Molecular Genetic Analysis of DNA Alterations in Irish Colorectal Cancer

A Thesis Presented to The National University of Ireland

for

The Degree of Doctor of Philosophy

by

Tanya M. Mulcahy

Department of Biochemistry
University College Cork.

December 1998

Research Supervisor: Dr. Tommie V. McCarthy.

Table of Contents *Pages*

Acknowledgements

Sincere thanks to my research supervisor, Dr. Tommie McCarthy for his advice, support and direction during this research and also for the numerous lab social events and home invasions he organised.

Thanks also to Prof. Kirwan at CUH and Drs Ciona Murphy and Eamonn Kavanagh and all who were involved in providing samples for me to work on!

A huge thank you to all the members of the molecular biology lab, new and old: Pat L, Annette, Marge, Colin, Kathy, Maureen, Cliona, Ruairi and especially Ger, Pat V and Linda for providing friendship, advice, a listening ear and most of all many laughs and memories, (Spalpin Fanach et al.,)!!

Thanks also to all the members of the Biochemistry department, staff, postgrads and technicians, especially the 'Thirsty Scholar Gang'.

A very big thanks to my good friends Tara and Karen, for the laughs and friendship over the years.

A special word of thanks to my family, my brothers Mark and Bruce and especially my mum, for her support, encouragement and friendship and most importantly for being so wonderful.

Finally, a huge thank you to Kevin for his encouragement, advice, laughter, hot ports and photocopying but most of all for being my very best friend.

This work was supported by the Health Research Board.

To my mum, June Mulcahy.

Father William

"You are old father William", the young surgeon said,

"and your colon from polyps is free.

Yet most of your siblings are known to be dead -

A really bad family tree."

" In my youth" father William replied with a grin,

"I was told that a gene had mutated,

That all who carried this dominant gene

To polyps and cancer were fated.

I sought advice from a medical friend,

Who sighed and said-"without doubt

Your only escape from an untimely end

Is to have your intestine right out".

" It seemed rather bad luck- I was then but nineteen-

So I went and consulted a quack,

Who took a firm grip on my dominant gene

And promptly mutated it back"

"This", said the surgeon, "is something quite new

And before we ascribe any merit

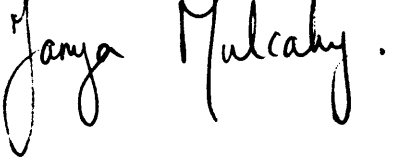
We must see if the claims of this fellow are true,

And observe what your children inherit".

Dukes C. Familial intestinal polyposis. Ann Royal Coll Surg Eng. 1952; 293-304.

Declaration

I hereby declare that this thesis for the degree of Doctor of Philosophy to the National University of Ireland, has not been previously presented for a higher degree in this or any other University. This thesis is of my own composition and may be available for consultation within the University library and may be photocopied or lent to other libraries for the purpose of consultation.

Signature: 

ABBREVIATIONS:

APC: Adenomatous Polyposis Coli.

FAP: Familial Adenomatous Polyposis.

HNPCC: Hereditary Non-Polyposis Colorectal Cancer.

hMSH2: Human *MutS* homolog 2.

hMLH1: Human *MutL* homolog 1.

hPMS1: Human Post Mitotic Segregation (yeast) homolog 1.

hPMS2: Human Post Mitotic Segregation (yeast) homolog 2.

TGF β: Transforming growth factor β.

MMR: Mismatch Repair.

MI: Microsatellite Instability.

RER: Replication error.

Taq: Thermus Aquaticus.

UDG: Uracil DNA Glycosylase.

GMPD: Glycosylase Mediated Polymorphism Detection.

DCC: Deleted in Colon Cancer.

GTBP: G-T Binding Protein

SSCP: Single Strand Conformation Polymorphism.

DGGE: Denaturing Gradient Gel Electrophoresis.

HGDB: Human Genome Database.

SSM: Slipped Strand Mispairing.

TS: Tumour Suppressor.

MTS: Muir-Torre syndrome.

PCR: Polymerase Chain Reaction.

CUH: Cork University Hospital.

Abstract

Colorectal cancer is the most common cause of death due to malignancy in non-smokers in the western world. In 1995 there were 1,757 cases of colon cancer in Ireland. Most colon cancer is sporadic, however ten percent of cases occur where there is a previous family history of the disease. In an attempt to understand the tumorigenic pathway in Irish colon cancer patients, a number of genes associated with colorectal cancer development were analysed in Irish sporadic and HNPCC colon cancer patients.

The hereditary forms of colon cancer include Familial adenomatous polyposis coli (FAP) and Hereditary Non-Polyposis Colon Cancer (HNPCC). Genetic analysis of the gene responsible for FAP, (the APC gene) has been previously performed on Irish families, however the genetic analysis of HNPCC families is limited. In an attempt to determine the mutation spectrum in Irish HNPCC pedigrees, the hMSH2 and hMLH1 mismatch repair genes were screened in 18 Irish HNPCC families. Using SSCP analysis followed by DNA sequencing, five mutations were identified, four novel and a previously reported mutation. In families where a mutation was detected, younger asyptomatic members were screened for the presence of the predisposing mutation (where possible). Detection of mutations is particularly important for the identification of at risk individuals as the early diagnosis of cancer can vastly improve the prognosis.

The sensitive and efficient detection of multiple different mutations and polymorphisms in DNA is of prime importance for genetic diagnosis and the identification of disease genes. A novel mutation detection technique has recently been developed in our laboratory. In order to assess the efficacy and application of the methodology in the analysis of cancer associated genes, a protocol for the analysis of the K-ras gene was developed and optimised. Matched normal and tumour DNA from twenty sporadic colon cancer patients was analysed for K-ras mutations using the Glycosylase Mediated Polymorphism Detection technique. Five mutations of the K-ras gene were detected using this technology. Sequencing analysis verified the presence of the mutations and SSCP analysis of the same samples did not identify any additional

mutations. The GMPD technology proved to be highly sensitive, accurate and efficient in the identification of K-ras gene mutations.

In order to investigate the role of the replication error phenomenon in Irish colon cancer, 3 polyA tract repeat loci were analysed. The repeat loci included a 10 bp intragenic repeat of the TGF-β-RII gene. TGF-β-RII is involved in the TGF-β epithelial cell growth pathway and mutation of the gene is thought to play a role in cell proliferation and tumorigenesis. Due to the presence of a repeat sequence within the gene, TGFB-RII defects are associated with tumours that display the replication error phenomenon. Analysis of the TGF-β-RII 10 bp repeat failed to identify mutations in any colon cancer patients. Analysis of the Bat 26 and Bat 40 polyA repeat sequences in the sporadic and HNPCC families revealed that instability is associated with HNPCC tumours harbouring mismatch repair defects and with 20% of sporadic colon cancer tumours. No correlation between K-ras gene mutations and the RER+ phenotype was detected in sporadic colon cancer tumours.

GENERAL INTRODUCTION

THE GENES INVOLVED IN COLORECTAL CANCER

(I) GENETICS OF COLORECTAL CANCER: GENERAL INTRODUCTION

Colorectal cancer is a common malignancy responsible for approximately 12% of all cancer deaths in Ireland each year (National Cancer Registry). Traditional diagnostic and therapeutic approaches have made little impact on survival over the last 30 years. Advances in molecular biology however, have greatly improved our understanding of the disease, and have had a major impact on early detection, treatment and prevention of colon cancer.

In recent years the cloning, identification and mutation analysis of a number of genes involved in colorectal cancer susceptibility and development has resulted in an unprecedented expansion in the knowledge of the fundamental genetic aspects of tumour biology. This general introduction briefly discusses the background of some of the most important colorectal cancer genes and their roles in both hereditary and sporadic colon cancer. A more detailed discussion of the topics of this thesis- HNPCC and the mismatch repair genes (chapter 1), K-ras (chapter 2), TGF-β and the replication error phenotype (chapter 3) is given in the subsequent chapters.

FAMILIAL ADENOMATOUS POLYPOSIS

Familial adenomatous polyposis (FAP) is an autosomal dominantly inherited cancer predisposition syndrome with reported incidence ranges from 1:17,000 to 1:5,000 (Utsonomiya and Lynch, 1990). Germline mutations in the APC gene are responsible for the disease and data suggest that somatic mutations in APC are ubiquitous in colorectal

adenomas and carcinomas (Miyoshi et al., 1992; Powell et al., 1992). The disease is characterised by the development of hundreds to thousands of adenomatous polyps in the colon and rectum with the inevitable development of malignancy in the absence of prophylactic colectomy (Bussey, 1975).

Initial localisation of the APC gene was suggested by cytogenic studies and further mapping by genetic linkage analysis, localised the gene to a region on the long arm of chromosome 5 (Bodmer et al., 1987). The APC gene was identified as the causative gene when mutations in this gene were shown to segregate with FAP in affected families (Nishisho et al., 1991). APC is a large gene consisting of 15 exons encoding a 2843 amino acid polypeptide, exons 1-14 are small whereas exon 15 accounts for 77% of the coding region (Kinzler et al., 1991). APC was classified as a tumour suppressor gene based on the frequently observed loss of heterozygosity at 5q21 in colon tumours (Miyoshi et al., 1992). To date more than 300 mutations in the APC gene have been described (Beroud et al., 1996), with 60% of all reported mutations occurring in the so-called mutation cluster region encompassing codons 1000 to 1600 at the 5' part of exon 15 (Miyoshi et al., 1992). More than 95% of mutations result in a premature stop codon with the result that a truncated protein is produced (Miyoshi et al., 1992). Within the mutation cluster region, 10% of mutations occur at codons 1061 and 1309 (Nagase et al., 1993). In the Irish population these mutations account for 6% and 11% of reported germline APC mutations, respectively (O'Sullivan et al., 1998).

The APC gene encodes a 310kD protein with cytoplasmic localisation. Insight into the function of the protein came when the α and β-catenins were shown to co-purify with the APC protein (Su et al., 1993). The catenins are a group of cytoplasmic proteins

that were identified primarily due to their association with the cell adhesion molecule E-cadherin. The association of these proteins with APC, suggests that the protein may play a role in cell adhesion. Analysis of the role of catenins in Drosophila suggest that the APC complex may regulate transmission of the contact inhibition signal into the cell (Peifer et al., 1993). This hypothesis is consistent with the observation that APC mutations are associated with the development of hyperplasia - an early event in tumorigenesis. The APC protein has also been shown to associate with microtubules and cell adhesion molecules to modify transcriptional activation and to alter cell cycle regulation (Baeg et al., 1995).

HEREDITARY NON-POLYPOSIS COLORECTAL CANCER.

Hereditary non-polyposis colorectal cancer (HNPCC) is an autosomal dominantly inherited cancer predisposition syndrome characterised by the lack of any phenotype other than the development of cancer at an early age. The disease has a high penetrance and accounts for approximately 10% of cases of colorectal cancer (Bodmer et al., 1994). Adenomatous polyps are found in HNPCC but the numbers are much smaller than in FAP (usually less than 10) which is comparable to the frequency of adenomas in the general population. There is a propensity for carcinomas to develop in the proximal colon and the age of onset is earlier than in the general population (Lynch et al., 1993). HNPCC can be inherited as a site specific colorectal cancer (Lynch Type I) or associated with extracolonic tumours such as stomach ovarian, uterine and other malignancies (Lynch Type II) (Watson and Lynch, 1994). Germline mutations in the human mismatch repair genes (hMSH2, hMLH1, hPMS1 and hPMS2) have been causally linked to the

disease (Leach et al., 1993; Papadopolous et al., 1994; Nicolaides et al., 1994). Genetic analysis of these genes in affected pedigrees has greatly improved the understanding of the disease and has facilitated in the detection of at risk individuals and in the early detection and prevention of tumour development in these families. The genetic analysis of HNPCC in the Irish population has been limited and the mutation spectrum in mismatch repair genes in Irish pedigrees has not been previously investigated. A more detailed discussion of HNPCC is given in section 1.1.

THE p53 GENE IN COLORECTAL CARCINOGENESIS.

Abnormalities of the p53 gene are the commonest genetic events associated with human cancers (Hollstein et al., 1991). p53 is a phosphoprotein which enters the nucleus of the cell during DNA synthesis, exerting control over the transcription of other genes and regulating the onset of DNA replication at the G1-S boundary of the cell cycle. (Vogelstein et al., 1992). Under conditions of DNA damage in the cell, p53 appears to block cell growth allowing DNA repair prior to cell division. In the absence of DNA repair apoptosis may be induced, thereby preventing the further growth and replication of a mutant cell (Yonish et al., 1991).

The p53 mutation spectrum exhibits tissue specificity. Most mutations in colorectal cancer are found in codons 175, 248, 278 and 282, these are most commonly single base transitions in CpG nucleotides causing missense mutations (Prives et al.,1993). These mutations result in the production of full length protein with abnormal conformation and function. p53 mutation is rare in colorectal adenomas, indicating that it

occurs late in the process of tumorigenesis (Carder et al., 1993) and is more likely involved in tumour progression rather than development.

DELETED IN COLON CANCER (DCC).

Deleted in colon cancer (DCC) has been associated with the cumulative mutation pathway in the development of colon cancer (Fearon and Vogelstein, 1990). Mutations in DCC are associated with more than 70% of colorectal carcinomas (Vogelstein et al., 1988). The gene encodes a receptor for netrin-1 a molecule involved in axon guidance (Keino-Masu et al., 1989). Recent investigations have shown that DCC induces apoptosis in the absence of ligand binding, but blocks apoptosis when engaged by netrin-1 (Mehlen et al., 1998). Mutation analysis of DCC suggests that the protein may function as a tumour suppressor gene by inducing apoptosis in cells that are not exposed to the ligand - as in metastatic or invasive cells. Mutations in DCC result in loss of function and enhances cell survival of cells outside the region of ligand availability. This is consistent with reports that DCC mutation is a late event in the tumorigenic pathway of colon cancer (Cho and Fearon, 1995).

THE K-RAS GENE.

The involvement of K-ras in colorectal tumorigenesis has been widely investigated. Mutations in K-ras have been identified in 50% of colon cancers (Bos et al., 1987; Forresster et al., 1987). K-ras is a member of the ras gene family of cellular oncogenes. These proetins play a role in signal transduction from G-protein coupled receptors. A signal from the receptor triggers the binding of GTP to the ras protein and

GTP-ras transmits the signal onwards to the cell. The GTPase activity of the ras proteins results in the rapid breakdown of the active GTP-ras to inactive GDP-ras (Barbacid et al., 1987). Mutations in the ras gene lead to reduced GTPase activity with the result that the signal remains active longer, leading to excessive response to the signal from the receptor and thus cellular proliferation (Barbacid et al., 1987). Mutations in K-ras have been detected in carcinomas, adenomas and also in the earliest detectable cancer lesions (aberrant crypt foci) of cancer patients (Shivapurkar et al., 1997; Forrester et al., 1987; Vogelstein et al., 1988). These findings suggest that K-ras mutations are an early event in the tumorigenic pathway in colon cancer.

TRANSFORMING GROWTH FACTOR β RECEPTOR II

The transforming growth factor β protein (TGF-β) is a member of the supergene family of factors that regulate cellular growth. TGF-β inhibits epithelial cell growth and as such it is believed that alteration of the normal activity of this protein may have a role in tumorigenesis (Roberts et al., 1990).

TGF-β forms a complex with two receptors; type I and type II, in the cell (Chen et al., 1993). Mutation analysis of both receptors identified mutations in a intragenic repeat region of the RII gene. These mutations were confined to tumours displaying microsatellite instability (Markowitz et al., 1995). The cumulative evidence from TGF-βRII studies reveal that 70-90% of RER+ tumours (sporadic or HNPCC) harbour mutations in this gene. It has been proposed that due to the nature of the intragenic repeat sequences of TGFβ-RII- the gene is a target for mutations, which escape correction due to mismatch repair defects. The loss of tumour suppressor activity of

TGFβ (through the inactivation of RII) provides a selective growth advantage to the cells thus advancing the tumorigenic pathway (Markowitz et al., 1995; Akiyama et al., 1996).

THE MULTISTEP PATHWAY OF COLORECTAL TUMORIGENESIS.

The molecular basis by which cancers arise has been investigated abundantly. Due to the availability and relative ease of obtaining tumours at early, late and intermediate stages of cancer development, colon cancer has become one of the most widely investigated human cancers.

Tumorigenesis is believed to be a multistep process in which malignant colorectal tumours arise from pre-existing benign adenomas. Histopathological and clinical assessment of cancer development suggest that most tumours arise from early adenomas, these progress in size, become dysplastic and eventually form an invasive tumour. This concept is known as the theory of clonal expansion (Nowell et al., 1976; Loeb et al., 1991). Various models have been proposed to explain the sequential development of colon cancer. The most widely accepted genetic model for colorectal tumorigenesis proposes that colorectal tumours arise as a result of mutational inactivation of tumour suppressor genes coupled to the mutational activation of cellular oncogenes. The model proposes that at least four genes are required for the formation of a tumour and that it is the overall accumulation of mutations in specific genes such as DCC, APC, K-ras, p53 and the human mismatch repair genes that give rise to a tumour. (Fearon et al., 1990).

The increased understanding of the molecular mechanisms underlying the development of colorectal tumourigenesis will greatly improve advances in predictive

testing, diagnosis, gene therapy, chemoprevention and quality of life of those affected with the disease. An understanding of the genetic mechanisms associated with colorectal cancer in the Irish population will have major implications for clinical practice, early diagnosis of at risk individuals and may play a role in reducing the death rate due to colon cancer in this country.

Lii. BIBLIOGRAPHY

Akiyama, Y., Iwanaga, R., Ishikawa, T., Sakamoto, K., Nishi, N., Nihei, Z., Iwama, T., Saitoh, K., Yuasa, Y. (1996). *Mutations of the transforming growth factor-beta type II receptor gene are strongly related to sporadic proximal colon carcinomas with microsatellite instability.* Cancer. 78(12), 2478-2484.

Baeg. G.H., Matsumine, A., Kuroda, T., Bhattacharjee, R.N., Miyashiro, I., Toyoshima, K., Akiyama, T. *The tumour suppressor gene product APC blocks cell cycle progression from Go-G1 to S phase.* (1995). EMBO J. 14, 5618-5625.

Barbacid, M. (1987). *Ras Genes.* Annu Rev Biochem. 56, 779-827.

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

Bodmer, W.F., Bailey, C.J., Bodmer, J., Bussey, H.J., Ellis, A., Gorman, P., Lucibello, F.C., Murday, V.A., Rider, S.H., Scambler, P. et al., (1987*). Localization of the gene for familial adenomatous polyposis on chromosome 5.* Nature. 328, 614-616.

Bodmer, W., Bishop, T. and Karran, P. (1994). *Genetic steps in colorectal cancer.* Nat Genet. 6, 217-219.

Bos, J.L., Fearon, E.R., Hamilton, S.R., Verlaan-de Vries, M., van Boom, J.H., van der Eb, A.J., Vogelstein, B. (1987). *Prevalence of ras Gene Mutations in Human Cancers.* Nature 327, 293-297.

Bussey, H.J.R. (1975). *Familial Polyposis Coli.* Baltimore, MD: Johns Hopkins University Press;47-58.

Carder, P., Wyllie, A.H., Purdie, C.A. (1993). *Stabilised p53 facilitates aneuploid clonal divergence in colorectal cancer.* Oncogene. 8:1397-1401.

Chen, R.H., Ebner, R., Derynck, R. (1993). *Inactivation of the type II receptor reveals two receptor pathways for the diverse TGF-beta activities.* Science. 260, 1335-1338.

Cho, K.R. and Fearon, E.R. (1995). *DCC:Linking tumour suppressor genes and altered cell surface interactions in cancer ?* Curr Opin Genet Dev. 5, 72-78.

Fearon, E.R. and Vogelstein, B. (1990). *A Genetic Model for Colorectal Tumorigenesis.* Cell, 61, 759-767.

Forrester, K., Almoguera, C., Han, K., Grizzle, W.E., Perucho, M. (1987). *Detection of High Incidence of K-ras Oncogenes During Human Colon Tumorigenesis.* Nature, 327, 298-303.

Hollstein, M., Sidransky, D., Vogelstein, B., Harris, C.C. (1991). *p53 mutations in human cancers.* Science 253, 49-53.

Keino-Masu, K., Masu, M., Hinck, L., Leonardo, E.D., Chan, S.S., Culotti, J.G., Tessier-Lavigne, M. (1996). *Deleted in Colorectal Cancer (DCC) encodes a netrin receptor.* Cell. 87,175-85.

Kinzler, KW., Nilbert, MC., Su, LK., Vogelstein, B., Bryan, TM., Levy, DB., Smith, KJ., Preisinger, AC., Hedge, P., McKechnie D, et al., (1991). *Identification of FAP locus genes from chromosome 5q21.* Science, 253, 661-665.

Leach, F.S., Nicolaides, N.C., Papadopoulos, N., Liu, B., Jen, J., Parsons, R., Peltomaki, P., Sistonen, P., Aaltonen, L.A., Nystrom-Lahti, M., et al. (1993). *Mutations of a mutS homolog in hereditary nonpolyposis colorectal cancer.* Cell 75(6), 1215-1225.

Loeb, L. (1991). *Mutator Phenotype May Be Required for Multistage Carcinogenesis.* Cancer Res. 51, 3075-3079.

Lynch, H.T., Smyrk, T.C., Watson, P., Lanspa, S.J., Lynch, J.F., Lynch, P.M., Cavalieri, R.J., Boland, C.R. (1993). *Genetics, natural history, tumor spectrum, and pathology of hereditary nonpolyposis colorectal cancer: an updated review.* Gastroenterology 104(5), 1535-1549.

Markowitz, S., Wang, J., Myeroff, L., Parsons, R., Sun, L., Lutterbaugh, J., Fan, R.S., Zborowska, E., Kinzler, K.W., Vogelstein, B., et al., (1995). *Inactivation of the type II TGF-beta receptor in colon cancer cells with microsatellite instability.* Science. 268, 1336-1338.

Mehlen, P., Rabizadeh, S., Snipas, S.J., Assa-Munt, N., Salvesen, G.S., Bredesen, D.E. (1998). *The DCC gene product induces apoptosis by a mechanism requiring receptor proteolysis.* Nature. 395, 801-804.

Miyoshi, Y., Nagase, H., Ando, H., Horii, A., Ichii, S., Nakatsuru, S., Aoki, T., Miki, Y., Mori, T., Nakamura, Y. (1992). *Somatic mutations in the APC gene in colorectal tumours: mutation cluster region in the APC gene.* Hum Mol Genet. 1, 229-233.

Nagase, H., Nakamura, Y. (1993). *Mutations of the APC gene.* Hum Mutat. 2, 425-434.

Nishisho, I., Nakamura, Y., Miyoshi, Y., Miki, Y., Ando, H., Horii, A., Koyama, K., Utsunomiya, J., Baba, S., Hedge, P. (1991). *Mutations of chromosome 5q21 genes in FAP and colorectal cancer patients* . Science. 253, 665-669.

Nicolaides, N.C., Papadopoulos, N., Liu, B., Wei, Y.F., Carter KC, Ruben, S.M., Rosen, C.A., Haseltine, W.A., Fleischmann, R.D., Fraser, C.M., et al., (1995). *Mutations of Two PMS Homologues in Hereditary Nonpolyposis Colon Cancer.* Nature 371(6492), 75-80.

Nowell, P.C. (1976). *The clonal evolution of tumor cell populations.* Science.194, 23-28.

O'Sullivan, M.J., Mulcahy, T.M., Cambell, J., O'Suilleabhain, C.B., Kirwan, W.O., Doyle, C.T., McCarthy, T.V. (1989). *Detection of five novel germline mutations of the APC gene in Irish familial adenomatous polyposis families.* Hum Mutat. S 1(6), S251-S253.

Papadopoulos, N., Nicolaides, N.C., Wei, Y.F., Ruben, S.M., Carter, K.C., Rosen, C.A., Haseltine, W.A., Fleischmann, R.D., Fraser, C.M., Adams, M.D, Venter, C., Hamilton, S., Petersen, G., Watson, P., Lynch, H.T., peltomaki, P., Mecklin, J-P., de la Chapelle, A., Kinzler, K., Vogelstein, B. (1994). *Mutation of a mutL homolog in hereditary colon cancer.* Science 263(5153), 1625-1629.

Peifer. M. (1993). *Cancer , catenins and cuticle pattern: a complex connection.* Science. 262, 1731- 1743.

Powell, S.M., Petersen, G.M., Krush, A.J., Booker, S., Jen, J., Girdiello, F.M., Hamilton, S.R., Vogelstein, B., Kinzler,K.W. (1993). *Molecular Diagnosis of Familial Adenomatous Polyposis.* N Engl J Med. 329, 1982-1987.

Prives, C., Manfredi, J.J. (1993). *The p53 tumour suppressor protein; meeting review.* Genes Dev. 7, 529-534.

Roberts, A.B and Sporn, M.B. (1990). *Peptide growth factors and their receptors.* Springer-Verlag. Berlin. pp 420-427.

Shivapurkar, N., Huang, L., Ruggeri, B., Swalsky, P.A., Bakker, A., Finkelstein, S., Frost, A, Silverberg, S. (1997). *K-ras and p53 mutations in aberrant crypt foci and colonic tumours from colon cancer patients.* Cancer Lett. 115, 39-46.

Su, L-K., Vogelstein, B., Kinzler, K. (1993). *Association of the APC tumour suppressor protein with catenins.* Science. 262, 1734-1737.

Utsinomiya, J. and Lynch, H.T. (1990). *Hereditary colon cancer.* Berlin, Germany: Springer-Verlag.

Vogelstein, B., Fearon, E.R., Hamilton, S.R., Kern, S.E,, Preisinger, A.C., Leppert, M., Nakamura, Y., White, R., Smits, A.M., Bos, J.L. (1988). *Genetic Alterations During Colorectal-Tumor Development.* N Engl J Med. 319(9), 525-532.

Vogelstein, B., Kinzler, K.W. (1992). *p53 function and dysfunction*. Cell. 70, 523-526.

Watson, P., Lynch, H.T. (1994). *The tumor spectrum in HNPCC.* Anticancer Res. 14(4B), 1635-1639.

Yonish-Rouach, E., Resnitzky, D., Lotem, J., Sachs, L., Kimchi, A., Oren, M. (1991). *Wild-type p53 induces apoptosis of myeloid leukaemic cells that is inhibited by interleukin-6.* Nature. 352, 345-347.

CHAPTER 1

MUTATION SCREENING OF THE hMSH2 AND hMLH1 GENES IN IRISH HNPCC FAMILIES

1.1 INTRODUCTION

1.1.(I) HEREDITARY NON-POLYPOSIS COLORECTAL CANCER.

HNPCC is an autosomal dominantly inherited colorectal cancer predisposition syndrome characterised by early onset (<45 years), colorectal tumours located predominantly in the proximal colon (70% of cases) and with an excess of synchronous and metachronous colonic cancers (Lynch *et al.*, 1993). HNPCC tumours display an unfavourable histology in that they are poorly differentiated, however prognosis tends to be better than in sporadic colon carcinomas. Prolonged survival in HNPCC patients is associated with the diploid status of tumour cells (Frei, 1992).

The HNPCC syndrome (originally cancer family syndrome) was first recognised by Warthin in 1895, when an aggregation of cancer of the colon, stomach and endometrium in a large pedigree known as family G was noted (Warthin, 1913). However, the hereditary significance of the disease was not elucidated until 35 years later, when Lynch and colleagues defined the clinical features and inheritance profile of the disease (Lynch *et al.*, 1993). HNPCC, also known as Lynch syndromes I and II, does not display the classical FAP like pre-malignant polyps (see 'General Introduction' for a more detailed account of FAP), rather there is what appears clinically to be a progression from normal non neoplastic mucosa to adenoma (pre-cancerous growth) and ultimately to tumour (Lanspa *et al.*, 1990). As there are no pre-neoplastic indicators of tumour development early diagnosis of HNPCC has proved difficult.

The disease is sub-classified as Lynch syndrome I or II based on the presence or absence of extra-colonic tumours. Lynch type I pedigrees display the characteristics outlined above, with the lack of extra-colonic tumours. Lynch II syndrome pedigrees display all the Lynch I type clinical features in addition to a significantly elevated

frequency of endometrial tumours and an associated modest increase in the frequency of uterine, ovarian, stomach, pancreas and small bowel tumours (Watson and Lynch, 1994). Recent investigations of tumour type and frequency in Lynch I and II cases suggest that there is discordance among researchers on the necessity of this subdivision. Analysis of the occurrence of non-colorectal tumours in 40 Finnish HNPCC pedigrees with and without endometrial tumours showed equal frequencies of extra-colonic tumours. The only significant difference between endometrial and non-endometrial tumour families was that there were more affected members and specifically more affected female members within the endometrial cancer families (Mecklin and Jarvinen, 1991).

INCIDENCE

The issue of incidence of HNPCC is actively debated with values varying greatly from 1 in 200 to as low as 1 in 15,000 in the Western population (Bodmer *et al.*, 1994). Epidemiological studies have demonstrated that 1-5% of all colorectal cancers fully satisfy the criteria for HNPCC (Mecklin, 1987; Ponz De Leon, 1989; Kee *et al.*, 1991(a)). The variation in incidence rates is most likely due to lack of objective diagnostic criteria and lack of accurate clinical data on family history of the disease. However, with the rapid advancement in HNPCC research and the ongoing establishment of national database networks, more accurate incidence values may soon be available.

HNPCC PEDIGREES: CRITERIA FOR SELECTION.

The minimum criteria for the selection of a HNPCC pedigree were agreed upon at the second meeting of The International Collaborative Group on HNPCC in

Amsterdam in 1990 (Vasen *et al.*, 1991). The minimum criteria known as 'The Amsterdam Criteria' are :

1)There should be at least three family members with a histologically verified colorectal cancer, one a first degree relative of the other two.

2)There must be at least two successive generations affected.

3)At least one individual must be less than 50 years of age at the time of diagnosis.

4)Familial adenomatous Polyposis Coli (FAP) must be ruled out.

These criteria serve to eliminate most cases of chance clustering. There is however, disagreement among researchers as to the selectivity of this criteria. Many believe it is too stringent as it does not take into account family members with endometrial and other extra-colonic tumours. Furthermore due to the small sizes of western families there is a limited probability of three cases of colorectal cancer occurring in these pedigrees and it excludes cases of reduced penetrance and *de Novo* germline mutations. This may lead to under-diagnosis of HNPCC and some research groups are now using less strict criteria for the selection of their families (Beck *et al.*, 1997). The Amsterdam criteria however, remain the standard criteria for pedigree selection in genetic based screening studies. A recent publication by the ICG-HNPCC analyses the mutations spectra of 202 HNPCC pedigrees of diverse ethnic and geographical backgrounds (Peltomaki, 1997). The majority of these pedigrees (82%) strictly satisfy the Amsterdam criteria and results from mutation screening in these families show that 85% of all identified mutations have been confined to hMSH2 and hMLH1. Although many mutations in hMSH2 and hMLH1 have been identified in HNPCC pedigrees which satisfy the Amsterdam criteria, little is known about the involvement of these genes in pedigrees which do not satisfy all the criteria. It has been suggested that the Amsterdam criteria confer a selection bias towards pedigrees harbouring hMSH2 and hMLH1 mutations and that smaller pedigrees with less affected members and later age of onset may

harbour mutations in other mismatch repair genes such as hPMS1, hPMS2 or in cell growth regulators such as TGFβ–RII and others (Wijnen *et al.*, 1997).

CLINICAL ASPECTS OF HNPCC

Double contrast barium enema is the standard method used to clinically diagnose colorectal carcinoma. Patients usually present with a change in bowel habit and in many cases, abdominal bleeding. A typical carcinoma appears as an obstruction of the colon and has a characteristic apple core-like appearance. In many cases this visual characteristic is not obvious and confirmation requires colonoscopy, biopsy and histological verification. Treatment is usually by colonic resection. Histological assessment of the carcinoma and mesenteric lymph nodes is used to stage the tumour and define prognosis. Dukes staging of colonic tumours is the most commonly used scheme. Dukes A tumours are confined to the bowel wall. Tumours that have infiltrated the bowel wall in the absence of lymph node metastases are classified as Dukes stage B. Once the regional lymph nodes become involved and liver metastases are observed the carcinoma is classed as Dukes C. Dukes A carcinomas carry a mean five year survival rate of 85%, Dukes B is associated with a decreased value of 65% after colonic resection while the prognosis of Dukes stage C with liver metastases is less than 1% (Pounder *et al.*, 1989).

GENETIC ASPECTS OF HNPCC

The HNPCC syndrome was first recognised by Warthin in 1895 when an aggregation of colon, endometrium and stomach cancer was noted in a large pedigree. Since its initial documentation, the clinical aspects of HNPCC have been well elucidated, predominantly through the work of Lynch and colleagues as early as 1966

(Lynch *et al.*, 1966). However, the first clue to the genetics of the disease was not discovered until as late as 1993, 80 years after Warthin described the first HNPCC pedigree.

Mutations in four human DNA mismatch repair genes have been shown to segregate with predisposition to colorectal carcinoma in HNPCC pedigrees. The first HNPCC susceptibility locus was mapped to chromosome 2p16 via linkage analysis in two large colon cancer pedigrees, a maximum lodscore for linkage of the disease to the microsatellite marker D2S123 of 6.39 at a recombination fraction of θ=0.00 was obtained (Peltomaki *et al.*, 1993(a)). Initial investigations into the putative tumour suppressor role of the gene product, combined with attempts to clone and characterise the gene resulted in the observation that DNA from tumours of HNPCC patients displayed widespread alterations in short repetitive DNA sequences (microsatellites) when compared with DNA from normal, non-neoplastic tissues in the same patient (Aaltonen *et al.*, 1993; Aaltonen *et al.*, 1994; Ionov , 1993).

Simultaneously, researchers involved in yeast and bacterial genetics noticed that the microsatellite instability observed in HNPCC tumours was similar to the characteristic instability exhibited by yeast and bacterial cells known to be defective in their DNA mismatch repair process (Fishel *et al.*, 1993; Strand *et al.*, 1993; Thibodeau *et al.*, 1993). Positional cloning identified the first human homologue of bacterial MutS and the yeast MSH2 gene - known as hMSH2 (human MutS homologue) (Fishel *et al.*, 1993; Leach *et al.*, 1993). The involvement of the gene in HNPCC susceptibility was elucidated based on the identification of inherited germline mutations which segregated with the disease within HNPCC kindreds (Leach *et al.*, 1993).

The identification of the hMSH2 gene resulted in a rapid search for human homologues of the other known bacterial and yeast mismatch repair genes. A homologue of MutL was identified by linkage in HNPCC pedigrees to chromosome 3p21 (Lindblom *et al.*, 1993). The gene was cloned based on its suspected homology to *E. coli* MutL and yeast MLH1 (Papadopoulos et *al*, 1994., Bronner *et al*, 1994).

The involvement of hMLH1 in HNPCC was proved as with hMSH2, based on the identification of heritable germline mutations which segregated with the disease in affected families (Papadopoulos *et al.*, 1994; Bronner *et al.*, 1994). Two further MutL homologues were identified based on analysis of expressed sequence tags (Papadopoulous *et al.*, 1994). The cloned genes were named hPMS1 (chromosome 2q31-33) and hPMS2 (chromosome 7p22) and displayed heritable mutations in HNPCC pedigrees (Nicolaides *et al.*, 1994). Table 1.1 details the characteristics of the human mismatch repair genes associated with HNPCC.

Other human homologues of the mismatch repair proteins have been identified, in addition to hMSH2 there are three other human MutS homologues. Duc 1 (hMSH3) is the product of a divergent transcript from the dihydrofolate reductase promoter and forms a complex with hMSH2 known as hMutS β (Fujii and Shimada, 1989) while hMSH4 is similar in sequence to *E. coli* MutS (Ross-Macdonald *et al.*, 1994). The function of these two homologues is unknown and neither has been implicated in the disease. GTBP a 160 kDa protein with mutS homology has been shown to bind to hMSH2 forming what is known as the hMutSα complex which participates in mismatch binding (Drummond *et al.*, 1995; Palombo *et al.*, 1995).

1.1.(II) MISMATCH REPAIR IN YEAST AND BACTERIA

Pairing errors in which Watson-Crick bases occur in a non-complementary order within a DNA helix are known as mismatches or mispairs. Mismatched bases arise as a result of incorrect recombination or DNA replication errors, and have major consequences for protein synthesis, cell viability and ultimately evolution. Enzyme systems which detect and correct mismatches have been identified in prokaryotes and eukaryotes. These repair systems have a major role in increasing the fidelity of replication and of preventing drastic alterations in subsequent cellular generations. The

GENE	LOCUS	NO. OF EXONS	PROTEIN SIZE	% MUTATIONS
hMSH2	2p16	16	934 amino acids	48/126 (38%)
hMLH1	3p21	19	756 amino acids	75/126 (60%)
hPMS1	2q32	unknown	932 amino acids	1/126 (0.8%)
hPMS2	7p22	15	862 amino acids	2/126 (1.6%)

Table 1.1. Human mismatch repair gene characteristics: Chromosome position, protein size and the number of exons in each gene are detailed. The total number of mutations was calculated based on a combination of information from the ICG-HNPCC and HGDB-Cardiff databases.

best studied mismatch repair process with respect to both genetics and biochemistry is the *E. coli* methyl directed Mut HLS system. The pathway detects and corrects mismatched nucleotides in nascent DNA by employing a number of enzymes, some of which have sequence homologies to the human mismatch repair enzymes associated with HNPCC (MutS and MutL).

MISMATCH REPAIR IN *E. COLI*

The identification and correction of mispaired nucleotides (from single mispairs to loops of up to 4 nucleotides) in newly replicated DNA, in *E. coli*, relies on a large number of cellular factors. At least 10 proteins have been implicated in the methyl directed repair process: Mut H, Mut S, Mut L, DNA helicase II (which plays a role in DNA unwinding), DNA Polymerase III(holoenzyme), DNA ligase, single stranded binding protein (SSB) and either Exonuclease I, Exonuclease VII or RecJ (exonuclease)-depending on the position of the mismatch from d(GATC) (Cooper *et al.*, 1993; Grilley *et al.*, 1993). Also required are the four dNTP's and the co-factors NAD and ATP (Lahue *et al.*, 1989; Modrich, 1991; Modrich *et al.*, 1994).

Initiation of mismatch repair in *E. coli* is activated by the presence of a mismatch in newly replicated DNA. MutS, a 97kD protein binds to mismatched bases independently of any other factors (Su and Modrich, 1986;Su *et al.*, 1988). MutH has been shown to interact at d(GATC) sites that are hemimethylated at adenine residues and to incise the unmethylated strand. This methylation sensitive selection process confers strand specificity and ensures that only the newly replicated strand is corrected (Welsh *et al.*, 1987). The specific activity of MutL remains to be clearly elucidated, however, it is thought to bind the MutS:DNA mismatch complex and to facilitate the interaction between this complex and MutH at the point of DNA incision (d(GATC)

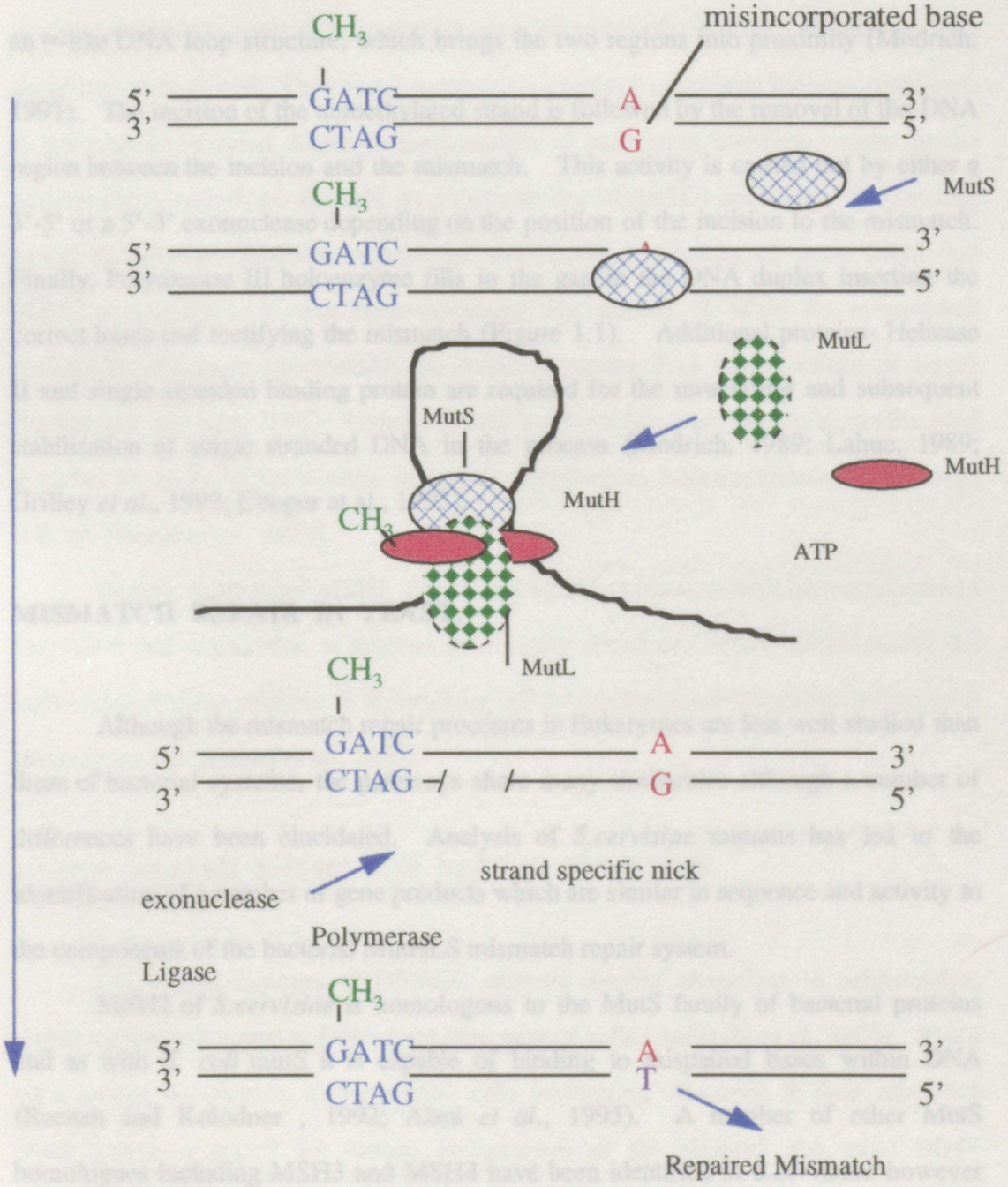


Figure 1.1 Diagrammatic representation of the mismatch repair mechanism in *E. coli*. The MutS protein binds to mismatched base pairs in an ATP-dependent reaction. A Mut S-Mut L-Mut H complex is formed, which interacts with hemimethylated GATC sites. Mut H incises the unmethylated strand, exonucleases, DNA polymerase and ligase strip back and repair the DNA. DNA helicase II and single stranded binding protein are also required for the unwinding and stabilisation of the single stranded DNA

sites). These incision sites may lie up to 1Kb from the mismatch position (Bruni *et al.*, 1988). The interaction of MutH with MutS is thought to occur due to the formation of an α-like DNA loop structure, which brings the two regions into proximity (Modrich, 1991). The incision of the unmethylated strand is followed by the removal of the DNA region between the incision and the mismatch. This activity is carried out by either a 3'-5' or a 5'-3' exonuclease depending on the position of the incision to the mismatch. Finally, Polymerase III holoenzyme fills in the gap in the DNA duplex inserting the correct bases and rectifying the mismatch (Figure 1.1). Additional proteins- Helicase II and single stranded binding protein are required for the unwinding and subsequent stabilisation of single stranded DNA in the process (Modrich, 1989; Lahue, 1989; Grilley *et al.*, 1993; Cooper at al., 1993).

