon cancer cell lines (Suter *et al.,* 2003) a phenomenon whose significance remains somewhat controversial. On the one hand, it was reported that most of the methylation of genes seen in cancer cell lines is present also in the primary carcinomas from which they are derived (Ueki *et al.,* 2001). On the other hand, methylation at most loci was observed at a significantly greater frequency than that expected from *in vivo* studies (Hawkins *et al.,* 2002). The reason for the excessive levels of methylation in colon cancer cell lines as opposed to their primary counterparts is unclear, but it suggests that colon cancer cell lines may be only representative of a small subset of real tumours, and this should be taken into account in the use of colon cancer cell lines for epigenetic studies. Thus it suggests that further methylation studies of primary colon and lung tumours is needed to comprehensively assess methylation status *in vivo*. Of course, such experiments will have attendant problems of mixed tumour (methylated) and normal (unmethylated) cells that would have to be overcome or borne in mind when analysing data.

Genes silenced by DNA methylation can be reactivated by treatment with 5-aza-2-deoxycytidine (5-Aza-dC), which is a well-established inhibitor of DNA methyltransferase (Jones and Taylor, 1980). Inactive chromatin mediated by histone deacetylation by HDAC is also known to be involved in gene silencing, affecting for example *p21Waf1*, *hTERT*, and *hLHR* (Sowa *et al.,* 1997; Takakura *et al.,* 2001; Zhang and Dufau, 2002). Trichostatin A (TSA), a histone deacetylase inhibitor, is reported to activate genes whose expression is silenced by such HDAC activity (Van

Lint *et al.,* 1996). Furthermore studies show that 5-Aza-dC and TSA can synergistically reactivate methylation-associated silenced genes (Cameron *et al.,* 1999). Thus, 5-Aza-dC and/or TSA should be used with the cell lines shown here to be silent and methylated for *SMO* to confirm the methylation effect in further experiments.

The other very interesting finding in the described experiments is that the lack of expression of normal *SMO* (either with or without expression of mutant *SMO*) is accompanied by absence of *GLI3*, independent of transcription of *SHH*, *PTC*, *GLI1* and *GLI2*. These results suggest that expression of wild type *SMO* is required for expression of *GLI3* in the first instance but also that regulation of *GLI3* by *SMO* can occur independent of intervening genes of the pathway according to current dogma.

The GLI proteins are large transcription factors of >1000 amino acids that bind DNA in a sequence-specific manner (Kinzler and Vogelstein, 1990; Vortkamp *et al.,* 1995), via the last three fingers of their five zinc-finger domain (Pavletich *et al.,* 1993). At least three forms of Ci have been shown or predicted to exist: a cytoplasmic protein, a nuclear repressor and a nuclear activator (Aza-Blanc *et al.,* 1997; Ohlmeyer and Kalderon.1998; Mèthot and Basler1999). C-terminally truncated forms existing in the nucleus have dominant negative activity over that of full length proteins and, additionally, C-terminally truncated GLI3 can act as a transcriptional repressor (Dai *et al.,* 1999; Ruiz 1999; Shin *et al.,* 1999). In the absence of SHH, GLI3 is proteolytically cleaved to the repressor (Wang *et al.,* 2000) that then inhibits SHH (Masuya *et al.,* 1995) and GLI dependent transcription (Wang

et al., 2000), reinforcing the latent state of the signalling pathway when SHH is not present. Consistent with the observation in this report of regulation of GLI3 outwith the canonical SHH pathway is the observation that during vertebrate limb bud development, GLI3 restricts expression of the transcription factor dHAND before SHH signalling is activated (Welscher *et al.,* 2002), indicating that GLI3 can function independent of Hh signalling, as well as being a component of the pathway. Interestingly, during somite formation in the chick embryo, GLI3 is subject to β-catenin mediated regulation by Wnt (Borycki *et al.,* 2000). The Wnț/β-catenin/APC signalling pathway is dysregulated in the majority of colon cancers (reviewed in Giles *et al.,* 2003), supporting the idea that dysregulation of the Hh pathway (in this case by silencing of SMO-GLI3) may play a role in such tumours. Hh signalling is a well documented initiator of proliferation in many vertebrate systems, including the epithelium of the mouse gut (Ramalho-Santos *et al.,* 2000). According to current hypotheses, aberrant expression of *SHH*, loss of function of *PTC*, or oncogenic mutation of *SMO* leading to hyperproliferation (through loss of GLI3 repression of transcription) may have serious implications for tumour progression in the gut. The data presented here for loss of expression of *SMO* and *GLI3* suggest further currently unknown functions of these genes that act in a contradictory manner.

There was one exception to the stated trend connecting *SMO* to *GLI3* in the cell line panel used in these experiments, that of SW480. In this case, *SMO* is expressed at very low levels only detectable by PCR while *GLI3* is expressed at significant levels, suggesting some other mechanism of activation of *GLI3*. If, according to the new hypothesis proposed here that loss of function of *GLI3*, caused by loss of activity of *SMO*, could play a role in tumourigenesis as for a tumour suppressor gene, then it

would be of interest to ascertain the status of *GLI3*-expressing cell lines. Perhaps SW480 and COLO320 have mutations in this gene making it dysfunctional.

Chapter 7

Summary and Further Work

1) *MBD3*. Few changes were found in a screen for mutations. Expression of the gene remained significant and there was no complete methylation of the putative promoter region. It was concluded that this gene plays little or no role in colon or lung carcinogenesis.

2) *MBD2*. Expression of the gene remained significant and there was no complete methylation of the putative promoter region. The data were consistent with one of the current hypotheses regarding *MBD2* that states that continued expression of the gene is required for tumourigenesis. An attempt was made to find a way to screen exon 1 for mutations as this was not done in an earlier study and the exon encodes half of the methyl-CpG binding domain. However, a suitable, reproducible method was not identified during the time available, and screening of this exon remains to be done.

3) *PTC*. The CGG repeat seen in exon 1 of this gene was found to represent a naturally occurring repeat length polymorphism that was stable even in microsatellite unstable colon tumours. Similarly, expression of the gene appeared normal in the cancer cell lines tested, with no complete methylation of promoter sequences. These data show no evidence that silencing of this gene plays a significant role in colon or lung tumourigenesis.

4) *SMO/GLI3*. Expression of *SMO* was found to be absent, significantly reduced or from an allele carrying a putative mutation, in a large proportion of colon and lung cancer cell lines. Silencing of wild type alleles was accompanied by

methylation of putative promoter sequences. Such changes suggested a tumour suppressor role of *SMO* in these cancers and were the first such observations for this gene. In order to rule out the possibility that these changes were simply an artifact of transformation of cells in culture, primary tumour samples need to be assayed. Furthermore, the role of methylation in the silencing of *SMO* should be tested by treatment of appropriate cell lines with the methylation-reversing drug 5azaC and/or the histone deacetylase inhibitor trichostatin A.