MISMATCH REPAIR IN YEAST.

Although the mismatch repair processes in Eukaryotes are less well studied than those of bacterial systems, the pathways share many similarities although a number of differences have been elucidated. Analysis of *S.cervisiae* mutants has led to the identification of a number of gene products which are similar in sequence and activity to the components of the bacterial MutHLS mismatch repair system.

MSH2 of *S.cervisiae* is homologous to the MutS family of bacterial proteins and as with *E. coli* mutS it is capable of binding to mispaired bases within DNA (Reenan and Kolodner , 1992; Alani *et al.*, 1995). A number of other MutS homologues including MSH3 and MSH4 have been identified in *S.cervisiae* however they have not been shown to play a major role in mismatch repair (New *et al.*, 1993; Ross- MacDonald *et al.*, 1994). PMS1 (postmitotic segregation gene), (Kramer *et al.*, 1989) and MLH1 (Strand *et al*, 1993) in yeast are both homologues of bacterial MutL

and can form a complex that binds DNA bound MSH2 in a similar manner to mutL and MutS in *E. coli.* (Prolla *et al.*, 1994). A significant difference between Eukaryotic and Prokaryotic mismatch repair is the lack of a MutH homologue in Eukaryotic systems - strand specificity in yeast as in humans is not governed by the methylation state of DNA sequences as in *E. coli.* The precise mechanism of strand specificity in these organisms is not well understood, however the presence of replication associated nicks in the nascent DNA strand are thought to substitute for d(GATC) sequences in directing strand specificity. Differences between the binding specificity of yeast MutS and MSH2 in humans have also been elucidated -the affinity of MutS for mismatched nucleotides decreases considerably in stretches greater than 4 nucleotides in length (Parker and Marinus, 1992). Human MSH2 however has been shown to readily bind stretches of up to 14 nucleotides in length. (known as insertion/deletion type loops IDL's), (Fishel *et al.*, 1994).

Finally, *S.cervisiae* MSH2, PMS1 and MLH1 mutants exhibit increased rates of expansion and contraction at microsatellite repeat sequences which demonstrates the importance of these proteins in maintaining correct transmission of genetic information from cell to cell (Strand *et al.*, 1993).

1.1.(III) HUMAN MISMATCH REPAIR PROTEINS

DNA from tumours of HNPCC patients display alterations in microsatellite repeat regions compared to DNA from their normal cells (Aaltonen *et al.*, 1993). Based on a similar observation in *S.cervisiae* cells harbouring mutations in their mismatch repair genes, a search for human homologues of the bacterial and yeast MMR (Mismatch Repair) proteins was initiated. Using a variety of methods including, linkage analysis (Lindblom *et al*, 1993) cloning -based on suspected sequence homologies to bacterial proteins, and the analysis of expressed sequence tags (Papadopoulos *et al.*, 1994) a number of homologues of bacterial and yeast mismatch

repair genes were identified in humans. Four human MutS homologues have been identified of which two (hMSH2 and hMLH1) are known to have a specific role in MMR. hMSH2 acts similarly to *E. coli* MutS and yeast MSH2 by binding mismatched nucleotides. As in yeast, binding of hMSH2 to mismatches is not directed by methylation of d(GATC) sequences rather pre-existing replication associated nicks in the DNA are thought to direct binding (Fang *et al.*, 1993). hMSH6, more commonly known as p160 or GTBP (for G/T binding protein) was identified based on its ability to bind G/T mismatches in vitro. This 160kD protein cannot however, bind G/T mismatches independently, the presence of hMSH2 is required (Palombo *et al.*, 1995). The hMSH2:GTBP complex known as hMutS α purifies as a 100kD:160kD activity which restores mismatch repair activity in LoVo cell extracts (cells deficient in hMSH2 activity), (Drummond *et al.*,1995). Mobility-shift experiments have demonstrated that the hMutS α complex can bind G/T or 1-2 nucleotide loops with high affinity, this however decreases by 1/10th with A/T homoduplex DNA and with larger mismatch loops (Karran *et al.*, 1995). Figure 1.2 outlines the possible mechanism for mismatch correction by MutS α.

Cell lines deficient in GTBP activity display alterations primarily in mononucleotide repeat tracts, which suggests that GTBP may have a role in simple base: base correction (Papadopoulos *et al.*, 1995). hMSH2 on the other hand, can independently bind loops of up to 14 nucleotides in length and is thought to be responsible for correcting more dramatic mismatches.

More recently another complex comprising of hMSH2 and hMSH3 (DUC-1) known as hMutS β has been shown to carry out mismatch correction activities similar to the hMUTSα complex (Polambo *et al.*, 1996). It would appear that the roles of GTBP and DUC-1 in MMR are overlapping and somewhat confined to small single or

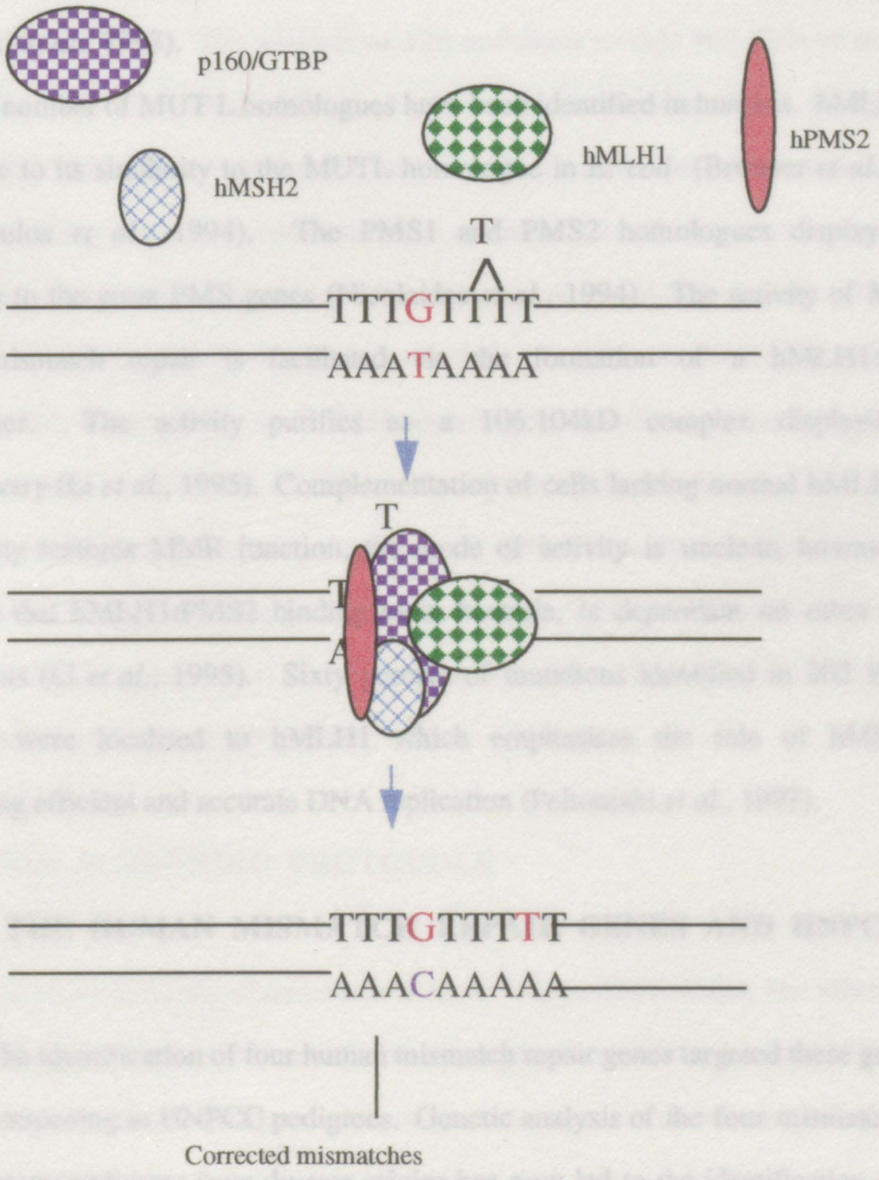


Figure 1.2. Possible mechanism for the correction of mispairs and displaced one or two base pair loops, generated during replication in humans (adapted from Karran et al., 1995).

dinucleotide loop repair, while the role of hMSH2 is more significant as it is required to repair larger more dramatic mismatch stretches. This fact is borne out somewhat by noting that no mutations in either GTBP or DUC-1 have been identified in HNPCC patients while nearly 38 of 126 identified mutations have been confined to hMSH2 (Peltomaki *et al.*, 1997).

A number of MUT L homologues have been identified in humans. hMLH1 was named due to its similarity to the MUTL homologue in *E. coli* (Bronner *et al.*, 1994; Papadopoulos *et al.*, 1994). The PMS1 and PMS2 homologues display closer homology to the yeast PMS genes (Nicolaides *et al.*, 1994). The activity of MutL in human mismatch repair is facilitated via the formation of a hMLH1:hPMS2 heterodimer. The activity purifies as a 106:104kD complex displaying 1:1 stoichiometry (Li *et al.*, 1995). Complementation of cells lacking normal hMLH1 with this activity restores MMR function, the mode of activity is unclear, however it is suspected that hMLH1:PMS2 binding as in bacteria, is dependant on other cellular components (Li *et al.*, 1995). Sixty percent of mutations identified in 202 HNPCC pedigrees were localised to hMLH1 which emphasises the role of hMLH1 in maintaining efficient and accurate DNA replication (Peltomaki *et al.*, 1997).

1.1.(IV) THE HUMAN MISMATCH REPAIR GENES AND HNPCC

The identification of four human mismatch repair genes targeted these genes for mutation screening in HNPCC pedigrees. Genetic analysis of the four mismatch repair genes in many pedigrees from diverse origins has now led to the identification of more than 120 germline mutations. Two mutations have been identified in hPMS2 and 1 in hPMS1, the low mutation rate of these genes in HNPCC may be due to the fact that for the most part, researchers have limited their screening to hMSH2 and hMLH1 and confined their analysis to pedigrees which strictly satisfy the Amsterdam selection

criteria (Peltomaki *et al.*,1997). Initially it was believed that mutations in hMSH2 were responsible for the majority of HNPCC cases with values of >60% for hMSH2 and only 30% for hMLH1 (Lui *et al.*, 1994; Han *et al.*, 1995). However, a more accurate evaluation of mutation spectra is documented in a recent report by the ICG-HNPCC which presents the results on the mutation analysis of 202 pedigrees of diverse ethnic and geographical origin. The analysis of 126 mutations reveals that 38% of mutations affect hMSH2 and 59% affect hMLH1. Further analysis of the mutation spectra of the 202 HNPCC kindreds reveals that the majority of mutations are point mutations with 59% representing single base substitutions as opposed to deletions or insertions (41%). Most mutations identified are unique with no obvious mutation hotspots, however in hMSH2 the most frequently altered exons include 12 (17% of hMSH2 mutations), 7 (15%), 5 (9%), and 6 (9%). In hMLH1 mutations occurred more frequently in exons 16 (15%), 9 and 13 (9%) and 19 (8%). The report represents the efforts of the ICG-HNPCC to establish an international database of HNPCC mutations. The database was established to provide information to researchers to facilitate evaluation of mutation spectra of HNPCC genes and to assess geographic and population variation in mutation spectra (Peltomaki *et al.*, 1997), (Table 1.2).

MUTATION SCREENING PROTOCOLS

Due to the diversity of mutations in terms of type and position, the identification of a causative mutation within an affected pedigree is a tedious process. The most widespread protocol for pedigree analysis involves the initial selection of pedigrees based on the Amsterdam criteria, followed by an evaluation of RER (replication error) phenotype via assessment of microsatellite instability. A positive RER phenotype suggests that the mismatch repair process in these patients may be defective. A number of screening methods have been employed in the identification of mutations in HNPCC pedigrees world-wide. Initially, SSCP and DGGE (Orita *et al.*, 1989a; Sheffield *et al.*,

GENE	MUTATION TYPE	MUTATIONS IDENTIFIED	TOTAL
hMSH2	Transitions	12	
	Transversions	9	
	Deletions	19	
	Insertions	8	48
hMLH1	Transitions	30	
	Transversions	19	
	Deletions	16	
	Insertions	8	74
hPMS1	Transition	1	1
hPMS2	Transition	1	
	Deletion	1	2

Table 1.2. Mutation profile of the human mismatch repair genes. Mutations have been identified in four human mismatch repair genes. Mutation type, number of mutations and total mutations are provided for each gene. The information used to construct this table was obtained from ICG-HNPCC database and the HGDB at Cardiff.

1989; Traystman *et al.*, 1990) analyses were the most commonly used mutation detection methods, however, when it was determined that a large proportion (>80%) of the mutations resulted in the generation of a stop codon and the subsequent production of a truncated protein product, many researchers chose to use the *in vitro* protein synthesis assay (IVPS) or protein truncation test (PTT) as the method of analysis (Roest *et al.*, 1993: Powell *et al.*, 1993). The employment of this method in HNPCC analysis however, has proved troublesome due to the fact that there are a number of alternatively spliced transcripts of the hMLH1 gene (Charbonnier, 1995). Analysis using the PTT results in false positive results, in that what is believed to be a defective truncated protein product is merely the product of a naturally occurring alternative transcript of the mRNA for this gene. This transcript is normally not translated in vivo, but because the PTT is designed to selectively amplify segments of mRNA, they are translated in the PTT. With the advancement in automated sequencing, direct sequencing of genes has also proved successful in identifying mutations within HNPCC genes.

THE HUMAN MISMATCH REPAIR GENES - STRUCTURE AND FUNCTION

The hMSH2 gene contains approximately 73kb of genomic DNA with a 2727bp open reading frame. The gene contains 16 exons and encodes a 934 amino acid protein with 41% homology to the yeast MSH2 gene (Figure 1.3). The intron-exon boundaries have also been determined (Fishel *et al.*, 1993; Kolodner *et al.*, 1994b).

The hMLH1 gene contains approximately 58kb of genomic DNA with an open reading frame of 2525 nucleotides. The gene encodes a 756 amino acid protein and contains 19 exons (Figure 1.4). As with hMSH2 the intron-exon boundaries have also been determined (Kolodner *et al.*, 1995).

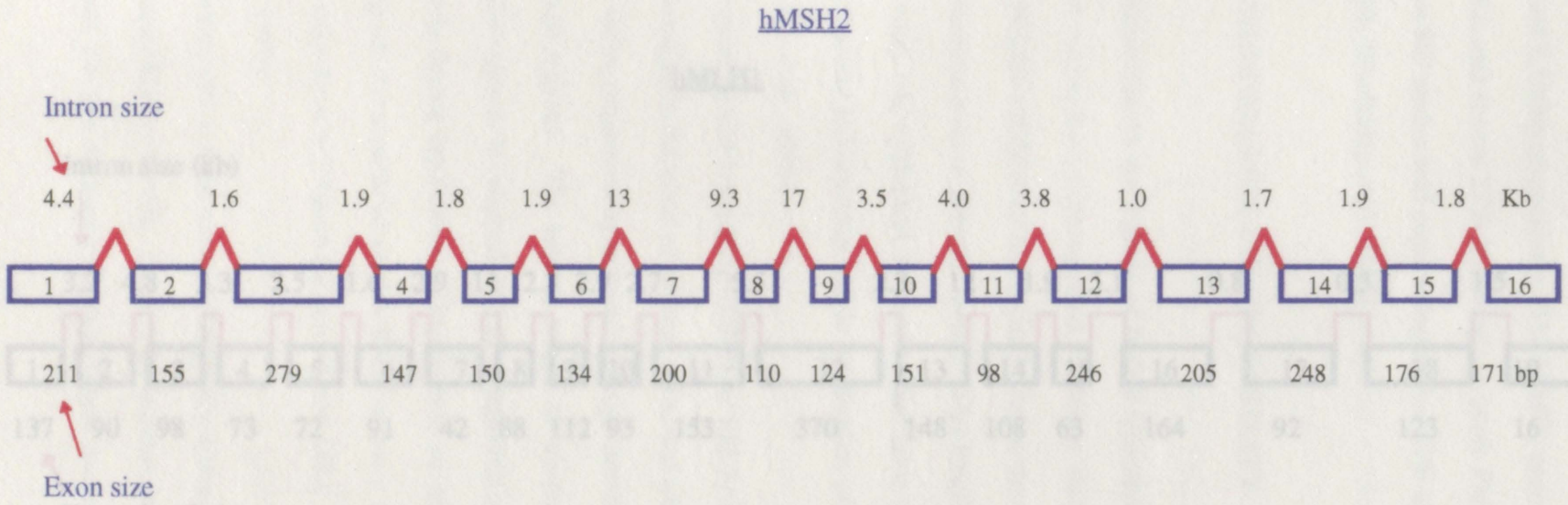


Figure 1.3. Diagram of the organisation of the hMSH2 gene. Boxes 1-16 represent the individual exons. The size is given below each exon, and the size of each intron is given above the region between the exons. Adapted from Kolodner et al., 1994.

Figure 1.4. Diagram of the organisation of the hMLH1 gene. Boxes 1-19 represent the individual exons. The size is given below each exon, and the size of each intron is given above the region between the exons. Adapted from Kolodner et al., 1995.

The open reading frame sequences of hPMS1 and hPMS2 have been determined as 2,795bp and 2,586bp respectively. The hPMS1 gene encodes a protein of 932 amino acids and shows 27% homology to yeast PMS1. hPMS2 contains 15 exons, encodes an 862 amino acid peptide and shows 32% identity to yeast PMS1 (Nicolaides *et al.*, 1994; Nicolaides *et al.*, 1995)

1.1.(V) MICROSATELLITE INSTABILITY

Microsatellites are regions of highly repetitive DNA sequences ((CA)n,(GT)n, and polyA tracts) found interspersed randomly throughout the genome. There are between 50,000-100,000 (CAn) repeats within the human genome, the majority are located within introns or between genes. The lengths of many of these repeats vary naturally from one individual to another and have proved very useful as a tool for following genetic transmission and in localisation of disease causing regions within a gene (Weber *et al.*, 1989).

Slipped strand mispairing (SSM) is widely believed to be the mechanism by which alterations in the sizes of repeats occur and plays a vital role in the evolution of repetitive DNA sequences by generating large numbers of short frameshift mutations within these regions. The tandem nature of simple repeat sequences is thought to render them prone to slipped strand mispairing and hence insertion or deletion mutagenesis during replication. The process of slipped strand mispairing involves mispairing of two strands of duplex DNA within a tandem repeat area, the mispaired duplexes may serve as primers for subsequent DNA synthesis, thus duplication occurs (Kunkel, 1993).

Examination of the mutation frequency within a Poly CA sequence inserted into bacteriophage M13 in *E. coli*, showed an increase in frameshift mutation frequency of greater than 1% in a 40 base pair tract, when compared to random non-repeat DNA.

This augmentation in frequency increases more than 13 fold in cells deficient in methyl-directed mismatch repair (MutS$^-$, MutL$^-$ hosts) suggesting that mismatch repair plays a vital role in correcting frameshift mutations (Levinson and Gutman, 1987).

MICROSATELLITE INSTABILITY IN HNPCC

Tumours from HNPCC patients display increased instability at microsatellite repeat sequences when compared to the DNA from their normal non-neoplastic cells (Thibodeau *et al.*, 1993; Aaltonen *et al.*, 1993). Simultaneous to the discovery of the hMSH2 gene and associated mutations in HNPCC, researchers observed the phenomenon of microsatellite instability in tumour cells (Thibodeau *et al.*, 1993; Aaltonen *et al.*, 1993). These alterations observed as shifts in the electrophoretic mobility of CA dinucleotide repeat fragments, suggested that frameshift mutations/replication errors had been introduced during tumour development. Up to 71% of HNPCC colorectal tumours studied, showed alterations with three (CA)n markers, while sporadic colorectal tumours studied showed far lower frequency of instability (Aaltonen *et al.*, 1993). Alterations have also been observed in trinucleotide repeats and at polyA tract regions (Aaltonen *et al.*, 1993; Parsons *et al.*, 1995).

The observations of instability in microsatellite repeats in association with mismatch repair defects in HNPCC led to the proposal of a number of mechanisms for carcinogenesis. The "mutator phenotype" hypothesis suggests that some cells in cancer patients undergo an early mutation event which renders them susceptible to the accumulation of multiple mutations. If these mutations occur in essential regulatory genes such as tumour suppressor genes, oncogenes or cell cycle regulators, the cells become tumorigenic (Loeb, 1991). In HNPCC tumours, this mutator phenotype manifests as microsatellite instability (termed the replication error phenotype/ RER+ or

MI) as a result of defects in any of the mismatch repair genes (Aaltonen *et al.*, 1993; Parsons *et al.*, 1993; Loeb *et al.*, 1994).

Evidence to support this theory stems from the observation that mutations in any of hMSH2, hMLH1 or hPMS1 lead to 100-200 fold increases in microsatellite instability in yeast cells whereas mutations that alter the proof reading effect of DNA polymerases have little consequence, suggesting that DNA instability arises due to an inability to correct mutations caused by strand slippage during replication- the role of the mismatch repair genes (Strand *et al.*, 1993).

Furthermore, microsatellite instability has been observed at more than one locus in the tumour cells of patients and this instability can often be observed in the pre-tumourigenic adenoma stage (Ionov *et al.*, 1993). A recent study shows that aberrant crypt foci, which are microscopic colonic lesions, thought to be the earliest identifiable precursors of colon cancer, also display microsatellite instability (Heinen *et al.*, 1996). This suggests that microsatellite instability occurs at an early stage during tumour development and lends evidence to support the mutator phenotype hypothesis.

MICROSATELLITE INSTABILITY IN NON COLON TUMOURS IN HNPCC

Microsatellite instability has been observed in an number of other cancer types other than colorectal cancer. Due to the association of endometrial carcinomas with HNPCC, tumours of this type have been studied for microsatellite instability by a number of different groups. Risinger *et al*, report that in a study of 36 sporadic endometrial tumours 17% showed instability at microsatellite loci, whereas 75% of endometrial tumours from HNPCC pedigrees display microsatellite instability (Risinger *et al.*, 1993). Katabuchi *et al.*, analysed sporadic endometrial tumours for instability and those that were positive were subsequently analysed for mutations in the four known mismatch repair genes. Two of nine microsatellite instability positive tumours

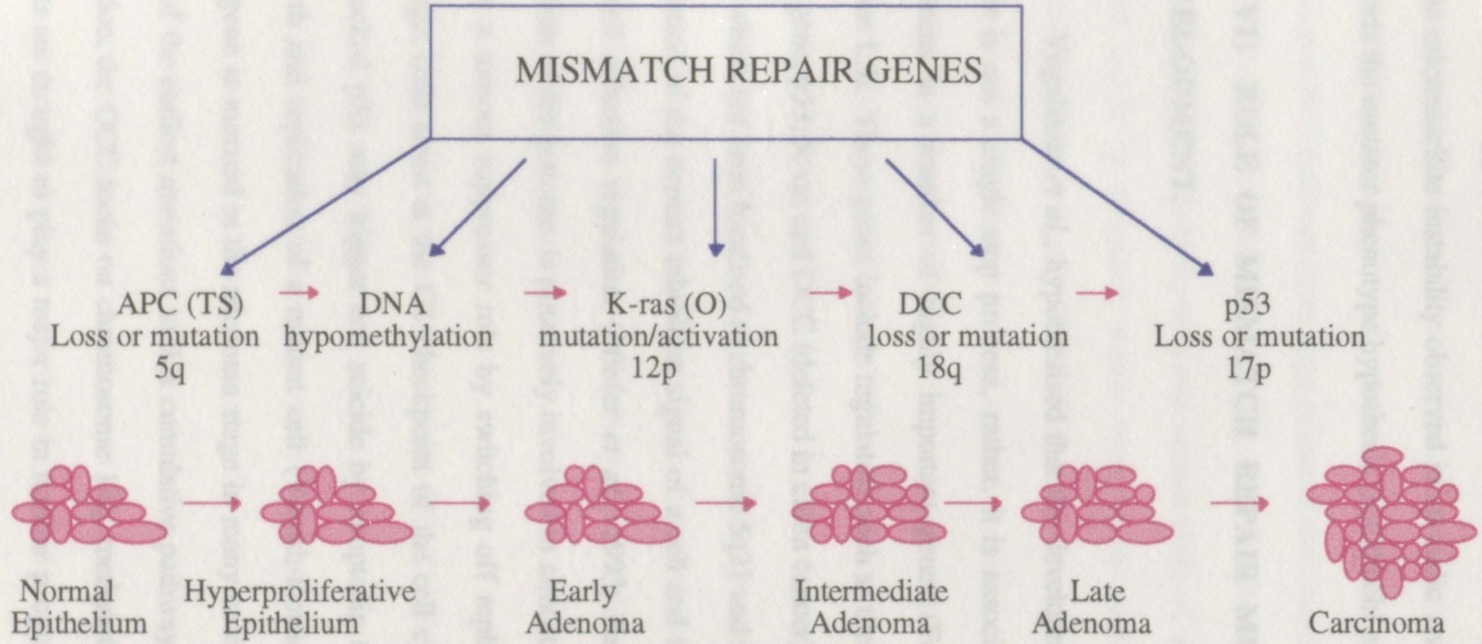


Figure 1.5: An adaptation of Fearon and Vogelstein's model for the development of colorectal cancer. (Fearon and Vogelstein, 1990). Gene mutations are displayed in parallel with histological changes. Mismatch repair genes are thought to play an indirect role in the tumorigenic pathway by increasing the overall mutation rate thus making each individual transition in the pathway more likely. Note: early adenomas; less than 1cm in size, intermediate; more than 1cm in size but without foci of carcinoma and late adenoma; more than 1cm but with foci of carcinoma. TS; tumour suppressor gene, O; oncogene.

showed mutations in hMSH2, these mutations were somatic and not germline mutations which would suggest that mutations in the mismatch repair genes are not responsible for the microsatellite instability observed in sporadic endometrial tumours and further supports the mutator phenotype hypothesis (Katabuchi *et al.*, 1995).

1.1.(VI) ROLE OF MISMATCH REPAIR MUTATIONS IN CANCER DEVELOPMENT.

Vogelstein *et al.*, hypothesised that the development and progression of colon cancer is not a single step process, rather, it is associated with the accumulation of mutations in a number of highly important genes (Fearon and Vogelstein, 1990), (Figure 1.5). These genes include regulators such as the APC (adenomatous polyposis coli) gene, p53, K-ras and DCC (deleted in colon cancer). APC is a tumour suppressor gene which has been localised to chromosome 5q21 and is believed to play a role in the regulation of the contact inhibition signal of a cell and is associated with microtubules and cell adhesion regulation (Pfiefer *et al.*, 1993; Baeg *et al.*, 1995). P53, the guardian of the genome is putatively involved in almost all cancers. Like APC it also exerts a tumour suppressor role by switching off replication during repair of DNA damage, cells arrest at the G1 checkpoint of the cell cycle, if DNA damage remains unchecked p53 may trigger cell suicide by apoptosis thereby preventing the further growth and replication of a mutant cell (Yonish-Rouach *et al.*, 1991). The K-ras oncogene is mutated at the adenoma stage in many colon cancers and is thought to be one of the earliest mutations in this cumulative pathway (Vogelstein *et al.*, 1988). In addition, the DCC locus on chromosome 18q, methylation status and other mutational events are thought to play a major role in tumour progression (Silverman *et al.*, 1989; Fearon *et al.*, 1990b). In FAP, germline mutations in the APC gene predispose individuals to the development of polyps and subsequently cancer. However, loss of

heterozygosity at the APC locus in addition to the accumulation of mutations in other cell cycle regulators are required for tumour development (Bodmer *et al.*, 1987). Similarly, an inherited mutation in one of the mismatch repair genes predisposes the individual to the development of colon cancer, however, a germline mutation alone is not directly involved in cancer development. HNPCC affected patients carry a germline mutation which inactivates one copy of one of their mismatch repair genes. Analysis of tissues in these individuals reveals that they are phenotypically normal (it has been demonstrated that cell lines, with one mutant allele of a mismatch repair gene exhibit normal levels of mismatch repair even under conditions of increased mutation frequency), (Branch *et al.*, 1993), they are however at increased risk of developing colorectal cancer. When a second somatic mutation in the MMR gene occurs, this inactivates both copies of the gene and mismatch repair activity becomes defective. A hypermutable state exists in these cells with the result that more and more mutations persist after each round of DNA replication. This hypermutable state is characterised by microsatellite instability, also known as an RER+ phenotype. This decreased ability to correct mutations is thought to result in an increased risk of accumulation of mutations in critical tumour suppressor genes, oncogenes and cell regulators with the subsequent development of a neoplastic growth and tumour development (Lazar *et al.*, 1994).

MISMATCH REPAIR GENES AND CELL CYCLE REGULATION

The development and progression of a tumour stems from an uncontrolled cell cycle. A mutant cell no longer grows, replicates or divides at the same rate as its normal neighbouring cells, furthermore, the apototic process which facilitates suicide and hence the elimination of the damaged cell may be defective or switched off, thus facilitating uncontrolled cell growth. Examination of the expression levels of MMR proteins during various stages of the cell cycle demonstrate that in normal cells these genes are expressed at all stages (constitutively expressed). Furthermore, the levels of

the hMLH1 protein increased by 50% at G1 and S phases -suggesting an upregulation in expression level of hMLH1. Similarly the level of hPMS1 increases at G1 and S phases while the hMSH2 level remains constant (Myers *et al.*, 1997). Sequence analysis of the hMSH2 promoter region revealed a site with homology to the p53 consensus DNA binding sequence. Mobility shift assays demonstrated that purified p53 binds to hMSH2 *in vitro* suggesting that p53 may play a role in the expression of hMSH2. This evidence of altered expression levels supports the hypothesis that MMR proteins may play a role in regulation of the cell cycle (Myers *et al.*, 1997).

HNPCC AND TOLERANCE TO ALKYLATING AGENTS

An increased resistance to alkylating agents such as n- methyl-nitroso-urea has been observed in cells lacking MMR activity (Branch *et al.*, 1993; Koi *et al.*, 1994). This observation suggests that cells defective in their mismatch repair activity can tolerate DNA adducts caused by these agents. Under normal cell conditions, these adducts if not corrected would result in cell arrest or death. The phenomenon of tolerance has been linked to the inactivation of hMLH1 in particular. HCT16 cell lines lacking hMLH1, display microsatellite instability, methylation tolerance and mismatch repair deficiency. The introduction of chromosome 3 (hMLH1 locus) restores normal functionality to the cell and destroys tolerance levels (Koi *et al.*, 1994). These findings may have important implications in chemotoxic treatment of HNPCC tumours, however the exact mechanism by which this tolerance occurs is still to be established.

CANCER RISK IN HNPCC PEDIGREES

A number of studies based on calculating the lifetime risk of developing cancer for a carrier of a HNPCC gene defect have been performed. The results are important in terms of surveillance and genetic counselling of patients. A recent study by Vasen *et*

al., 1996) assessed the age specific cancer risk in a large series of patients with known mutations in either hMSH2 or hMLH1. A total of 210 gene carriers were analysed. Risk assessment was carried out on the basis of carriers developing cancer from birth to death. Data was evaluated using survival analysis methods and survival curves were constructed with use of Kaplan-Meier estimates. Their results show that the lifetime risk of colorectal cancer was 80% and was the same whether the mutation was in hMSH2 or hMLH1. They also found that there was a very high relative risk of generating cancer of the small bowel and carriers of hMSH2 gene mutations had a higher risk of developing cancer of the urinary tract, stomach and ovaries, than hMLH1 carriers.

Dunlop *et al.*, (1997), suggested that while the Vasen study was significant to known HNPCC pedigrees it did not take into account families carrying mismatch repair gene mutations which were not strictly 'Amsterdam Families'. They used the RER status of tumours from HNPCC and sporadic cancer patients to determine whether patients carried defects in their mismatch repair genes. The calculated risk of developing any cancer by age 70 was 91% for males and 69% for females. The risk of colon cancer was higher in males than in females (74% versus 30%) and the risk of uterine cancer exceeded that of colorectal cancer in females (42% versus 30%). These results have particular significance in terms of screening protocols, especially in terms of uterine screening in female carriers of HNPCC pedigrees.

MUIR-TORRE SYNDROME

Muir-Torre syndrome (MTS) is an autosomal dominant disorder characterised by the development of sebaceous gland tumours and skin cancers such as keratoacanthomas and basal cell carcinomas (Muir *et al.*, 1967). In addition affected individuals may develop a number of internal malignancies including colorectal, endometrial, upper gastrological and urinary tract cancers (Cohen *et al.*, 1991). As in

HNPCC, tumours in these patients display microsatellite instability, they are usually synchronous and metachronous and prognosis is usually favourable. On the basis of similarities between Muir-Torre syndrome and HNPCC it was suggested that there is a common genetic link (Lynch *et al.*, 1985). Analysis of a number of Muir-Torre families has led to the identification of mutations in the hMSH2 gene in 6 unrelated Muir Torre pedigrees (Kolodner *et al*, 1994b). This finding has important implications at the clinical level as the presentation of sebaceous gland and skin tumours implies that the patient may have mismatch repair gene defects and is at risk of developing other HNPCC type tumours.

GENETIC DIAGNOSIS OF HNPCC

The traditional clinical procedure for the diagnosis of HNPCC was based on an early onset colorectal cancer associated with the lack of FAP type symptoms. On presentation of such a patient a family history was obtained to determine if there was an associated familial component. However due to small family sizes and lack of accurate knowledge pertaining to other family members, most cases of colorectal cancer were defined as sporadic. The lack of definitive diagnostic criteria has made it difficult in the past to accurately diagnose HNPCC in a family. The identification of four genes associated with HNPCC has greatly improved the diagnosis and outcome in HNPCC pedigrees. The large number and heterogeneous nature of causative mutations in HNPCC means that genetic diagnosis is a rather tedious process, however once a causative mutation has been identified within a particular family, subsequent screening of younger members has major implications for surveillance and treatment of these patients.

1.2. MATERIALS AND METHODS

1.2.(I) DNA ISOLATION FROM PERIPHERAL BLOOD

Blood was collected in tubes containing EDTA as an anti-coagulant and stored at -20°C. DNA Isolation was performed using one of the two methods outlined below. Method 1 is a modification of the Kunkle *et al.*, (1978) protocol. This method is more labour intensive and time consuming than method two, however the isolated DNA is more robust, free from residual proteins (average A260/A280 ratio of 1.7), with an average yield of 500ng of DNA from 10mls of blood and remains stable when stored at -20° C for at least four years. Method two is less tedious, less time consuming and does not require the use of phenol-chloroform, however the yield is lower (average yield 300ng from 10 mls blood) and the quality of the DNA is poorer- samples which have been stored for more than 1 year at -20°C are poor PCR templates and residual proteins must be removed by using DNA clean-up systems such as Wizard ™ DNA Clean- Up system (Promega).

DNA ISOLATION METHOD 1:

Ten mls of blood was combined with 90mls of lysis buffer (0.32M sucrose, 10mM Tris-HCl, pH 7.5, 5mM $MgCl_2$, 1% Triton X-100) at 4°C. Lysis was spontaneous. Nuclei were collected by centrifugation at 12,000g for 10 minutes in a sorvall RC-2B centrifuge maintained at 4°C. The nuclear pellet was re-suspended in 4.5ml of suspension buffer (24mM EDTA pH8, 72mM NaCl) and transferred to a 30ml centrifuge tube. To this, 0.5ml of 5% SDS (final concentration 0.5%) and Proteinase K to final concentration of 200µg/µl were added. Samples were incubated at 55°C

overnight. Following incubation, 5ml of Tris saturated phenol pH8.0 was added to the solution, mixed by inversion and centrifuged at 12,000g for 10 minutes at 4°C in a Sorvall RC-2B centrifuge (the top aqueous phase contains the DNA, whereas the protein remains at the interface and in the organic phase). The aqueous layer was retained and phenol extraction was repeated twice to ensure maximum removal of residual proteins. 5ml Chloroform: isoamylalcohol (24:1) were added to the aqueous phase, mixed by inversion and centrifuged at 12,000g for 10 minutes as before. The aqueous layer was retained and chloroform extraction repeated. The aqueous phase was removed, added to a solution of 0.5ml Sodium Acetate (3.0M) pH 4.8 and 11ml of 100% ethanol. The tube was inverted until DNA precipitated. The DNA was transferred to a solution of 70% ethanol for 30 seconds and transferred to a clean tube and allowed to dry. The DNA was re-suspended in 0.5ml TE buffer (10mM Tris-Cl pH 8.0, 1mM EDTA pH 8.0) and stored at 4°C for 3 days to ensure complete re-suspension. The DNA concentration was estimated from the A260 readings obtained using a spectrophotometer. The average yield of 500ng DNA was obtained from 10mls blood.

DNA ISOLATION METHOD 2:

The method outlined is for 300µl of blood, this was proportionally increased for larger volumes.

Cell lysis was induced by adding 300µl of whole blood to 900µl Red Blood Cell (RBC) Lysis solution (155mM Ammonium Chloride, 10mM Potassium Hydrogen Carbonate, 1mM EDTA). Tubes were inverted to mix and incubated at room temperature for 10 minutes. Samples were centrifuged for 20 seconds at 16,000g in a bench top centrifuge. The supernatant was removed to leave a visible white pellet (contains the white blood cells) these were vortexed vigorously to resuspend the white blood cells in the remaining residual liquid. 300µl of cell lysis solution (25mM EDTA,

2% SDS) was added and pipetted up and down to lyse the cells. Samples were incubated at room temperature for 30 minutes or until the solution was homogeneous. Samples were incubated with 1U RNase A at 37°C for 15 minutes and subsequently cooled to room temperature. 100µl of Protein precipitation solution (10M ammonium Acetate) was added to the cell lysate and vortexed vigorously for 20 seconds. Samples were centrifuged at 16,000g for 3 minutes as before (a tight dark pellet of proteins formed). The supernatant containing the DNA was removed to a new tube containing 300µl of 100% isopropanol. The samples were mixed by inverting gently 50 times until white threads of DNA became visible. DNA was removed by centrifugation at 16,000g for 1 minute. The supernatant was removed and DNA was drained by inverting tube on clean absorbent paper. 300µl of 70% ethanol was added to wash the DNA sample, after centrifugation at 16.000g for 1 minute the ethanol was removed and DNA was allowed to air dry. DNA was re-hydrated by adding 50-100µl of TE buffer and stored at 4°C for 3 days to ensure complete re-suspension.

1.2.(II) PCR AMPLIFICATION OF HNPCC SAMPLES

The hMSH2 gene contains 16 exons and the hMLH1 gene contains 19 exons. Each exon was amplified using published primer sequences (Kolodner *et al.*, 1994a, 1995). Primer sequences are listed in table 1.3 and 1.4.of Section 1.3. (VIII).

Amplification reactions were initially performed in a 25µl reaction volume containing 60ng genomic DNA, 50ng of each primer, 0.2mM dNTP's, 1U *Taq* DNA Polymerase, *Taq* DNA Polymerase buffer (50mM KCl, 10mM TrisCl pH 9.0, 0.1% Triton X-100, 1.5mM $MgCl_2$). Samples were overlaid with mineral oil .

Cycling parameters: Amplification was performed on an MJ PTC-100 thermal cycler. All samples were heated to 95°C for 5 minutes (Hot-start) to ensure complete denaturation and subjected to 30 cycles of 95°C for 1 min, 55°C for 1 min, 72°C for 1 min followed by a final extension step of 72°C for 10 minutes. PCR optimisation was performed on primer pairs which did not give good PCR yield under these conditions. Optimisation involved performing the PCR reactions according to the above conditions however, either the concentration of $MgCl_2$ or the annealing temperature was altered. The optimum conditions for each primer pair are detailed in table 1.5. Section 1.3.(VII).

PCR AMPLIFICATION FOR SSCP ANALYSIS

Gene segments for SSCP analysis were amplified essentially as described above, however the reaction volume was reduced to 10µl and the dNTP mix concentration was reduced to 0.02mM. In addition 2µCi of [α^{32}P]-dCTP (3000Ci/mmol) was included in the reaction.

1.2.(III) SSCP ANALYSIS

Following PCR amplification in the presence of [α^{32}P]-dCTP, samples were diluted 1:100 in formamide dye (95% formamide, 20mM EDTA, 0.05% Bromophenol blue, 0.05% Xylene cyanol). Samples were denatured at 95°C for 5 minutes and individual DNA strands were separated on a 6% polyacrylamide gel under three conditions. (1) with 5% glycerol at room temperature (2) with 5% glycerol at 4°C and

(3) 0% glycerol at 4°C. Constant power of 60 Watts was maintained and samples were separated for approximately 4 hours. Gels were dried and subjected to autoradiography.

1.2.(IV) DNA SEQUENCING

Nucleotide sequence analysis of the HNPCC samples was performed either by the direct sequencing of PCR products or by sequencing the cloned PCR product.

Cloning of PCR products:

The pTag cloning vector (R&D Systems) exploits the additional T residue that Taq DNA polymerase adds on to the end of PCR products. Ligation of the PCR product with the pTag plasmid was performed by incubating 2µl of PCR product (removed directly from a 10µl PCR reaction tube) with 2U T4 DNA ligase, ligase buffer, 5mM DTT, 0.5mM ATP and 50ng of pTag plasmid in a 10µl reaction. Samples were incubated overnight at 16°C and subsequently transformed into competent *E.coli* cells as described by Sambrook *et al.*, 1989. Transformants were plated on LB-AMP-IPTG-X-gal plates and incubated overnight at 37°C. Recombinant clones were selected based on antibiotic resistance and insertional inactivation of the β-Galactosidase gene. Plasmid DNA was isolated by minipreparation of the plasmid as outlined in: Molecular Cloning a Laboratory Manual, Sambrook *et al.*, (1989). Plasmid inserts were sequenced in both directions using either a radiolabelled primer (see below) and the T7 DNA sequencing kit (Pharmacia) or direct cycle sequencing of the PCR product using [α^{33}-P] labelled dideoxy terminators and Thermosequenase as outlined in the cycle sequencing kit (Amersham Life Sciences).

Sequencing products were diluted with an equal volume of Formamide loading dye (95% formamide, 20mM EDTA, 0.05% Bromophenol blue, 0.05% Xylene cyanol) denatured and separated on a 6% denaturing (7M urea) polyacrylamide gel. Analysis of

the DNA sequence data was performed by comparing with the known published sequences of the genes.

1.2.(V) OLIGONUCLEOTIDE LABELLING

Primers used in the sequencing reactions were end labelled by incubating 60pmol primer with 1U of T4 Polynucleotide Kinase (New England Biolabs), the appropriate buffer and 50uCi [γ-^{32}P]-ATP (3000Ci/mmol) at 37°C for 30 minutes. One volume of 4M sodium acetate and two volumes of ethanol was added in order to precipitate the primers and to remove unlabelled nucleotide. Primers were re-suspended in water to the required concentration.

1.3 RESULTS

1.3.(I) OBJECTIVES

The objective of the research included in this chapter was to collect as many Irish HNPCC pedigrees as possible in the given period and to analyse the hMSH2 and hMLH1 genes for causative mutations. It was envisaged that the results would provide information on the Irish HNPCC mutation spectrum, facilitate comparison with other ethnic HNPCC populations and would provide information pertaining to the susceptibility of individuals within affected pedigrees.

Mutations in four human mismatch repair genes are known to segregate with HNPCC (hMSH2, hMLH1, hPMS1 and hPMS2), of all the mutations identified, over 80% are confined to hMSH2 and hMLH1 (Peltomaki *et al.*, 1997). In an attempt to determine the mutation spectrum of HNPCC genes in the Irish population, eighteen HNPCC pedigrees were ascertained, sampled and analysed for mutations in the hMSH2 and hMLH1 mismatch repair genes. Figure 1.6 displays a diagrammatic representation of the methods employed and the order of analysis in this study. Briefly, Irish HNPCC patients were selected based on presentation at Cork University Hospital, with colorectal carcinoma in one individual in association with a strong family history of colorectal cancer and in some cases extra-colonic HNPCC associated tumours. Genomic DNA was extracted from peripheral blood leukocytes from at least one affected member of each family. The hMSH2 gene contains 16 exons while hMLH1 contains 19 exons, each individual exon of both genes was amplified via PCR, optimisation of each PCR amplicon was performed and each fragment was subsequently re-amplified from genomic DNA template in the presence of a radiolabelled precursor, using the optimised PCR conditions. Radiolabelled PCR fragments were subsequently analysed by SSCP under varying conditions. Any regions displaying aberrant electrophoretic patterns were sequenced to identify

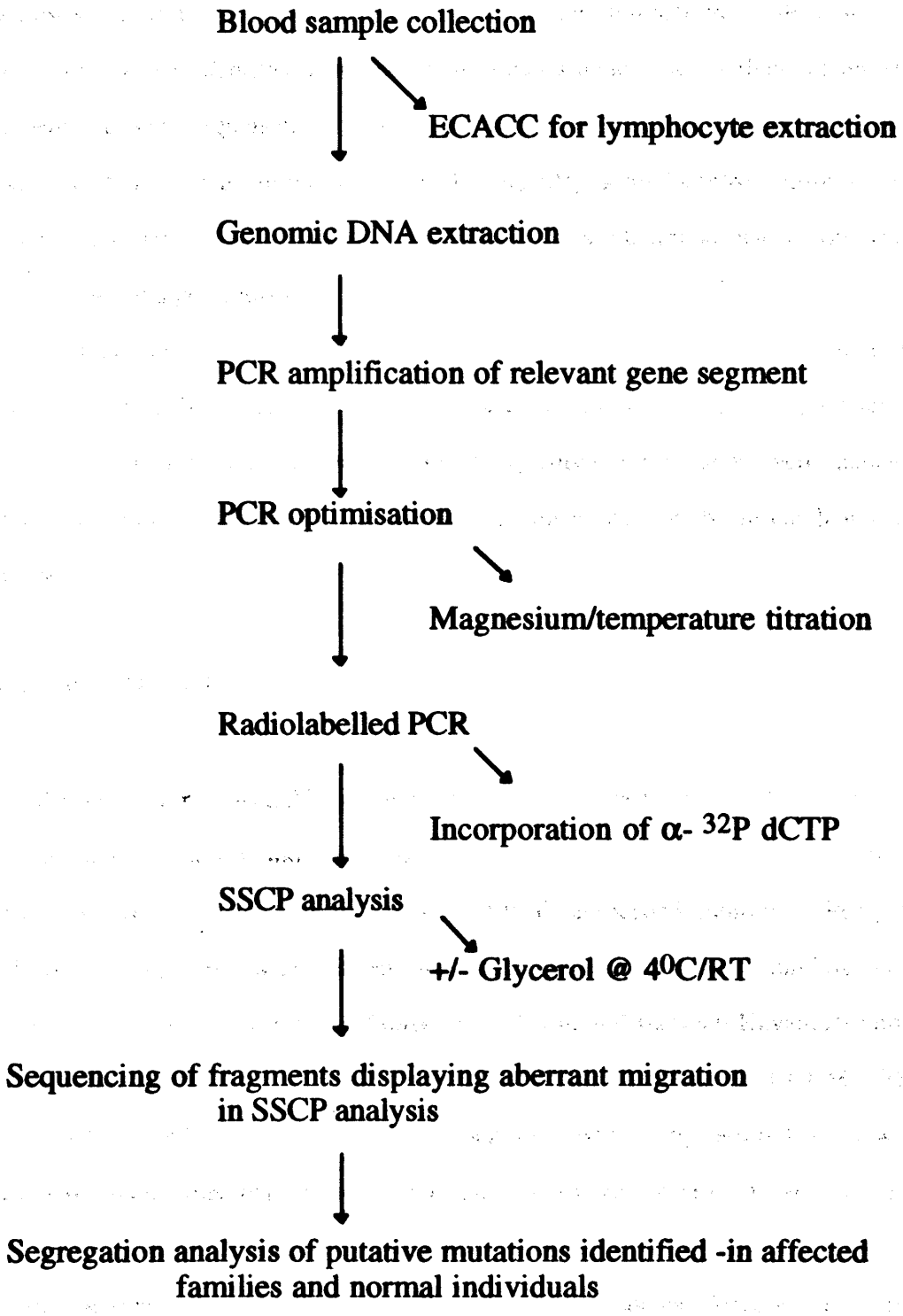


Figure 1.6. Schematic representation of the protocol for the mutation screening of the hMSH2 and hMLH1 genes in Irish HNPCC families.

mutations/polymorphisms responsible for the altered migration pattern. In the case where a mutation was identified in an individual, other affected and unaffected members of the pedigree were sequenced to determine whether the putative causative mutation segregates with the disease in the pedigree. Finally fifty normal control samples were screened for the presence of the mutation to determine the likelihood that the observed mutation could be a polymorphism.

Each step of the analysis required careful appraisal of available techniques to determine which method would best suit this study. Furthermore, careful optimisation at each stage of analysis was performed. Finally quality control checks were included to avoid discrepancies and errors during analysis. Each stage of the research will be discussed in detail.

1.3.(II) PATIENTS

Pedigrees were selected based on presentation at Cork University Hospital, with colorectal carcinoma in one individual, in association with a strong family history of colorectal and in some cases extra-colonic HNPCC associated tumours. Pedigree identification, diagnosis, ascertainment of family history and blood sampling were performed by Drs. Cliona Murphy, Maureen O'Sullivan and Eamonn Kavanagh under the supervision of Professor. Liam Kirwan of the Dept. of Surgery at CUH. The selection criteria used in this study were: (1) There must be a strong family history with at least two affected members in two generations. (2) Early onset (< 50 years) colon cancer in at least one individual (3) Localisation of tumours to the proximal colon (4) FAP must be ruled out. These criteria are less stringent than the Amsterdam Criteria which state that there must be at least three affected members in two successive generations. However the Amsterdam guidelines had not been published at the commencement of this project. The criteria were chosen to compensate for the lack of documentation of family history in medical charts and the fact that most of this evidence

was provided by the patient themselves and in many cases incomplete histories were provided. Many but not all of the pedigrees included in this study satisfy all the Amsterdam criteria however, tumours of the proximal colon, evidence of family history, and the absence of FAP is true for all pedigrees. It is important to note at this stage that genetic screening was carried out blind with respect to whether the pedigrees satisfied all the Amsterdam Criteria or not. Pedigree details were only examined after all families were screened and mutations identified. Details of each pedigree included in this study are detailed below, for ease of reference, determination of whether pedigrees are 'typical' Amsterdam (satisfy all the Amsterdam criteria) or not has been included with the family details.

DNA from forty unrelated unaffected Irish DNA samples (collected for another unrelated study), and ten unaffected "married in" individuals from the HNPCC pedigrees were employed as the control population sample.

1.3.(III) PEDIGREE DETAILS

Figure 1.7 details symbols used in pedigree diagrams.

Pedigree B1 (Figure 1.8)

In this pedigree the proband, individual IV:I presented with cancer of the proximal colon at 50 years of age. His brother, individual IV:5 previously diagnosed at 48 years and his sister IV:7 diagnosed at 50 years were deceased at time of study. Three other siblings presented with colorectal carcinoma IV:4, IV:8 and IV:9. This pedigree is a typical Amsterdam HNPCC pedigree in that all the criteria laid out by the International Collaborative Group on HNPCC are satisfied. There is a strong family history of cancer with 14 members affected with colorectal cancer in three generations (maternal and paternal) and two cases of breast carcinoma on the maternal side. Individual IV:1 has two as yet asymptomatic children and one who presented with a colonic lesion which was clear on histological analysis. (see Section 1.3.(XII) for the genetic analysis of this individual). Samples included for analysis in this study are IV:1, IV:3, IV:4,

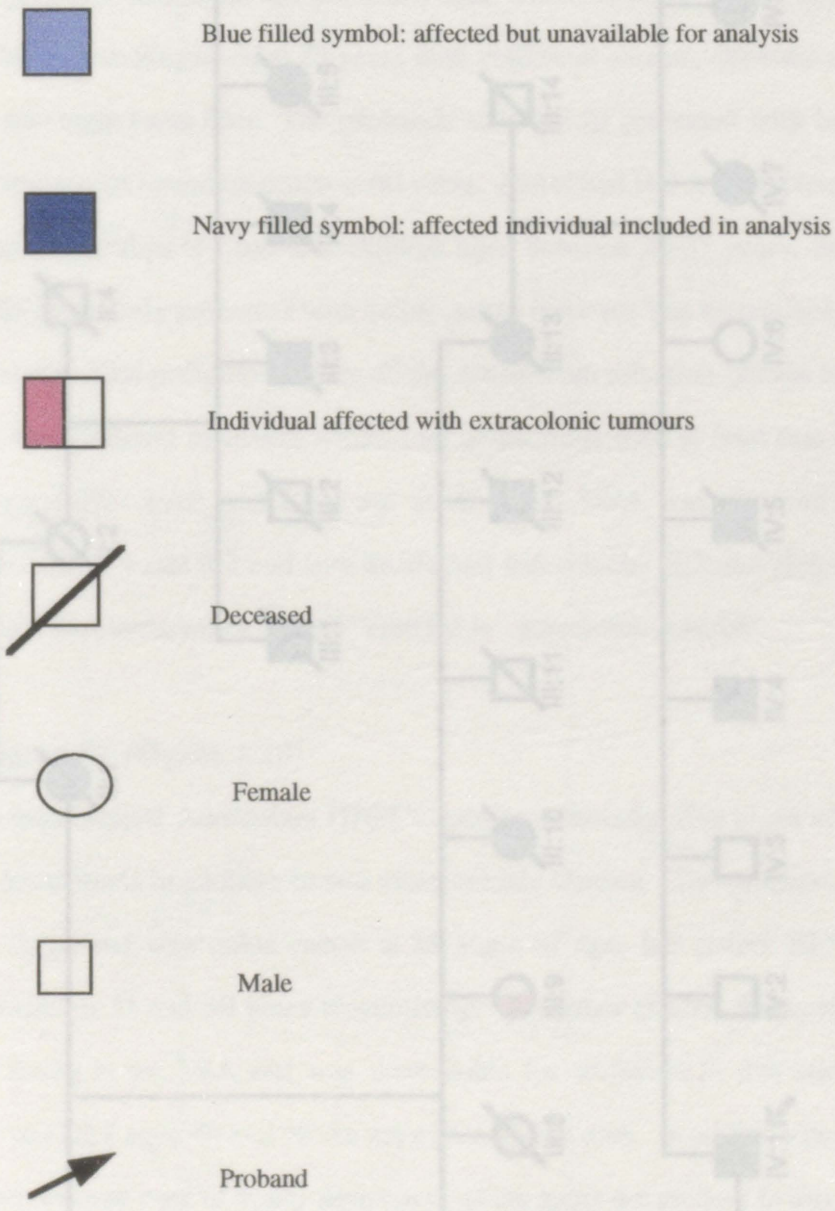


Figure 1.7 . Symbols used in pedigree diagrams

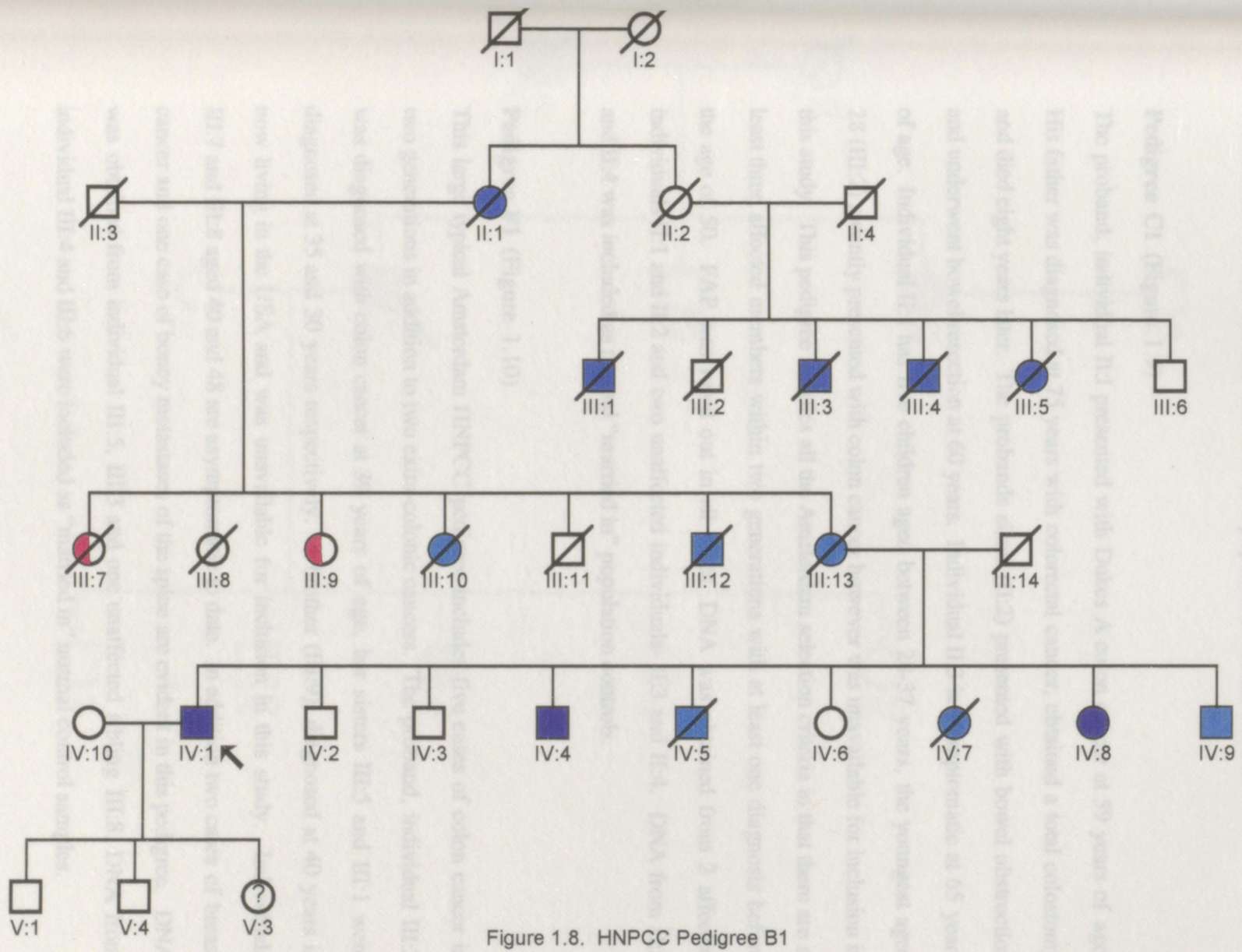


Figure 1.8. HNPCC Pedigree B1

IV:8, V:3 and V:1. In addition DNA was obtained from individual IV:10 an unaffected "married in" individual, and was used as a population normal control.

Pedigree C1 (Figure 1.9)

The proband, individual II:I presented with Dukes A colon cancer at 59 years of age. His father was diagnosed at 75 years with colorectal cancer, obtained a total colostomy and died eight years later. The probands sister (II:2) presented with bowel obstruction and underwent bowel resection at 60 years. Individual II:3 is asymptomatic at 65 years of age. Individual II:1 has five children aged between 28-37 years, the youngest aged 28 (III:2) recently presented with colon cancer however was unavailable for inclusion in this study. This pedigree satisfies all the Amsterdam selection criteria in that there are at least three affected members within two generations with at least one diagnosis before the age of 50. FAP was ruled out in all cases. DNA was obtained from 2 affected individuals II:1 and II:2 and two unaffected individuals- II:3 and II:4. DNA from II:5 and II:4 was included as normal "married in" population controls.

Pedigree F1 (Figure 1.10)

This large typical Amsterdam HNPCC pedigree includes five cases of colon cancer in two generations in addition to two extra-colonic cancers. The proband, individual III:3 was diagnosed with colon cancer at 39 years of age, her sisters III:5 and III:1 were diagnosed at 35 and 30 years respectively. A brother (III:9), diagnosed at 40 years is now living in the USA and was unavailable for inclusion in this study. Individuals III:7 and III:8 aged 40 and 48 are asymptomatic to date. In addition two cases of breast cancer and one case of boney metastases of the spine are evident in this pedigree. DNA was obtained from individual III:5, III:3 and one unaffected sibling III:8. DNA from individual III:4 and III:6 were included as "married in" normal control samples.

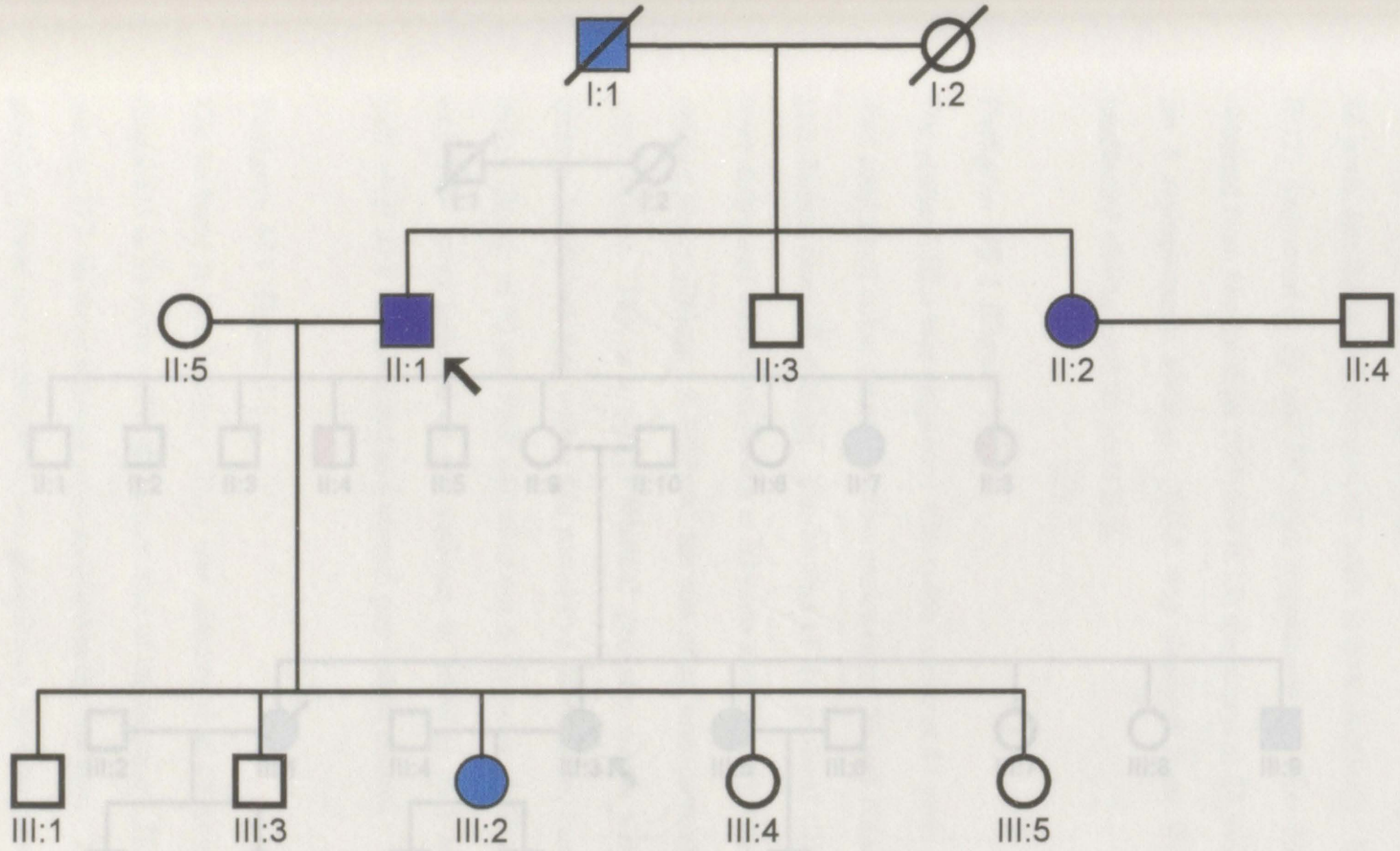


Figure 1.9 HNPCC Pedigree C1

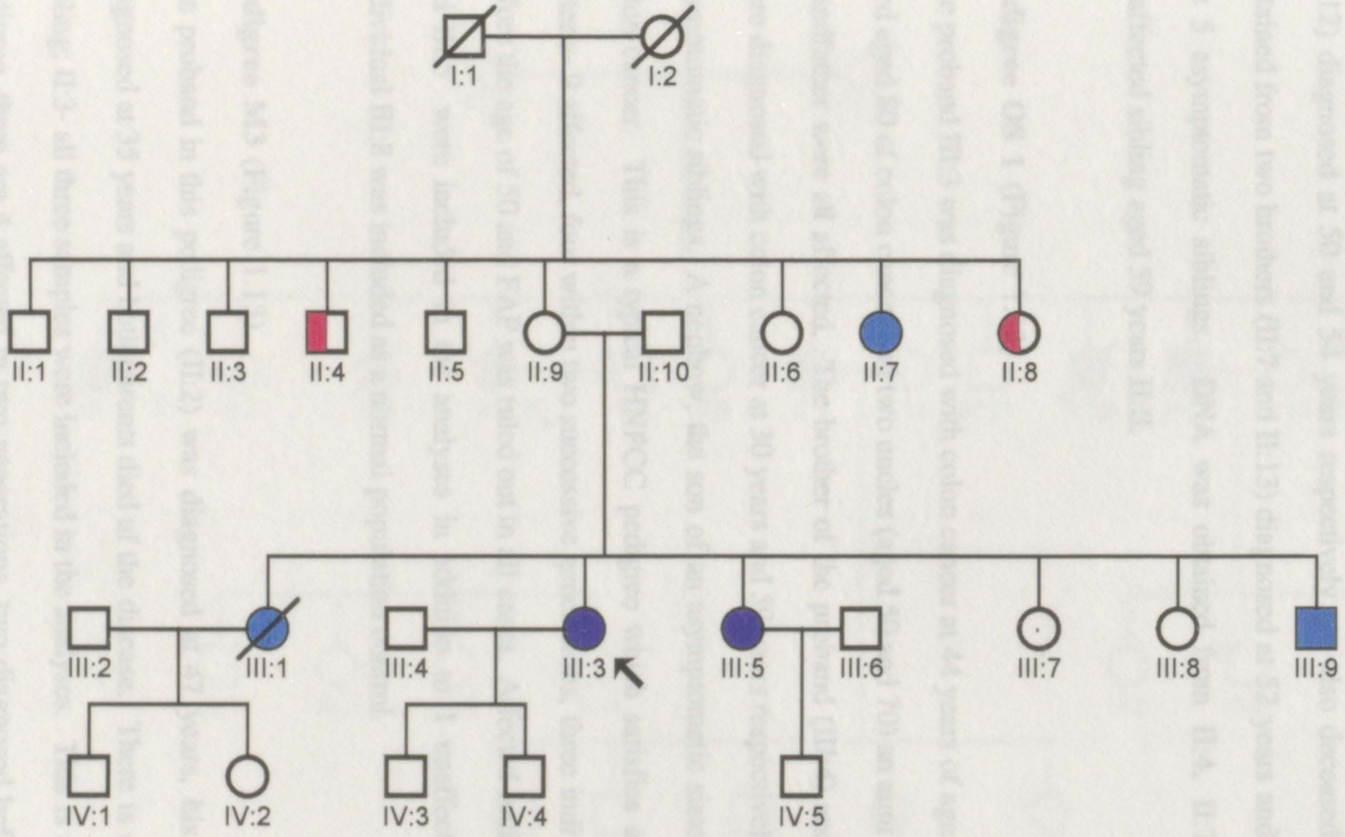


Figure 1.10 HNPCC Pedigree F1

Pedigree H1 (Figure 1.11)

There are six affected members in two successive generations in this pedigree. Furthermore, there is one case of uterine cancer and one nephroblastoma in an individual of 1.5 years of age. The proband (II:4) was diagnosed with colon cancer at 48 years, her father diagnosed at 67 years is now deceased. Her two sisters (II:1 and II:12) diagnosed at 50 and 54 years respectively are also deceased. DNA was also obtained from two brothers (II:7 and II:13) diagnosed at 52 years and 47 years. There are 5 asymptomatic siblings. DNA was obtained from II:4, II:7, II:13 and one unaffected sibling aged 59 years II:II.

Pedigree OS 1 (Figure 1.12)

The proband III:3 was diagnosed with colon cancer at 44 years of age. His father (II:1) died aged 80 of colon cancer and two uncles (aged 50 and 70) an aunt (aged 59) and his grandfather were all affected. The brother of the proband (III:4) and his sister (III:5) were diagnosed with colon cancer at 30 years and 59 years respectively. There are four asymptomatic siblings. A nephew, the son of an asymptomatic sister died aged 34 of colon cancer. This is a typical HNPCC pedigree which satisfies all the Amsterdam criteria- 9 affected, four within two successive generations, three individuals diagnosed before the age of 50 and FAP was ruled out in all cases. Affected individuals III:3, II:4 and III:5 were included in the analyses in addition to 1 unaffected sibling: III:7. Individual III:8 was included as a normal population control.

Pedigree M3 (Figure 1.13)

The proband in this pedigree (II:2) was diagnosed at 47 years, his brother II:1 was diagnosed at 35 years and both parents died of the disease. There is one asymptomatic sibling; II:3- all three samples were included in the analyses. This is a true Amsterdam pedigree, there are 4 affected in two generations, two diagnosed before the age of 50 and FAP was ruled out.

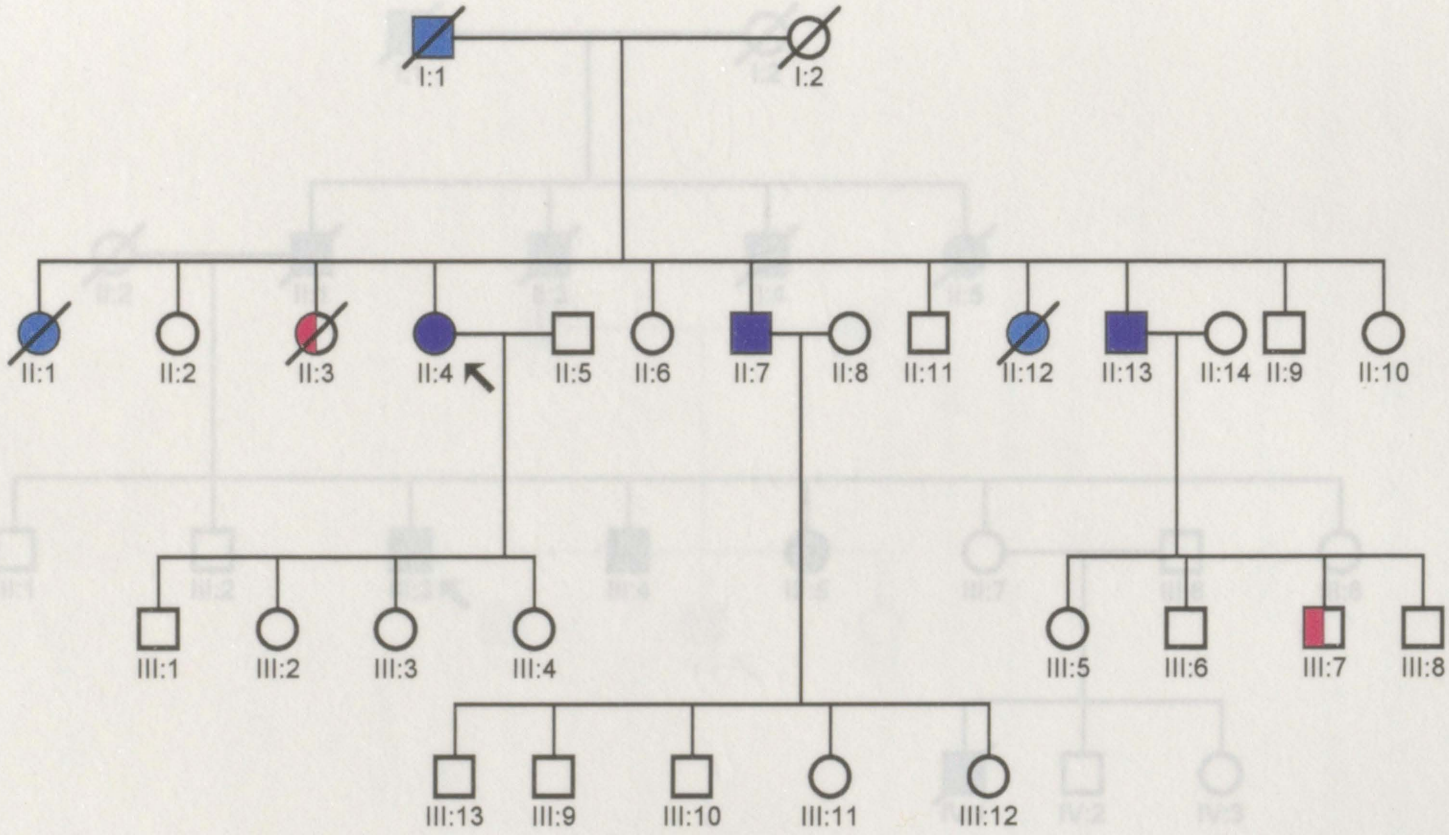


Figure 1.11 HNPCC Pedigree H1

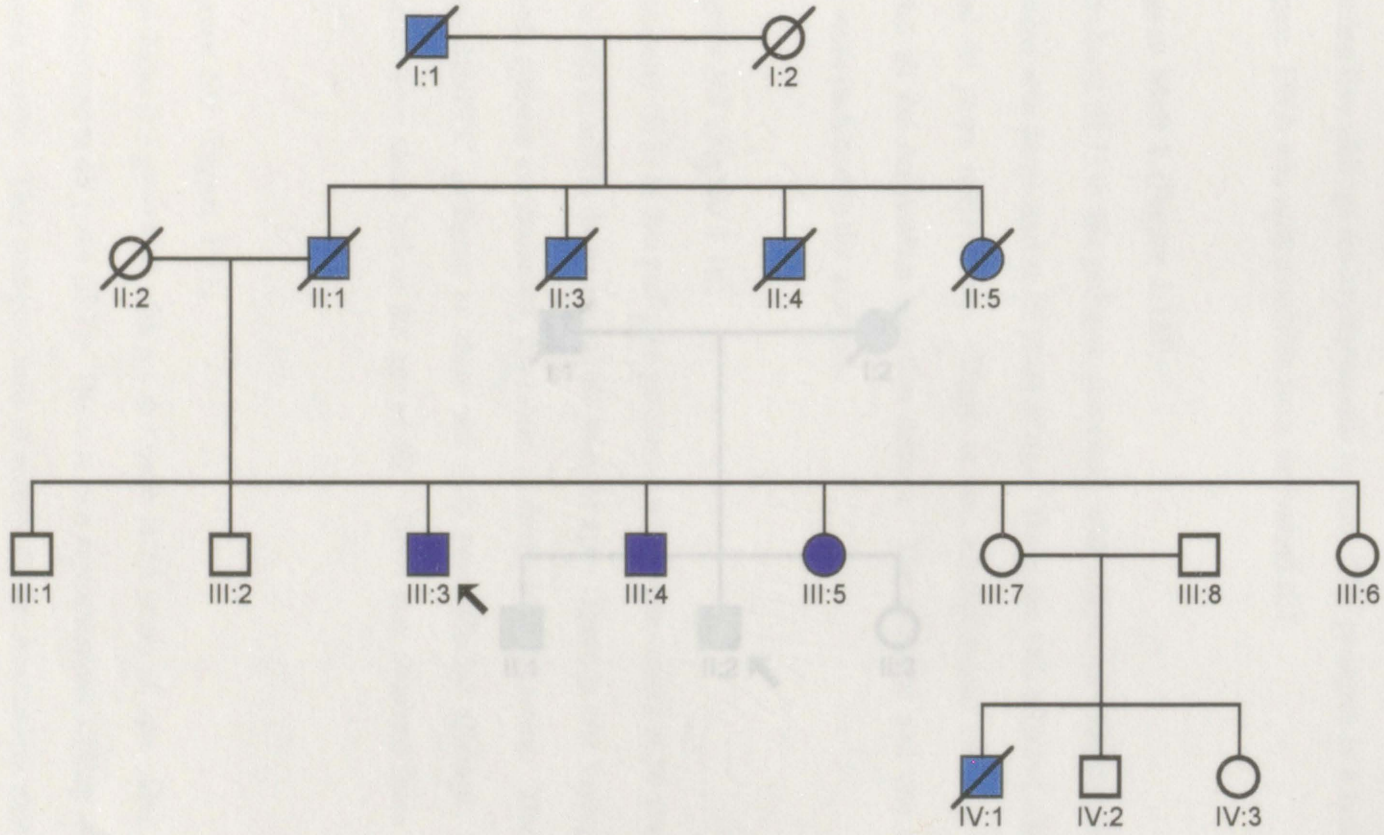


Figure 1.12 HNPCC Pedigree OS1

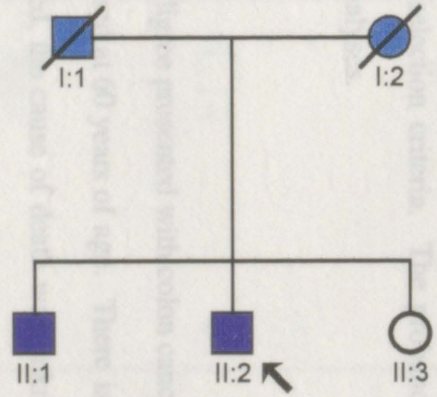


Figure 1.13 HNPCC Pedigree M3

Pedigree W3 (Figure 1.14)

The proband (II:1) presented with colorectal cancer at 62 years of age, her brother (II:3) diagnosed at 42 years of age is deceased. The proband has seven children, two have been diagnosed with colon cancer III:4 at 26 years and III:7 at 31 years of age. The remaining five siblings are asymptomatic to date. This pedigree is a typical Amsterdam pedigree. DNA was only available from individual II:1.

Pedigree McS 1 (Figure 1.15)

The proband (II:1) in this pedigree presented with colorectal cancer at 69 years of age. His father was diagnosed at 49 years of age. There are two affected sibs, diagnosed at 60 and 40 years respectively. There is one asymptomatic sibling. This pedigree satisfies all the Amsterdam selection criteria. The proband and one affected sibling (II:2) were included in the analyses.

Pedigree C2 (Figure 1.16)

The proband (II:2) in this pedigree presented with colon cancer at 50 years of age. Her brother was similarly diagnosed at 60 years of age. There is one asymptomatic sibling and both parents are deceased, the cause of death was unknown. This pedigree is an atypical HNPCC pedigree as there are only two affected siblings, furthermore no diagnoses were made before the age of 50. DNA was obtained from individuals II:2 and II:3.

Pedigree M1 (Figure 1.17)

The proband II:2 presented with colon cancer at 54 years of age. One affected sibling was diagnosed at 48 years (II:1). There is one asymptomatic sibling and a single case of gastric cancer. This pedigree fails to satisfy all the Amsterdam criteria in that there are only two affected members in two generations, there is however a strong family history of the disease, a diagnosis before 50 years of age , evidence of extra-colonic

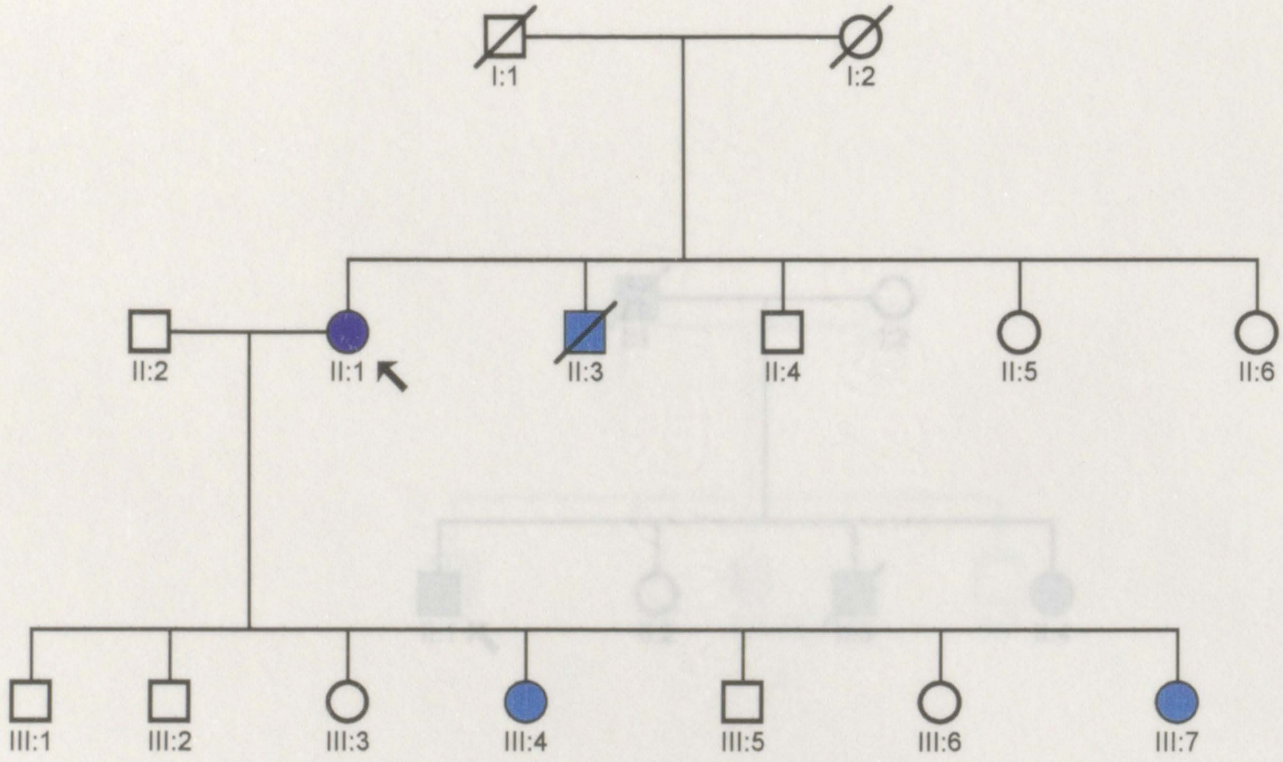


Figure 1.14 HNPCC Pedigree W3

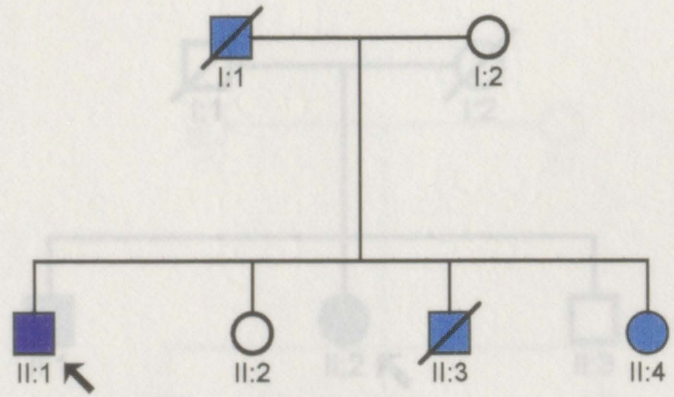


Figure 1.15 HNPCC Pedigree MCS1

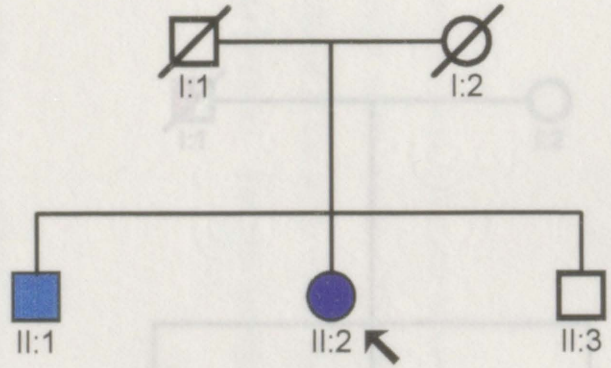


Figure 1.16 HNPCC Pedigree C2

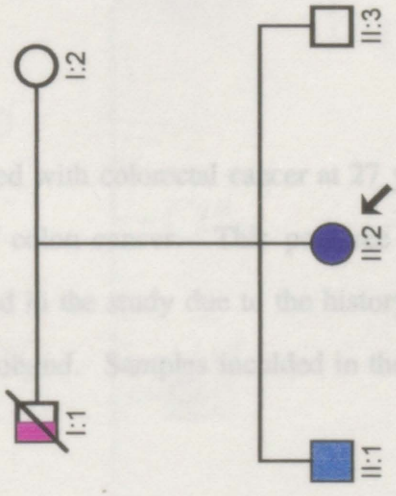


Figure 1.17 HNPCC Pedigree M1

tumours and FAP was ruled out in all cases.