Absence of expression of *GLI3* was also observed in association with loss of *SMO* or expression of only *SMO* message containing a putative mutation, independent of the other genes of the Hh pathway tested in this study. There are two main implications of these results. Firstly, that *GLI3* is regulated by *SMO* in a manner outside the normal sequence of steps currently thought to comprise the Hh pathway. Loss of *GLI3* expression may be due either to absence of *SMO*-related activation or may also be mediated by tumour-associated methylation and silencing of its promoter. The promoter sequence of *GLI3* should therefore be tested for methylation, and re-expression could be tested with 5azaC and/or TSA. Experiments should also be done to try to demonstrate re-expression of *GLI3* in association with re-expression of *SMO*, for example by transfection of a *SMO* cDNA expression construct. Secondly that *GLI3* may also have a role in tumourigenesis as the effector gene regulated by *SMO* and/or as a tumour suppressor gene function in its own right. If the latter is true, one might expect the cases where *SMO* is absent or

present (LOVO or COLO320) but *GLI3* still expressed to have a mutation in *GLI3*. Therefore it would be interesting to screen those cell lines for mutations, as well as a set of primary tumours.

The possible activity of *SMO* and *Gli3* as a tumour suppressor can be examined by a tumourigenesis assay in nude mice. For example, HT29 and exogenous *SMO* transfected HT29 can be transplanted into nude mice to assay tumour growth rate. Also *Smo* and *Gli3* targeted disruption mice can be made to study their roles in hyperplasia and increased epithelial proliferation.

Chapter 8

Reference

Aaltonen LA *et al.* (1993). Clues to the pathogenesis of familial colorectal cancer. *Science* **260,**812–816.

Amano K, Nomura Y, Segawa M and Yamakawa K. (2000). Mutational analysis of the *MECP2* gene in Japanese patients with Rett syndrome. *Hum. Genet.* **45,**231–236.

Amir RE *et al.* (1999). Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat. Genet.* **23(2),** 185-188.

Avizienyte E *et al.* (1998). Somatic mutations in *LKB1* are rare in sporadic colorectal and testicular tumours. *Cancer Res.* **58,**2087-2090.

Aza-Blanc P *et al.* (1997). Proteolysis that is inhibited by Hedgehog targets Cubitus interruptus protein to the nucleus and converts it to a repressor. *Cell* **89,** 1043–1053.

Bader S *et al.* (1999). Somatic frameshift mutations in the MBD4 gene of sporadic colon cancers with mismatch repair deficiency. *Oncogene* **8(56),** 8044-8047.

Bader S *et al.* (2003). *MBD1*, *MBD2* and *CGBP* genes at chromosome 18q21 are infrequently mutated in human colon and lung cancers. *Oncogene* **22(22),**3506-3510.

Baeg GH *et al.* (1995). The tumour suppressor gene product APC blocks cell cycle progression from G0/G1 to S phase. *EMBO J.* **14,**5618–5625.

Baker SJ *et al.* (1989). Chromosome 17 deletions and p53 gene mutations in colorectal carcinomas *Science* **244,** 217-221.

Baker SJ, Markowitz S, Fearon ER, Willson JK and Vogelstein B. (1990). Suppression of human colorectal carcinoma cell growth by wild- type p53. *Science* **249,** 912-915.

Baker SJ *et al.* (1990). *p53* gene mutations occur in combination with 17p allelic deletions as late events in colorectal tumorigenesis. *Cancer Res.* **50,** 7717-7722.

Baylin SB. (1998). Alterations in DNA methylation: A fundamental aspect of neoplasia. In: G. Klein and G.F. Van de Woude Editors, *Advances in Cancer Research* Academic Press, pp. 141–196.

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

Behrens J *et al.* (1998). Functional interaction of an axin homolog, conductin, with beta-catenin, APC, and GSK3beta. *Science* 280, 596–599.

Berman DM *et al.* (2003). Widespread requirement for Hedgehog ligand stimulation in growth of digestive tract tumours. *Nature* **425(6960),** 846-851.

Beroud C and Soussi T. (1996). *APC* gene: database of germline and somatic mutations in human tumors and cell lines. *Nucleic Acids Res.* **24,** 121–124.

Bestor T, Laudano A, Mattaliano R and Ingram V. (1988). Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases. *J. Mol. Biol.***203 (4),** 971-83.

Bhattacharya SK, Ramchandani S, Cervoni N and Szyf M. (1999). A mammalian protein with specific demethylase activity for mCpG DNA. *Nature* **397,** 579–583.

Bienvenu T et al. (2000). *MECP2* mutations account for most cases of typical forms of Rett syndrome. *Hum. Mol. Genet.* **9,**1377–1384.

Billard LM, Magdinier F, Lenoir GM, Frappart L and Dante R. (2002). *MeCP2* and *MBD2* expression during normal and pathological growth of the human mammary gland. *Oncogene* 21(17), 2704-2712.

Bird AP. CpG-rich islands and the function of DNA methylation.(1986). *Nature* **321(6067),** 209-13.

Bird A, Taggart M, Frommer M, Miller O J and Macleod D. (1985). A fraction of the mouse genome that is derived from islands of nonmethylated CpG-rich DNA. *Cell* **10,**91-99

Bisgaard ML, Fenger K, Bulow S, Niebuhr E and Mohr J. (1994). Familial adenomatous polyposis (FAP): frequency, penetrance, and mutation rate. *Hum. Mutat.* **3,** 121–125.

Bodmer WF *et al.* (1987). Localization of the gene for familial adenomatous polyposis on chromosome 5. *Nature* **328,** 614–616.

Boeke J, Ammerpohl O, Kegel S, Moehren U and Renkawitz R. (2000). The minimal repression domain of MBD2b overlaps with the methyl-CpG-binding domain and binds directly to Sin3A. *J. Biol. Chem.* **275,** 34963-34967

Borycki A, Brown AM and Emerson CP Jr. (2000). Shh and Wnt signaling pathways converge to control Gli gene activation in avian somites. *Development* **127(10),**2075-2087.

Browne SJ, Williams AC, Hague A, Butt AJ and Paraskeva C. (1994). Loss of APC protein expressed by human colonic epithelial cells and the appearance of a specific low molecular weight form is associated with apoptosis in vitro. *Int. J. Cancer.***59,** 56–64.

Broca PP. (1866). *Traite des Tumerus.*

Burchill SA. (1994). The tumour suppressor APC product is associated

with cell adhesion. *Bioessays* **16,**225–227.

Cagle PT, el-Naggar AK and Xu HJ. (1997). Differential retinoblastoma protein expression in neuroendocrine tumors of the lung. Potential diagnostic implications. *Am. J. Pathol.* **150,** 393 –400.