Pedigree W1 (Figure 1.18)

The proband II:2 presented with colon cancer at 31 years. Her father died of pancreatic cancer at 47 years. There is one affected sibling (II:4), diagnosed at the age of 30 but unavailable for inclusion in this study. There are two asymptomatic siblings. This family fails to satisfy the Amsterdam criteria in that there are only two affected members in two successive generations. DNA was obtained from individuals IIII:2 (affected, II:3 (unaffected), furthermore, individual II:5 was included as a "married in" normal control sample.

Pedigree W4 (Figure 1.19)

The proband individual II:2 presented with colorectal cancer at 27 years. One affected sibling died at 58 years of age of colon cancer. This pedigree fails to satisfy the Amsterdam criteria but was included in the study due to the history of the disease and the very early age of onset in the proband. Samples inculded in the analyses were II:1 and II:2.

Pedigree D1 (Figure 1.20)

Individual II:3, was diagnosed with colon cancer at the age of 61, her sister (II:2) now deceased, was similarly diagnosed at the age of 51 years. Her brother (II:1) also had a positive diagnosis of colon cancer however the age of diagnosis was unknown. The father of the proband presented with cancer of the bowel at 62 years of age. This pedigree does not satisfy the Amsterdam criteria as there was no proof of diagnosis before the age of 50 in any affected member. However as there are three affected members in two generations the pedigree was included in the analyses. DNA was obtained from individuals II:3 and II:1.

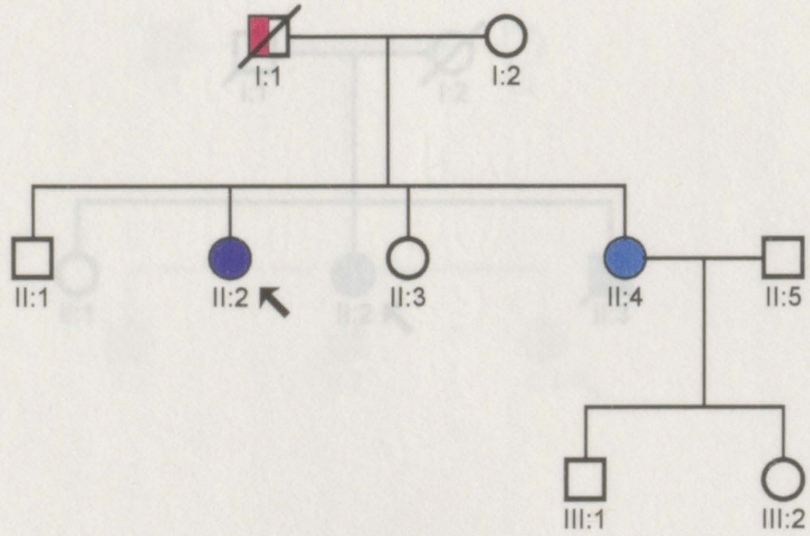


Figure 1.18 HNPCC Pedigree W1

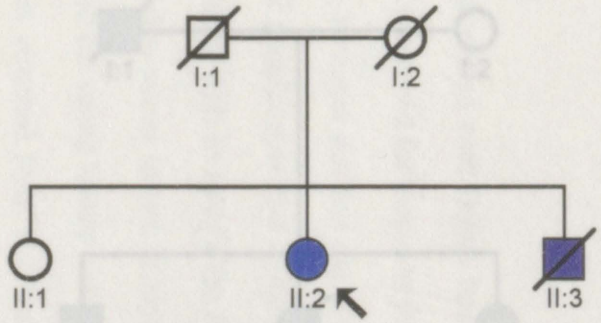


Figure 1.19 HNPCC Pedigree W4

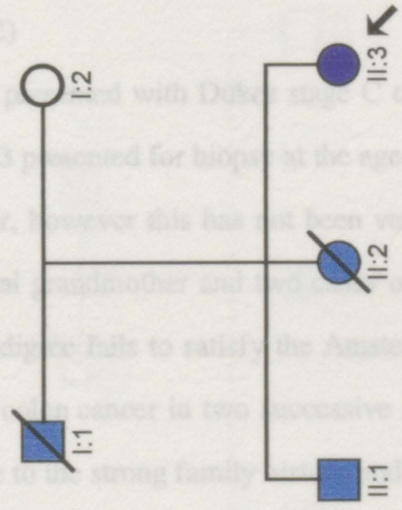


Figure 1.20 HNPCC Pedigree D1

Pedigree D3 (Figure 1.21)

In this pedigree there are three affected members in two successive generations, individual II:1 had a positive diagnosis of colon cancer at the age of 55. His sister now deceased, was diagnosed at 57 years of age. Individual I:1 the affected mother died at 60 years of age as a result of colon cancer. This is an atypical HNPCC pedigree as there is no diagnosis before the age of 50 in any individual. There is however, a strong family history of the disease, FAP was ruled out and there are three affected members within two generations. DNA was obtained from peripheral blood of individual II;1 and from parraffin embedded tissue from individual II:2.

Pedigree MG1 (Figure 1.22)

The proband in this pedigree presented with Dukes stage C colon cancer at the age of 37. Her brother individual II:3 presented for biopsy at the age of 34. The mother (I:1) died of suspected colon cancer, however this has not been verified. There is evidence of colon cancer in the maternal grandmother and two cases of non-colonic tumours in two paternal uncles. This pedigree fails to satisfy the Amsterdam criteria in that there are only two verified cases of colon cancer in two successive generations. This family was included in this study due to the strong family history and the early age of onset in affected individuals. DNA was obtained from II:2 and II:3 (affected) and from one unaffected individual II:4.

Pedigree D5 (Figure 1.23)

The proband in this atypical HNPCC pedigree presented with colon cancer at 35 years of age. Her brother individual III:3 was diagnosed at 34 years of age. The father of these siblings is reported to have died aged 42 of colon cancer however this has not been clinically verified. There is evidence of colon cancer in a paternal aunt and her son. DNA was available from individuals III:2, III:3 and III:1. DNA from individual II:2 and II:4 was included as normal control samples.

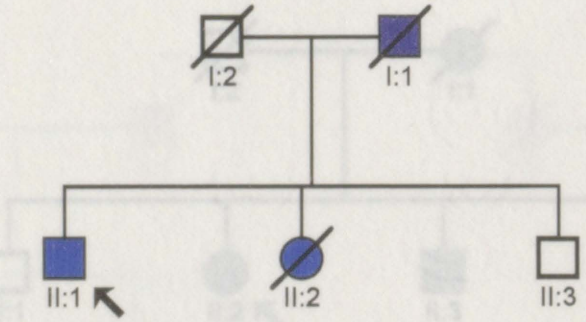


Figure 1.21 HNPCC Pedigree D3

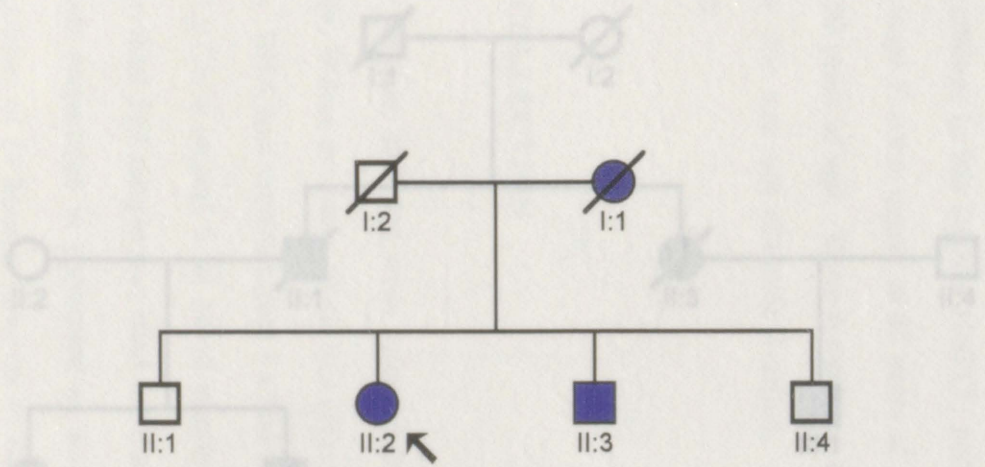


Figure 1.22 HNPCC Pedigree MG1

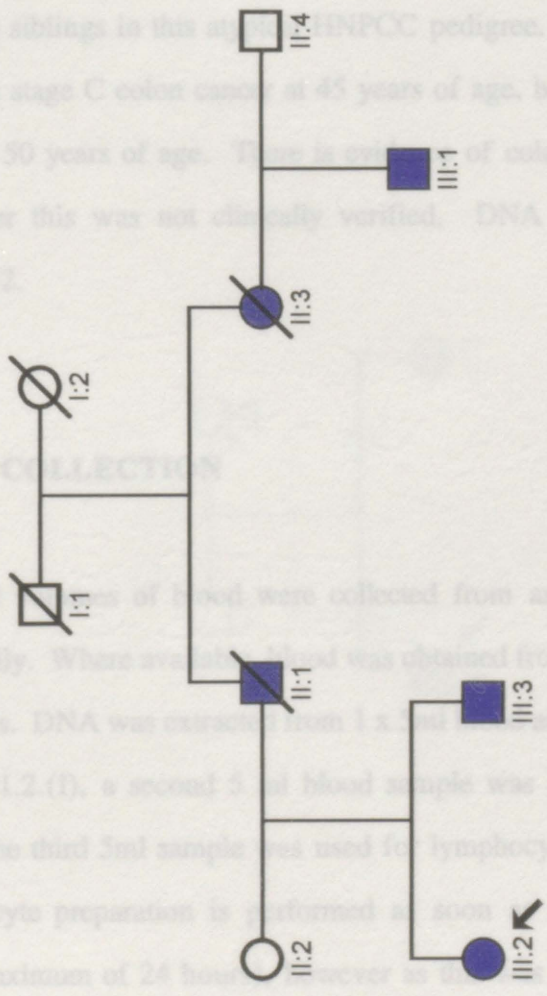


Figure 1.23 HNPCC Pedigree D5

Pedigree D6 (Figure 1.24)

The proband in this pedigree presented with Dukes stage B colon cancer at the age of 40. His sister now deceased was similarly diagnosed at 50 years of age. There is evidence of colon cancer on the maternal side however this had not been verified. DNA was obtained from samples II:1 and II:2.

Pedigree W5 (Figure 1.25)

There are two affected siblings in this atypical HNPCC pedigree. Individual II:1 was diagnosed with Dukes stage C colon cancer at 45 years of age, his brother individual II:2 was diagnosed at 50 years of age. There is evidence of colorectal cancer on the paternal side, however this was not clinically verified. DNA was obtained from individuals II:1 and II:2.

1.3.(IV) SAMPLE COLLECTION

Three, five ml volumes of blood were collected from an affected individual (proband) in each family. Where available, blood was obtained from other affected and an unaffected members. DNA was extracted from 1 x 5ml blood as outlined in Material and Methods Section 1.2.(I), a second 5 ml blood sample was stored at -70°C as a back-up sample and the third 5ml sample was used for lymphocyte preparation. It is essential that lymphocyte preparation is performed as soon as possible after blood sampling (within a maximum of 24 hours), however as this was not compatible with our sampling procedure (many samples had to be posted or sampling involved travelling long distances) blood was sent by courier to the ECACC facility at Porton Down in the UK. Lymphocytes were transformed on site and stored until required. This step was taken as a precautionary measure to ensure that a constant supply of DNA/RNA/cells were available should the need arise.

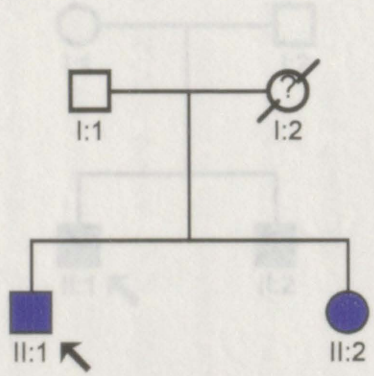


Figure 1.24 HNPCC Pedigree D6

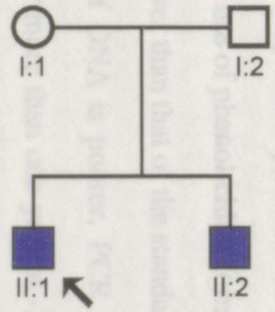


Figure 1.25 HNPCC Pedigree W5

1.3.(V) DNA EXTRACTION

Genomic DNA was extracted from peripheral blood leukocytes using two different methods (see Section 1.2.(I)). The first method, a rather laborious and time consuming (2-3 days) method is based on cell lysis and extraction using the standard phenol-chloroform method (Kunkle *et al.*, 1978). This method provided, high yield (average 500ng/ul), good quality DNA with an average A260/A280 ratio of 1.7. An A260/A280 reading of 1.8 is indicative of a pure preparation of DNA. (Sambrook *et al.*, 1989). DNA extracted by this method is robust free from residual proteins and remains stable when stored at -20°C for at least 4 years. The second method is based on the Puregene method by gentra systems, USA, it is a less tedious procedure, which can be performed on large and very small quantities of blood (<100ul). Furthermore the process does not require the use of phenol-chloroform and can be completed within 1 day. However the yield is lower than that of the standard method (average 300ng/ul), more importantly the quality of DNA is poorer, PCR efficiency decreases as DNA sample ages, samples stored for more than one year at -20°C are poor PCR templates. Re-extraction of DNA using DNA clean-up columns greatly improved the efficiency of PCR suggesting interference from protein residues in the sample. The second method is suitable for situations where once off analysis of a sample is required, however for long term use and storage, the more time consuming phenol chloroform method provides highest quality DNA.

1.3.(VI) MUTATION DETECTION METHODS

HNPCC is usually caused by mutations in one of four mismatch repair genes- hMSH2, hMLH1, hPMS1 and hPMS2 (Leach *et al.*, 1993; Papadopoulos *et al.*, 1994,;

Nicolaides *et al.*, 1994). More than 90% of reported mutations have been confined to hMSH2 and hMLH1. The aim of this study was to screen both these genes for causative mutations in the 18 Irish HNPCC families collected.

There are many methods currently employed to screen for mutations in genes. The most common of these being, Denaturing gradient gel electrophoresis (Sheffield *et al.*, 1989), Chemical cleavage, RNase A protection analysis (Myers *et al.*, 1985), single strand conformation polymorphism analysis (SSCP) (Orita *et al.*, 1989b), Protein truncation test (PTT) or IVPS (in vitro protein synthesis), (Powell *et al.*,1993 Roest *et al.*, 1993) and direct sequencing.

The RNase A protection assay (Myers *et al.*, 1985) involves the PCR amplification of the region of interest from the affected individual. A radiolabelled wild type probe is hybridised to the PCR amplified DNA. If the PCR amplified DNA fragment contains a mutation, a mismatch occurs between the DNA:RNA duplexes and complementary annealing does not occur at this position. The resulting heteroduplex are subjected to cleavage by RNase A, which digests only unbound RNA, hence it will only cleave RNA at a mismatch position. The resulting fragments are analysed by polyacrylamide gel electrophoresis, the size of the fragment indicates the position of the mismatch. One drawback of this technique is the fact that some DNA:RNA heteroduplexes are resistant to cleavage by RNase A, hence mutations may be missed in the analysis. Secondly the use of RNA can often prove troublesome as it requires an RNase A free environment.

Denaturing Gradient gel electrophoresis (DGGE) (Sheffield *et al.*, 1989; Traystman *et al.*, 1990) exploits the small differences in melting temperatures (Tm) of two DNA molecules differing by a single base. The technique involves the electrophoresis of DNA fragments through a gel containing increasing amounts of denaturant (usually a chemical gradient, however temperature gradients may also be used). As the DNA migrates in the gel, it encounters increasing concentrations of denaturant and at some point, it becomes partially or totally single stranded. The

position in the gel where the DNA becomes denatured is determined by its nucleotide sequence and composition. A single base pair difference between a normal and a mutant DNA duplex is sufficient to cause them to migrate to different positions in the gel. The technique is potentially highly sensitive and reproducible however it requires the attachment of GC-sequences (clamps) to the 5' ends of all primers to provide this sensitivity which makes it an expensive procedure when large numbers of fragments are being analysed. This method is more suitable to diagnostic work where the mutation is already known, rather than for screening large samples for unknown mutations.

Direct sequencing of genes is effective, however it is tedious and expensive, furthermore a high percentage of spurious bands are observed on sequencing hence necessitating replication of sequencing for verification. This method is essential in identifying mutations in regions containing putative alterations, however it is not efficient for broad scale screening of large regions or sample numbers.

The protein truncation test (PTT) or *in vitro* protein synthesis assay (IVPS) is based on the use of RT-PCR in combination with transcription and translation of genomic PCR products in vitro (Roest *et al.*, 1993; Powell *et al.*, 1993). The resulting radiolabelled peptides are separated by polyacrylamide gel electrophoresis and compared to a wild-type full length peptide to identify any truncated products. DNA from patients displaying truncated products are further analysed to determine the exact molecular nature of the truncation. This method is highly successful in the identification of translation terminating mutations or large deletions, and it has been instrumental in the identification of hundreds of mutations in genes where early terminations are common, such as in the APC gene in Familial Adenomatous polyposis coli. (O'Sullivan *et al.*, 1998). Despite its effectiveness, we chose not to use this technique in our analysis as it relies on the use of RNA as a template. RNA preparation is tedious, requires an RNase free environment and RNA is less robust than DNA, in addition the identification of a number of alternative RNA transcripts of the hMLH1 gene makes it difficult to determine whether observed truncated peptides are the result

of alternative transcripts or true mutations.

SSCP analysis (Orita *et al.*, 1989a; 1989b) involves electrophoresis under non-denaturing conditions, of single stranded DNA fragments. The single stranded DNA adopts a folded conformation which is stabilised by intrastrand interactions, so that the conformation and therefore the mobility is dependent on the sequence. Therefore the allelic variants of DNA segments which differ by a single base substitution, deletion or insertion exhibit a mobility shift upon denaturation and electrophoresis in a neutral polyacrylamide gel. Initial screening of the hMSH2 and hMLH1 genes in this study was carried out using SSCP as it is a simple, rapid and efficient method and has been used routinely for mutation screening in our laboratory. A more comprehensive description of SSCP is given in "SSCP analysis" Section 1.3.(VIII).

1.3.(VII) PCR AMPLIFICATION OF hMSH2 and hMLH1

The hMLH1 and hMSH2 genes contain 19 and 16 exons respectively with an average size of 150 base pairs (Kolodner *et al.*,1994, 1995). DNA oligonucleotide primers previously published by Kolodner *et al.*, (1994, 1995) were used to amplify each of these exons for subsequent SSCP analysis. The primers were designed to flank each individual exon of the genes and to include the intron-exon boundaries (Table 1.3 and 1.4).

PCR is sensitive to changes in reaction conditions-the ramping rate of the thermal cycler, reagents used, the thermostable DNA Polymerase used, magnesium concentration and annealing temperature may all effect the efficiency of the reaction. In order to test the compatibility of the published PCR amplification conditions with our reagents and equipment, PCR was initially performed on a test sample (three HNPCC DNA samples, and a negative control (no DNA). Inefficient amplification in addition to non-specific spurious amplicons were observed for most primer pairs, under the published conditions (results not shown). In order to ensure efficient accurate and

EXON		SEQUENCE 5'-3'	EXON		SEQUENCE 5'-3'
1	F	tcgcgcattttcttcaacc	9	F	gtctttacccattatttatagg
	R	gtccctccccagcacgc		R	gtatagacaaaagaattattcc
2	F	gaagtccagctaatacagtgc	10	F	ggtagtaggtatttatggaatac
	R	cttcacatttttatttttctactc		R	catgttagagcatttaggg
3	F	gcttataaaatttttaaagtatgttc	11	F	cacattgcttctagtacac
	R	gcctttcctaggcctggaatctcc		R	ccaggtgacattcagaac
4	F	ttcatttttgctttttcttattcc	12	F	attcagtattcctgtgtac
	R	atatgacagaaatatccttc		R	cgttacccccacaaagc
5	F	ccagtggtatagaaatcttcg	13	F	cgcgattaatcatcagtg
	R	ccaatcaacatttttaaccc		R	ggacagagacatacatttctatc
6	F	gttttcactaatgagcttgcc	14	F	taccacattttatgtgatgg
	R	gtggtataatcatgtggg		R	ggggtagtaagtttccc
7	F	gacttacgtgcttagttg	15	F	ctcttctcatgctgtccc
	R	gtatatattgtatgagttgaagg		R	atagagaagctaagttaaac
8	F	gatttgtattctgtaaaatgagatc	16	F	taattactcatgggacattc
	R	ggcctttgctttttaaaaataac		R	taccttcattccattactgg

Table 1.3. Primer sequences used to amplify each exon of the hMSH2 gene. Sequences are written 5'-3'. F denotes forward primer, R denotes reverse primer.

EXON		SEQUENCE 5'-3'	EXON		SEQUENCE 5'-3'
1	F	aagcactgaggtgattggc	11	F	gggctttttctcccccctccc
	R	tcgtagcccttaagtgagc		R	aaaatctgggctctcacg
2	F	aatatgtacattagagtagttg	12	F	aattatacctcatactagc
	R	cagagaaaggtcctgactc		R	gttttattacagaataaaggagg
3	F	agagatttggaaaatgagtaac	13	F	tgcaacccacaaaatttggc
	R	acaatgtcatcacaggagg		R	ctttctccatttccaaaacc
4	F	aacctttccctttggtgagg	14	F	tggtgtctctagttctgg
	R	gattactctgagacctaggc		R	cattgttgtagtagctctgc
5	F	gattttctctttttccccttggg	15	F	cccatttgtcccaactgg
	R	caaacaaagcttcaacaatttac		R	cggtcagttgaaatgtcag
6	F	gggttttatttttcaagtacttctag	16	F	catttggatgctccgttaaagc
	R	gctcagcaactgttcaatgtatgagc		R	cacccggctggaaatttttatttg
7	F	ctagtgtgtgtttttggc	17	F	ggaaaggcactggagaaatggg
	R	cataaccttatctccacc		R	ccctccagcacacatgcatgtaccg
8	F	ctcagccatgagacaataaatcc	18	F	taagtagtctgtgatctccg
	R	ggttcccaaataatgtgatgg		R	atgtatgaggtcctgtcc
9	F	caaaagcttcagaatctc	19	F	gacaccagttgatgttgg
	R	ctgtgggtgttcctgtgagtgg		R	gagaaagaagaacacatccc
10	F	catgactttgtgtgaatgtacacc			
	R	gaggagagcctgatagaacatctg			

Table 1.4. Primer sequences used to amplify each exon of the hMLH1 gene. Sequences are written 5'-3' F denotes forward primer, R denotes reverse primer.

specific PCR products, optimisation assays were performed for each primer set (table 1.5 outlines the optimised PCR conditions).

PCR OPTIMISATION

Hot- Start

In order to ensure complete denaturation of double stranded template, and to avoid the risk of introducing non-specific amplicons at an early stage, an initial "Hot-start" was performed for all PCR reactions. The 'Hot Start ' technique involves the deliberate exclusion of at least one essential reagent from the reaction mixture. The missing component (in this case Taq DNA polymerase) is added only after the reaction mixture has reached a temperature of at least 70°C. In the optimisation of samples, Taq DNA polymerase was only added after the PCR reaction mixture had undergone an initial heating step to 94°C for 5 minutes. The effect of "hot- start" is that it prevents the synthesis of extraneous products arising from non-specific primer/template hybridisation at temperatures between 30°C and 60°C, a range at which Taq polymerase retains partial activity. (Ruano *et al.*, 1991). These non-specific products could be synthesised during the first cycle of the PCR if the entire reaction was prepared at room temperature and ramped to 94°C. In subsequent cycles, spurious products could serve as templates resulting in a heterogeneous population of amplified products.

Magnesium

Changes in the PCR reaction buffer will usually effect the outcome of the reaction. In particular changes in magnesium concentrations profoundly effect the efficiency and specificity of the reaction. Excess magnesium usually results in the

accumulation of non-specific amplification products while insufficient magnesium will reduce the yield. Furthermore, deoxynucleotide triphosphates quantitatively bind magnesium therefore the concentration of dNTPs in the reaction will have bearing on the amount of free magnesium available.

In order to ensure the most efficient reaction conditions for this study, a magnesium titration curve was performed for each primer set. In the magnesium optimisation procedure, five samples were amplified (three HNPCC DNA samples, a standard lab control DNA sample, and a negative control (no DNA)). The reactions were performed at a range of magnesium concentrations, all other components remained constant. Figure 1.26 demonstrates the effect magnesium titration may have on a particular amplification reaction. Table 1.5 outlines the optimum PCR reaction conditions obtained for each primer set.

Temperature

A temperature optimisation reaction was performed on any primer sets where residual non-specific products were observed or inefficient amplification occurred post-magnesium optimisation. The annealing temperature at which a PCR is performed depends on the length and base content of the primers. Generally, primers with high GC content require higher annealing temperatures than those with high AT sequence however a close approximation to the exact annealing temperature can be calculated based on the sequence content of the primers. An annealing temperature permitting binding of primer to a matched template results in a highly specific amplification and reduction of non-specific amplifications. An annealing temperature above the optimum often results in the partial or complete loss of the amplification product as the primers can no longer anneal to the DNA. Conversely, an annealing temperature which is much lower than the optimum, results in the amplification of non-specific fragments, as the primers randomly anneal to the DNA at points of incomplete complimentarity. A range

hMSH2

	Magnesium concentration	Annealing temperature
EXON 1	1.5mM	58°C
EXON 2	2.5mM	55°C
EXON 3	2.5mM	55°C
EXON 4	1.5mM	55°C
EXON 5	2.5mM	55°C
EXON 6	1.5mM	56°C
EXON 7	3.5mM	55°C
EXON 8	2.5mM	55°C
EXON 9	2.0mM	55°C
EXON 10	3.0mM	56°C
EXON 11	2.5mM	54°C
EXON 12	3.0mM	55°C
EXON 13	2.0mM	56°C
EXON 14	1.5mM	55°C
EXON 15	2.5mM	55°C
EXON 16	2.5mM	57°C

Table 1.5(a). Optimised PCR conditions for the amplification of the 16 exons of the hMSH2 gene.

hMLH1

	Magnesium concentration	Annealing temperature
EXON 1	2.5mM	55°C
EXON 2	2.0mM	55°C
EXON 3	2.5mM	55°C
EXON 4	2.5mM	55°C
EXON 5	3.0mM	55°C
EXON 6	3.0mM	55°C
EXON 7	2.5mM	55°C
EXON 8	1.5mM	55°C
EXON 9	1.5mM	55°C
EXON 10	2.0mM	55°C
EXON 11	2.5mM	55°C
EXON 12	1.5mM	55°C
EXON 13	3.0mM	55°C
EXON 14	3.0mM	55°C
EXON 15	3.0mM	55°C
EXON 16	2.5mM	55°C
EXON 17	2.5mM	55°C
EXON 18	2.0mM	55°C
EXON 19	1.5mM	55°C

Table 1.5(b). Optimised PCR conditions for the amplification of the 19 exons of the hMLH1 gene.

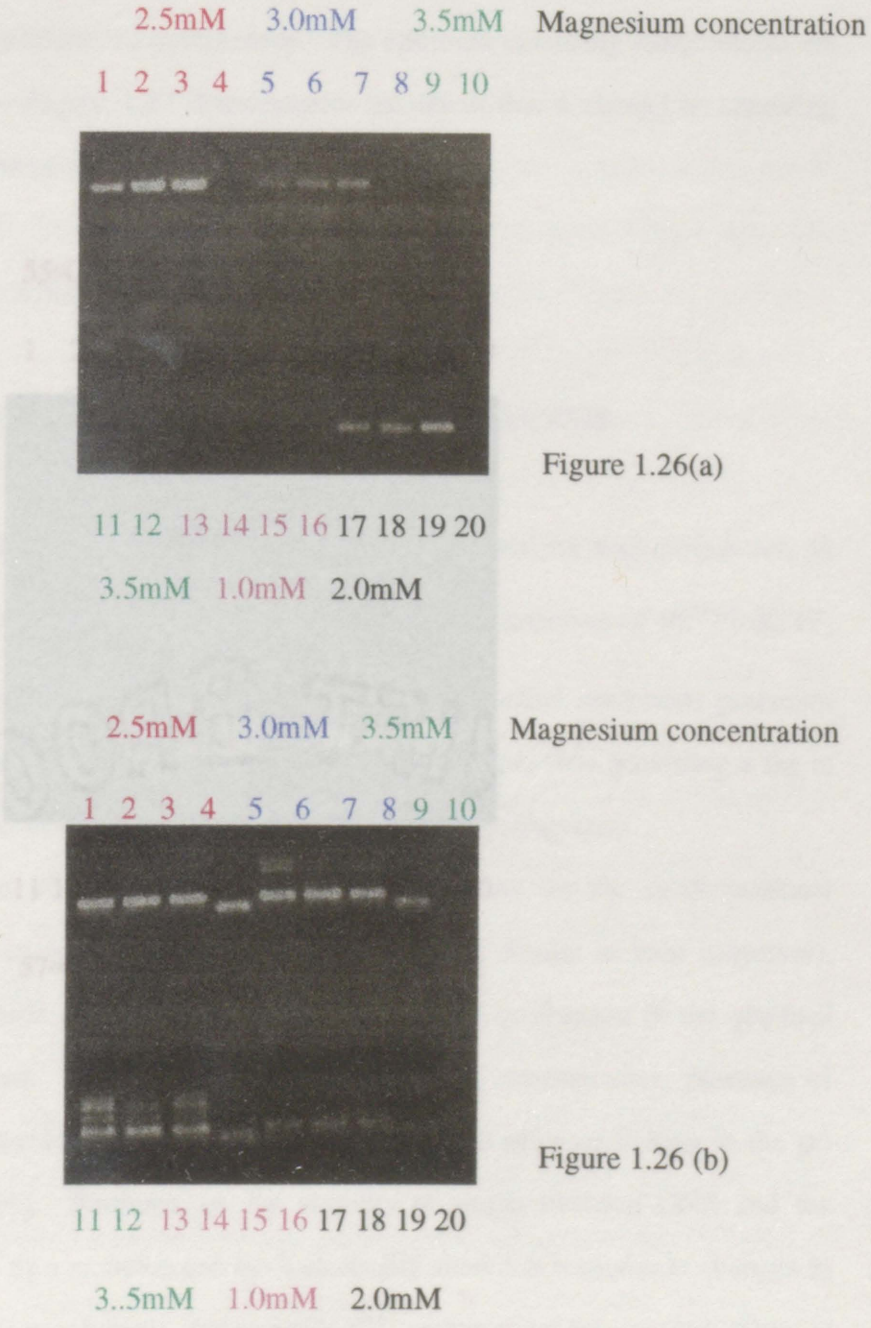


Figure 1.26. Agarose gel analysis of PCR products amplified from genomic DNA using primers specific for exon 5 of the hMSH2 gene (a) and exon 17 of the hMLH1 gene (b). Four samples (2 HNPCC samples, 1 laboratory control and a negative control (no DNA)) were amplified in PCR reactions containing a range of magnesium concentrations (indicated above lane numbers).

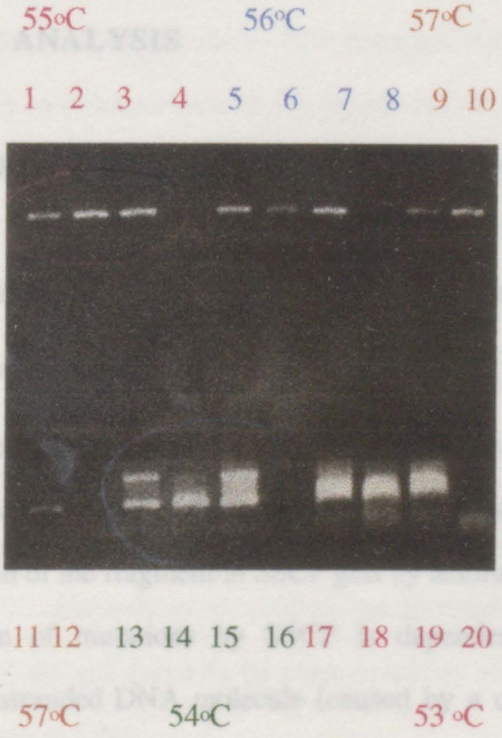


Figure 1.27. Agarose gel analysis of products amplified from genomic DNA using primers specific for exon 8 of the hMSH2 gene. PCR amplification was performed on 4 samples (2 HNPCC samples, 1 laboratory control and a negative control (no DNA) at five different annealing temperatures (as shown above la ne numbers.).

of annealing temperatures close to the calculated annealing temperature were employed in optimising the amplifications in this study. The optimum annealing temperatures are shown in Table 1.5, Figure 1.27 demonstrates the effect that a change in annealing temperature may have on a PCR reaction.

1.3.(VIII) SSCP ANALYSIS

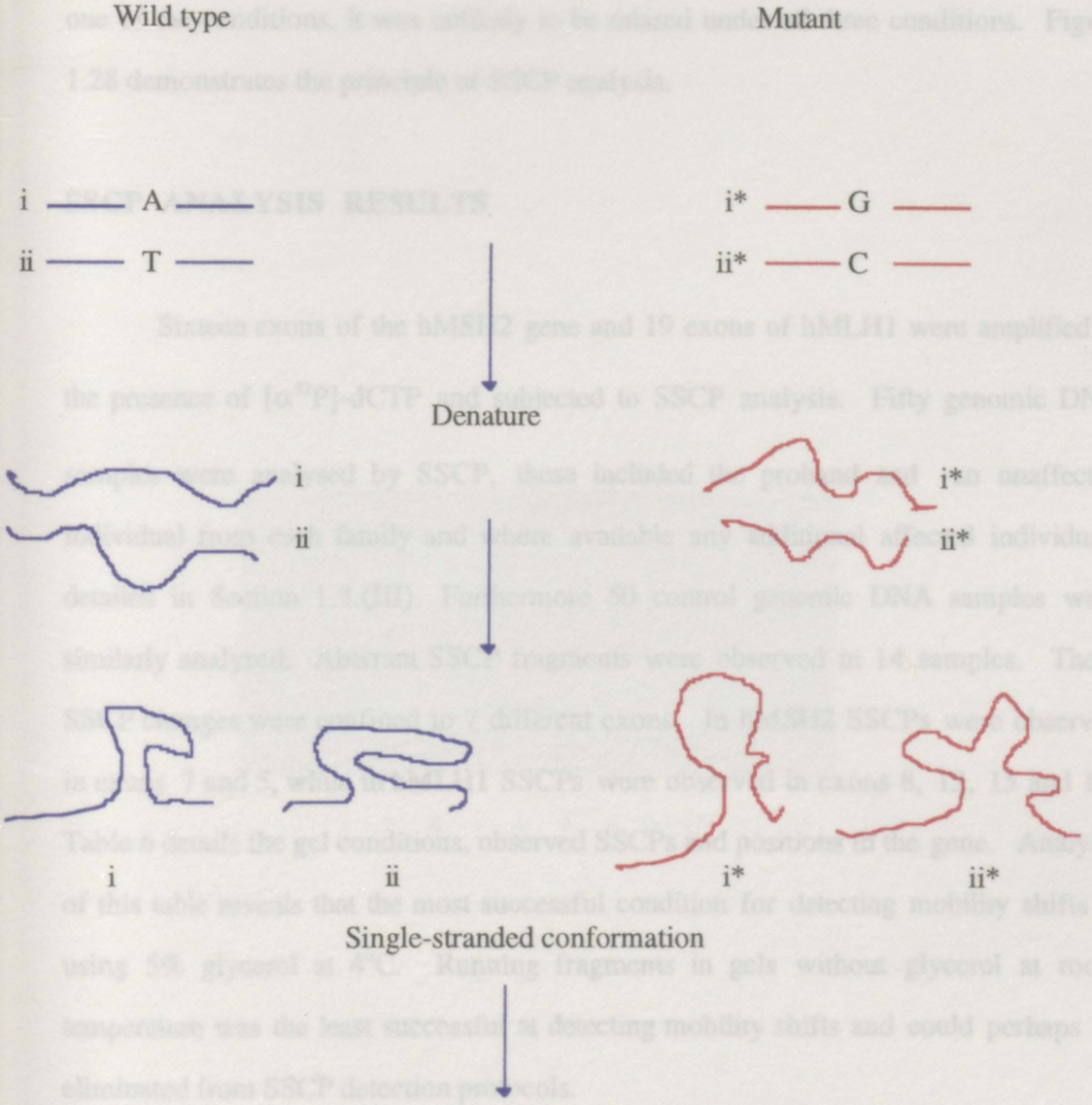
AMPLIFICATION OF SAMPLES FOR SSCP ANALYSIS

Once the optimum PCR conditions had been determined for each primer set, all DNA samples to be analysed by SSCP were amplified in the presence of [α^{32}P]-dCTP, using original genomic DNA as the template. The radiolabelled nucleotide precursor was incorporated into the amplified PCR product at C positions thus providing a tag to facilitate the detection of the fragment in SSCP gels by autoradiography.

The detection of mutations by SSCP is dependent on the conformational changes of a single stranded DNA molecule (caused by a change in base sequence), within a polyacrylamide gel, hence the method is sensitive to changes in the physical environment of the gel. This includes, temperature, ionic concentration, presence of solvents such as glycerol (Orita *et al.*, 1989a) and the extent of cross-linking in the gel (Hayashi *et al.*, 1991). Furthermore, the mobility of single stranded DNA and the mobility shift caused by a mutation can be dramatically altered in response to changes in the temperature of electrophoresis, for example 4°C versus room temperature (Orita *et al.*, 1989b). Changes in the concentration of the running buffer are also reported to affect the mobility shift (Orita *et al.*, 1989b). Reports indicate that running the gels in 0.5 X TBE (45mMTris-Borate, 1mM EDTA) rather than 1X TBE (90mM Tris-borate,1mM EDTA) may result in sharper bands (Spinardi *et al.*, 1991) whereas other

reports indicate no significant change in band separation (Iizuka *et al.*, 1992). The presence of low concentrations (5-10%) of solvents, such as glycerol in the gel can improve separation of mutated sequences (Orita *et al.*, 1989a). The weak denaturant action of glycerol on nucleic acids is thought to partially open the folded structure of single stranded DNA and increase the probability of altering the mobility of the strand with locally confined structural differences caused by the mutation (Hayashi *et al.*, 1991). Raising the glycerol content above 10% does not result in improved separation (Spinardi *et al.*, 1991) - a concentration of 5% glycerol in the gel results in the detection of the majority of mutations (Hayashi *et al.*, 1991). Increasing the polyacrylamide concentration above 6% does not significantly improve detection of the mobility shift (Spinardi *et al.*, 1991) It has been reported that the extent of cross-linking (ratio of acrylamide to N,N'- methylenebisacrylamide) within the gel may have an effect on detection efficiency. In one report, electrophoresis was performed in either 5% polyacrylamide gels with an acrylamide:bisacrylamide ratio of 29:1 or in 6% polyacrylamide gels with a ratio of acryalmide:bisacrylamide of 19:1 with similar results (Michaud *et al.*, 1992). Hayashi *et al*, (1991), recommend a ratio of 99:1 since the increased pore size of the gel seems to be more sensitive with respect to detection of conformational changes in DNA .

On analysis of all the reported SSCP conditions, SSCP in this project was performed using 6% polyacrylamide gels with an acrylamide:N,N'-methylenebisacrylamide ratio of 99:1. The electrophoresis buffer used was 0.5X TBE. Constant temperature was maintained during electrophoresis using thin gels (0.4mm), a gel apparatus with an aluminium plate to maintain an even temperature distribution and a fan was used to cool the gel. Electrophoresis was performed under three different conditions (1) with 5% glycerol at room temperature (2) with 5% glycerol at 4°C (3) without glycerol at 4°C. This was to ensure that if a mobility shift was missed under



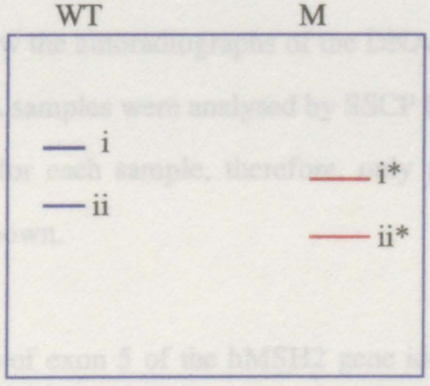


Figure 1.28. The principle of SSCP analysis. Wild type (WT) and mutant (M) DNA fragments are amplified by PCR, denatured and applied to a non-denaturing polyacrylamide gel, where they assume a secondary conformation which is dependent on their sequence. The presence of the G and C residues in strands i* and ii* cause them to assume secondary conformations which differ from the wild type conformation (i and ii) and thus to migrate in a different manner on the polyacrylamide gel.

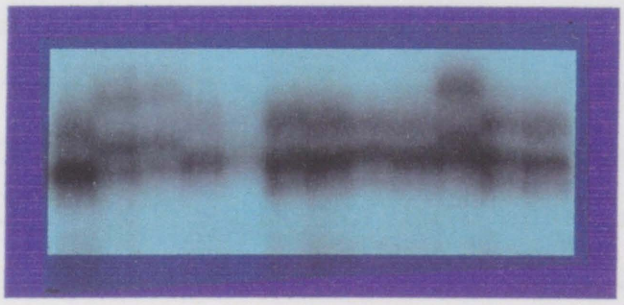
one of the conditions, it was unlikely to be missed under all three conditions. Figure 1.28 demonstrates the principle of SSCP analysis.

SSCP ANALYSIS RESULTS

Sixteen exons of the hMSH2 gene and 19 exons of hMLH1 were amplified in the presence of [α^{32}P]-dCTP and subjected to SSCP analysis. Fifty genomic DNA samples were analysed by SSCP, these included the proband and an unaffected individual from each family and where available any additional affected individuals detailed in Section 1.3.(III). Furthermore 50 control genomic DNA samples were similarly analysed. Aberrant SSCP fragments were observed in 14 samples. These SSCP changes were confined to 7 different exons. In hMSH2 SSCPs were observed in exons 7 and 5, while in hMLH1 SSCPs were observed in exons 8, 12, 15 and 17. Table 6 details the gel conditions, observed SSCPs and positions in the gene. Analysis of this table reveals that the most successful condition for detecting mobility shifts is using 5% glycerol at 4°C. Running fragments in gels without glycerol at room temperature was the least successful at detecting mobility shifts and could perhaps be eliminated from SSCP detection protocols.

Figures 1.29-1.33 show the autoradiographs of the DNA subjected to SSCP analysis. As 100 genomic DNA samples were analysed by SSCP for 35 exons it was not feasible to show a gel image for each sample, therefore, only protions of the gels where an SSCP is visible are shown.

SSCP analysis of exon 5 of the hMSH2 gene identified altered SSCP mobility in five fragments. **Figure 1.29** shows the autoradiograph of the SSCP analysis of Exon 5 of the hMSH2 gene under conditions of 5% glycerol at 4°C. Altered migration



1 2 3 4 5 6 7 8 9 10 11 12

Figure 1.29. Autoradiograph of SSCP analysis of exon 5 of the hMSH2 gene. Exon 5 of the hMSH2 gene was amplified from the genomic DNA of 50 affected and unaffected individuals from Irish HNPCC pedigrees (not all shown). Radiolabelled PCR products were separated by electrophoresis on a 6% polyacrylamide gel containing 5% glycerol. The gel was run at 60W and a constant temperature of 4^0C was maintained. Altered SSCP banding patterns were observed in individual IV:1 of pedigree B1(lane 2), II:2 of MCS1 (lane 3), III:7 of pedigree OS1 (lane 6), II:II of H1 (lane 7), and III:3 of pedigree F1 (lane 10)

patterns are visible in individual IV:1 of pedigree B1 (lane 2), individual II:2 of pedigree MCS1 (lane 3), Individual III:7 of pedigree OS1 (lane 6), individual II:II of pedigree H1 (lane 7) and individual III:3 of pedigree F1(lane 10). Alterations were observed under all three electrophoretic conditions (0% glycerol at 4°C, 5% glycerol at room temperature and 5% glycerol at 4°C) however, using the latter conditions, alterations were more pronounced.

Individual IV:1 of B1 and III:3 of pedigree F1 are affected HNPCC patients, individuals III:7 (OS1), II:II (H1) and II:2 (MCS1) are unaffected individuals from HNPCC pedigrees.

Figure 1.30 shows the SSCP banding pattern of exon 7 of the hMLH1 gene in 2 affected HNPCC probands under the most stringent electrophoretic conditions of 5% glycerol at 4°C. Of the 50 genomic DNA samples analysed a single altered SSCP fragment was observed in individual II:1 of pedigree D3 (lane 2). This SSCP is not highly obvious and was not observed under the two other electrophoretic conditions. However as the SSCP was not detected in the 50 normal controls analysed, the sample was sequenced to determine if it is a rare polymorphism or a disease causing mutation in this pedigree.

Figure 1.31 shows the SSCP analysis of exon 12 of hMSH2. Exon 12 of the hMSH2 gene was amplified from 50 genomic DNA samples from the Irish HNPCC families. Alterations in the SSCP migration pattern was observed in two samples- individual IV:1 of pedigree B1 (lane 5) and individual II:3 of pedigree C2 (lane 6). This SSCP is likely to be an innocuous polymorphism as the sequence change does not segregate with the disease in these families. Individual IV:1 of pedigree B1 is affected and displays the SSCP pattern, however an affected sibling IV:4 does not exhibit the same SSCP alteration in this exon (lane 8). Furthermore, individual II:3 of pedigree C2 is an unaffected member of a HNPCC pedigree. Analysis of the normal control population identified a corresponding SSCP of exon 12 in 3/50 control samples (not shown). These results are consistent with the SSCP representing a polymorphism

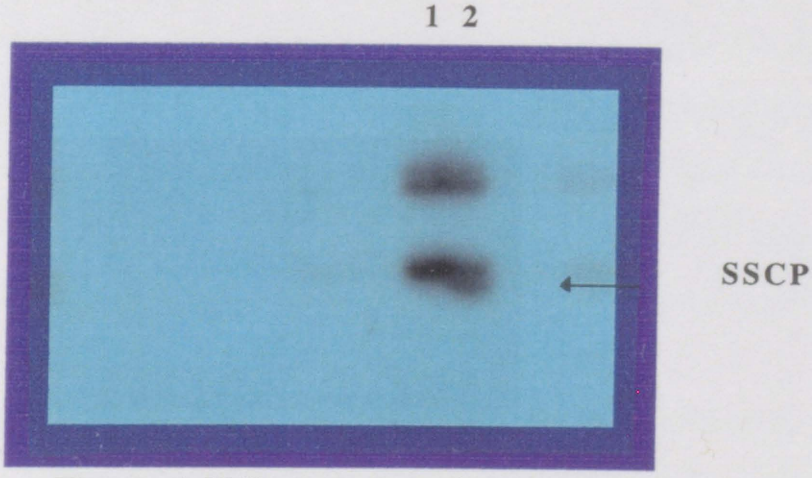


Figure 1.30(i)

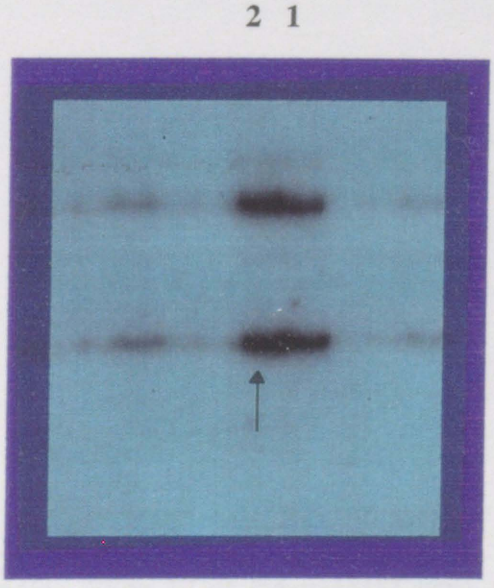


Figure 1.30(ii)

Figure 1.30. (i) Autoradiograph of SSCP analysis of exon 7 of hMLH1 in two samples; individual II:3 of pedigree C1, (lane 1) and II:1 of pedigree D3 (lane 2). Exon 7 of the hMLH1 gene was amplified from 50 genomic DNA samples from 18 Irish HNPCC pedigrees (not all shown). Radiolabelled PCR fragments were electrophoresed in a 6% polyacrylamide gel containing 5% glycerol. Analysis was performed at a temperature of 4^0C. There is a visible migration difference between samples 1 and 2. (ii) SSCP analysis of the same fragments in a 6% polyacrylamide gel in the absence of glycerol. The difference in migration is less obvious in the absence of glycerol.

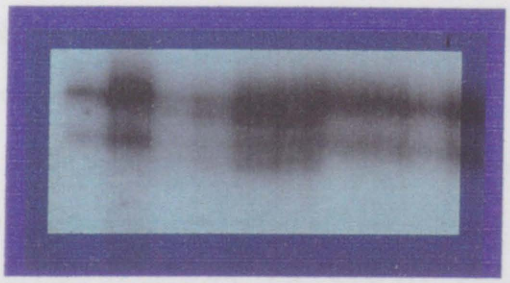


Figure 1.31. Autoradiograph of SSCP analysis of exon 12 of hMSH2 in 10 Irish HNPCC samples. Exon 12 of the hMSH2 gene was amplified from genomic DNA from 50 individuals (affected and unaffected) from HNPCC pedigrees. The PCR products were radiolabeled with α^{32}P-dCTP, separated by electrophoresis in a 6% polyacrylamide gel containing 5% glycerol. A constant temperature of 4^{o}C was maintained during electrophoresis. Alterations in migration patterns were observed in samples IV:1 an affected member of pedigree B1 (lane 5) and II:3 of pedigree C2 (unaffected) (lane 6).

rather than a disease causing mutation.

Exon 8 of the hMLH1 gene was amplified from 50 genomic DNA samples from the HNPCC families and analysed by SSCP. The SSCP analysis of exon 8 of the hMLH1 gene in 8 samples is shown in **Figure 1.32.** An altered SSCP migration pattern was observed in three samples- individual II:2 of pedigree C1 (lane 20), individual II:1 of pedigree W3 (lane 22) and individual II:1 of pedigree C1 (lane 23). This altered SSCP pattern is likely to represent a mutation in pedigree C1 as it was identified in two affected members of this family (II:1 and II:2) however the corresponding SSCP was not detected in two unaffected members of this pedigree, or in the 50 normal control samples.

SSCP analysis of exon 17 of the hMLH1 gene was performed on 50 genomic DNA samples from the 18 Irish HNPCC pedigrees. Altered SSCP electrophoresis fragments were observed in three affected individuals from pedigree H1 -individual II:4 (lane 4), individual II:7 (lane 5) and individual II:13 (lane 6). No alteration in SSCP pattern was observed in an unaffected individual of the same pedigree (II:II, lane 7). The SSCPs are likely to represent mutations in this exon in this HNPCC pedigree as the SSCP segregates with the disease in the family and is absent from an unaffected member. (Figure 1.33).

Table 1.6. details the results of the SSCP analysis of the Irish HNPCC population.

1.3.(IX) SEQUENCE ANALYSIS

In order to determine the nature of the DNA sequence changes responsible for the mobility shift observed during SSCP analysis, any fragments displaying an altered migration pattern were subsequently sequenced. Sequencing was carried out using two different methods.

Method 1 involved the PCR amplification of the product, and the subsequent cloning of the PCR fragment into the TA cloning vector. The vector was subsequently

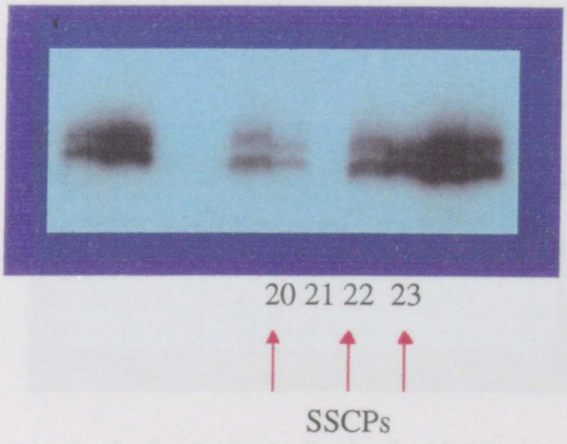
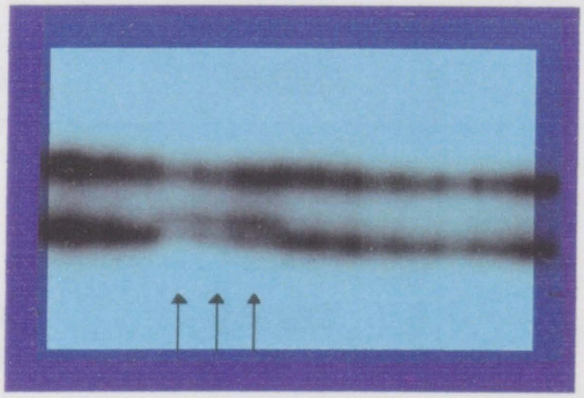


Figure 1.32. Autoradiograph of SSCP analysis of exon 8 of the hMLH1 gene in 8 samples from the Irish HNPCC families. Exon 8 of hMLH1 was amplified from 50 individuals from HNPCC pedigrees (affected and unaffected), (not all shown). Samples were PCR amplified in the presence of a radioactive precursor and analysed in a 6% polyacrylamide gel containing 5% glycerol at a temperature of 4^{o}C. Aberrent SSCP banding patterns are visible in lanes 20, 22 and 23. Lane 20 contains DNA from an affected individual II:2 of pedigree C1. Lane 23 contains DNA from individual II:1, an affected sibling, both fragments display a similar mobility shift. Lane 22 contains DNA from the proband of pedigree W3.



1 2 3 4 5 6 7 8 9 10

Figure 1.33. Autoradiograph of the SSCP analysis of exon 17 of hMLH1. Exon 17 of the hMLH1 gene was amplified from genomic DNA from 50 (affected and unaffected) individuals from HNPCC pedigrees. Samples were electrophoresed in a 6% polyacrylamide gel containing 5% glycerol, maintained at a temperature of 4°C. Altered migration patterns are visible in individual II:4 (lane 4), II:7 (lane 5) and II:13 (lane 6), all affected members of pedigree H1. Lane 7 contains DNA from an unaffected sibling (II:II) of this pedigree.

Electrophoretic conditions	Gene	Exon	SSCP observed
RT, 5% G		5	5 samples
4°C, 5% G	hMSH2	5	5 samples
4°C, 0% G		5	5 samples
RT, 5% G		7	0 sample
4°C, 5% G	hMSH2	7	1 sample
4°C, 0% G		7	0 sample
RT, 5% G		12	2 samples
4°C, 5% G	hMSH2	12	2 samples
4°C, 0% G		12	0 sample
RT, 5% G		8	2 samples
4°C, 5% G	hMLH1	8	2 samples
4°C, 0% G		8	0 sample
RT, 5% G		17	*3 samples
4°C, 5% G	hMLH1	17	3 samples
4°C, 0% G		17	0 sample

Table 1.6. Results of the SSCP analysis of hMSH2 and hMLH1 in 18 Irish HNPCC pedigrees. The conditions of analysis and observed SSCPs are detailed. *; A similar SSCP was observed in 3 affected patients within the same family, RT; room temperature, G; glycerol.

transformed into *E. coli* and colonies putatively bearing the PCR insert were selected based on an insertional-inactivation, chromogenic assay. Plasmids were subsequently isolated and sequenced using T7 DNA Polymerase with dideoxy terminator incorporation and [γ^{32}-P] labelled primers.

The second method of analysis employs the "cycle sequencing" protocol which employs the thermostable enzyme 'thermosequenase' and facilitates PCR coupled sequencing with the incorporation of [α^{33}-P] labelled dideoxy terminators.

Controls

To rule out the possibility that the identified mutations may be common polymorphisms fifty unrelated, unaffected population 'normals' were analysed. The normal population included 40 unaffected Irish samples (collected for another unrelated study) and 10 'married in' members of the HNPCC pedigrees. The 'married in' samples were obtained from individuals with no family history of the colon cancer and who were over 50 years of age. The corresponding SSCPs were not identified in any of the controls.

SEQUENCING RESULTS:

All fragments bearing altered migration patterns under SSCP conditions were sequenced to identify any mutations/polymorphisms present. Sequencing identified five mutations, three splice site and two point mutations. Of these five mutations, four are novel and one is a previously reported common HNPCC mutation. In addition a number of polymorphisms were identified.

Causative Mutations:

Mutations were identified in five pedigrees, exon 5 of hMSH2 was mutated in two pedigrees and hMLH1 was mutated in three families- two mutations in exon 8 and one in exon 17.

hMSH2 mutations:

Altered SSCP mobilities of exon 5 of the hMSH2 gene were detected in 5 samples from 5 independent families- individual IV:1 of pedigree B1, individual II:3 of pedigree F1, individual II:2 of pedigree MCS1, individual III:7 of pedigree OS1 and individual II:II of pedigree H1. DNA sequence analysis was performed on each sample-two different splice site mutations of exon 5 in hMSH2 were identified in pedigrees B1 and F1. The SSCPs detected in the other samples were identified as sequence polymorphisms in these patients.

Pedigree B1, exon 5 hMSH2 mutation:

Sequence analysis of the proband IV:1 of pedigree B1 (Figure 1.34) revealed an A to T transversion at the +3 position of the splice donor site consensus sequence. This mutation has been previously reported in 14 other pedigrees and appears to be one of the more common hMSH2 mutations (Lui *et al.*, 1994) first identified this mutation in three US pedigrees, analysis of RNA transcripts from these patients revealed that the mutation results in the alteration of the splice site, with the result that exon five is skipped and a defective, truncated protein product is produced (Lui *et al.*, 1994). Sequence analysis of two other affected members in this pedigree (IV:4 (fig.1.35a) and IV:8) identified the presence of a similar mutation. Subsequent to the identification of this mutation two undiagnosed children of the proband were screened for the mutation. Individual V:1 proved not to harbour the mutation, however the mutation was identified in individual V:3 (see Section 1.3.(XII) for a more detailed discussion). The mutation was identified in four affected members of this pedigree but was not identified in 2

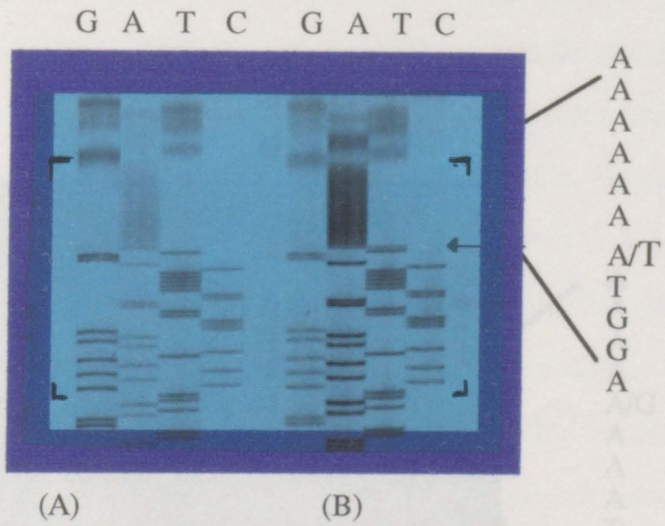


Figure 1.34. Nucleotide sequence analysis of exon 5 of hMSH2 in individual IV:1 of pedigree B1. Sequencing of this exon was carried out directly on a PCR product, using the cycle sequencing protocol with the incorporation of $\alpha^{33}P$ labelled dideoxy terminators. Lanes were loaded as indicated, sample (A) represents the nucleotide sequence of a normal unaffected individual of pedigree B1-IV:3. Sample (B) shows the mutant sequence of the affected proband of pedigree B1. An A-T transversion was observed at the +3 splice donor site (indicated by the arrow). A similar mutation was observed in two other affected members of this family and an undiagnosed individual V:3.

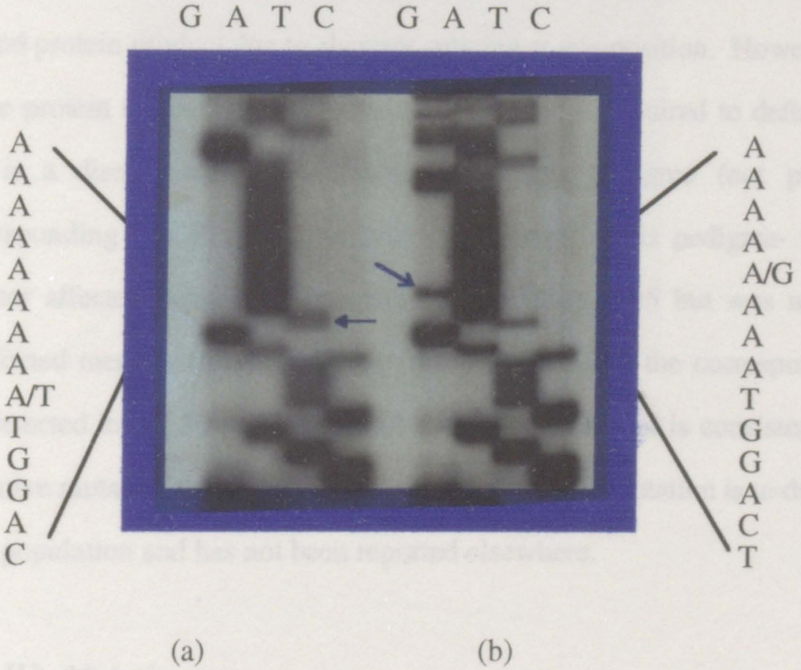


Figure 1.35.(a) Nucleotide sequence analysis of exon 5 of hMSH2 in individual IV:4 an affected sibling of individual IV:1 in pedigree B1. An A-T transition was detected in this individual at the +3 splice donor sequence position. **(b)** Nucleotide sequence analysis of the proband (III:3) of pedigree F1. An A-G transition at the +6 splice donor site in this individual was detected. Sequence analysis in both samples was performed on PCR products using the T7 Sequenase protocol with the incorporation of α^{32}P-dCTP label.

unaffected members (V:1 and IV:3), this is consistent with the base change being a causative mutation within the pedigree.

Pedigree F1 exon 5 hMSH2 mutation:

SSCP analysis of individual III:3 of pedigree F1 revealed an altered migration pattern similar to that observed in pedigree B however sequence analysis revealed an A to G transition at the +6 position of the splice donor site in pedigree F1 (Figure 1.35b). This base change occurs within the region associated with the consensus sequence of the intron. Therefore, it is likely that the mutation could result in the production of an altered protein product due to aberrant splicing at this position. However RNA analysis or the protein truncation test analysis of this gene is required to definitively prove that this is a disease causing mutation within this pedigree (not performed). The corresponding SSCP segregates with the disease in this pedigree- it was present in another affected member of the pedigree-individual III:5 but was not detected in an unaffected member of this pedigree (III:8) furthermore the corresponding SSCP was not detected in the 50 normal control samples therefore it is consistent with it being a causative mutation within this pedigree. This putative mutation is to date, specific to the Irish population and has not been reported elsewhere.

hMLH1 Mutations:

Three novel mutations were identified in hMLH1 in three unrelated families.

Pedigree C1 and W3, exon 8 mutations:

SSCP analysis of exon 8 identified altered migration patterns in 2 affected individuals II:1 and II:2 of pedigree C1 and the proband (II:1) of pedigree W3. Sequence analysis of individuals II:1 and II:2 of pedigree C1 revealed a GCC-CCC change at codon 210 (**Figure 1.36**). This heterozygous point mutation results in the

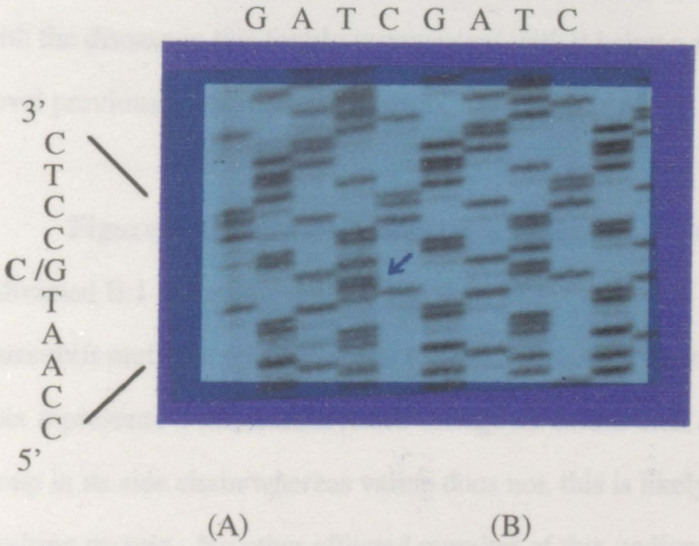


Figure 1.36. Nucleotide sequence analysis of exon 8 of the hMLH1 gene in individual II:1 of pedigree C1. Analysis was performed on PCR products using cycle sequencing with the incorporation of $\alpha^{33}P$ dideoxy terminators (both normal and mutant sequence are visible in sample A). A G-C alteration was observed in sample (A) at codon 210 (indicated by the arrow). Note: The sequence given is written in 3'-5' direction in this instance. This point mutation substitutes a proline residue for an alanine. (B) The DNA sequence of exon 8 of hMLH1 from an unaffected individual of this pedigree; II:3.

substitution of a proline residue for an alanine. Alanine is one of the more simple amino acids with a single methyl side chain, proline on the other hand, an imino acid has a cyclical structure, usually found at the bends of folded proteins, it is not averse to water and markedly influences protein architecture. Thus this amino acid change is likely to have a significant effect on the protein structure and hence affect it's activity. This sequence change was present in two affected members of this pedigree however, no corresponding SSCPs were observed in 2 unaffected individuals from this pedigree (II:3 and III:4) or in the 50 normal control individuals. The segregation of the mutation with the disease in this family is consistent with it being a mutation. This mutation is a novel previously unreported mutation.

Figure 1.37 shows the nucleotide sequence of exon 8 of the hMLH1 gene in individual II:1 of pedigree W3. The loss of a T at codon 224 was identified, this is a frameshift mutation that replaces a valine residue with a serine at the 3' end of the exon. This represents a non-conservative change in amino acid as serine contains an amino group in its side chain whereas valine does not, this is likely to effect the function of the resulting protein. No other affected member of this pedigree was available for analysis however the corresponding SSCP was not identified in 50 normal control samples. This mutation has not been previously reported.

Pedigree H1, exon 17, hMLH1:

Figure 1.38 shows the nucleotide sequence of exon 17 of the hMLH1 gene in individuals II:4, II:7, II:13 and II:11 of pedigree H1. Three affected individuals (II:4, II:7 and II:13) displayed altered SSCP patterns, whereas individual II:11 (sample A) an unaffected sibling, did not show any changes in migration pattern. Sequence analysis revealed the presence of a CCC-CTC change at codon 648 in the three affected siblings. A normal sequence pattern is clearly seen in individual II:11. This previously

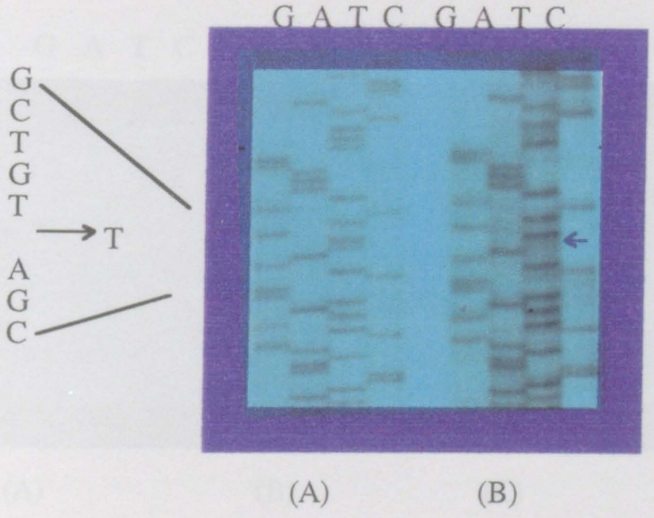


Figure 1.37. Nucleotide sequence analysis of exon 8 of the hMLH1 gene in individual II:1 of pedigree W3. The deletion of a single T residue at codon 224 is indicated by the arrow in sequence B. Sample A represents the nucleotide sequence from a normal unaffected individual. The frameshift mutation results in the non-conservative substitution of a valine residue with a serine.

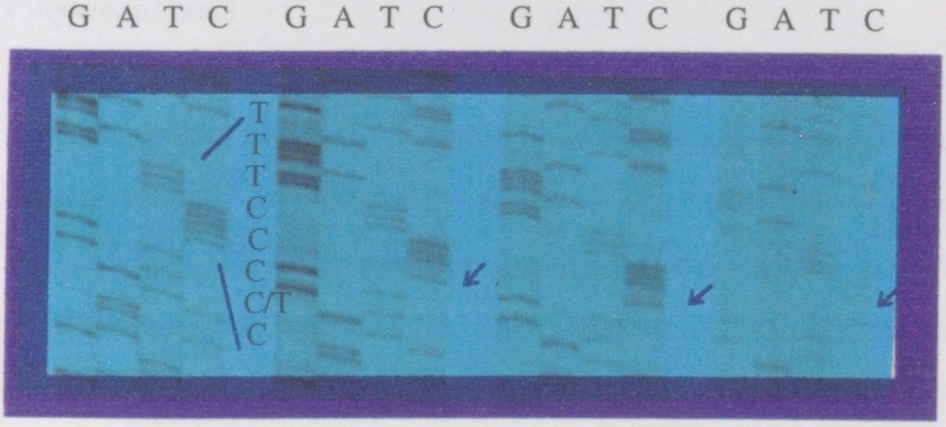


Figure 1.38. Nucleotide sequence analysis of exon 17 of the hMLH1 gene in pedigree H1. Sequence analysis was performed directly on PCR products using cycle sequencing with the incorporation of $\alpha^{33}P$ dideoxy terminators (both normal and mutant sequence are visible). Three affected and one unaffected member of this pedigree were sequenced. Individual II:II is unaffected and the normal sequence of this exon is shown in sequence A. Individuals II:4, II:7 and II:13 are affected siblings and their sequences are shown in samples B, C and D. In sequences B,C and D a C-T change observed at codon 648 is indicated by the arrow. This mutation results in the substitution of a leucine residue for a proline.

unreported mutation substitutes a proline residue with a leucine. This represents a significant amino acid change as proline an imino acid has a cyclical structure, is hydrophyllic in nature and is usually found at the folds of proteins. Leucine on the other hand, has a basic hydrophobic side chain therefore the alteration is likely to affect the protein function. The corresponding SSCP was not observed in the 50 normal control samples.

Polymorphisms

In addition to the mutations identified, a large number of polymorphisms were detected, most of these corresponding to alterations within intronic regions. The common hMSH2 exon 5 polymorphisms were identified in four samples as shown (Figure 1.39). This highly polymorphic polyA tract is altered in individual II:2 (MCS1)-sequence A, individual III:7 (OS1) sequence B, II:II (H1) sequence C, and individual II:3 (W1) sequence D. Sequence analysis revealed alterations in the number of A residues in the polyA tracts of each sample, however it proved difficult to count the exact number of As spanning this tract. The base changes are unlikely to be causative mutations as they were identified in unaffected individuals of the HNPCC pedigrees. Furthermore this region is reported to be highly polymorphic in the general population and is a repeat sequence frequently employed in the analysis of replication error phenotype (at the BAT 25 locus) **Figure 1.39** displays the differences in this region in sequences A, B, C and D. Individual II:2 (MCS1) displays an extra GAG in the polyA tract in addition to an increase in the number of A's (sequence A).

Figure 1.40 shows the sequence analysis of exon 7 (hMSH2) from individual II:1 in pedigree D3. A T-C alteration was observed at the downstream -7 position, however this base change lies outside the splice site consensus sequence range and is unlikely to affect splicing. Furthermore the corresponding SSCP was not observed in one other affected member within this family; II:2, (this individual was deceased at the time of the study, however paraffin embedded archival tissue was

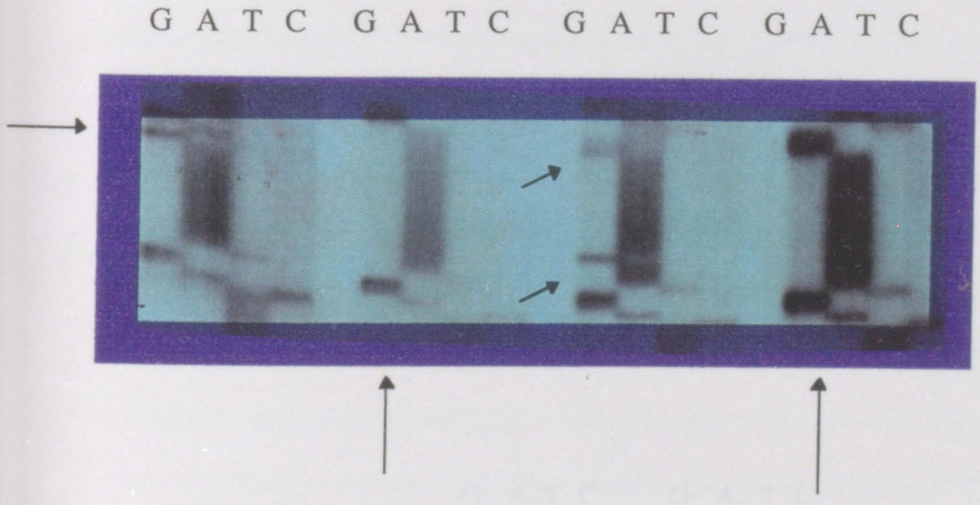


Figure 1.39. Nucleotide sequence analysis of exon 5 of hMSH2 in samples which displayed fragment alterations on SSCP analysis. Sequence analysis was perfomed on cloned PCR products. A number of polymorphisms were identified in this region of poly A sequence, however due to the large number of A residues in the run it was difficult to determine the exact sequence change. This region has been shown to be highly polymorphic in the general population. This sequence shows polymorphisms identified in 4 Irish samples; individual II:2 (MCS1)-sample A, III:7 (OS1)-sample B, II:II (H1)-sample C and II:3 (W1)-sample D.

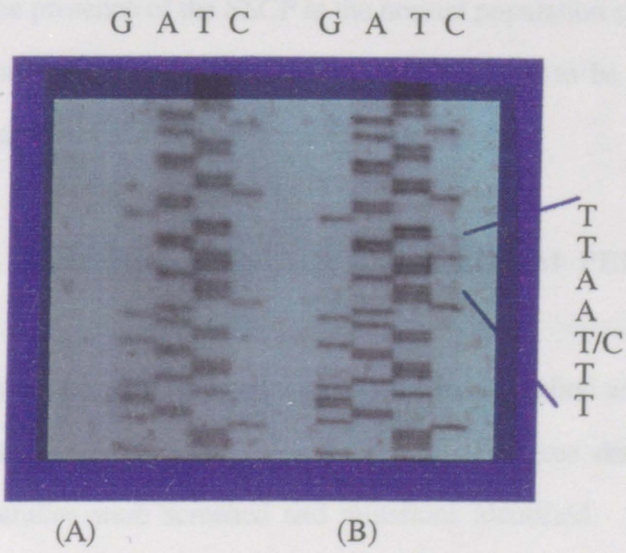


Figure 1.40. Nucleotide sequence analysis of exon 7 of the hMSH2 gene in individual II:1 of pedigree D3 (sample A). Cycle sequencing of a PCR product identified a T-C alteration at the -7 position of the splice site. Sample A represents the sequence from a normal unaffected individual.

available and was used to obtain genomic DNA for analysis). RNA or protein analysis is required to prove the presence or absence of a splice site mutation in this individual (not performed).

Alterations in SCCP migration patterns were also observed in exon 12 of hMSH2 in two samples. An SSCP of exon 12 of hMSH2 was detected in individual IV:4 of pedigree B1. This is an affected member of a HNPCC family and this sample was previously shown to contain a mutation of exon 5 of the same gene. The exon 5 mutation segregates with the disease in this pedigree however the SSCP of exon 12 was not present in three other affected members of this pedigree (IV:8, IV:1 and V:3). Furthermore, the corresponding exon 12 SSCP was identified in 3/ 60 normal control samples which is consistent with a polymorphism rather than a causative mutation. An SSCP of exon 12 was also seen in individual II:3 of pedigree C2. This individual is an unaffected member of this pedigree and although this individual may develop colon cancer later in life, the presence of the SSCP in the normal population suggests that it is most likely a polymorphism. As these SSCPs were believed to be polymorphisms sequence analysis was not performed on these samples.

1.3.(X) TYPICAL VERSUS ATYPICAL AMSTERDAM PEDIGREES

As previously stated, genetic screening was carried out blind as to whether the pedigrees satisfied all the Amsterdam Criteria or not. Pedigree details were only examined after all families were screened and mutations identified. Details of each pedigree included in this study are provided in Section 1.3.(III) and for ease of reference, determination of whether pedigrees are 'typical Amsterdam' (satisfy all the Amsterdam criteria) or 'atypical' (not satisfying all the criteria) has been included with the family details. Correlation of pedigrees with mutations identified and 'Amsterdam' status reveals that all the identified mutations are confined to typical Amsterdam pedigrees. SSCP analysis can fail to detect up to 5% of mutations and our mutation

screening did not include promoter regions which are potential mutation sites, however this does not adequately explain the lack of mutations in the Non-Amsterdam pedigrees. Low frequencies of hMSH2 and hMLH1 mutations in atypical HNPCC pedigrees was recently reported by Wijnen *et al.*, (1997), in their analysis hMSH2 and hMLH1 gene mutations were identified in 50% of families fully complying with the Amsterdam selection criteria, whereas, disease causing mutations were only identified in 8% of families which did not satisfy the criteria. The screening protocol in the Irish pedigrees focused on hMSH2 and hMLH1 only, mutations were detected in 5 of 8 Amsterdam pedigrees (62%). This is in keeping with reported detection rates for typical HNPCC pedigrees (Peltomaki *et al.*, 1997), however, the lack of mutations in these genes in the atypical-Amsterdam pedigrees suggests that perhaps some other genes are responsible for colon cancer in these families. Table 1.7 details the mutation profile of the Irish HNPCC pedigrees screened in this study.

1.3.(XI) SEGREGATION ANALYSIS OF IDENTIFIED MUTATIONS IN THE HNPCC PEDIGREES.

Mutations in the mismatch repair genes were identified in five Irish HNPCC pedigrees. Following the identification of the mutation in the proband, other available affected and unaffected members of the pedigree were screened for the corresponding mutation or SSCP. This analysis was performed in order to determine if the mutation segregates with the disease in the family.

Segregation analysis was possible in four of the five families in which a mutation was identified. No other members of pedigree W3 were available for analysis therefore it cannot be definitively concluded that this is the disease causing mutation in the pedigree. However, as the corresponding SSCP was not detected in the 50 normal control samples and because the frameshift mutation results in the non-conservative alteration of an amino acid residue of the protein, which is likely to affect

FAMILY	MUTATION	CHANGE	NOVEL?	AMSTERDAM STATUS
B1	hMSH2-Exon 5 A-T @ 3' SD site	+3 Splice donor site	Previously reported in 14 pedigrees	Typical
F1	hMSH2-Exon 5 A-G @ 3' SD site	+6 Splice donor site	Novel	Typical
C1	hMLH1-Exon 8 GCC-CCC @ codon 210	Ala- Pro	Novel	Typical
H1	hMLH1-Exon 17 CCC-CTC @ codon 648	Pro-Leu	Novel	Typical
W3	hMLH1-Exon 8 T deletion @ codon 224	Val-Ser	Novel	Typical

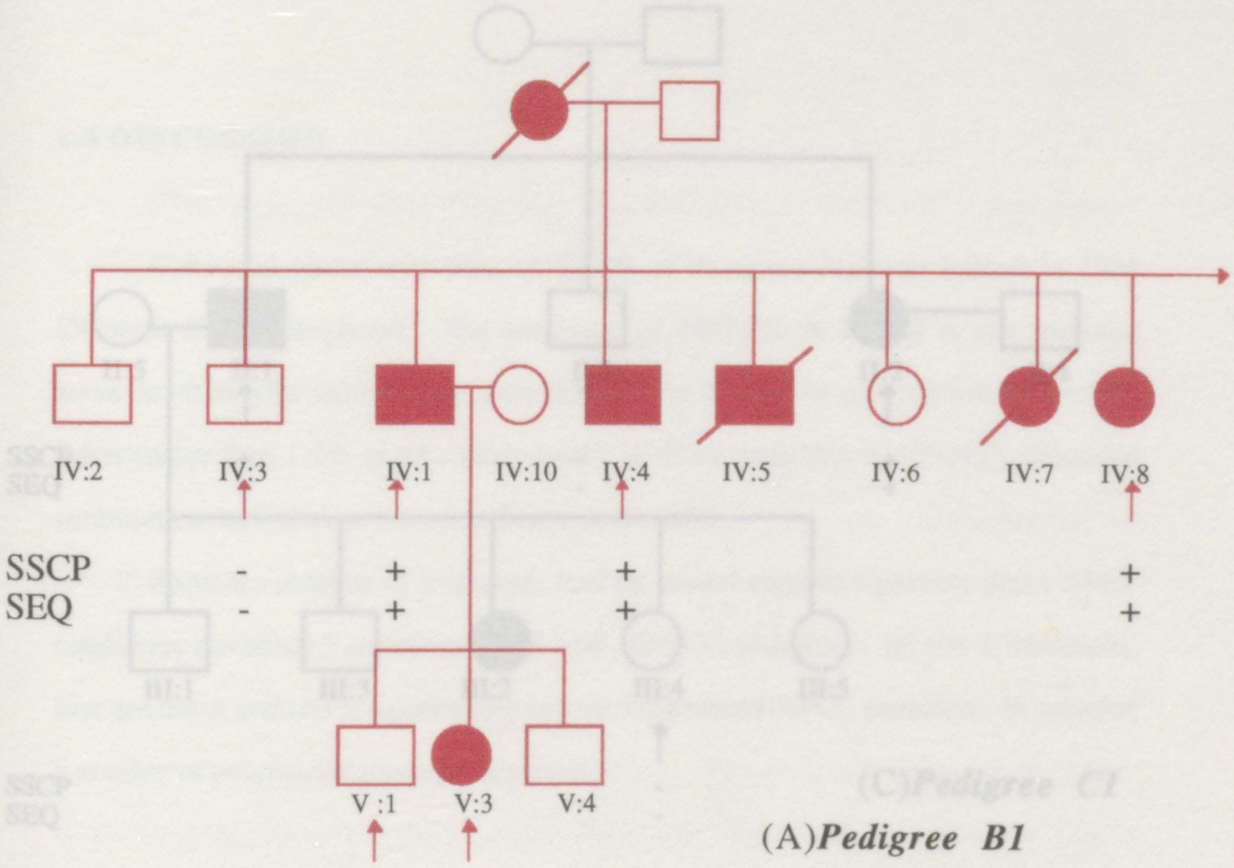
FAMILY	MUTATION	CHANGE	NOVEL?	AMSTERDAM STATUS
OS1	None detected	N/A	N/A	Typical
M3	None detected	N/A	N/A	Typical
MCS1	None detected	N/A	N/A	Typical
C2	None detected	N/A	N/A	Atypical
M1	None detected	N/A	N/A	Atypical
W2	None detected	N/A	N/A	Atypical
W3	None detected	N/A	N/A	Atypical
D1	None detected	N/A	N/A	Atypical
D2	None detected	N/A	N/A	Atypical
MG	None detected	N/A	N/A	Atypical
D5	None detected	N/A	N/A	Atypical
D6	None detected	N/A	N/A	Atypical
W5	None detected	N/A	N/A	Atypical

Table 1.7. Mutation profile of the Irish HNPCC pedigrees analysed in this study.