Cameron EE, Bachman KE, Myohanen S, Herman JG and Baylin SB. (1999). Synergy of demethylation and histone deacetylase inhibition in the re-expression of genes silenced in cancer. *Nat. Genet.* **21(1),** 103-107.

Cheadle JP *et al.* (2000). Long-read sequence analysis of the *MECP2* gene in Rett syndrome patients: correlation of disease severity with mutation type and location. *Hum. Mol. Genet.***9,** 1119–1129.

Chen RZ, Pettersson U, Beard C, Jackson-Grusby L and Jaenisch R. (1998). DNA hypomethylation leads to elevated mutation rates. *Nature* **395**(6697)**,** 89-93.

Chen X, Mariappan SV, Moyzis RK, Bradbury EM and Gupta G. (1998). Hairpin induced slippage and hyper-methylation of the fragile X DNA triplets. *J. Biomol. Struct. Dyn.* **15,** 745–756.

Cho KR *et al.* (1994). The *DCC* gene: structural analysis and mutations in colorectal carcinomas. *Genomics* **19(3),** 525-531.

Chuang LS *et al.* (1997). Human DNA- (cytosine-5) methyltransferase-PCNA complex is a target for $p21^{Waf1}$. *Science* **277,**1996-2000.

Cohen SN and JA Shapiro. (1980). Transposable genetic elements. *Sci. Am.* **242(2),** 40-49.

Cottrell S, Bicknell D, Kaklamanis L and Bodmer WF. (1992). Molecular analysis of *APC* mutations in familial adenomatous polyposis and sporadic colon carcinomas . *Lancet* **340,** 626–630.

Craig NL. (1990). P element transposition. *Cell* **62(3),** 399-402.

Cross SH, Meehan RR, Nan X and Bird A. (1997). A component of the transcriptional repressor MeCP1 shares a motif with DNA methyltransferase and HRX proteins. *Nat. Genet.* **16(3),** 256-259.

Cummings CJ and Zoghbi HY. (2000). Fourteen and counting Unraveling trinucleotide repeat diseases. *Hum. Mol. Genet.* **9,** 909–916.

Cunningham JM et al. (1998). Hypermethylation of the *hMLH1* promoter in colon cancer with microsatellite instability. *Cancer Res.* **58(15),** 3455-3460.

Dai P *et al.* (1999) Sonic Hedgehog-induced activation of the *Gli1* promoter is mediated by GLI3. *J. Biol. Chem.***274,** 8143–8152.

Di Cristofano A *et al.* (1999). Impaired Fas response and autoimmunity in *Pten*$^{+/-}$ mice. *Science* **285,**2122-2125

Dong SM *et al.* (1998). Frequent somatic mutations in *Serine/Threonine Kinase 11/Peuta-Jeghers Syndrome* gene in left-sided colon cancer. *Cancer Res.* **58,** 3787-3790.

Dosaka-Akita H *et al.* (1997). Altered retinoblastoma protein expression in nonsmall cell lung cancer: its synergistic effects with altered ras and p53 protein status on prognosis. *Cancer* **79,** 1329 -1337.

El-Deiry WS *et al.* (1993). WAF1, a potential mediator of p53 tumor suppression. *Cell* **75,** 817-825.

Eliyahu D, Raz A, Gruss P, Givol D and Oren M. (1984) Participation of p53 cellular tumour antigen in transformation of normal embryonic cells *Nature* **312,** 646-649.

Esteller M *et al.* (1999). Detection of aberrant promoter hypermethylation of tumor suppressor genes in serum DNA from non-small cell lung cancer patients. *Cancer Res.* **59,** 67-70.

Esteller M *et al.* (2000). Epigenetic inactivation of *LKB1* in Primary tumors associated with the Peutz-Jeghers syndrome. *Oncogene* **19,**164-168.

Fearon ER, Hamilton S, and Vogelstein B. (1987). Clonal analysis of human colorectal tumors. *Science* **238,** 193-197.

Fearon ER and Vogelstein B. (1990). A genetic model for colorectal tumorigenesis. *Cell* **61,**759-767.

Feng Q and Zhang Y. (2001). The MeCP1 complex represses transcription through preferential binding, remodeling, and deacetylating methylated nucleosomes. *Genes Dev.* **15,**827-832.

Fero ML, Randel E, Gurley KE, Roberts JM and Kemp CJ. (1998). The murine gene *p27Kip1* is haplo-insufficient for tumour suppression. *Nature* 396, 177-180

Fishel R *et al.* (1993). The human mutator gene homolog *MSH2* and its association with hereditary nonpolyposis colon cancer. *Cell* **75,**1027–1038.

Ford JM and Hanawalt PC. (1995). Li-Fraumeni syndrome fibroblasts homozygous for *p53* mutations are deficient in global DNA repair but exhibit normal transcription-coupled repair and enhanced UV resistance. *Proc. Natl. Acad. Sci.USA.* **92,** 8876-8880.

Fournier C *et al.* (2002). Allele-specific histone lysine methylation marks regulatory regions at imprinted mouse genes. *EMBO J.* 21(23), 6560-6570.

Gailani MR *et al.* (1996). Relationship between sunlight exposure and a key genetic alteration in basal cell carcinoma. *J. Natl. Cancer Inst.* **88,** 349-354.

Gao AC *et al.* (1999). Suppression of the tumourigenicity of prostatic cancer cells by gene(s) located on human chromosome 19p13.1-13.2. *Prostate* **38(1),** 46-54.

Gazzeri S *et al.* (1998). The human p19/ARF protein encoded by the beta transcript of the p16/INK4a gene is frequently lost in small cell lung cancer. *Cancer Res.* **58,** 3926 -3931.

Giardiello FM *et al.* (1987). Increased risk of cancer in the Peutz-Jeghers syndrome. *N. Engl. J. Med.* **316,**1511-1514.

Gillies S *et al.* (1997). Mutations in patched, the human homologue of *Drosophila* patched in nevoid basal cell carcinoma syndrome. *Dev. Bio.***186,** A90.

Giles RH, van Es JH and Clevers H. (2003). Caught up in a Wnt storm**:** Wnt signaling in cancer. *Biochim. Biophys. Acta.* **1653(1),** 1-24.

Goelz SE, Vogelstein B, Hamilton SR and Feinberg AP. (1985). Hypomethylation of DNA from benign and malignant human colon neoplasms. *Science* **228,** 187–190.

Goodrich LV and Scott MP. (1998). Hedgehog and patched in neural development and disease. *Neuron* **21,** 1243-1257.

Grachtchouk M *et al.* (2000). Basal cell carcinomas in mice overexpressing *Gli2* in skin. *Nat. Genet.* **24,**216-217.

Grachtchouk V et al. (2003). The magnitude of hedgehog signaling activity defines skin tumor phenotype. *EMBO J.* **22(11),** 2741-2751.