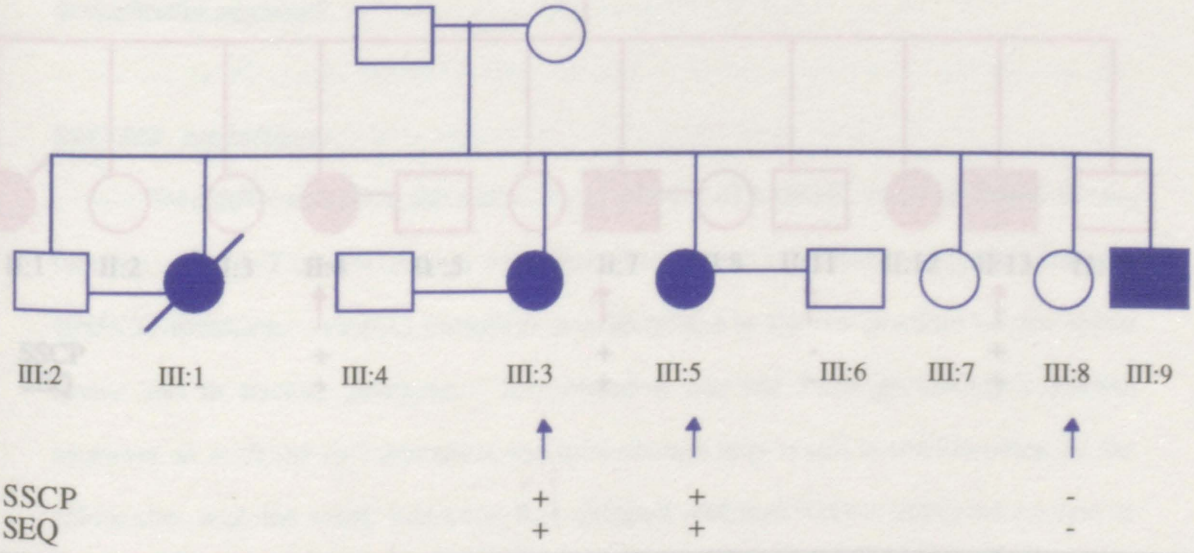
the protein function, it is likely that this is the disease causing mutation within this individual (Figure 1.41).

1.3.(XII) MUTATION SCREENING- FAMILY SIGNIFICANCE:

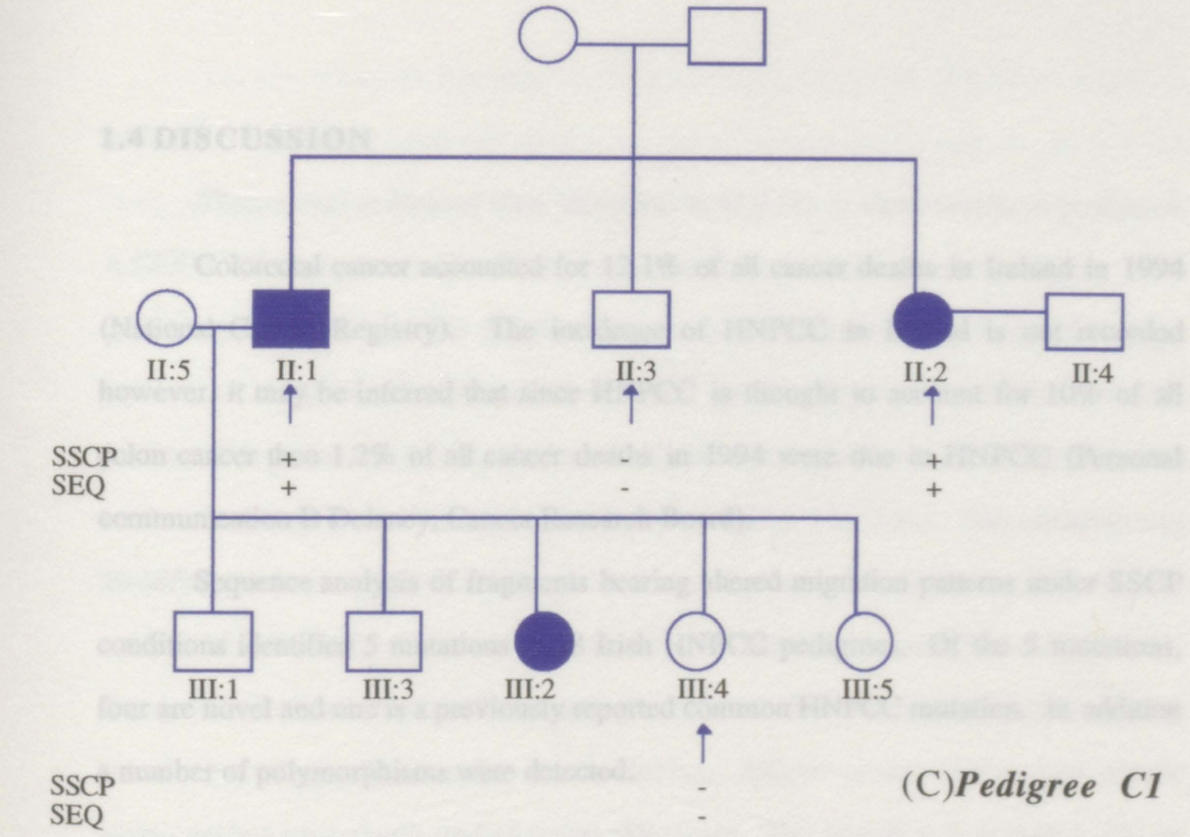
As previously stated, a A-T transversion at the +3 splice donor site was identified in exon 5 of hMSH2 in pedigree B1. This mutation results in the alteration of the splice donor site consensus sequence and results in the subsequent production of a non-functional, truncated protein product (Lui *et al.*, 1994). At the time of analysis four members of this pedigree were available for screening, IV:1, IV:4 and IV:8 (all affected) and IV:3 (unaffected). The mutation was identified only in the affected individuals and was not present in individual IV:3. Subsequent to the identification of the mutation within this pedigree, individual V:3 (daughter of IV:1) presented for colonoscopy at the age of 31yrs. A small colonic lesion was identified, however histological analysis was clear. Analysis of genomic DNA revealed the presence of the predisposing mutation in this individual. Her brother, individual V:1 volunteered for analysis and proved clear of any predisposing mutation. Identification of predisposing mutations prior to tumour development is important to assist at clinical level in terms of surveillance, diagnosis and follow-up.



(A)***Pedigree B1***



(B)***Pedigree F1***



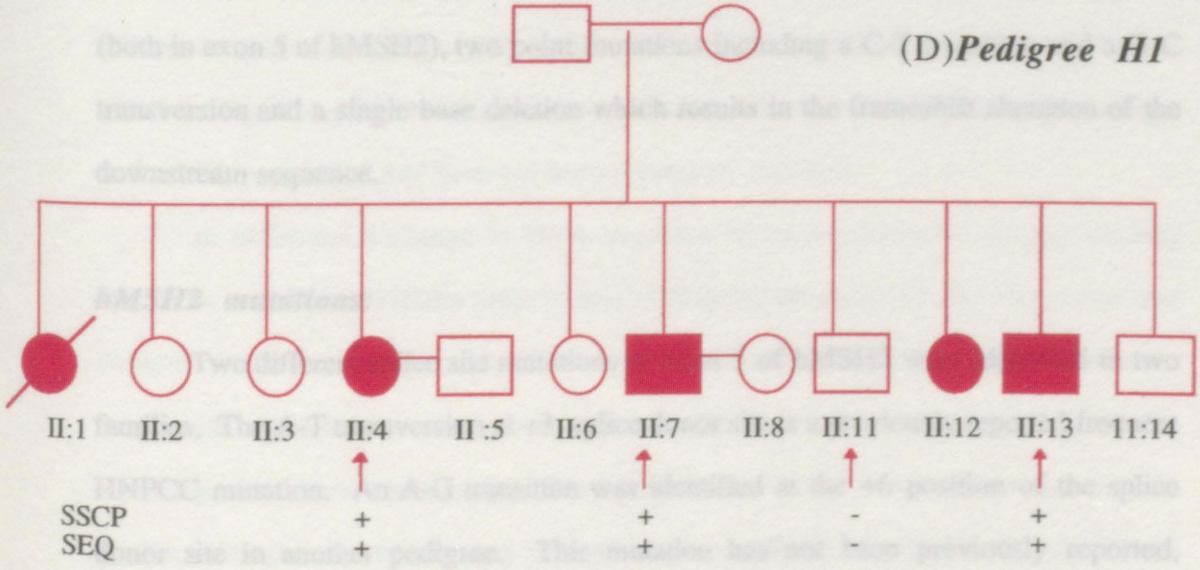


Figure 1.41. Pedigree diagrams showing the results of segregation analysis of the mutation and corresponding SSCP with the disease in pedigree B1(A), F1(B), C1(C) and H1(D). The arrows indicate samples which were analysed, under each analysed individual a + or - symbol indicates the presence or absence of the SSCP or mutation. SEQ denotes sequence analysis.

1.4 DISCUSSION

Colorectal cancer accounted for 12.1% of all cancer deaths in Ireland in 1994 (National Cancer Registry). The incidence of HNPCC in Ireland is not recorded however, it may be inferred that since HNPCC is thought to account for 10% of all colon cancer then 1.2% of all cancer deaths in 1994 were due to HNPCC (Personal communication D Delaney, Cancer Research Board).

Sequence analysis of fragments bearing altered migration patterns under SSCP conditions identified 5 mutations in 18 Irish HNPCC pedigrees. Of the 5 mutations, four are novel and one is a previously reported common HNPCC mutation. In addition a number of polymorphisms were detected.

1.4.(I) IRISH HNPCC MUTATIONS

The mutations identified in the Irish population include two splice site mutations (both in exon 5 of hMSH2), two point mutations including a C-T transition and a G-C transversion and a single base deletion which results in the frameshift alteration of the downstream sequence.

hMSH2 mutations:

Two different splice site mutations of exon 5 of hMSH2 were identified in two families. The A-T transversion at +3, splice donor site is a previously reported frequent HNPCC mutation. An A-G transition was identified at the +6 position of the splice donor site in another pedigree. This mutation has not been previously reported, however as with the A-T mutation, the base change may result in the alteration of the splice site, with the result that exon 5 is skipped and a defective truncated protein is produced.

hMLH1 mutations:

Three novel mutations were identified in hMLH1 in three unrelated pedigrees. A CCC-CTC change was identified at codon 648 in exon 17 in one pedigree. This mutation substitutes a proline residue with a leucine. This is a significant change as proline, an imino acid has a cyclical structure usually found at the bends of folded proteins, it is not averse to water and markedly influences protein architecture. Leucine on the other hand has a basic hydrophobic hydrocarbon side chain. The mutation was identified in three affected members of this pedigree and was not present in one unaffected individual. Two novel mutations were identified in exon 8. A GCC-CCC change was detected at codon 210 in a single pedigree, this point mutation results in the substitution of a proline residue for an alanine. Alanine is one of the more simple amino acids with a single methyl group side chain. The loss of a T at codon 224 in another pedigree was also observed. This frameshift mutation substitutes a valine residue with a serine. This represents a non-conservative change in amino acid as serine contains an amino group in its side chain whereas valine does not, this is likely to effect the function of the resulting protein. The mutations in hMLH1 are to date, unique to the Irish population and have not been previously reported.

In order for a change in DNA sequence to be considered a disease causing mutation, a number of criteria must be met. The mutation must (1) alter the amino acid sequence in the resulting protein. This may be at the exact site of the DNA alteration or may occur elsewhere in the protein as a result of a sequence change (splice-site alteration, insertion/deletion) (ii) The mutation must be absent from the normal unaffected population (iii) it must segregate with the disease in the family. Additional evidence of a disease causing mutation is provided by (iv) the conservation of the amino acid across species and (v) the presence of the mutation in more than one affected family.

The five mutations identified in this study satisfy all of the first three criteria, in all cases the resulting protein is altered, the mutation segregates with the disease and none of the mutations were identified in the normal population. Analysis of amino acids across a number of species performed using a BLAST search reveals that in the case of mutations that cause a specific amino acid change the mutations were identified in residues which are conserved across different species (Figure 1.42). The splice site mutations result in the skipping of exon 5 of the hMSH2 gene, this entire exon is highly conserved across species therefore deletion of the exon would have major implications for protein function (Figure.1.43).

The mutation at codon 648 of exon 17 of hMLH1 causes a proline residue to be replaced by a leucine. Peptide sequence analysis demonstrates that this residue is conserved in *R norvegicus* (Rat) and *S. cervisiae.* The importance of this residue is reiterated by noting that the sequence immediately preceeding the proline is not conserved between human and *S. cervisiae*, however the proline *is* conserved perhaps because it is essential for the functionality of the protein (Figure 1.43).

The mutation at codon 210 of exon 8 in hMLH1 results in the substitution of an alanine with a proline. The alanine residue is conserved across species including (mutL) *R. norvegicus* . The two exon 5 mutations in hMSH2 result in the loss of the entire exon. Comparison of the exon sequence across species reveals the high level of conservation between mouse (*M. musculus*), rat (*R. norvegicus*) and homo sapiens. Only 1 of 50 amino acids in this peptide vary between the species (Figure 1.43).

Of the mutations identified only the +3, splice donor site of exon 5 in hMSH2 was identified in other pedigrees. The exon 5 hMSH2 splice site mutation in pedigree B has been previously identified in 14 other pedigrees of diverse geographical background (Lui *et al.*, 1994; Peltomaki *et al.*, 1997). The mutation has not been detected in any other Irish HNPCC pedigrees however future analysis of more Irish families may identify it

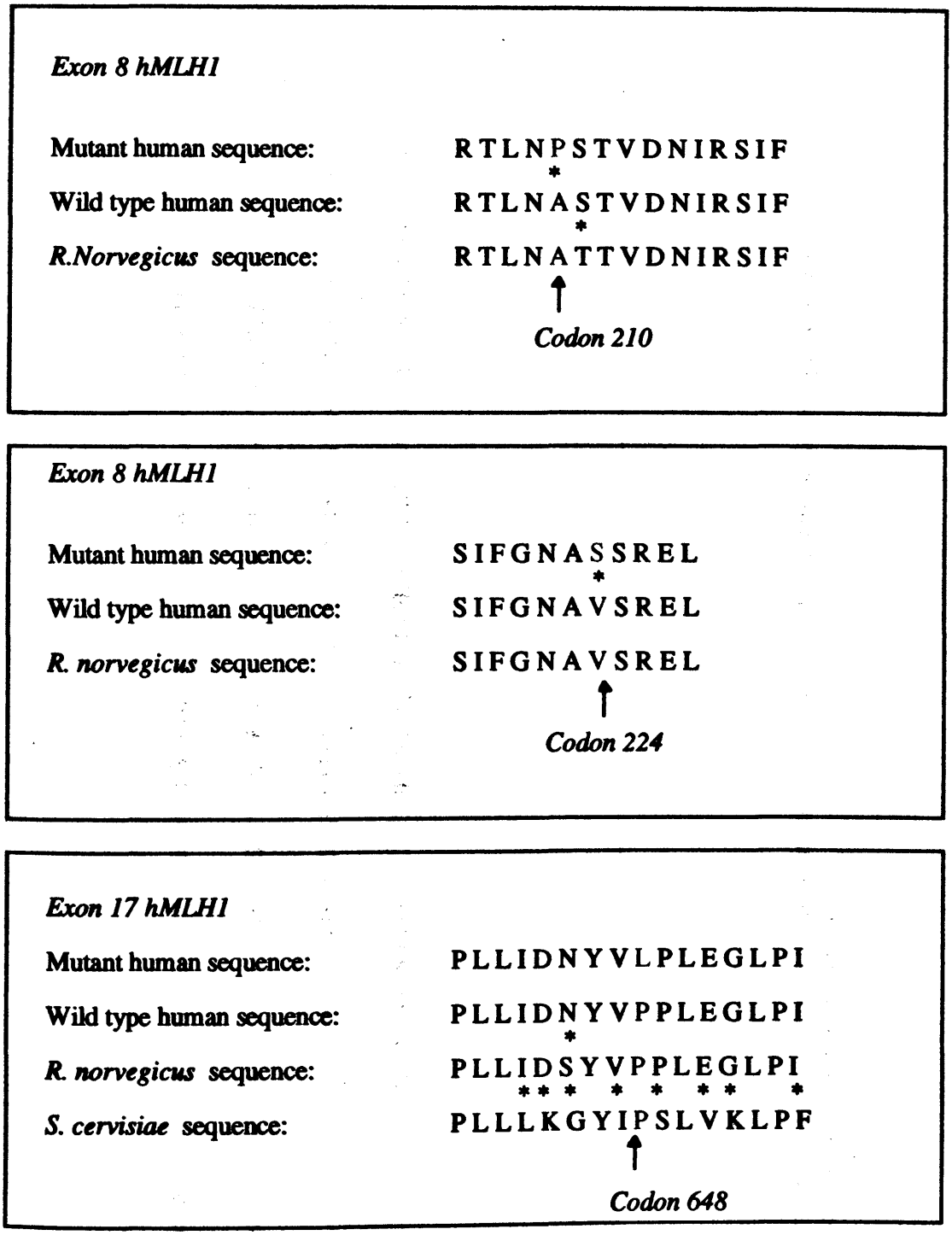


Figure 1.42. Blast amino acid sequence alignments between the wild type human MLH1 gene, mutant hMLH1 gene sequences identified in Irish HNPCC pedigrees and the *R. norvegicus* and *S. cervisiae* (exon 17 only) protein. * denotes amino acid differences between sequences.

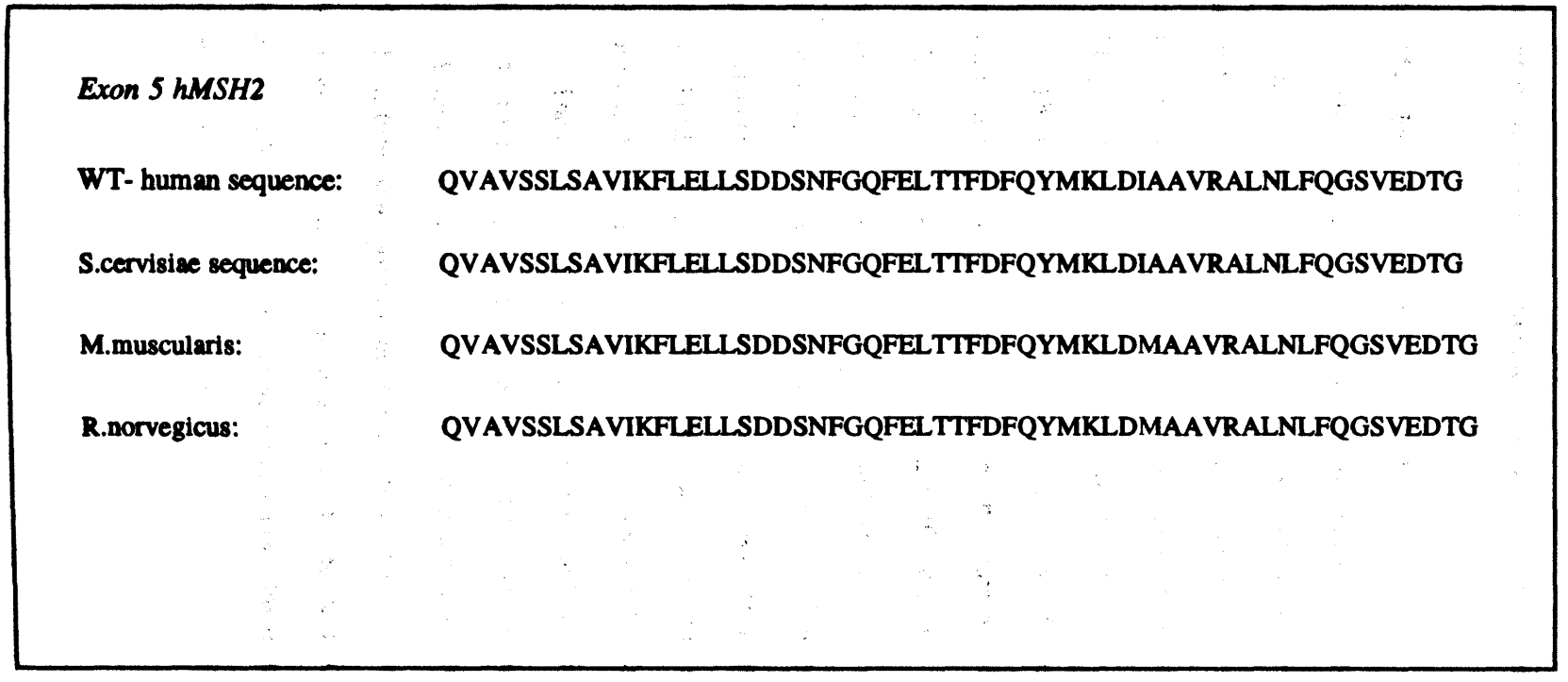


Figure 1.43. Blast amino acid sequence alignment exon 5 of the human MSH2 gene and the equivalent sequence in *S.cervisiae*, *M. muscularis* and *R. norvegicus*.

as a common HNPCC mutation in this population. The mutation has been identified in three other pedigrees in the USA (Lui *et al.*, 1994) and is responsible for the disease in 12% of affected UK kindreds (Peltomaki *et al.*, 1997). Haplotype analysis of all pedigrees harbouring this mutation may allow us to trace the origins of this mutation.

1.4.(II) FOUNDER MUTATIONS.

The 'founder effect' refers to mutations which have arisen once and subsequently have been passed through successive generations to present day carriers and patients. Some founder mutations are extremely widespread and account for a large proportion of some of today's major diseases such as the delta F508 mutation of the CTFR gene in cystic fibrosis (Morral *et al.*, 1994).

The majority of mutations identified in HNPCC are unique except for a hMLH1 mutation identified in 25 Finnish families, the exon 5 hMSH2 splice site mutation (14 families) and a hMLH1 codon 616 mutation (9 families). These may represent founder mutations. The exon 16 mutation of hMLH1 was detected initially in five Finnish families, geographical analysis of these families showed that they have a common geographical origin and share a common ancestor born in 1505 (Nystrom-Lahti *et al.*, 1994). Further analysis of other families which have this mutation showed that there is a common shared haplotype of alleles around hMLH1 in these families.

The exon 5 splice site mutation detected in Irish pedigree B1 has been identified in 14% of UK pedigrees (Peltomaki *et al.*, 1997) and in three North-American families. The American families share no obvious recent ancestry however, they may share Northern European origin (Lui *et al.*, 1994). This mutation may represent a founder mutation originating from an individual of UK or Irish origin, haplotype analysis around the hMSH2 region in families which have this mutation may allow the mutation to be traced back to a single individual.

1.4.(III) MISMATCH REPAIR MUTATIONS AND COLON CANCER DEVELOPMENT

HNPCC affected patients carry a germline mutation which inactivates one copy of one of their mismatch repair genes. Analysis of tissues in these individuals reveals that they are phenotypically normal (it has been demonstrated that cell lines, with one mutant allele of a mismatch repair gene exhibit normal levels of mismatch repair even under conditions of increased mutation frequency), (Branch *et al.*, 1993), they are however at increased risk of developing colorectal cancer. When a second somatic mutation in the MMR gene occurs, this inactivates both copies of the gene and mismatch repair activity becomes defective. A hypermutable state exists in these cells with the result that more and more mutations persist after each round of DNA replication. This hypermutable state is characterised by microsatellite instability, or the RER+ phenotype. This decreased ability to correct mutations is thought to result in an increased risk of accumulation of mutations in critical tumour suppressor genes, oncogenes and cell regulators with the subsequent development of a neoplastic growth and tumour development (Lazar *et al.*, 1994), (Figure 1.5)

1.4.(IV) IRISH MUTATION SPECTRUM VERSUS INTERNATIONAL SPECTRA

In an effort to establish an international database of HNPCC mutations, the international collaborative group on HNPCC (ICG-HNPCC) collected information on 126 predisposing mutations occurring in 202 kindreds of varied geographical and ethnic origin (Peltomaki *et al.*, 1997). The aim of this endeavour was to provide information to researchers, to facilitate evaluation of mutation spectra of HNPCC genes and to assess geographic and population variation in mutation spectra. Analysis of the mutation spectra of the 202 kindreds revealed that 48 mutations were identified in

hMSH2 (38%), 75 mutations affected the hMLH1 gene (59%), one mutation was identified in PMS1 and two were identified in PMS2. There are no obvious mutation 'hot-spots' however in hMSH2 the most frequently altered exons include exon 12 (17%) of hMSH2 mutations, 7 (15%), 5 (9%) and 6 (9%). In hMLH1 mutations occurred more frequently in exons 16 (15%), 9 and 13 (9%) and 19 (8%). All but two mutations were point mutations, usually single base substitutions (59%) however a large number of deletion/insertions were also identified (41%). The majority of mutations identified are unique except for a hMLH1 mutation identified in 25 Finnish families, the exon 5 hMSH2 splice site mutation (14 families) and a hMLH1 codon 616 mutation (9 families) which may represent founder mutations, (see previous Section).

Comparison of the mutations identified in the Irish HNPCC population reveal that mutation frequencies (although small sample size) are consistent with those of the ICG-HNPCC report, with 2/5 of Irish mutations confined to hMSH2 (versus 38%) and 3/5 confined to hMLH1 (versus 59%). All mutations identified in the Irish population are point mutations, two affecting splice sites, one single base deletion, one transition and one transversion. No mutation hotspots were identified and there is an absence of mutations in pedigrees which do not satisfy the all the Amsterdam Criteria (table 1.7).

The results of this study will be added to the ICG-HNPCC database so that the Irish population may be included in all future analyses of HNPCC mutation trends. The results of this study provide some information on the mutation spectrum in Irish HNPCC pedigrees, may facilitate in the analyses of mutation spectra across different ethnic backgrounds and may assist in the development of efficient screening strategies for mutation detection in affected families.

1.4.(V) THE AMSTERDAM SELECTION CRITERIA: TYPICAL VERSUS ATYPICAL IRISH HNPCC PEDIGREES.

Due to the absence of presymptomatic signs and a definitive clinical test to determine HNPCC susceptibility, the minimum criteria for the identification of a HNPCC pedigree was agreed upon at the second meeting of The International Collaborative Group on HNPCC in Amsterdam in 1990 (Vasen *et al.*, 1991). The minimum criteria known as 'The Amsterdam Criteria' are :

1)There should be at least three family members with a histologically verified colorectal cancer, one a first degree relative of the other two.

2)There must be at least two successive generations affected.

3)At least one individual must be less than 50yrs of age at the time of diagnosis.

4)Familial adenomatous Polyposis Coli must be ruled out.

The criteria used to select families for inclusion in this study are less stringent than the Amsterdam criteria however strong family history of the disease is evident in all pedigrees (see 'Pedigree details' Section 1.3.(III)). Analysis of hMSH2 and hMLH1 in all pedigrees was carried out blind to affection status of each individual and to whether pedigrees satisfied all the Amsterdam criteria. Pedigree details were only examined post genetic screening and mutation identification. Analysis of the 18 Irish HNPCC pedigrees included in this study identified two distinct groups- those pedigrees which strictly satisfied all the Amsterdam criteria and atypical HNPCC pedigrees which do not satisfy all the Amsterdam criteria but do have a strong family history of the disease. Eight pedigrees were classed as Typical Amsterdam pedigrees while 10 pedigrees failed to satisfy the criteria based on either (1) that there were less than 3 affected individuals within two successive generations or (ii) none of the affected were diagnosed before the age of 50. However, all the atpyical pedigrees did have at least 2

affected members in two generations, a strong family history of the disease and FAP was ruled out in all cases.

Correlation of mutation identification with Amsterdam status revealed that the five mutations identified were confined to 'Typical' pedigrees and no mutations were detected in the 'atypical' Amsterdam families. SSCP as a method of analysis can fail to detect up to 5% of mutations and can be affected by changes in a number of experimental conditions (Hayashi *et al.*, 1991; Sheffield *et al.*, 1993). Comparison of mutation detection rate in other studies with this study suggests a high detection failure rate when all pedigrees are taken into account, however if only true Amsterdam pedigrees are looked at the rate of mutation detection using SSCP (63%) is in keeping with reported detection rates of 70% (Lui *et al.*, 1996; Peltomaki *et al.*, 1997). The complete lack of mutations in hMSH2 and hMLH1 in atypical pedigrees cannot be explained by shortfalls in mutation detection rate alone. It is possible that the mutation spectrum in atypical HNPCC pedigrees differs from that of Typical Amsterdam pedigrees and that by using the Amsterdam criteria as an inclusion process there is a bias towards selection of hMSH2 and hMLH1 mutation bearing pedigrees. Mutations in other genes including hPMS1, hPMS2 have been detected in HNPCC pedigrees (Lui *et al.*, 1995), furthermore a number of recent studies suggest that when atypical HNPCC pedigrees are analysed for hMSH2 and hMLH1 mutations, the mutation identification rates in these pedigrees are much lower than in 'typical Amsterdam' families (Wijnen *et al.*, 1997). Perhaps mutations in other genes are responsible for predisposition to tumour development in these kindreds. The Amsterdam criteria are very stringent and were chosen to eliminate the likelihood of chance mutation clustering. It may be that mutations in hMSH2 and hMLH1 have a more "aggressive" phenotype within a pedigree and that mutations in other less "aggressive" genes may result in a predisposition to colon cancer with HNPCC symptoms but with a later age of onset and lower penetrance rate, as in atypical HNPCC pedigrees.

The majority of mutations identified in HNPCC pedigrees are in hMSH2 and hMLH1, however families included in these mutations studies are for the most part typically HNPCC in nature. Screening protocols are frequently set up based on analysis of published literature. Three important factors established from the literature have major implications for protocol plan (i) analysis of the literature suggests that of all mutations identified in HNPCC pedigrees more than 90% were detected in hMSH2 and hMLH1. (ii) The Amsterdam criteria is the preferred selection criteria and (iii) There is always a certain element of bias towards choosing genes which have the most potential for providing results. Taking these points into account most studies of HNPCC pedigrees focus on hMSH2 and hMLH1 alone, this reduces the likelihood of detecting mutations in other potentially predisposing genes and increases the mutation spectrum of the most commonly analysed genes.

The lack of mutations in atypical Irish HNPCC pedigrees suggests that perhaps some other genes such as hPMS1, hPMS2 (mismatch repair genes) or TGF-BRII (involved in epithelial cell growth regulation) (see chapter 3) may be involved in predisposing to colon cancer in these pedigrees.

1.4.(VI) THE SIGNIFICANCE OF MUTATION SCREENING -A FAMILY FOLLOW-UP:

The identification of a HNPCC gene mutation within an affected pedigree makes presymptomatic diagnosis of at-risk relatives possible. The early identification of predisposition to the development of colon cancer in an asymptomatic individual is particularly relevant at the clinical level in terms of diagnosis and surveillance. With the advancement of risk assessment and genetic counselling facilities, individuals within a family may be made aware of their risk and helped to understand the genetics of their disease this may allow them to make informed judgements on their own situations.

This has significant implications on decision making, choice of prophylactic therapy and ultimately on the quality of life for the individual.

The impact of mutation identification within a HNPCC pedigree was elucidated in the analysis of individuals in the Irish HNPCC pedigree B. A mutation at the +3 splice donor site of exon 5 of the hMSH2 gene was identified in three affected first degree relatives in this pedigree. The mutation results in the skipping of exon 5 and the production of a truncated protein product. Subsequent to the identification of the mutation in this pedigree, individual IV:3 presented for routine colonoscopy at the age of 31. A small lesion was detected, however histological examination did not provide evidence of dysplasia. Analysis of genomic DNA from this individual identified the predisposing mutation carried by her father and two other affected members of the family. An asymptomatic brother of this patient was also analysed but no evidence of the predisposing mutation was found. This information was passed on to the clinician involved in the case, for patient management.

The identification of a tumour at an early stage greatly increases the survival rate of the patient, thus identification of an increased risk of developing a tumour facilitates careful monitoring of the patient and increase the chance of detecting the tumour at an early stage. In terms of colon cancer, there is a greater chance that the patient will lead a normal life on prophylactic removal of a tumour at an early stage. Genetic screening of individuals in an at risk family has major implications at clinical level in terms of surveillance, diagnosis and prophylactic treatment and has the potential to increase the prognosis and quality of life for the patients themselves.

In conclusion, 18 pedigrees were analysed in this study, 5 mutations were identified in 5 Irish HNPCC pedigrees. These results are significant on a number of levels (i) This is the first analysis of the Irish HNPCC population where a large number of pedigrees have been analysed (ii) Four novel mutations were identified. The mutations will be added to the international HNPCC mutation database so that the Irish

population may be included in all future analyses of HNPCC mutation trends. (iii) The exon 5 splice site mutation identified in the Irish population provides further evidence to suggest that this is a HNPCC mutation hotspot. (iv) The lack of identification of mutations in atypical HNPCC pedigrees lends support to the suggestion that HNPCC pedigrees with a strong family history but which do not satisfy all the Amsterdam criteria should be analysed for mutations in a wider range of genes to determine the true mutation spectrum in a broader HNPCC selection category. Futhermore, the identification of a mutation within a HNPCC pedigree facilitates the screening of asymptomatic members in order to facilitate risk assessment and to assist at the clinical level in terms of surveillance, diagnosis and prophylactic therapy. Finally, these results suggest that perhaps there is a case for the setting up of a national HNPCC screening programme, the results of which can be added to a national HNPCC database, this may serve to increase our knowledge of the disease, the mutation spectrum and the true incidence in the Irish population and most importantly it may assist in reducing the death rate due to colon cancer in Irish HNPCC families.

1.5. BIBLIOGRAPHY

Aaltonen, L. A., Peltomaki, P., Leach, F. S., Sistonen, P., Pylkkanen, L., Mecklin, J-P., Jarvinen, H., Powell, S. M., Jen,J., hamilton, S.R., Petersen, G.M., Kinzler, K.W., Vogelstein, B., de la Chapelle, A. (1993). *Clues to the Pathogenesis of Familial Colorectal Cancer.* Science 260, 812-819.

Aaltonen, L.A. and Peltomaki, P. (1994). *Genes Involved in hereditary Nonpolyposis Colorectal Carcinoma.* Anticancer Res 14, 1657-1660.

Alani, E., Chi, N.W., Kolodner, R. (1995). *The Saccharomyces cerevisiae Msh2 protein specifically binds to duplex oligonucleotides containing mismatched DNA base pairs and insertions.* Genes Dev. 15;9 (2), 234-247.

Baeg. G.H., Matsumine, A., Kuroda, T., Bhattacharjee, R.N., Miyashiro, I., Toyoshima, K., Akiyama, T. *The tumour suppressor gene product APC blocks cell cycle progression from Go-G1 to S phase.* (1995). EMBO J. 14, 5618-5625.

Beck, N.E., Tomlinson, I. P., Homfray, T., Hodgson, S.V., Harocopos, C. J., Bodmer, W.F. (1997). *Genetic testing is important in families with a history suggestive of hereditary non-polyposis colorectal cancer even if the Amsterdam criteria are not fulfilled.* Br J Surg. 84 (2), 233-237.

Bodmer, W., Bishop, T. and Karran, P. (1994). *Genetic steps in colorectal cancer.* Nat Genet. 6, 217-219.

Bodmer, W.F., Bailey, C.J., Bodmer, J., Bussey, H.J., Ellis, A., Gorman, P., Lucibello, F.C., Murday, V.A., Rider, S.H., Scambler, P. (1987). *Localization of the gene for familial adenomatous polyposis on chromosome 5.* Nature 328, 614-616.

Branch, P., Aquilina, G., Bignami, M., Karran, P. (1993). *Defective mismatch binding and a mutator phenotype in cells tolerant to DNA damage.* Nature 362, 652-654.

Bronner, C.E., Baker, S.M., Morrison, P.T., Warren, G., Smith, L.G., Lescoe, M.K., Kane, M., Earabino, C., Lipford, J., Lindblom, A., (1994). *Mutation in the DNA mismatch repair gene homologue hMLH1 is associated with hereditary non-polyposis colon cancer.* Nature 368 , 258-261.

Bruni, R., Martin, D., Jiricny, J. (1988). *d(GATC) sequences influence Escherichia coli mismatch repair in a distance-dependent manner from positions both upstream and downstream of the mismatch.* Nucleic Acids Res.16 (11), 4875-4890.

Charbonnier, F., Martin, C., Scotte, M., Sibert, L., Moreau, V., Frebourg, T. (1995). *Alternative splicing of MLH1 messenger RNA in human normal cells.* Cancer Res. 55 (9), 1839-1841.

Cohen, P.R., Kohn, S.R., Kurzrock, R. (1991). *Association of sebaceous gland tumours and internal malignancy: The Muir-Torre Syndrome.* Am J Med. 90: 606-613.

Cooper, D.L., Lahue, R.S., Modrich, P. (1993). *Methyl-Directed Mismatch Repair is Bidirectional.* J Biol Chem. 268, 11823-11829.

Drummond, J.T., Li, G.M., Longley, M.J., Modrich, P. (1995). *Isolation of an hMSH2-p160 heterodimer that restores DNA mismatch repair to tumor cells.* Science 268 (5219), 1909-1912.

Dunlop, M.G., Farrington, S.M., Carothers, A.D., Wyllie, A.H., Sharp, L., Burn, J., Liu, B., Kinzler, K.W., Vogelstein, B. (1997). *Cancer risk associated with germline DNA mismatch repair gene mutations.* Hum Mol Genet. 6 (1), 105-110.

Fang, W. and Modrich, P. (1993). *Human Strand-specific Mismatch Repair Occurs by a Bidirectional Mechanism Similar to That of the Bacterial Reaction.* J Biol Chem. 268, 11838-11844.

Fearon, E.R., Vogelstein, B. (1990a). *A genetic model for colorectal tumorigenesis.* Cell, 61(5), 759-767.

Fearon, E.R., Cho, K.R., Nigro, J.M., Kern, S.E., Simons, J.W., Ruppert, J.M., Hamilton, S.R., Preisinger, A.C., Thomas, G., Kinzler, K.W. (1990b). *Identification of a chromosome 18q gene that is altered in colorectal cancers.* Science 247(4938), 49-56

Fishel, R., Ewel, A., Lee, S., Lescoe, M.K., Griffith, J. (1994). *Binding of mismatched microsatellite DNA sequences by the human MSH2 protein.* Science, 266 (5189), 1403-1405.

Fishel, R., Ewel, A., Lescoe, M.K. (1994). *Purified human MSH2 protein binds to DNA containing mismatched nucleotides.* Cancer Res. 54 (21), 5539-5542.

Fishel, R., Lescoe, M.K., Rao, M.R., Copeland, N.G., Jenkins, N.A., Garber, J., Kane, M., Kolodner, R.D. (1993). *The Human Mutator Gene Homolog MSH2 and its Association with Hereditary Nonpolyposis Colon Cancer.* Cell 75(5), 1027-1038.

Frei, J.V. (1992). *Hereditary Nonpolyposis Colorectal Cancer (Lynch Syndrome II).* Cancer 69 (5), 1108-1111.

Fujii, H., Shimada, T. (1989). *Isolation and characterization of cDNA clones derived from the divergently transcribed gene in the region upstream from the human dihydrofolate reductase gene.* Biol Chem 264(17), 10057-10064.

Grilley, M., Griffith, J., Modrich, P. (1993). *Bidirectional Excision In Methyl-Directed Mismatch Repair.* J Biol hem. 268, 11830-11837.

Han HJ, Maruyama M, Baba S, Park JG, Nakamura Y. (1995). *Genomic structure of human mismatch repair gene, hMLH1, and its mutation analysis in patients with hereditary non-polyposis colorectal cancer (HNPCC).* Hum Mol Genet. 4 (2), 237-242.

Han, H. J., Akio, Y., Yo, K., Jae-Gahb, P., Nakamura, Y. (1993). *Genomic Instability in Pancreatic Cancer and Poorly Differentiated Type of Gastric Cancer.* Cancer Res. 53,5087-5089.

Hayashi, K. (1991). *PCR-SSCP: a simple and sensitive method for detection of mutations in the genomic DNA.* PCR Methods Appl. 1 (1), 34-38.

Heinen, C.D., Shivapurkar, N., Tang, Z., Groden, J., Alabaster, O. (1996). *Microsatellite instability in aberrant crypt foci from human colons.* Cancer Res. 56 (23), 5339-5341.

Iizuka, M., Mashiyama, S., Oshimura, M., Sekiya T, Hayashi, K. (1992). *Cloning And Polymerase Chain Reaction-Single-Strand Conformation Polymorphism Analysis Of Anonymous Alu Repeats On Chromosome 11.* Genomics 12(1), 139-146.

Ionov, Y., Peinado, M.A., Malkhosyan, S., Shibata, D., Perucho, M. (1993). *Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis.* Nature 363,(6429), 558-561.

Karran, P. (1995). *Appropriate Partners Make Good Matches.* Science 268, 1857-1858.

Katabuchi, H., van Rees, B., Lambers, A.R., Ronnett, B.M., Blazes, M.S., Leach, F.S., Cho, K.R., Hedrick, L. (1995). *Mutations in DNA mismatch repair genes are not responsible for microsatellite instability in most sporadic endometrial carcinomas.* Cancer Res. 55(23), 5556-5560.

Kee, F., Collins, B.J. (1991(a)). *How Prevalent is cancer family syndrome?* Gut 32, 509-512.

Kee, F., Collins, B. J., Patterson, G.C. (1991(b)). *Prognosis in familial non-polyposis colorectal cancer.* Gut 32, 513-516.

Koi, M., Umar, A., Chauhan, D.P., Cherian, S.P., Carethers, J.M., Kunkel, T.A., Boland, C.R. (1994). *Human chromosome 3 corrects mismatch repair deficiency and microsatellite instability and reduces N-methyl-N'-nitro-N-nitrosoguanidine tolerance in colon tumor cells with homozygous hMLH1 mutation.* Cancer Res. 54 (16), 4308-4312.