Greenblatt MS, Bennett WP, Hollstein M, and Harris CC. (1994). Mutations in the *p53* tumor suppressor gene**:** clues to cancer etiology and molecular pathogenesis *Cancer Res.* **54,** 4855-4878.

Groden J *et al.* (1991). Identification and characterization of the familial adenomatous polyposis coli gene. *Cell* **66,** 589-600.

Gruber SB *et al.* (1998). Pathogenesis of adenocarcinoma in Peutz-Jeghers syndrome. *Cancer Res.* **58,**5267-5270.

Guan RJ *et al.* (2000). *Drg-1* as a differentiation-related, putative metastatic suppressor gene in human colon cancer. *Cancer Res.* **60(3),** 749-755.

Hahn SA *et al.* (1996). *DPC4*, a candidate tumor suppressor gene at human chromosome 18q21.1. *Science* **271,** 350-353

Hall JM *et al.* (1990). Linkage of early-onset familial breast cancer to chromosome 17q21 *Science* **250,** 1684-1689.

Harper JW, Adami GR, Wei N, Keyomarsi K, Elledge SJ. (1993). The p21/Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases. *Cell* **75,** 805-816.

Harris CC. (1996). Structure and function of the *p53* tumor suppressor gene: clues for rational cancer therapeutic strategies. *J. Natl. Cancer Inst.* **88(20),** 1442-1455.

Hartl DL, Lozovskaya ER, Nurminsky DI and Lohe AR. (1997). What restricts the activity of mariner-like transposable elements. *Trends Genet.* **13(5),** 197-201.

Hawkins N et al. (2002). Sporadic colorectal cancers with CpG island methylation: clinicopathological features and relationship to microsatellite instability. *Gastroenterology* **122,** 1376–1387.

Hemminki A *et al.* (1997). Localization of a susceptibility locus for Peutz-Jeghers syndrome to 19p using comparative genomic hybridization and targeted linkage analysis. *Nat. Genet.* **15,**87-90.

Hendrich B and Bird AP. (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 VA and Bird A. (2001). Closely related proteins MBD2 and MBD3 play distinctive but interacting roles in mouse development. *Genes Dev.* **15,**710-723.

Herman JG *et al.* (1995). Inactivation of the *CDKN2/p16/MTS1* gene is frequently associated with aberrant DNA methylation in all common human cancers. *Cancer Res.* **55(20),** 4525-4530.

Hermeking H *et al.* (1997). 14-3-3 sigma is a p53-regulated inhibitor of G2/M progression. *Mol. Cell.* **1,** 3-11.

Herrera L, Kakati S, Gibas L, Pietrzak E and Sandberg AA. (1986). Gardner syndrome in a man with an interstitial deletion of 5q. *Am. J. Med. Genet.* 25, 473–476.

Hibi K *et al.* (1992). Three distinct regions involved in 3p deletion in human lung cancer. *Oncogene* **7,** 445-449.

Howe JR *et al.* (1998). Mutations in the *SMAD4/DPC4* gene in juvenile polyposis. *Science* **280**, 1086-1088.

Hsu S, Huang F, Ossowski L and Friedman E. (1995). Colon carcinoma cells with inactive nm23 show increased motility and response to motility factors. *Carcinogenesis* **16(9),** 2259-2262.

Ingham P W. (1998). Transducing hedgehog: the story so far. *EMBO J.* **17,** 3505-3511.

Ingham PW and McMahon AP. (2001). Hedgehog signaling in animal development: paradigms and principles. *Genes Dev.***15,** 3059–3087.

Ishizuya-Oka A *et al.* (2001). Thyroid hormone-induced expression of sonic hedgehog correlates with adult epithelial development during remodeling of the *Xenopus* stomach and intestine. *Differentiation.* **69(1),** 27-37.

Iwao K *et al.* (1998). Activation of the *beta-catenin* gene by interstitial deletions involving exon 3 in primary colorectal carcinomas without adenomatous polyposis coli mutations. *Cancer Res.* **58,**1021-1026.

Jaenisch R. (1997). DNA methylation and imprinting: why bother? *Trends Genet.* **13,**23-329.

Jenkins JR, Rudge K, and Currie GA. (1984). Cellular immortalization by a cDNA clone encoding the transformation-associated phosphoprotein p53. *Nature* **312,** 651-654.

Jiang CL *et al.* (2002). MBD3L1 and MBD3L2, two new proteins homologous to the methyl-CpG-binding proteins MBD2 and MBD3: characterization of MBD3L1 as a testis-specific transcriptional repressor. *Genomics* **80(6),** 621-629.

Ji W *et al.* (1997). DNA demethylation and pericentromeric rearrangements of chromosome 1. *Mutat. Res.* **379(1),** 33-41.

Jones KL. (1997). Smith's recognizable patterns of human malformation, 5th ed. WB Saunders, Philadelphia.

Jones PL *et al.* (1998). Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nat. Genet.* **19,** 187-191.

Jones PA and Laird PW. (1999). Cancer epigenetics comes of age. *Nat. Genet.* **21,** 163–166.

Jones PA and Taylor SM. (1980). Cellular differentiation, cytidine analogs and DNA methylation. *Cell* **20(1),** 85-93.

Kalderon D. (2000). Transducing the hedgehog signal. *Cell* **103,** 371–374.

Kalff-Suske M *et al.* (1999). Point mutations throughout the *GLI3* gene cause Greig cephalopolysyndactyly syndrome. *Hum. Mol. Genet.* **8(9):**1769-1777.

Kamb A et al. (1994). Analysis of the *p16* gene (*CDKN2*) as a candidate for the chromosome 9p melanoma susceptibility locus. *Nat. Genet.* **8(1):**23-26.

Kang S, Graham JM Jr, Olney AH and Biesecker LG. (1997). *GLI3* frameshift mutations cause autosomal dominant Pallister-Hall syndrome. *Nat. Genet.* **15,**266-268.

Katayam M et al. (2002). *Patched* and *smoothened* mRNA expression in human astrocytic tumors inversely correlates with histological malignancy. *J. Neurooncol.* **59(2),** 107-115.

Kinzler KW *et al.* (1987). Identification of an amplified, highly expressed gene in a human glioma. *Science* **236,** 70-73.

Kinzler KW *et al.* (1991). Identification of FAP locus genes from chromosome 5q21. *Science* **253,** 661–665.

Kinzler LW and Vogelstein B. (1990). The *GLI* gene encodes a nuclear protein which binds specific sequences in the human genome. *Mol. Cell Biol.* **10,** 634–642

Kinzler KW and Vogelstein B. (1996). Lessons from hereditary colorectal cancer. *Cell* **87,**159-170.

Knudson AG. (1971). Mutation and cancer**:** statistical study of retinoblastoma. *Proc. Natl. Acad. Sci. USA.* **68,** 820-823.