Kolodner, R.D., Hall, N.R., Lipford, J., Kane, M.F., Rao, M.R., Morrison, P., Wirth, L., Finan, P.J., Burn, J., Chapman, P. (1994(b)). *Structure of the human MSH2 locus and analysis of two Muir-Torre kindreds for msh2 mutations.* Genomics. 24(3), 516-526.

Kolodner, R.D., Hall, N.R., Lipford, J., Kane, M.F., Morrison, P.T., Finan, P.J., Burn, J., Chapman, P., Earabino, C., Merchant, E., Bishop, T. (1995). *Structure of the human MLH1 locus and analysis of a large hereditary nonpolyposis colorectal carcinoma kindred for mlh1 mutations.* Cancer Res 55 (2), 242-248.

Kolodner, R.D., Hall, N.R., Lipford, J., Kane, M.F., Morrison, P.T., Finan, P.J., Burn, J., Chapman, P., Earabino, C., Merchant, E., Bishop, T., Garber, J., Bronner, C.E., Baker, S.M., Warren, G., Smith, L.G., Lindblom, A., Tannergard, P., Bollag, R.J., Godwin, A.R., Ward, D.C., Liskay, R.M., Copeland, N., Jenkins, N., Lescoe, M.K., Ewel, A., Lee,S., Griffith, J., Fishel, R. (1994(a)). *Human Mismatch Repair Genes and their Association with HNPCC.* Cold Spring Harb Symp Quant Biol. 59:331-338.

Kramer, W., Kramer, B., Williamson, M.S., Fogel, S. (1989). *Cloning and nucleotide sequence of DNA mismatch repair gene PMS1 from Saccharomyces cerevisiae: homology of PMS1 to procaryotic MutL and HexB.* J Bacteriol. 171(10), 5339-5346.

Kunkel, T.A. (1993). *Nucleotide repeats. Slippery DNA and diseases.* Nature 365(6443), 207-208.

Kunkle, M., Longo, F.J., Myrun, B.E. (1978). *Nuclear protein changes in the maternally and paternally derived chromatin at fertilization.* J Exp Zool.203(3), 371-380.

Lahue, R.S., Au, K.G., Modrich, P. (1989). *DNA Mismatch Correction in a Defined System.* Science 245(4914), 160-164.

Lanspa, S.J., Lynch, H.T., Smyrk, T.C., Strayhorn, P., Watson, P., Lynch, J.F., Jenkins, J.X., Appelman, H.D. (1990). *Colorectal Adenomas in the Lynch Syndromes. Results of a Colonoscopy Screening Program.* Gastroenterology 98(5 Pt 1), 1117-1122.

Lazar, V., Grandjouan, S., Bognel, C., Couturier, D., Rougier, P., Bellet, D., Bressac-de Paillerets, B. (1994). *Accumulation of multiple mutations in tumour suppressor genes during colorectal tumorigenesis in HNPCC patients.* Hum Mol Genet. 3(12), 2257-2260.

Leach, F.S., Nicolaides, N.C., Papadopoulos, N., Liu, B., Jen, J., Parsons, R., Peltomaki, P., Sistonen, P., Aaltonen, L.A., Nystrom-Lahti, M., et al. (1993). *Mutations of a mutS homolog in hereditary nonpolyposis colorectal cancer.* Cell 75(6), 1215-1225.

Levinson, G., Gutman, G.A. (1987). *High frequencies of short frameshifts in poly-CA/TG tandem repeats borne by bacteriophage M13 in Escherichia coli K-12.* Nucleic Acids Res. 15(13), 5323-5338.

Li, G.M. and Modrich, P. (1995). *Restoration of mismatch repair to nuclear extracts of H6 colorectal tumor cells by a heterodimer of human MutL homologs.* Proc Natl Acad Sci U S A. 92(6), 1950-1954.

Lindblom, A., Tannergard, P., Werelius, B., Nordenskjold, M. (1993). *Genetic mapping of a second locus predisposing to hereditary non-polyposis colon cancer.* Nat Genet. 3, 279-282.

Loeb, L.A. (1991). *Mutator phenotype may be required for multistage carcinogenesis.* Cancer Res. 51(12), 3075-3079.

Loeb, L.A. (1994). *Microsatellite instability: marker of a mutator phenotype in cancer.* Cancer Res. 54 (19), 5059-5063.

Lui, B., Parsons, R., Hamilton, S., Petersen, G., Lynch, H.T., Watson, P., Markowitz, S., Willson, J., Green, J., de la Chapelle, A., Kinzler, K., Vogelstein, B. (1994). *HMSH2 Mutations in hereditary Nonpolyposis Colorectal Cancer Kindreds.* Cancer Research 54, 4590-4594.

Lui, B., Parsons, R., Papadopoulos, N., Nicolaides, N.C., Lynch, H.T., Watson, P., Jass, J.R., Dunlop, M., Wyllie, A., Peltomaki, P., de la Chapelle, A., Hamilton, S.R., Vogelstein, B., Kinzler, K.W. (1996). *Analysis of Mismatch repair genes in hereditary Nonpolyposis Colorectal cancer Patients.* Nat Med. 2, 169-174.

Lui, B., Nicolaides, N.C., Markowitz, S., Wilson, J.K., Parsons, R.E., Jen, J., Papadopoulos, N., Peltomaki, P., de la Chapelle, A., Hamilton, S.R., Kinzler, K.W., Vogelstein, B. (1995). *Mismatch Repair Gene Defects in Sporadic Colorectal cancers with Microsatellite instability.* Nat Genet. 1, 48-55.

Lynch, H.T., Fusaro, R.M., Roberts, L., Voorhees, G.J., Lynch, J.F. (1985). *Muir-Torre syndrome in several members of a family with a variant of the Cancer Family Syndrome.* Br J Dermatol. 113 (3), 295-301.

Lynch, H.T., Smyrk, T.C., Watson, P., Lanspa, S.J., Lynch, J.F., Lynch, P.M., Cavalieri, R.J., Boland, C.R. (1993). *Genetics, natural history, tumor spectrum, and pathology of hereditary nonpolyposis colorectal cancer: an updated review.* Gastroenterology 104(5), 1535-1549.

Lynch, H.,T., Shaw, M.W., Magnuson, C.W., Larsen, A.L., Krush. A.J. (1966). *Hereditary factors in cancer. Study of two large midwestern kindreds.* Arch Intern Med. 117(2), 206-212.

Mecklin, J.P. (1987). Frequency of Hereditary Colorectal Carcinoma. Gastroenterology 93(5),1021-1025.

Mecklin, J.P., Jarvinen, H.J. (1991). *Tumor spectrum in cancer family syndrome (hereditary nonpolyposis colorectal cancer).*Cancer 68(5), 1109-1112.

Michaud, J. Brody, L.C., Steel, G., Fontaine, G., Martin, L.S., Valle, D., Mitchell, G. (1992). *Strand Separating Conformational Polymorphism Analysis: Efficacy Of*

Detection Of Point Mutations In The Human Ornithine Delta Amino Transferase Gene. Genomics 2, 389-394.

Modrich, P. (1989). *Methyl-directed DNA mismatch correction.* Biol Chem. 264(12), 6597-6600.

Modrich, P. (1991). *Mechanisms and biological effects of mismatch repair.* Annu Rev Genet. 25, 229-253

Modrich, P. (1994). *Mismatch repair, genetic stability, and cancer.* Science. 266, 1959-60.

Morral, N., Bertranpetit, J., Estivill, X., Nunes, V., Casals, T. (1994). *The Origin of The Major Cystic Fibrosis Mutation (DF508) in European Populations.* Nature Genet. 7, 169-175.

Muir, E.G., Bell, A.J., Barlow, K.A. (1967). *Multiple primary carcinomata of the colon, duodenum, and larynx associated with kerato-acanthomata of the face.* Br J Surg. 54(3), 191-195.

Myers, R.M., Larin, Z., Maniatis, T. (1985). *Detection of Single Base Substitutions by Ribonuclease Cleavage at Mismatches in RNA:DNA Duplexes.* Science 230, 1242-1246.

Meyers, M., Theodosiou, M., Acharya, S., Odegaard, E., Wislon, T., Lewis, J.E., Davis, T.W., Wilson-van patten, C., Fishel, R., Boothman, D.A. (1997). *Cell Cycle Regulation of the Human DNA Mismatch Repair Genes hMSH2, hMLH1 and hPMS2.* Cancer Res. 57, 206-208.

New, L., Liu, K., Crouse, G.F. (1993). *The Yeast MSH3 Gene Defines A New Cllass of Eukaryotic MutS Homologues.* Mol. Gen. Genet. 239, 97-108.

Nicolaides, N.C., Papadopoulos, N., Liu, B., Wei, Y.F., Carter KC, Ruben, S.M., Rosen, C.A., Haseltine, W.A., Fleischmann, R.D., Fraser, C.M., et al., (1994). *Mutations of Two PMS Homologues in Hereditary Nonpolyposis Colon Cancer.* Nature 371(6492), 75-80.

Nicolaides, N.C., Carter, K.C., Shell, B.K., Papadopoulos, N., Vogelstein, B., Kinzler, K.W. (1995). *Analysis of the 5' region of PMS2 reveals heterogeneous transcripts and a novel overlapping gene.* Genomics. 29(2), 329-334.

Nystrom-Lahti, M., Sistonene, P., Mecklin, J-P., Pylkkanen, L., Aaltonen, L., Jarvinen, H., Weissenbach, J., de la Chapelle, A., Peltomaki P. (1994). *Close Linkage to Chromosome 3p and Conservation of Ancestral Founding Haplotype in Hereditary Nonpolyposis Colorectal Cancer Families.* Proc. Natl. Acad. Sci. USA. 91, 6054-6058.

O'Sullivan, M.J., Mulcahy, T.M., Cambell, J., O'Suilleabhain, C.B., Kirwan, W.O., Doyle, C.T., McCarthy, T.V. (1989). *Detection of five novel germline mutations of the*

APC gene in Irish familial adenomatous polyposis families. Hum Mutat. S 1(6), S251-S253.

Orita, M., Iwahana, H., Kanazawa, H., Hayashi, K., Sekiya, T. (1989(a)). *Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms.* Proc Natl Acad Sci USA 6(8), 2766-2770.

Orita, M., Suzuki, Y., Sekiya, T., Hayashi, K. (1989(b)). *Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction.* Genomics 5 (4), 874-879.

Palombo, F., Gallinari, P., Iaccarino, I., Lettieri, T., Hughes, M., D'Arrigo, A., Truong, O., Hsuan, J..J, Jiricny, J. (1995). *GTBP, a 160-kilodalton protein essential for mismatch-binding activity in human cells.* Science 268(5219), 1912-1914.

Palombo, F., Iaccarino, I., Nakajima,E., Ikemjima, M., Shimada, T, Jiricny, J. (1996). *HMUTSbeta, a heterodimer of hMSH2 and hMSH3, binds to insertion/deletion loops in DNA.* Curr Biol. 6(9), 1181-1184.

Papadopoulos, N., Nicolaides, N.C., Wei, Y.F., Ruben, S.M., Carter, K.C., Rosen, C.A., Haseltine, W.A., Fleischmann, R.D., Fraser, C.M., Adams, M.D, Venter, C., Hamilton, S., Petersen, G., Watson, P., Lynch, H.T., peltomaki, P., Mecklin, J-P., de la Chapelle, A., Kinzler, K., Vogelstein, B. (1994). *Mutation of a mutL homolog in hereditary colon cancer.* Science 263, 1625-1629.

Papadopoulos, N., Nicolaides, N., Liu, B., Parsons, R., Lengauer, C., Palombo., F., D'Arrigo, A., Markowitz, S., Willson, J., Kinzler, K., Jiricny, J., Vogelsein, B. (1995). *Mutations of GTBP I genetically Unstable cells.* Science 268, 1915-1917.

Parker, B.O., Marinus, M.G. (1992). *Repair of DNA heteroduplexes containing small heterologous sequences in Escherichia coli.* Proc Natl Acad Sci U S A.89(5), 1730-1734.

Parsons, R., Myeroff, L.L., Liu, B., Willson, J.K., Markowitz, S.D., Kinzler, K.W., Vogelstein ,B. (1995). *Microsatellite instability and mutations of the transforming growth factor beta type II receptor gene in colorectal cancer.* Cancer Res 55(23), 5548-5550.

Parsons, R., Li, G.M., Longley, M.J., Fang, W.H., Papadopoulos, N., Jen ,J., de la Chapelle, A., Kinzler, K.W., Vogelstein. B.,,Modrich, P. (1993). *Hypermutability and mismatch repair deficiency in RER+ tumor cells.* Cell 75(6),1227-1236.

Peifer. M. (1993). *Cancer , catenins and cuticle pattern: a complex connection.* Science. 262, 1731- 1743.

Peltomaki, P., Lothe, R.A., Aaltonen, L.A., Pylkkanen, L., Nystrom-Lahti, M., Seruca, R., David, L., Holm, R., Ryberg, D., Haugen, A. (1993b). *Microsatellite instability is*

associated with tumors that characterize the hereditary non-polyposis colorectal carcinoma syndrome. Cancer Res 53(24), 5853-5855.

Peltomaki, P., Aaltonen, L.A., Sistonen, P., Pylkkanen, L., Mecklin, J.P., Jarvinen, H., Green, J.S., Jass, J.R., Weber, J.L., Leach, F.S. (1993a). *Genetic mapping of a locus predisposing to human colorectal cancer.* Science 260(5109), 810-812.

Peltomaki, P., Vasen, H.F.A., and the International Collaborative Group on Hereditary Nonpolyposis Colorectal Cancer: *Database and Results of a Collaborative Study.* (1997). Gastroenterology, 113, 1146-1158.

Ponz de Leon, M., Sassatelli, R., Sacchetti, C., Zanghieri, G., Scalmati, A., Roncucci, L. (1989). *Familial aggregation of tumors in the three-year experience of a population-based colorectal cancer registry.* Cancer Res 49(15), 4344-4348.

Pounder , R.E, Allison, M.C., Dhillon, A.P. (1989). *A Color Atlas of the Digestive System.* Wolfe Medical Publications Ltd. Pp 136-151.

Powell, S.M., Petersen, G.M., Krush, A.J., Booker, S., Jen, J., Girdiello, F.M., Hamilton, S.R., Vogelstein, B., Kinzler,K.W. (1993). *Molecular Diagnosis of Familial Adenomatous Polyposis.* N Engl J Med. 329, 1982-1987.

Prolla, T.A., Pang, Q., Alani, E., Kolodner, R.D., Liskay, R.M. (1994). *MLH1, PMS1, and MSH2 interactions during the initiation of DNA mismatch repair in yeast.* Science 265(5175), 1091-1093.

Reenan, R.A., Kolodner, R.D. (1992). *Characterization of insertion mutations in the Saccharomyces cerevisiae MSH1and MSH2 genes: evidence for separate mitochondrial and nuclear functions.* Genetics 32(4), 975-985.

Risinger, J.I., Berchuck, A., Kohler, M.F., Watson, P., Lynch, H.T., Boyd, J. (1993). *Genetic instability of microsatellites in endometrial carcinoma.* Cancer Res. 53(21), 5100-5103.

Roest, P.A., Roberts, R.G., Sugino, S., van Ommen, G.J., den Dunnen. J,T. (1993). *Protein truncation test (PTT) for rapid detection of translation-terminating mutations.* Hum Mol Genet. 2(10), 1719-1721.

Ross-Macdonald, P., Roeder, G.S. (1994). *Mutation of a meiosis-specific MutS homolog decreases crossing over but not mismatch correction.* Cell. 79(6), 1069-1080.

Ruano. G., Brash, D.E., Kidd, K.K. (1991). *Amplifications-A forum for PCR users.* 7, 1-4.

Sambrook, J., Fritsch,E.F., Maniatis, T.(1989). *Molecular Cloning a Laboratory Manual. 2nd. Edition. Cold Spring Harbour Laboratory Press.*

Sheffield, V.C., Beck, J.S., Kwitek, A.E., Sandstrom, D.W., Stone, E.M. (1993). *The Sensitivity Of Single-Strand Conformation Polymorphism Analysis For The Detection Of Single Base Substitutions.* Genomics 16(2), 325-332.

Sheffield, V.C., Cox, D.R., Lerman, L.S., Myers, R.M. (1989). *Attachment of a 40-base-pair G + C-rich sequence (GC-clamp) to genomic DNA fragments by the polymerase chain reaction results in improved detection of single-base changes.* Proc Natl Acad Sci U S A.86(1), 232-236.

Silverman, A.L., Park, J.G., Hamilton, S.R., Gazdar, A.F., Luk, G.D., Baylin, S.B. (1989). *Abnormal Methylation Of The Calcitonin Gene In Human Colonic Neoplasms.* Cancer Res 49(13), 3468-3473.

Smith, K.J., Johnson, K.A., Bryan, T.M., Hill, D.E., Markowitz, S., Willson, J.K., Paraskeva, C., Petersen, G.M., Hamilton, S.R.,Vogelstein, B. (1993). *The APC gene product in normal and tumor cells.* Proc Natl Acad Sci U S A 90(7), 2846-2850.

Spinardi, L., Mazars, R., Theillet, C. (1991). *Protocols for an improved detection of point mutations by SSCP.* Nucleic Acids Res. 19(14), 4009.

Strand, M., Prolla, T.A., Liskay, R.M., Petes, T.D. (1993). *Destabilization of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair.* Nature 365(6443), 274-276.

Su, S.S., Modrich, P. (1986). *Escherichia coli mutS-encoded protein binds to mismatched DNA base pairs.* Proc Natl Acad Sci U S A 83(14), 5057-5061.

Su, S.S., Lahue, R.S., Au, K.G., Modrich, P. (1988). *Mispair specificity of methyl-directed DNA mismatch correction in vitro.* J Biol Chem 263(14), 6829-6835.

Thibodeau, S.N., Bren, G., Schaid , D. (1993). Microsatellite instability in cancer of the proximal colon. Science 260 (5109), 816-819.

Traystman, M.D., Higuchi, M., Kasper, C.K., Antonarakis, S.E., Kazazian, H.H. Jr. (1990). *Use Of Denaturing Gradient Gel Electrophoresis To Detect Point Mutations In The Factor VIII Gene.* Genomics (2), 293-301.

Vasen, H.F., Wijnen, J.T., Menko, F.H., Kleibeuker, J.H., Taal, B.G., Griffioen, G., Nagengast, F.M., Meijers-Heijboer, E.H, Bertario, L., Varesco, L., Bisgaard, M.L., Mohr, J., Fodde, R., Khan, P.M. (1996). *Cancer Risk In Families With Hereditary Nonpolyposis Colorectal Cancer Diagnosed By Mutation Analysis.* Gastroenterology 110(4), 1020-1027.

Vasen, H.F., Mecklin, J.P., Khan, P.M., Lynch, H.T. (1991). *The International Collaborative Group on Hereditary Non-Polyposis Colorectal Cancer (ICG-HNPCC).* Dis Colon Rectum. 34(5), 424-425.

Vogelstein, B., Fearon, E.R., Hamilton, S.R., Kern, S.E,, Preisinger, A.C., Leppert, M., Nakamura, Y., White, R., Smits, A.M., Bos, J.L. (1988). *Genetic Alterations During Colorectal-Tumor Development.* N Engl J Med. 319(9), 525-532.

Warthin, A.S. (1913). *Heredity with Reference to Carcinoma.* Arch. Int. med. 12, 546-555.

Watson, P., Lynch, H.T. (1993). *Extracolonic Cancer In Hereditary Nonpolyposis Colorectal Cancer* . Cancer 71(3), 677-685.

Watson, P., Lynch, H.T. (1994). *The tumor spectrum in HNPCC.* Anticancer Res. 14(4B), 1635-1639.

Weber, J.L., May, P.E. (1989). *Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction.* Am J Hum Genet. 44(3), 388-396.

Welsh, K.M., Lu, A.L,, Clark, S., Modrich, P. (1987). *Isolation and characterization of the Escherichia coli mutH gene product.* J Biol Chem. 262(32), 15624-15629.

Wijnen, J., Khan, P.M., Vasen, H., van der Klift, H., Mulder, A., van Leeuwen-Cornelisse, I., Bakker, B., Losekoot, M., Moller, P, Fodde, R. (1997). *Hereditary nonpolyposis colorectal cancer families not complying with the Amsterdam criteria show extremely low frequency of mismatch-repair-gene mutations.* Am J Hum Genet. 61(2), 329-335.

Yonish-Rouach, E., Resnitzky, D., Lotem, J., Sachs, L., Kimchi, A., Oren, M. (1991). *Wild-type p53 induces apoptosis of myeloid leukaemic cells that is inhibited by interleukin-6.* Nature. 352, 345-347.

CHAPTER TWO

APPLICATION OF GMPD FOR THE ANALYSIS OF K-RAS IN SPORADIC COLON CANCER

2.1. INTRODUCTION

2.1.(I) SPORADIC COLORECTAL CANCER

Colorectal cancer is the most common cause of death due to malignancy in non-smokers in the western world. A recent report by the National Cancer Registry states that in 1995 there were 1,757 cases of colorectal carcinoma in Ireland (Nat. Cancer. Reg., 1995 D. Delaney, personal communication). Ninety percent of all colon cancer cases occur in individuals where there is no obvious family predisposition to the disease (Lynch *et al.*, 1993).

CLINICAL ASPECTS OF COLORECTAL CANCER

Patients usually present with a change in bowel habit and in many cases, rectal bleeding. Double contrast barium enema is the standard method used to clinically diagnose colorectal carcinoma. A typical carcinoma appears as an obstruction of the colon and has a characteristic apple core-like appearance. In many cases this visual characteristic is not obvious and confirmation requires colonoscopy, biopsy and histological verification. Treatment is usually by colonic resection. Histological assessment of the carcinoma and mesenteric lymph nodes is used to stage the tumour and define prognosis. Dukes staging of colonic tumours is the most commonly used scheme. Dukes A tumours are confined to the bowel wall. Tumours that have infiltrated the bowel wall in the absence of lymph node metastases are classified as Dukes stage B. Once the regional lymph nodes become involved the carcinoma is classed as Dukes C. Tumours with associated liver metasteses are classed as Dukes D. Dukes A carcinomas carry a mean five year survival rate of 85%, Dukes B is associated with a decreased value of 65% after colonic resection while the prognosis of Dukes stage C with liver metastases is less than 1% (Pounder *et al.*, 1989).

Sporadic colon cancer differs from HNPCC and FAP in that the age of onset is greater (over 60 years) compared to hereditary colon cancer (40-45 years). Localisation of tumours in the colon is varied with only one third being confined to the proximal colon (compared to two thirds in HNPCC). Deviation from diploidy is higher in sporadic cancer tumours, indicating a loss or addition of chromosomes. Prognosis in sporadic cancers is poorer than in hereditary cancers. FAP-like polyposis is not observed in sporadic cancers and there is no association with cancers of multiple organs as in HNPCC (de la Chapelle, 1995).

GENETIC ASPECTS OF COLORECTAL CANCER.

The molecular basis by which cancers arise has been investigated abundantly. Due to the availability and relative ease of obtaining tumours at early intermediate and advanced stages of cancer development, colon cancer has become one of the most widely investigated human cancers.

Tumorigenesis is believed to be a multistep process in which malignant colorectal tumours arise from pre-existing benign adenomas. Histopathological and clinical assessment of cancer development suggest that most tumours arise from early adenomas, these progress in size, become dysplastic and eventually form an invasive tumour. This concept is widely known as the theory of clonal expansion (Nowell *et al.*, 1976; Loeb *et al.*, 1991). Various models have been proposed to explain the sequential development of colon cancer. The most widely accepted genetic model for colorectal tumorigenesis proposes that colorectal tumours arise as a result of mutational inactivation of tumour suppressor genes coupled to the mutational activation of cellular oncogenes. The model proposes that at least four genes are required for the formation of a tumour and that it is the overall accumulation of mutations in specific genes rather than the mutation order which is more important (Fearon *et al.*, 1990).

2.1.(II) GENES INVOLVED IN COLORECTAL CANCER.

TUMOUR SUPPRESSOR GENES

Tumour suppressor genes encode proteins which regulate cellular growth and differentiation. They exert their role by down-regulating cellular proliferation when required and thus have a vital role in preventing neoplastic progression. Loss of heterozygosity at a specific locus in tumour DNA is interpreted as evidence that this locus codes for a tumour suppressor protein and is based on Knudsons "two hit hypothesis" which proposes that the inactivation of both alleles of a gene is required to affect the tumour suppressor role of a protein. (Knudson, 1985)

A number of tumour suppressor genes have been linked to colorectal cancer. The APC gene on chromosome 5q plays a role in the regulation of transmission of the contact inhibition signal in cells and is associated with α and β catenin, microtubules and other cell adhesion molecules (Pfeifer *et al.*, 1993; Baeg *et al.*, 1995). Germline mutations in the APC gene are associated with a predisposition to familial adenomatous polyposis colon cancer (FAP) (Bodmer *et al*, 1987; O'Sullivan *et al.*, 1997). FAP, a subset of the hereditary colon cancers is characterised by the development of hundreds of colonic polyps at an early age. Progression to carcinoma is inevitable unless polyps are surgically removed. In FAP, germline mutations in the APC gene predispose individuals to the development of polyposis, however loss of heterozygosity (a somatic mutation) at the APC locus in addition to the accumulation of mutations in other cell regulators are required for tumour development. (Bodmer *et al.*, 1987).

The 'Guardian of the genome' p53, is putatively involved in almost all cancers and mutations in this gene have been detected in more than 75% of colorectal carcinomas (Vogelstein *et al.*, 1988). It has been demonstrated that allelic loss of the p53 locus on 17p is associated with the progression of tumours from adenoma to carcinoma (Fearon *et al.*, 1987; Vogelstein *et al.*, 1988). p53 exerts its tumour suppressor effect by switching off replication during repair of DNA damage, cells arrest

at the G1 checkpoint of the cell cycle, if the DNA remains un-repaired p53 may trigger cell suicide by apoptosis, thereby preventing further growth and replication of the mutant cell (Yonish- Rouach *et al*, 1991).

The DCC (Deleted in Colon Cancer) tumour suppressor gene was identified by observations of allelic loss on chromosome 18q and is associated with more than 70% of colorectal carcinomas (Vogelstein *et al.*, 1988). The gene encodes a receptor for netrin-1 a molecule involved in axon guidance (Keino-Masu *et al.*, 1996). Recent studies indicate that DCC may function as a tumour suppressor by inducing apoptosis in settings where ligand is unavailable - for example during metastasis or tumour growth outside the region of local blood supply (Mehlen *et al.*, 1998).

ONCOGENES

Proto-oncogenes are genes whose action promotes cellular proliferation in a regulated manner in a normal cell. The mutant version - the activated oncogene is overactive and leads to the disrupted regulation of proliferation and ultimately to tumour development. Analysis of cancer related oncogenes demonstrates that these genes are involved in cellular processes which if not accurately regulated would lead to cellular proliferation. Five oncogenic classes have been defined including secreted growth factors, cell surface receptors (e.g. ERB), components of the intracellular signal transduction pathways (e.g. Ras and ABL), DNA binding nuclear proteins (e.g. Myc and Jun) and components of the cyclins network including cyclin dependant kinases and inhibitors (Kamb 1995; Muller, 1995).

Of particular importance in colorectal cancer are the ras oncogenes. Mutations in ras genes have been detected in more than 50% of colorectal carcinomas (Bos *et al.*, 1987; Forrester *et al.*, 1987). Due to the presence of mutations in early adenomas it is believed that ras gene mutation may be one of the initiating events in the sporadic colorectal cancer pathway.

A number of models linking the mutation of cancer genes with the stages of colorectal tumorigenesis have been proposed. The most widely accepted hypothesis

suggests that a mutation in the APC gene leads to the generation of hyper-proliferative epithelium, K-ras activation and possibly hypomethylation result in the generation of an adenoma, this coupled to the accumulation of mutations in genes such as DCC and p53 result in the subsequent progression from adenoma to carcinoma and ultimately to metastases (Fearon *et al.*, 1990; Spandidos, 1995).

2.1.(III) THE RAS GENE FAMILY

The ras family constitutes a group of closely related genes which encode 21KD (p21) proteins, which are localised to the inner plasma membrane of the cell. The ras proteins play a functional role in signal transduction from G-protein coupled receptors. A signal from the receptor triggers the binding of GTP to the ras protein and GTP ras transmits the signal onwards to the cell. Ras proteins have GTPase activity and GTP-ras is rapidly converted to the inactive GDP-ras form. Point mutations in ras genes cause amino acid substitutions which decrease the GTPase activity of the ras protein. As a result the GTPase ras signal is inactivated more slowly leading to excessive cellular response to the signal from the receptor and thus, cellular proliferation (Barbacid, 1987).

RAS GENES AND TUMORIGENESIS

The involvement of K-ras in tumorigenesis has been widely investigated. It has been demonstrated using cultured cell lines that the substitution of a single amino acid at particular positions in any of the ras genes results in the activation of oncogenic activity in these proteins (Forrester *et al.*,1987). Analysis of K-ras genes at various stages of colorectal tumorigenesis suggests that K-ras mutations are most likely an early event in the progression pathway. Mutations in K-ras have been identified in 50% of colorectal carcinomas (Bos *et al.*, 1987; Forrester *et al.*, 1987) and a similar percentage of adenomas (Vogelstein *et al.*, 1988). Furthermore, K-ras mutations have been detected in aberrant crypt foci (ACF) of cancer patients. ACF are microscopic lesions which are detectable (by histological staining) in "normal" looking mucosa of cancer patients.

ACF are believed to be the initial lesions which progress to carcinoma. The presence of K-ras mutations in ACF of cancer patients corroborates the belief that K-ras gene mutations are an early event in colorectal tumorigenesis (Shivapurkar *et al.*, 1997).

The identification of K-ras mutations at various stages of tumour development stimulated the analysis of mutation type and position correlated with stage and tumour grade. Mutations in the K-ras gene have been confined to codons 12 and 13 of exon 1 and codon 61. The most common mutations are G-T and G-A mutations at the first and second G positions in both codons (Cooper *et al.*, 1982; Land *et al.*, 1983;) with codon 12 mutations predominating (Bos *et al.*, 1987), (Figure 2.2). Analysis of K-ras mutations at the various stages of tumour progression reveals a correlation between mutation position and tumour stage. Moerkerk *et al.*, 1994, have demonstrated that Dukes B stage tumours harbour predominantly G-A (first position) codon 12 mutations, and a smaller number of G-T mutations, whereas, Dukes C stage tumours harbour predominantly G-T mutations with a marked absence of G-A mutations. These results suggest that metastatic potential is related to the nature of genetic alterations in codon 12 and that cells with G-A mutations may have reduced metastatic potential. Further analysis of tumour stage and associated mutation is required in order to determine the true involvement of mutation type with metastatic potential.

The ability to detect K-ras mutations in a rapid and sensitive manner and to correlate mutation type with stage and aggressiveness of tumours may have important implications for diagnosis and prognosis in cancer patients. More than 90% of patients with adenomas or carcinomas confined to the mucosa can be treated with surgical methods, whereas less than 10% of patients with distant metastases can be treated successfully (Rich *et al.*, 1983). This highlights the significant importance that early identification of metastatic potential may have on diagnosis, treatment and prognosis of colon cancer patients.

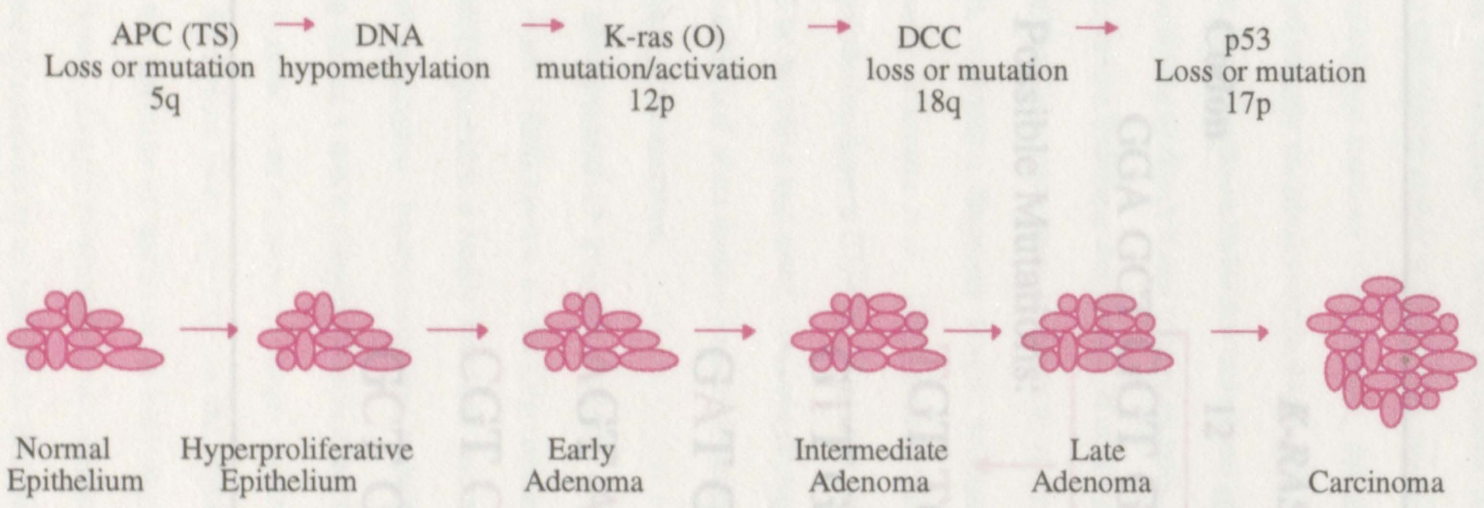


Figure 2.1: An adaptation of Fearon and Vogelstein's model for the development of colorectal cancer. (Fearon and Vogelstein, 1990). Gene mutations are displayed in parallel with histological changes. Note: early adenomas- less than 1cm in size, intermediate- more than 1cm in size but without foci of carcinoma and late adenoma- more than 1cm but with foci of carcinoma. TS- tumour suppressor gene, O-oncogene.

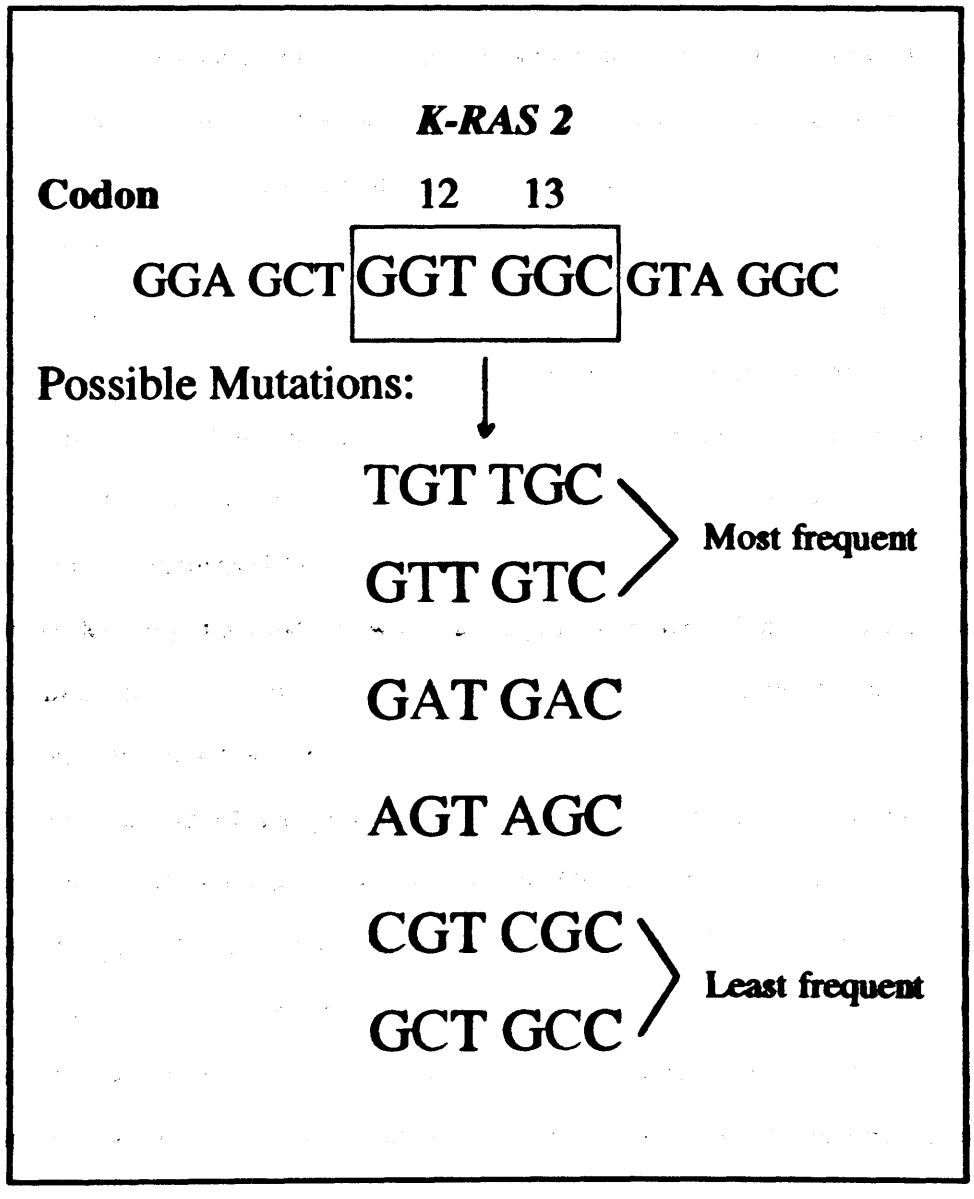


Figure 2.2. Mutations of codon 12 and 13 of the K-ras 2 gene. G to T mutations are the most frequently observed mutations whereas G to C are observed at a much lower frequency.

2.1.(IV) MUTATION DETECTION

The identification of K-ras mutations may provide useful biomarkers for identifying individuals at risk of developing colon cancer, particularly in the analysis of early pre-malignant tissues. Furthermore, the ability to identify mutations in stool samples and colonic washings may facilitate non-invasive analysis and assist at clinical level in the early diagnosis, treatment and cancer prevention.

The ability to detect K-ras mutations with sensitivity and accuracy relies on the mutation detection technique employed. A number of methods have been applied to the analysis of K-ras mutations including SSCP, (Orita *et al.*,1989 (a and b); Moerkerk, *et al.*, 1994), restriction fragment length polymorphism (RFLP), RNase mismatch cleavage assay (Forrester *et al.*, 1987) mutant allele enrichment (MAE) and denaturing gradient gel electrophoresis (DGGE) (Zhu *et al.*, 1997). These techniques have proved successful in detecting mutations, however all require a number of manipulations, are time consuming and often require nucleotide sequence analysis in order to identify the exact position of the mutation.

A new method of analysis recently developed at the laboratory in UCC (Vaughan *et al.*, 1998) known as the Glycosylase Mediated Polymorphism Detection assay (GMPD) provides a highly sensitive accurate and reproducible method for the analysis of mutations. Furthermore the technique offers the specificity of direct sequencing but is a much simpler process than those mentioned above as almost all mutations may be detected using a single enzyme. In addition the process is not sequence dependant, does not rely on the formation of heteroduplex DNA and it facilitates the detection of mutations in homozygous and heterozygous samples with equal efficiency. Simple manipulation of components of the reaction allows its use in the scanning of mutations in an area where mutation sites are unknown (GMPD-scan), or it may be used to determine the presence or absence of a mutation in a sequence with a known mutation site (GMPD-check).

2.1.(V) THE GMPD ASSAY

The GMPD process employs the use of a highly specific DNA glycosylase enzyme- in this case uracil DNA glycosylase. UDG is a highly specific DNA repair enzyme which removes uracil from DNA. Uracil is generated *in vivo* by either the misincorporation of dUTP by DNA Polymerase or by the deamination of cytosine residues (Lindahl *et al.*, 1977). UDG cleaves the N-glycosidic bond between a uracil base and the deoxyribose sugar in a DNA molecule, an apyrimidinic (AP) site is produced and this is usually cleaved *in vivo* by an associated AP endonuclease. In vitro cleavage of the AP site may be induced by the addition of NaOH or specific enzymes. (Lindahl *et al.*, 1972; Price *et al.*, 1991), (Figure 2.3).