Kolodner RD *et al.* (1999). Germ-line *msh6* mutations in colorectal cancer families. *Cancer Res.* **59,**5068–5074.

Lagna G, Hata A, Hemmati-Brivanlou A and Massagué J. (1996). Partnership between DPC4 and SMAD proteins in TGF-beta signalling pathways. *Nature* **383,** 832-836

Lam CW *et al.* (1999). A frequent activated smoothened mutation in sporadic basal cell carcinomas. *Oncogene* **18,**833–836.

Lamlum H *et al.* (1999). The type of somatic mutation at *APC* in familial adenomatous polyposis is determined by the site of the germline mutation**:** a new facet to Knudson's 'two-hit' hypothesis. *Nat. Med.* **5,** 1071–1075.

Lane D and Benchimol S. (1990). *p53***:** oncogene or anti-oncogene? *Genes Dev.* **4,**1-8.

Latif F et al. (1992). Molecular characterization of a large homozygous deletion in the small cell lung cancer cell line U2020**:** a strategy for cloning the putative tumor suppressor gene. *Genes Chromosomes Cancer* **5,** 119-127.

Launonen V *et al.* (2000). No evidence of Peutz-Jeghers syndrome gene *LKB1* involvement in left-sided colorectal carcinomas. *Cancer Res.* **60,** 546-548.

Leach FS *et al.* (1993). Mutations of a *mutS* homolog in hereditary nonpolyposis colorectal cancer. *Cell* **75,**1215–1226.

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

Lewis JD *et al.* **(**1992). Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. *Cell* **69,**905–914.

Li E, Beard C and Jaenisch R. (1993). Role for DNA methylation in genomic imprinting. *Nature* **366,** 362-365.

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

Li H *et al.* (2003). *SLC5A8*, a sodium transporter, is a tumor suppressor gene silenced by methylation in human colon aberrant crypt foci and cancers. *Proc. Natl. Acad. Sci. USA.* **100(14),** 8412-**8417.**

Lin X and Nelson WG. (2003). Methyl-CpG-binding domain protein-2 mediates transcriptional repression associated with hypermethylated *GSTP1* CpG islands in MCF-7 breast cancer cells.
*Cancer Res.***63,** 498-504.

Lindblom A, Tannergard P, Werelius B and Nordenskjold M. (1993). Genetic mapping of a second locus predisposing to hereditary non-polyposis colon cancer. *Nat. Genet.* **5,** 279-282.

Litingtung Y, Lei L, Westphal H and Chiang C. (1998). Sonic hedgehog is essential to foregut development. *Nat. Genet.* **20(1),** 58-61.

Liu B *et al.* (1995). Mismatch repair gene defects in sporadic colorectal cancers with microsatellite instability. *Nat. Genet.***9,** 48–55.

Liu W *et al.* (2000). Mutations in *AXIN2* cause colorectal cancer with defective mismatch repair by activating beta-catenin/TCF signalling. *Nat. Genet.* **26,**146-147.

Lukeis R, Irving L, Garson M and Hasthorpe S. (1990). Cytogenetics of non-small cell lung cancer**:** analysis of consistent non-random abnormalities. *Genes Chromosomes Cancer.* **2,** 116-124.

Lynch HT and Lynch JF. (1998). Genetics of colonic cancer. *Digestion* **59,**481-492.

Mady HH and Melhem MF. (2002). FHIT protein expression and its relation to apoptosis, tumor histologic grade and prognosis in colorectal adenocarcinoma**:** an immunohistochemical and image analysis study. *Clin. Exp. Metastasis.* **19(4),** 351-358.

Magdinier F and Wolffe AP. (2001). Selective association of the methyl-CpG binding protein MBD2 with the silent p14/p16 locus in human neoplasia. *Proc. Natl. Acad. Sci. USA.* **98,** 4990-4995.

Marchetti A *et al.* (1997). Alterations of P16 (MTS1) in node-positive non-small cell lung carcinomas. *J. Pathol.* **181,** 178 -182.

Marhold J *et al.* (2002). Stage-specific chromosomal association of Drosophila dMBD2/3 during genome activation. *Chromosoma* **111(1),** 13-21.

Matise MP and Joyner AL. (1999). *Gli* genes in development and cancer. *Oncogene* **18(55),** 7852-7859.

McMahon A P. (2000). More surprises in the hedgehog signaling pathway. *Cell* **100,** 185-188.

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

Merlo A *et al.* (1995). 5'-CpG island methylation is associated with transcriptional silencing of the tumour suppressor p16/CDKN2/MTS1 in human cancers. *Nat. Med.* **1,** 686 –692.

Mèthot N and Basler K. (1999). Hedgehog controls limb development by regulating the activities of distinct transcriptional activator and repressor forms of Cubitus interruptus. *Cell* **96,** 819–831.

Mirabelli-Primdahl L *et al.* (1999). *Beta-catenin* mutations are specific for colorectal carcinomas with microsatellite instability but occur in endometrial carcinomas irrespective of mutator pathway. *Cancer Res.* **59,**3346-3351.

Miyaki M *et al.* (1999). Frequent mutation of *beta-catenin* and *APC* genes in primary colorectal tumors from patients with hereditary nonpolyposis colorectal cancer. *Cancer Res.* **59,**4506-4509.

Miyaki M *et al.* (1999). Higher frequency of *SMAD4* gene mutation in human colorectal cancer with distant metastasis. *Oncogene* **18**, 3098-3103

Miyaki M *et al.* (2001). Alterations of repeated sequences in 5' upstream and coding regions in colorectal tumors from patients with hereditary nonpolyposis colorectal cancer and Turcot syndrome. *Oncogene* **20,** 5215-5218.

Miyaki M *et al.* (1994). Characteristics of somatic mutation of the adenomatous polyposis coli gene in colorectal tumours. *Cancer Res.* **54,**3011–3020.

Miyashita T and Reed JC. (1995). Tumor suppressor p53 is a direct transcriptional activator of the human bax gene. *Cell* **80,** 293-299.

Miyoshi Y *et al.* (1992). Somatic mutations of the *APC* gene in colorectal tumors: mutation cluster region in the *APC* gene. *Hum. Mol. Genet.***1,** 229–233.

Moffat AS. (2000). Genetics. Transposons help sculpt a dynamic genome. *Science* **289(5484),** 1455-1457.

Momparler RL. (2003). Cancer epigenetics. *Oncogene* **22(42),** 6479-6483.

Morrone A et al. (1997). RNA metabolism in myotonic dystrophy: Patient muscle shows decreased insulin receptor RNA and protein consistent with abnormal insulin resistance. *J. Clin. Invest.* **99,** 1691–1698.

Moroni MC *et al.* (2001). *Apaf-1* is a transcriptional target for E2F and p53. *Nat. Cell. Biol.* **3,** 552-558.

Muenke M and Beachy PA. in *The Metabolic and Molecular Bases of Inherited Disease* (eds Scriver, C., Beaudet, A., Sly, W. & Valle, D.) 6203-6230 (McGraw-Hill, New York, 2001).

Mullor JL, Sanchez P and Altaba AR. (2002). Pathways and consequences: Hedgehog signaling in human disease. *Trends Cell Biol.* **12(12),** 562-569.

Müller-Tidow C et al. (2001) Loss of expression of HDAC-recruiting methyl-CpG-binding domain proteins in human cancer. *Br J. Cancer.* **85**(8), 1168–1174.

Munemitsu S *et al.* (1994). The *APC* gene product associates with microtubules in vivo and promotes their assembly in vitro. *Cancer Res.* **54,**3676–3681.

Murone M, Rosenthal A and de Sauvage FJ. (1999). Sonic hedgehog signaling by the patched–smoothened receptor complex. *Curr. Biol.* **9,** 76–84.

Ng HH, Jeppesen P and Bird A. (2000). Active repression of methylated genes by the chromosomal protein MBD1. *Mol. Cell. Biol.* **20,**1394-1406.

Ng HH *et al.* (1999). MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex. *Nat. Genet.* **23(1),** 58-61.

Nigro JM *et al.* (1989). Mutations in the *p53* gene occur in diverse human tumour types. *Nature* 342, 705-608.

Nüslein-Volhard C. and Wieschaus E. (1980). Mutations affecting segment number and polarity in *Drosophila*. *Nature* **287,** 795–801.

Ohlmeyer JT and Kalderon D. (1998). Hedgehog stimulates maturation of Cubitus interruptus into a labile transcriptional activator. *Nature* **396,** 749–753.

Okamoto A *et al.* (1994). Mutations and altered expression of *p16/INK4* in human cancer. *Proc. Natl. Acad. Sci. USA.* **91,** 11045 -11049.

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

Okano M, Bell DW, Haber DA and Li W. (1999). DNA methyltransferases *Dnmt3a* and *Dnmt3b* are essential for *de novo* methylation and mammalian development. *Cell* **99,** 247-257.

Oro AE *et al.* (1997). Basal Cell Carcinomas in Mice Overexpressing Sonic Hedgehog. *Science* **276,** 817-821.

Owen JS *et al.* (2003). Deficiency of Mbd2 suppresses intestinal tumorigenesis. *Nat. Genet.* **34(2),** 145-147.

Parada LF, Land H, Weinberg RA, Wolf D and Rotter V. (1984). Cooperation between gene encoding p53 tumour antigen and ras in cellular transformation. *Nature* **312,** 649-651.

Park HL *et al.* (2000). Mouse *Gli1* mutants are viable but have defects in SHH signaling in combination with a *Gli2* mutation. *Development* **127,**1593-1605.

Patra SK, Patra A, Zhao H, Carroll P and Dahiya R. (2003). Methyl-CpG–DNA binding proteins in human prostate cancer: expression of CXXC sequence containing *MBD1* and repression of *MBD2* and *MeCP2*. *Biochemical and Biophysical Research Communications* **302(4),** 759-766.

Pavletich N and Pabo CO. (1993). Crystal structure of a fivefinger GLI–DNA complex: new perspectives on zinc fingers.*Science* 261, 1701–1707.

Polakis P. (1999). The oncogenic activation of beta-catenin. *Curr. Opin. Genet. Dev.* **9,** 15–21.

Polyak K, Xia Y, Zweier JL, Kinzler KW, Vogelstein B. (1997). A model for p53-induced apoptosis. *Nature* **389,** 300-305.

Powell SM, Harper JC, Hamilton SR, Robinson CR and Cummings OW. (1997). Inactivation of *SMAD4* in gastric carcinomas. *Cancer Res.* **57,** 4221-4224

Powell SM *et al.* (1992). *APC* mutations occur early during colorectal tumorigenesis. *Nature* **359,**235–237.

Ramalho-Santos M, Melton DA and McMahon AP. (2000). Hedgehog signals regulate multiple aspects of gastrointestinal development. *Development* **127,** 2763-2772.

Reddy PS and Housman DE. (1997). The complex pathology of trinucleotide repeats. *Curr. Opin. Cell. Biol.* **9(3),** 364-372.

Reifenberger J *et al.* (1998). Missense mutations in *SMOH* in sporadic basal cell carcinomas of the skin and primitive neuroectodermal tumors of the central nervous system. *Cancer Res.* **58,** 1798–1803.

Reissmann PT *et al.* (1993). Inactivation of the retinoblastoma susceptibility gene in non-small- cell lung cancer. The Lung Cancer Study Group. *Oncogene* **8,** 1913 – 1919.

Resta N *et al.* (1998). *STK11* Mutations in Peutz-Jeghers syndrome and sporadic colon cancer. *Cancer Res.* 1998; **58,** 4799-4801.

Riccio A *et al.* (1999). The DNA repair gene *MBD4 (MED1)* is mutated in human carcinomas with microsatellite instability. *Nat. Genet.* **23,** 266-268.

Richardson GE and Johnson BE. (1993). The biology of lung cancer. *Semin. Oncol.* **20,** 105-127.

Riggins GJ, Kinzler KW, Vogelstein B and Thiagalingam S. (1997) Frequency of *SMAD* gene mutations in human cancers. *Cancer Res.* **57,** 2578-2580.

Roberts DJ, Smith DM, Goff DJ and Tabin CJ. (1998). Epithelial-mesenchymal signaling during the regionalization of the chick gut. *Development* **125(15),** 2791-801.

Robertson KD *et al.* (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.

Robles AI, Bemmels NA, Foraker AB and Harris CC. (2001). *APAF-1* is a transcriptional target of p53 in DNA damage-induced apoptosis. *Cancer Res.* **61,** 6660-6664.

Rous P. (1911). A sarcoma of the fowl transmissable by an agent separable from the tumor cells. *J. Exp. Med.* **13,** 397.

Rubinfeld B *et al.* (1993). Association of the *APC* gene product with beta catenin. *Science* **262,**1731–1734.

Ruiz IA. (1999). Gli proteins encode positive and dominant negative functions: implications for development and disease. *Development* **126,** 3205–3216.

Ruiz IA, Nguyên V and Palma V. (2003). The emergent design of the neural tube: prepattern, SHH morphogen and GLI code. *Curr. Opin. Genet. Dev.* **13(5),** 513-521.

Ruppert JM, Vogelstein B and Kinzler KW. (1991). The zinc finger protein GLI transforms primary cells in cooperation with adenovirus E1A. *Mol. Cell. Biol.* **11,**1724-1728

Rusin MR *et al.* (1996). Intragenic mutations of the *p16 (INK4), p15 (INK4B)* and *p18* genes in primary non-small-cell lung cancers. *Int. J. Cancer* **65,** 734 –739.

Saito M *et al.* (1999). Expression of DCC protein in colorectal tumors and its relationship to tumor progression and metastasis. *Oncology* **56,**134–141.

Saldanha G. (2001). The Hedgehog signalling pathway and cancer. *J. Pathol.* **193(4),** 427-432.

Sanchez-Cespedes M *et al.* (2001). Chromosomal alterations in lung adenocarcinoma from smokers and nonsmokers. *Cancer Res.* **61,**1309-13.

Sanders BM, Jay M, Draper GJ and Roberts EM. (1989). Non-ocular cancer in relatives of retinoblastoma patients. *Br. J. Cancer.* **60,** 358 –365.

Serrano M, Hannon GJ and Beach D. (1993). A new regulatory motif in cell-cycle control causing specific inhibition of cyclin D/CDK4. *Nature* **366,** 704-707.

Sasaki H, Hui C, Nakafuku M and Kondoh H. (1997). A binding site for Gli proteins is essential for *HNF-3* floor plate enhancer activity in transgenics and can respond to Shh in vitro. *Development* **124,**1313-1322.

Shapiro GI *et al.* (1995). Multiple mechanisms of p16/INK4A inactivation in non-small cell lung cancer cell lines. *Cancer Res* **55,** 6200 –6209.

Shibata D *et al.* (1996). The DCC protein and prognosis in colorectal cancer. *N. Engl. J. Med.* **335,**1727–1732.

Shin SH. et al. (1999). *GLI3* mutations in human disorders mimic Drosophila Cubitus interruptus protein functions and localization. *Proc. Natl. Acad. Sci. U S A.* **96,** 2880–2884.

Shoemaker AR *et al.* (1998). A resistant genetic background leading to incomplete heterozygosity in $Apc^{min/+}$ mice. *Proc. Natl. Acad. Sci. USA.* **95,** 10826-10831.

Slack A. et al. (2002). Antisense *MBD2* gene therapy inhibits tumorigenesis. *J. Gene. Med.* **4, 381**-389.

Solomon E *et al.* (1987). Chromosome 5 allele loss in human colorectal carcinomas. *Nature* **328,**616–619.

Sowa Y et al. (1997). Histone deacetylase inhibitor activates the *WAF1/Cip1* gene promoter through the Sp1 sites. *Biochem. Biophys. Res. Commun.* **241(1),** 142-150.

Sparks AB, Morin PJ, Vogelstein B and Kinzler KW. (1998). Mutational analysis of the APC/beta-catenin/Tcf pathway in colorectal cancer. *Cancer Res.* **58,**1130-1134.

Stehlin D, Varmus HE, Bishop JM, and Vogt PK. (1976). DNA related to the transforming gene(s) of avian sarcoma viruses is present in normal avian DNA. *Nature* **b260,** 70-173.

Suh ER, Ha CS, Rankin EB, Toyota M and Traber PG. (2002). DNA methylation down-regulates *CDX1* gene expression in colorectal cancer cell lines. *J. Biol. Chem.* **277(39),** 35795-35800

Su JY, Erikson E and Maller J. (1996). Cloning and characterization of a novel serine/threonine protein kinase expressed in early *Xenopus* embryos. *J. Biol. Chem.* **271,**14430-14437.

Su LK, Vogelstein B and Kinzler KW. (1993). Association of the APC tumor suppressor protein with catenins. *Science* **262,** 1734–1737.

Sukegawa A *et al.* (2000). The concentric structure of the developing gut is regulated by Sonic hedgehog derived from endodermal epithelium. *Development* **127(9),** 1971-1980.

Suter CM et al. (2003). CpG island methylation is a common finding in colorectal cancer cell lines. *Br. J. Cancer.* **88,** 413 – 419.

Takagi Y *et al.* (1996). Somatic alterations of the *DPC4* gene in human colorectal cancers in vivo. *Gastroenterology* **111,** 1369-1372.

Takahashi T *et al.* (1989). *p53*, a frequent target for genetic abnormalities in lung cancer. *Science* **246,** 491-494.

Takakura M *et al.* (2001). Telomerase activation by histone deacetylase inhibitor in normal cells. *Nucleic Acids Res.* **29(14),** 3006-3011.

Taipale J *et al.* (2000). Effects of oncogenic mutations in *Smoothened* and *Patched* can be reversed by cyclopamine. *Nature*, **406,** 1005–1009.

Tatematsu KI, Yamazaki T and Ishikawa F. (2000). MBD2-MBD3 complex binds to hemi-methylated DNA and forms a complex containing DNMT1 at the replication foci in late S phase. *Genes Cells.* **5(8),** 677-688.

te Welscher P et al. (2002). Progression of vertebrate limb development through SHH-mediated counteraction of GLI3. *Science* **298(5594),** 827-830.

Thiagalingam S *et al.* (1996). Evaluation of candidate tumour suppressor genes on chromosome 18 in colorectal cancers. *Nat. Genet.* **13,** 343-346.

Tomita N. *et al.* (2002). Effects of triplet repeat sequences on nucleosome positioning and gene expression in yeast minichromosomes. *Nucleic Acids Res. Suppl.***(2),**231-232.

Tomlinson IPM, Novelli MR and Bodmer WF. (1996). The mutation rate and cancer. *Proc. Natl. Acad. Sci. U S A.***93,** 14800–14803.

Trojan J, Brieger A, Esteller M and Zeuzem S. (2000). 5'-CpG island methylation of the *LKB1/STK11* promoter and allelic loss at Chromosome 19p13.3 in sporadic colorectal cancer. *Gut* **47,** 272-276.

Trojan J *et al.* (1999). Peutz-Jeghers syndrome**:** molecular analysis of a three-generation kindred with a novel defect in the serine threonine kinase gene STK11. *Am. J. Gastroenterol.* **94,**257-261.

Ueki T et al. (2001). Identification and characterization of differentially methylated CpG islands in pancreatic carcinoma. *Cancer Res.* **61(23),** 8540-8546.

Van Lint C, Emiliani S and Verdin E. (1996). The expression of a small fraction of cellular genes is changed in response to histone hyperacetylation. *Gene Expr.***5 (4-5),** 245-53.

Vasen HFA, Mecklin JP, Meera Kahn P and Lynch HT. (1991). Hereditary non-polyposis colorectal cancer. *Lancet* **17710,**26478 -19244.

Venkatachalam S *et al.* (1998). Retention of wild-type p53 in tumors from p53 heterozygous mice**:** reduction of p53 dosage can promote cancer formation. *EMBO J.* **7,** 657-4667.

Virmani AK *et al.* (1998). Allelotyping demonstrates common and distinct patterns of chromosomal loss in human lung cancer types. *Genes Chromosomes Cancer* **21,**308-319.

Vogelstein B *et al.* (1988). Genetic alterations during colorectal-tumor development. *N. Engl. J Med.* **319,** 525-532.

Vortkamp A. *et al.* (1995). Identification of optimized target sequences for the GLI3 zinc finger protein. *DNA Cell Biol.***14,** 629–634.

Wade PA *et al.* (1999). Mi-2 complex couples DNA methylation to chromatin remodelling and histone deacetylation. *Nat. Genet.* **23,** 62-66.

Wade PA, Jones PL, Vermaak D and Wolffe AP. (1998). A multiple subunit Mi-2 histone deacetylase from Xenopus laevis cofractionates with an associated Snf2 superfamily ATPase. *Curr. Biol.* **8,** 843-846.

Wang B, Fallon JF and Beachy PA. (2000). Hedgehog-regulated processing of Gli3 produces an anterior/posterior repressor gradient in the developing vertebrate limb. *Cell* **100(4),** 423-434.

Wang XW *et al.* (1995). p53 modulation of TFIIH-associated nucleotide excision repair activity. *Nat. Genet.* **10,** 188-195.

Wang ZJ, Taylor F, Churchman M, *et al.* (1998). Genetic pathways of colorectal carcinogenesis rarely involve the *PTEN* and *LKB1* genes outside the inherited hamartoma syndromes. *Am. J. Pathol.* **153,**363-366.

Watkins DN *et al.* (2003). Hedgehog signalling within airway epithelial progenitors and in small-cell lung cancer. *Nature* **422(6929),** 313-317.

Webster MT *et al.* (2000). Sequence variants of the axin gene in breast, colon, and other cancers: an analysis of mutations that interfere with GSK3 binding. *Genes Chromosomes Cancer* **28,**443-453.

Wetmore C, Eberhart DE and Curran T. (2000). The Normal *patched* allele is expressed in medulloblastomas from mice with heterozygous germ-line mutation of *patched*. *Cancer Res.* **60(8),** 2239-2246.

Whyte P *et al.* (1988). Association between an oncogene and an anti-oncogene: the adenovirus E1A proteins bind to the retinoblastoma gene product. *Nature* **334,**124-129

Wicking C *et al.* (1997). *De novo* mutations of the *Patched* gene in nevoid basal cell carcinoma syndrome help to define the clinical phenotype. *Am. J. Med. Genet.* **73,** 304-307.

Wong NA and Harrison DJ. (2001). Colorectal neoplasia in ulcerative colitis-recent advances. *Histopathology* **39(3):**221-234.

Wooster R *et al.* (1995). Identification of the breast cancer susceptibility gene *BRCA2*. *Nature* **378,** 789-792.

Wu GS *et al.* (1997). *KILLER/DR5* is a DNA damage-inducible p53-regulated death receptor gene. *Nat. Genet.* **17,** 141-143.

Yamada T *et al.* (2002). Frameshift mutations in the *MBD4/MED1* gene in primary gastric cancer with high-frequency microsatellite instability. *Cancer Lett.* **181(1),** 115-120.

Zalcenstein A *et al.* (2003). Mutant p53 gain of function: repression of *CD95* (*Fas/APO-1*) gene expression by tumor-associated p53 mutants. *Oncogene* **22(36),** 5667-5676.

Zhang J, Rosenthal A, Sauvage FJ and Shivdasani RA. (2001). Downregulation of hedgehog signaling is required for organogenesis of the small intestine in *Xenopus*. *Developmental Biology* **229(1),** 188-202.

Zhang Y and Dufau ML. (2002). Silencing of transcription of the human luteinizing hormone receptor gene by histone deacetylase-mSin3A complex. *J. Biol. Chem.* **277(36),** 33431-33438.

Zhang Y, LeRoy G, Seelig HP, Lane WS and Reinberg D. (1998). The dermatomyositis-specific autoantigen Mi2 is a component of a complex containing histone deacetylase and nucleosome remodeling activities. *Cell* **95,**279-289.

Zhang Y *et al.* (1999). Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation. *Genes Dev.* **13,** 1924-1935.

Zhou XP et al. (2002). *PTEN* mutational spectra, expression levels, and subcellular localization in microsatellite stable and unstable colorectal cancers. *Am. J. Pathol.***161 (2),** 439-447.

Zöchbauer-Müller S *et al.* (2001). Aberrant promoter methylation of multiple genes in non-small cell lung cancers. *Cancer Res.* **61,** 249 –255.

Zochbauer-Muller S *et al.* (2001). 5' CpG Island Methylation of the *FHIT* Gene Is Correlated with Loss of Gene Expression in Lung and Breast Cancer. *Cancer Res.* **61,** 3581-3585.

Abbreviations

A	Adenosine
Amp	Ampicillin
APC	Adenomatous polyposis coli
APS	Ammonium Persulphate
BCC	Basal cell carcinoma
bp	Base pair(s)
C	Cytosine
C-	Carboxy-terminal
cDNA	DNA complementary to RNA
cpm	Counts per minute
CpG	Cytosine and guanine dinucleotide
DCC	Deleted in colon cancer gene
DNA	Deoxyribose nucleic acid
DMSO	Dimethysulphocide
dNTP	Deoxynucleoside triphosphate
DTT	Dithiothreitol
EDTA	Ethylenediaminetetra acetic acid
EtBr	Ethidium bromide
FAP	Familial adenomatous polyposis
FCS	Foetal calf serum
G	Guanine

HNPCC	Hereditary non-polyposis colorectal cancer
IPTG	Isopropyl-[beta]-D-thiogalactopyranoside
LB	Luria broth
LOH	Loss of heterozygote
MOPS	3-[*N*-morpholino]propane sulphonic acid
MBD	Methyl-CpG-binding domain
MMR	Mismatch repair
mRNA	Messenger Ribonucleic Acid
MSI	Microsatellite instability
MSS	Microsatellite stability
N-	Amino terminal
HNPCC	Nonpolyposis colorectal cancer
NSCLC	Non-small cell lung cancer
P	Short arm of chromosome
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PJS	Peutz-Jeghers syndrome
Poly A	Poly Adenylate
q	Long arm of chromosome
RNA	Ribonucleic Acid
RNaseA	Ribonuclease A
rpm	Revolutions per minute
SCLC	Small cell lung cancer
SDS	Sodium Dodecyl

SNP	Single-nucleotide polymorphism
SSC	Saline sodium citrate
SSCP	Single Stranded Conformational Polymorphism
T	Thymine
TAE	Tris-acetate EDTA
Taq	*Thermus aquaticus* DNA polymerase
TBE	Tris-borate EDTA
TE Tris	EDTA
TEMED	N,N,N',N'-tetramethylethylenediamine
T_m	Melting Temperature
Tris	2-amino-2-(hydromethyl) propane-1, 3 diol
X-Gal	5-Bromo-4-chloro-indoyl-[beta]-D galactoside