GMPD-CHECK

The principle of the GMPD assay is based on the ability of UDG to detect the presence or absence of a uracil residue in a specific sequence within a DNA molecule. A target region of DNA is amplified using primers designed to amplify only the region of interest. PCR is carried out in the presence of dUTP, dGTP, dCTP and dATP. This ensures that at every T position a U residue is incorporated. The upper or lower primer may be labelled depending on the strand of interest (known as the diagnostic primer). The primers are designed to ensure that during PCR amplification, the first U residue is incorporated at the mutation site if the mutation is present, or distal to the site if the mutation is absent. Cleavage of the resulting DNA molecule with UDG generates AP sites which may be cleaved enzymatically or chemically. The resulting digestion fragments are separated by polyacrylamide gel electrophoresis and a specific pattern of fragments is generated. If the mutation is present the labelled strand is cleaved at the mutation position generating a fragment of specific length (X, see fig 2.4) However in the absence of the mutation UDG does not recognise the site and cleavage occurs at the next U residue 3' to the mutation, generating a fragment of larger size (Y). Analysis of

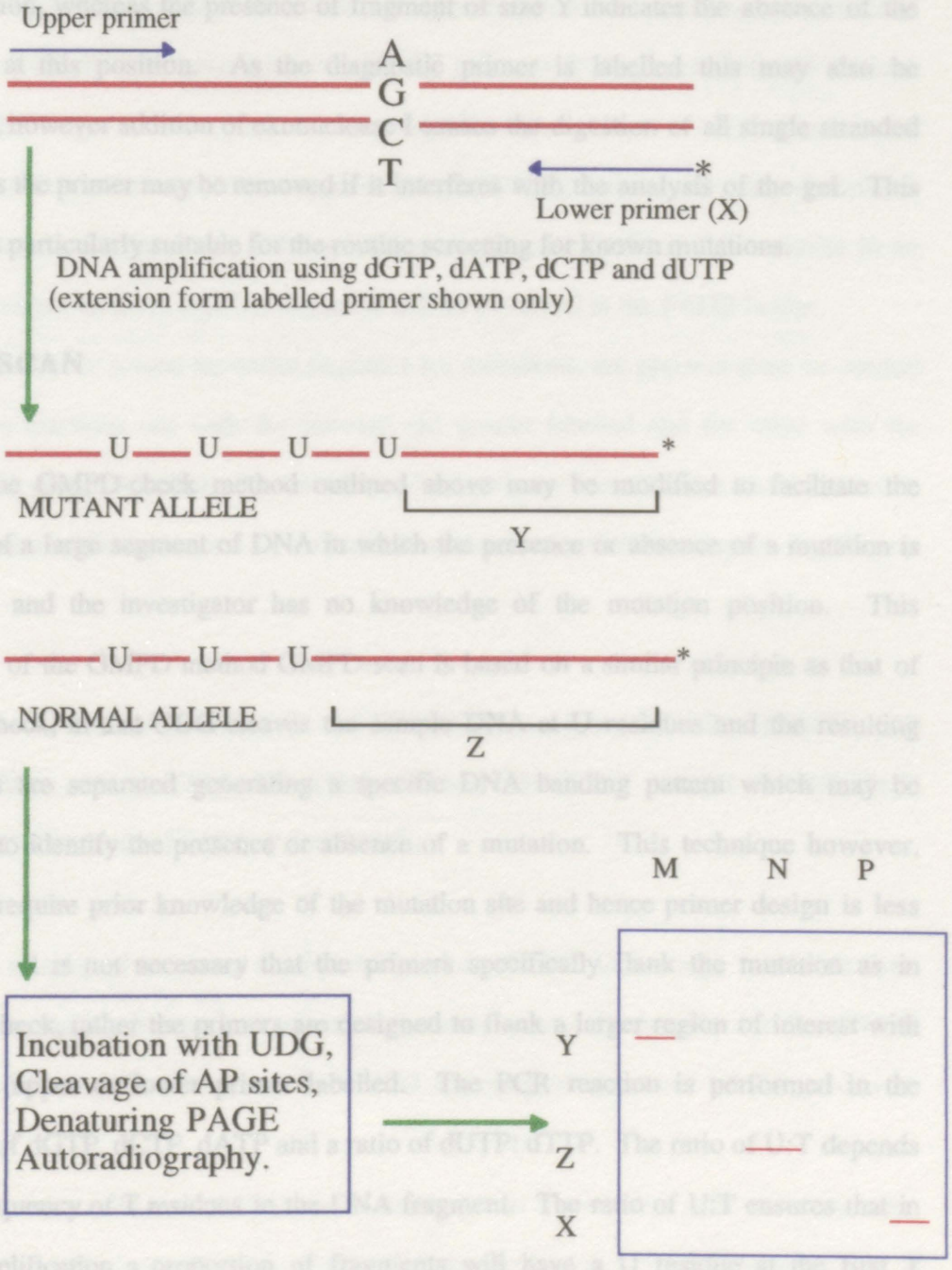


Figure 2.4. Schematic representation of the GMPD-check method. The labelled diagnostic primer is designated X, the mutant allele produces a fragment of size Y and the normal allele produces a digestion product of size Z. A heterozygote will contain fragments of size Y and Z. The presence or absence of these fragments in a GMPD digestion profile indicates the presence or absence of a mutation in that sample.

the electrophoretic separation profile of the digestion fragments identifies the presence or absence of the mutation, the presence of fragment of size X indicates the presence of the mutation, whereas the presence of fragment of size Y indicates the absence of the mutation at this position. As the diagnostic primer is labelled this may also be observed, however addition of exonuclease I causes the digestion of all single stranded DNA thus the primer may be removed if it interferes with the analysis of the gel. This method is particularly suitable for the routine screening for known mutations.

GMPD-SCAN

The GMPD-check method outlined above may be modified to facilitate the analysis of a large segment of DNA in which the presence or absence of a mutation is unknown and the investigator has no knowledge of the mutation position. This extension of the GMPD method GMPD-scan is based on a similar principle as that of GMPD-check, in that UDG cleaves the sample DNA at U residues and the resulting fragments are separated generating a specific DNA banding pattern which may be analysed to identify the presence or absence of a mutation. This technique however, does not require prior knowledge of the mutation site and hence primer design is less stringent. It is not necessary that the primers specifically flank the mutation as in GMPD- check, rather the primers are designed to flank a larger region of interest with either the upper or lower primer labelled. The PCR reaction is performed in the presence of dGTP, dCTP, dATP and a ratio of dUTP: dTTP. The ratio of U:T depends on the frequency of T residues in the DNA fragment. The ratio of U:T ensures that in every amplification a proportion of fragments will have a U residue at the first T position only, some fragments at the second T position only and so on so that the resulting test sample will provide a heterogeneous mix of DNA fragments with U residues at all possible T positions. The resulting fragments are cleaved by UDG to generate AP sites which are digested chemically or enzymatically. The resulting fragments are analysed by PAGE. Analysis of the resulting banding pattern allows the

identification and localisation of a mutation. For example, analysis of a normal DNA fragment containing 10 T residues with GMPD-scan will yield a PAGE ladder of 10 labelled fragments. Analysis of a similar fragment with a T-C mutation at the third T residue will result in the generation of a PAGE ladder of 9 fragments with the absence of a labelled fragment at the position of the third T residue. Uracil is not incorporated at the mutation site, cleavage does not occur so a DNA fragment of this particular size is not generated. Alternatively, if the sequence contains a mutation which results in an extra T residue, an extra labelled fragment will be observed in the PAGE ladder.

In order to screen the entire sequence for mutations, the process must be carried out in two reactions one with the forward (F) primer labelled and the other with the reverse (R) primer labelled. The technique facilitates the identification and localisation of a mutation without the need for multiple manipulations and DNA sequencing (Figure 2.5).

2.1.(VI) APPLICATION OF GMPD-SCAN AND GMPD-CHECK TO MUTATION ANALYSIS OF THE K-RAS GENE IN IRISH SPORADIC COLORECTAL CANCER PATIENTS.

In an attempt to assess the suitability of the GMPD technology for the analysis of K-ras mutations in tumours of cancer patients, protocols specific for the GMPD analysis of codon 12 and 13 and the GMPD scanning of exon one of the gene were designed.

Previous analyses of DNA from tumour cells shows that cells are very often mixed with populations of non-mutated cells (Moerkerk *et al.*, 1994). The specificity of the GMPD technology (particularly GMPD-check) should facilitate the detection of mutant DNA even in a heterogeneous cell population.

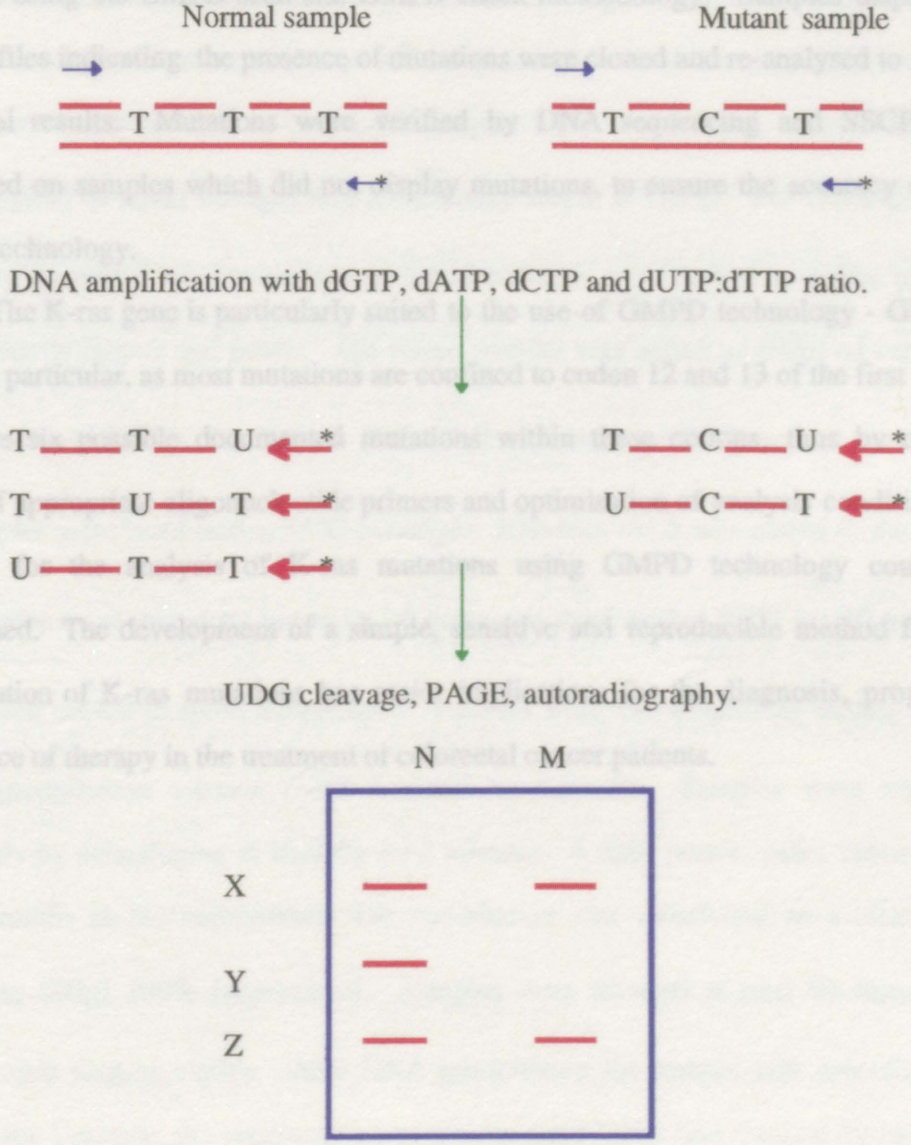


Figure 2.5. Schematic representation of the GMPD-scan method of mutation analysis. The absence of fragment size Y signifies a mutation at the second T position in the mutant sample. The GMPD digestion profile signifies the presence or absence of a mutation at particular positions in the target sequence.

This chapter presents the protocols established for both methods and the results of the application of these techniques to the detection of K-ras mutations in human colorectal cancer samples. Briefly 20 matched normal and tumour samples were analysed using the GMPD-scan and GMPD-check methodology. Samples displaying scan profiles indicating the presence of mutations were cloned and re-analysed to verify the initial results. Mutations were verified by DNA sequencing and SSCP was performed on samples which did not display mutations, to ensure the accuracy of the GMPD technology.

The K-ras gene is particularly suited to the use of GMPD technology - GMPD-check in particular, as most mutations are confined to codon 12 and 13 of the first exon. There are six possible documented mutations within these codons, thus by careful design of appropriate oligonucleotide primers and optimisation of analysis conditions, a protocol for the analysis of K-ras mutations using GMPD technology could be established. The development of a simple, sensitive and reproducible method for the identification of K-ras mutations has major implications for the diagnosis, prognosis and choice of therapy in the treatment of colorectal cancer patients.

2.2 MATERIALS AND METHODS

2.2. (I) DNA ISOLATION

DNA used in the GMPD analysis of the K-ras gene was isolated from blood, tumour tissue, normal mucosa and paraffin embedded archival tissues. DNA isolation from blood was performed as outlined in chapter 1 (Section 1.2.(I)). Tissue samples were supplied in liquid nitrogen and immediately stored at -70°C. For DNA isolation 10-20 mg of tissue was maintained in liquid nitrogen and crushed to a fine powder using a sterile mortar and pestle. The tissue powder was added to 600µl of cell lysis solution (25mM EDTA, 2% SDS) at 4°C. 3 µl of Proteinase K (20mg/ml) was added and samples were incubated at 55°C overnight. RNase A (1U) was added to the lysate and samples were mixed by inverting 25 times and incubated at 37°C for 60 minutes. Tubes were cooled to room temperature. Proteins were precipitated by adding 200µl protein precipitation solution (10M Ammonium Acetate). Samples were vortexed vigorously by centrifuging at 16,000g for 3 minutes. A tight protein pellet forms - (the DNA remains in the supernatant) The supernatant was transferred to a clean tube containing 600µl 100% Isopropanol. Samples were inverted at least 50 times until DNA strands became visible. After DNA precipitation the sample was centrifuged at 16,000g for 1 minute, the supernatant was removed and DNA was drained by inverting on clean absorbent paper. 600µl 70% ethanol was added to wash the DNA pellet. The sample was again centrifuged at 16,000g for 1 minute ethanol was removed and DNA allowed to air-dry. DNA samples were re-hydrated by adding 50-100µl TE buffer (TrisHCl 10mM pH 8.0, 1mM EDTA) and stored at 4°C for 4 days until dissolved.

DNA ISOLATION FROM PARAFFIN EMBEDDED TISSUES (PET):

DNA extraction from paraffin sections was performed using the Nucleon HT DNA Isolation kit from Amersham Life Science. 20-30 micron sections were sliced from PET tumour slides at the histology laboratory at CUH.

Tissue sections were placed in a 1.5ml tube (one section per tube) covered with Xylene and incubated at 37°C for 20 minutes (Xylene dissolves the paraffin). Samples were centrifuged at 1300g for 5 minutes and xylene was removed. Samples were again incubated in xylene for 2 minutes at room temperature, centrifuged and the xylene was removed. The samples were re-hydrated by washing consecutively in 100% ethanol, 75% ethanol, 50% ethanol and 25% ethanol and finally sterile water. Samples were centrifuged at maximum speed for 3 minutes between each wash. The water was removed from the pellet after the final wash and 0.35ml of reagent B (supplied in the Nucleon HT kit) was added. 18µl of Proteinase K (20ng/µl) was added and samples were incubated overnight at 55°C. Proteins were removed by adding 100µl of 5M sodium perchlorate solution and inverting. 600µl of chloroform was added and tubes were inverted again. 150µl of Nucleon resin (supplied) was added to the tubes and samples were centrifuged for 1 minute at 350g. The upper aqueous phase (containing the DNA) was removed to a clean tube, two volumes of 100% ethanol were added tubes were mixed by inversion and incubated at -20°C for 1-2 hours to allow DNA precipitation to occur. Tubes were centrifuged at top speed for 15 minutes and the supernatant discarded. The remaining DNA pellet was washed by adding 1ml cold 70% ethanol, inverting, centrifuging and removing the supernatant. The DNA was air dried and re-suspended in 50µl of TE (TrisHCl 10mM pH 8.0, 1mM EDTA) and incubated at 4°C until dissolved.

2.2.(II) GLYCOSYLASE MEDIATED POLYMORPHISM DETECTION ANALYSIS.

Oligonucleotide Labelling:

Nucleotide sequences of all the primers employed in the GMPD analysis of the K-ras gene are listed in table 2.1.

Primers (60pmol) were end labelled by incubation with 1U of T4 Polynucleotide Kinase (New England Biolabs), the appropriate buffer and 50uCi [γ-^{32}P]-ATP (3000Ci/mmol) at 37°C for 30 minutes. One volume of 4M sodium acetate and two volumes of ethanol were added in order to precipitate the primers and to remove unlabelled nucleotide. Primers were incubated overnight at -20°C, centrifuged for 20 minutes at high speed to precipitate, the supernatent was removed and primers were allowed to air dry. Primers were re-suspended in water to the required concentration.

PCR AMPLIFICATION OF SAMPLES FOR GMPD-SCAN ANALYSIS.

Selective amplification of the K-ras 2 functional gene:

In order to selectively amplify the functional K-ras 2 gene and not the pseudogene, an initial PCR was performed using primer pair U+2. PCR amplification was performed in a 10µl reaction volume containing 40ng genomic DNA, 40ng each primer, 0.2mM dNTPs, 1U Taq DNA polymerase and *Taq* DNA polymerase buffer (50mM KCl, 10mM TrisCl pH 9.0, 0.1% Triton X-100 and 1.5mM $MgCl_2$).

Cycling conditions: Reactions were performed on an MJ PTC-100 thermal cycler. Samples were initially denatured by heating to 95°C for 5 minutes (hot-start) before the addition of the *Taq* DNA polymerase. Reactions were subjected to 35 cycles of 95°C

for 1 minute, 55°C for 1 minute and 72°C for 1 minute. To ensure complete extension a final 10 minute incubation at 72°C was performed.

Preparation of samples for GMPD-scan:

Exon 1 of the K-ras 2 gene was amplified for GMPD analysis using primer pair F+R. PCR amplification was performed in a 10µl reaction volume containing 2µl of 1/100 dilution of the U+2 amplified template, 20ng of each primer (with either the F or R primer labelled), 1U *Taq* DNA polymerase, *Taq* DNA polymerase buffer (as before) and 0.2mM dNTP mix containing 0.2mM dGTP, dCTP, dATP and 0.2mM dCTP:dTTP at a ratio of 1: 10. Cycling parameters: Samples were subjected to a hot start denaturation step of 95°C for 5 minutes followed by 30 cycles of 95°C for 1 minute, 55°C for 1 minute and 72°C for 1 minute, a final incubation step of 72°C for 10 minutes was also included.

PCR AMPLIFICATION OF SAMPLES FOR GMPD-CHECK ANALYSIS.

The primers used to detect mutations at codons 12 and 13 of the k-ras 2 gene included combinations of primers A, B and C. The primer sequences are listed in table 2.1 and the primer pair combinations and target mutations are outlined in Figure 2.12. One labelled primer was included with an unlabelled primer in each reaction as appropriate.

PCR amplification using primer pair A+C was performed in a 10µl reaction volume containing 2µl of 1/100 dilution of the U+2 amplified template, 20ng of each primer, 1U Taq DNA polymerase and buffer (as before) and 0.2mM dNTP mix including

dATP, dGTP, dCTP and dUTP. No dTTP was added, this ensures that all the PCR products are digested with UDG at the first U residue.

Cycling Parameters: Amplification was performed on an MJ PTC-100 thermal cycler. Samples were subjected to a hot start denaturation step followed by 35 cycles of 95°C for 1 minute, 62°C for 1 minute, 72°C for 1 minute and a final extension step of 72°C for 10 minutes.

PCR amplification using primer set B + C was performed as above, however an annealing temperature of 65°C was required to ensure optimum amplification.

GLYCOSYLASE MEDIATED CLEAVAGE OF DNA:

The radiolabelled PCR products, containing U residues were treated with Exonuclease I (ExoI) (Amersham Life Sciences), to digest any primers unused in the amplification reaction. 5µl of PCR product was incubated with 1U of ExoI at 37°C for 30 minutes. Samples were subsequently heated to 80°C for 15 minutes in order to heat inactivate the exonuclease. 1U of Uracil DNA Glycosylase (UDG), (*E. coli*) was then added and the incubation was continued at 37°C for 30 minutes. The UDG cuts the DNA at the U residues, the resulting apyrimidinic sites in the DNA samples were cleaved to completion by adding NaOH to a final concentration of 0.05M and heating to 95°C for 15 minutes (cleavage occurs at the 5' side of each AP site). The reaction was neutralised by adding Tris base to a final concentration of 30mM . The dilution buffer for UDG and ExoI contains 70mM HEPES-KOH, pH 8.0, 1mM EDTA, 1mM DTT and 50% Glycerol.

PREPARATION OF SAMPLES FOR GEL ANALYSIS.

Samples were diluted with an equal volume of formamide loading dye (95% formamide, 0.025% bromophenol blue, 0.025% xylene cyanol) and denatured by heating at 85°C for 5 minutes. Samples were subsequently loaded onto a 20% denaturing (7M urea) polyacrylamide gel and electrophoresis was carried out for 3-4 hours at 60 Watts. Following electrophoresis autoradiography was performed by exposing gels to X-ray film for 2-12 hours (depending on the signal intensity) at -70°C.

2.2. (III) SEQUENCE ANALYSIS OF K-RAS GENE PRODUCTS

Samples which appeared to harbour mutations in the GMPD analysis were cloned, re-analysed by GMPD-check and sequenced.

CLONING OF PCR PRODUCTS

The pTag cloning vector (R&D systems) exploits the A residue that Taq DNA polymerase adds on to the end of PCR products. Ligation of the PCR product with the pTag plasmid was performed by incubating 2µl of PCR product (removed directly from a 10µl PCR reaction tube) with 2U T4 DNA ligase, ligase buffer, 5mM DTT, 0.5mM ATP and 50ng of pTag plasmid in a 10µl reaction. Samples were incubated overnight at 16°C and subsequently transformed into competent *E. coli* cells as described by Sambrook *et al.*, 1989. Transformants were plated on LB-AMP-IPTG-X-gal plates and incubated overnight at 37°C. Recombinant clones were selected based on antibiotic resistance and insertional inactivation of the β-Galactosidase gene.

Sixty colonies from each plate were re-analysed by GMPD-check in order to determine the presence or absence of the mutant fragment. Analysis was performed on

products amplified directly from colonies (colony PCR). Samples which proved positive for mutations on GMPD-check re-analysis were sequenced either by cycle sequencing incorporating α^{33}-P labelled dideoxy terminators and the Thermosequenase Cycle Sequencing Kit (Amersham Life Sciences) or by fluorescent analysis using Cy5 labelled primers in a cycle sequencing reaction incorporating unlabelled dideoxy terminators.

DNA sequencing was performed either directly on the colony PCR product (PCR was performed using the U+2 primer pair- sequencing was performed using the GMPD check primers A, B or C), or on plasmid DNA isolated using the minipreparation protocol as outlined in Sambrook *et al.*, 1989. PCR products for sequencing were prepared by incubating 5µl PCR product with 0.1U ExoI, 0.2U Shrimp alkaline phosphatase at 37°C for 30 minutes followed by 85°C for 15 minutes. This removes unincorporated dNTPs and primers.

Analysis of the radioactively labelled sequencing products was performed by electrophoresis in a 6% denaturing polyacrylamide gel whereas the fluorescent analysis was performed on an ALF -express automated DNA sequencer (Pharmacia).

Primer name:	Primer sequence	Size
U	5' TAG TCA CAT TTT CAT TAT TTT TAT TAT AA 3'	29mer
2	5' GCA TAT TAA AAC AAC TAA ATG GAG A 3'	24mer
F	5' TAA GGC CTG CTG AAA ATG AC 3'	20mer
R	5' ATT GTT GGA TCA TAT TCG TC 3'	20mer
A	5' AAC TTG TTG TAG TTG GAG CT 3'	20mer
B	5' AAA CTT GTG GTA GTT GGA GCT GGT 3'	24mer
C	5' TAG CTG TAT CGT CAA GGC ACT CTT GCC TA 3'	29mer

Table 2.1. Nucleotide sequence analysis of oligonucleotide primers used in the GMPD analysis of the K-ras 2 gene.

2.3 RESULTS

2.3.(I) PATIENTS and SAMPLES

Tumour samples and matched normal colonic mucosa were obtained from twenty sporadic colon cancer patients who presented at Cork University Hospital. Sporadic colon cancer cases where chosen based on the absence of FAP symptoms (polyposis), HNPCC characteristics (extracolonic tumours and early onset) and evidence for a family history of the disease. The samples set included thirteen females and seven males all over 50 years of age who had been diagnosed with cancer of the colon. Matched normal and tumour DNA from these patients were analysed for K-ras mutations using the GMPD-scan followed by the more specific GMPD-check method.

2.3.(II) GMPD-SCAN

Primer selection and protocol design:

Hybridisation analysis using genomic DNA coupled to the characterisation of molecular clones has identified two structurally distinct human K-ras genes known as K-ras 1 and K-ras 2 (Chang *et al.*, 1982). Sequence analysis of these genes suggests that the K-ras 2 gene is more complex than K-ras 1 which shares homology with the viral ras gene. Sequence comparisons identified K-ras 1 as a pseudogene which resembles a processed mRNA molecule rather than an intact gene and as such is a functionless genetic entity (McGrath *et al.*, 1983).

In order to specifically target the K-ras 2 functional gene, primers were designed to selectively anneal to and amplify regions unique to the K-ras 2 gene. Figure 2.6 details DNA sequences unique to K-ras 2 and indicates the position of the specific PCR primer. In order to provide good quality, high quantity, clean DNA for analysis, a nested PCR amplification reaction was performed. The initial amplification

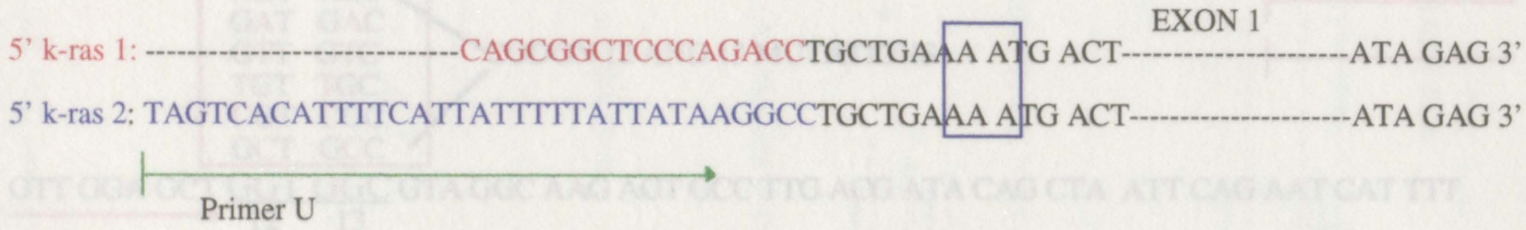


Figure 2.6. Comparison of the DNA sequence of human K-ras 2 functional gene with the K-ras 1 pseudogene. Regions of difference are indicated in colour. The 5' untranslated region of exon 1 shows a high degree of difference between the two genes. In order to selectively amplify only the K-ras 2 gene, primer U was designed to anneal to the K-ras 2 specific sequence.

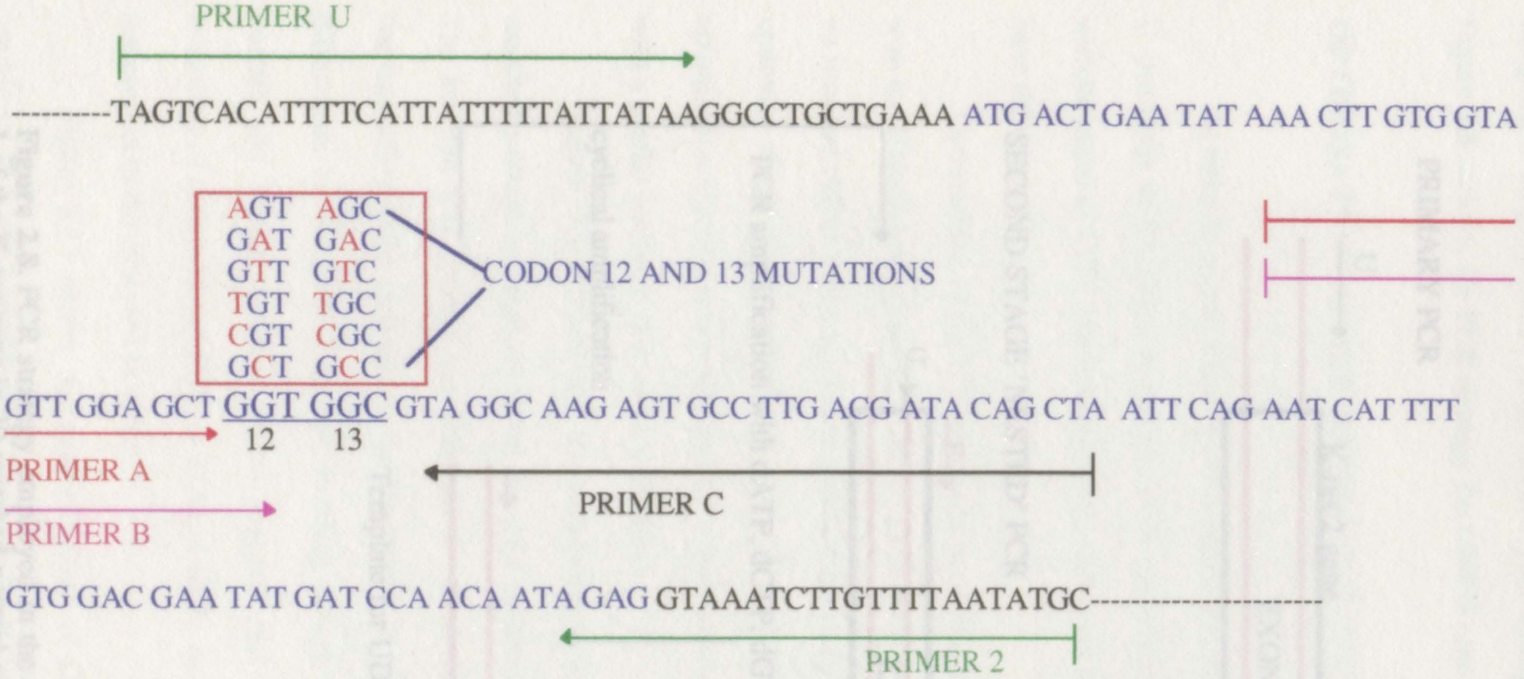


Figure 2.7. Target sequence for the GMPD analysis of exon 1 of K-ras 2.
The sequence flanking exon 1 is indicated in black, the coding sequence is highlighted in blue. The primer binding sites for the various primers are indicated by the arrows and primer names are provided below each primer site. Initial PCR amplification with primer pair U and 2 ensures that only K-ras 2 DNA is amplified as these primers do not bind to the K-ras 1 pseudogene. Combinations of primers AC and BA under various labelling options facilitates the analysis of all G-T and G-A mutations of codon 12 and 13 of exon 1 of the gene. The mutations of codon 12 and 13 of the gene are shown in red.

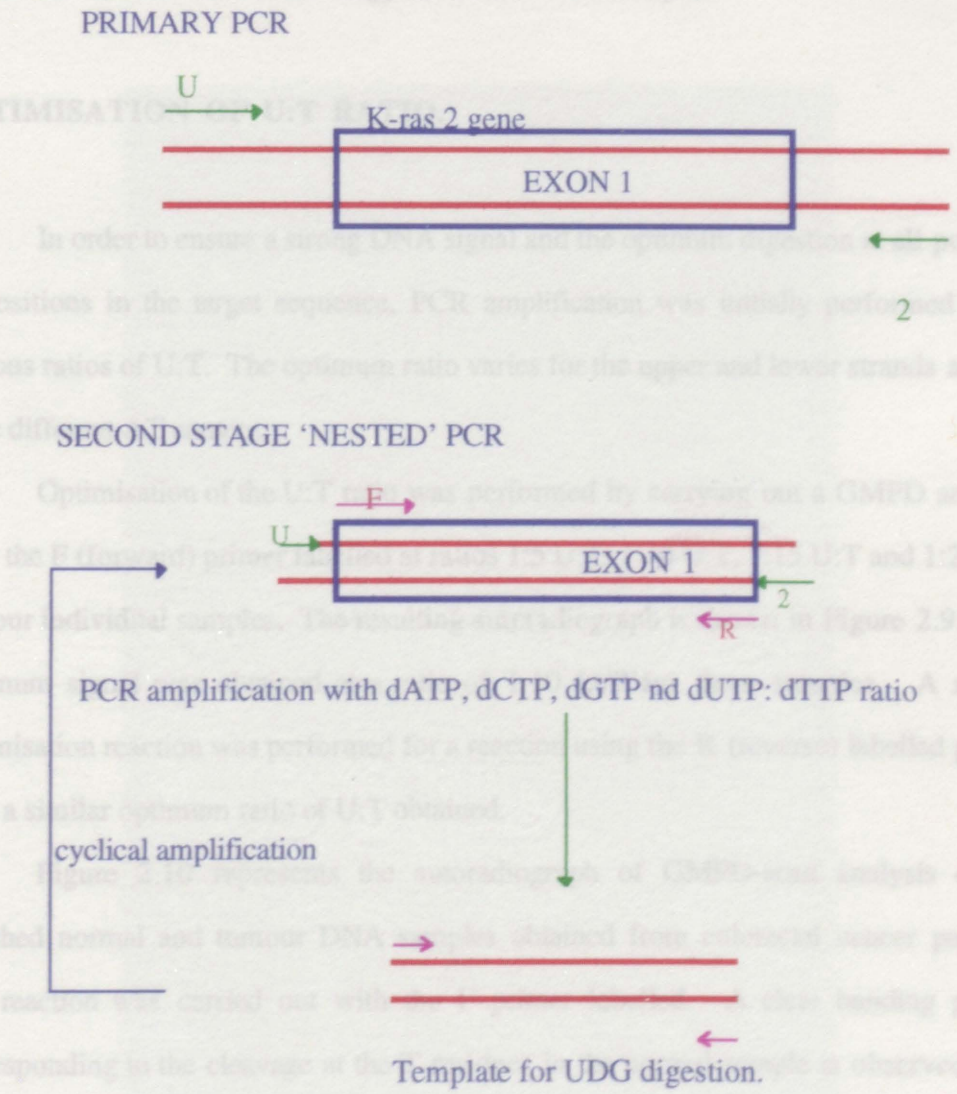


Figure 2.8. PCR strategy employed in the GMPD-scan analysis of exon 1 of the K-ras gene in 20 matched normal and tumour tissue samples.The initial U and 2 primer set selectively amplifies the K-ras 2 gene and not the K-ras 1 pseudogene. The second stage nested PCR provides good quality and high yield template for UDG digestion.

product of the K-ras 2 target sequence (selection primers designated U and 2 see Figure 2.7) was subsequently used as a template in a second round PCR reaction using primers designed to amplify exon 1 of the K-ras 2 gene (designated F and R primers see fig 2.7) Figure 2.8 outlines the PCR strategy for GMPD scan analysis.

OPTIMISATION OF U:T RATIO.

In order to ensure a strong DNA signal and the optimum digestion at all possible U positions in the target sequence, PCR amplification was initially performed using various ratios of U:T. The optimum ratio varies for the upper and lower strands as they have different AT content.

Optimisation of the U:T ratio was performed by carrying out a GMPD analysis with the F (forward) primer labelled at ratios 1:5 U:T, 1:10 U:T, 1:15 U:T and 1:20 U:T on four individual samples. The resulting autoradiograph is shown in Figure 2.9. The optimum signal was obtained at a ratio of 1:10 U:T for these samples. A similar optimisation reaction was performed for a reaction using the R (reverse) labelled primer with a similar optimum ratio of U:T obtained.

Figure 2.10 represents the autoradiograph of GMPD-scan analysis on 15 matched normal and tumour DNA samples obtained from colorectal cancer patients. The reaction was carried out with the F primer labelled. A clear banding pattern corresponding to the cleavage at the T residues in the normal sample is observed. No differences in the DNA fragment banding pattern were observed between matched normal and tumour samples. This suggests that GGT to TGT or GGT to GTT mutations of codon 12 or GGC to TGC or GGC to GTC mutations of codon 13 were not present in the analysed samples.

Figure 2.11 shows the autoradiograph of GMPD-scan analysis on the same sample set as above however in this reaction the R (reverse) primer was labelled. As in the previous assay, a clear banding pattern corresponding to the number of U incorporation sites in the sample DNA is observed. However in this case a number of

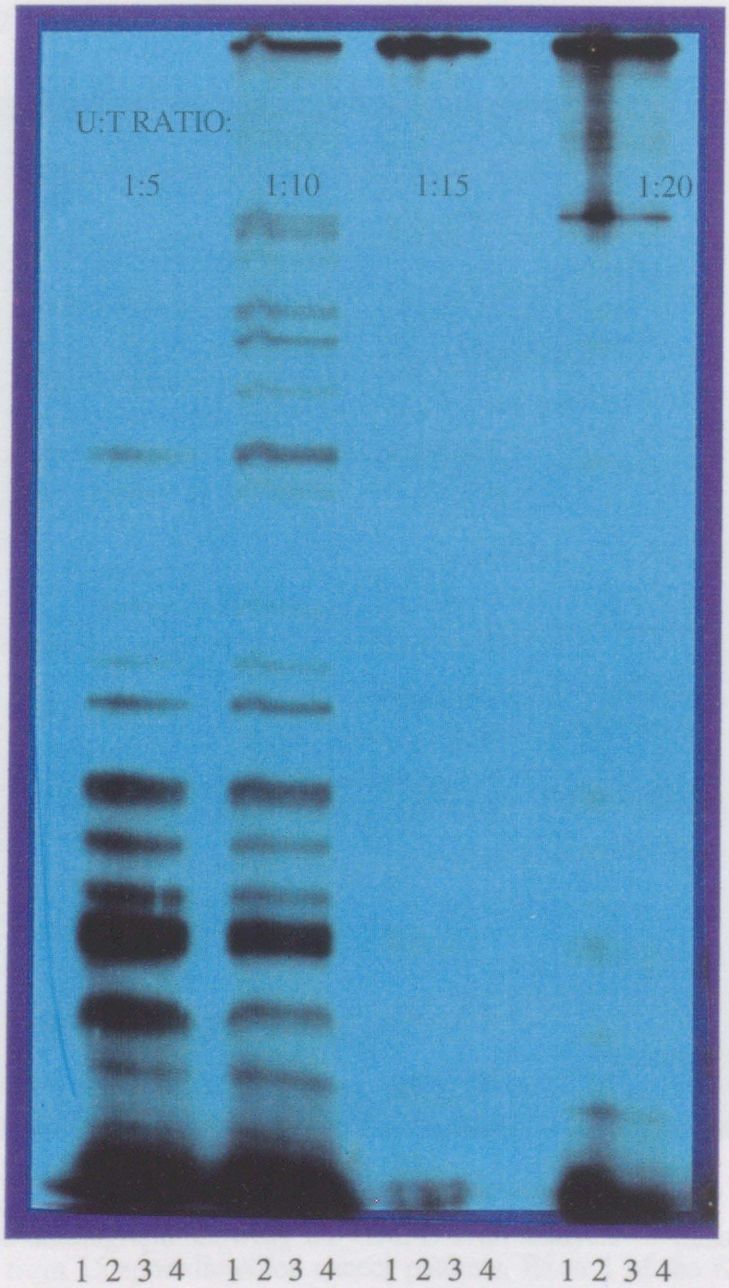


Figure 2.9. Optimisation of the U:T ratio for GMPD-scan analysis. Exon 1 of the k-ras 2 gene was amplified from 4 DNA samples in a PCR reaction containing $\gamma^{32}P$ labelled primer F and unlabelled primer R. The reaction was performed with U:T ratios of 1:5, 1:10, 1:15, 1:20 as indicated. Following amplification, GMPD analysis was performed on each sample, digestion products were separated in a 10% denaturing polyacrylamide gel and visualised by autoradiography.

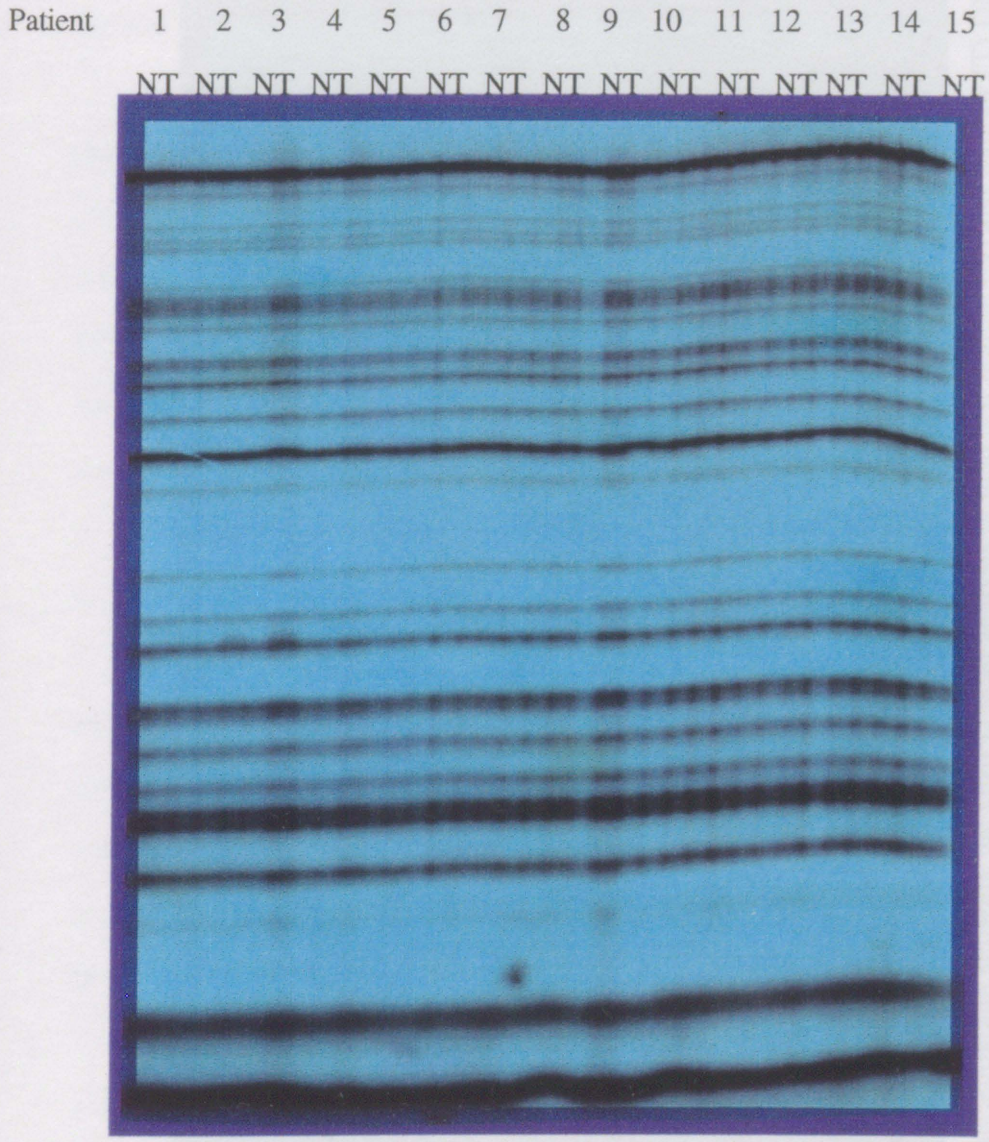


Figure 2.10. Autoradiograph showing the GMPD-scan analysis of matched normal and tumour DNA from 15 sporadic colon cancer patients. Exon 1 of the K-ras 2 gene was amplified from 20 matched normal and tumour DNA samples (not all shown), in a reaction containing unlabelled primer R and γ^{32} labelled primer F. Following amplification, GMPD analysis was performed on all samples and digestion products were separated in a 10% denaturing polyacrylamide gel. Normal (N) and tumour (T) DNA samples were loaded as indicated.

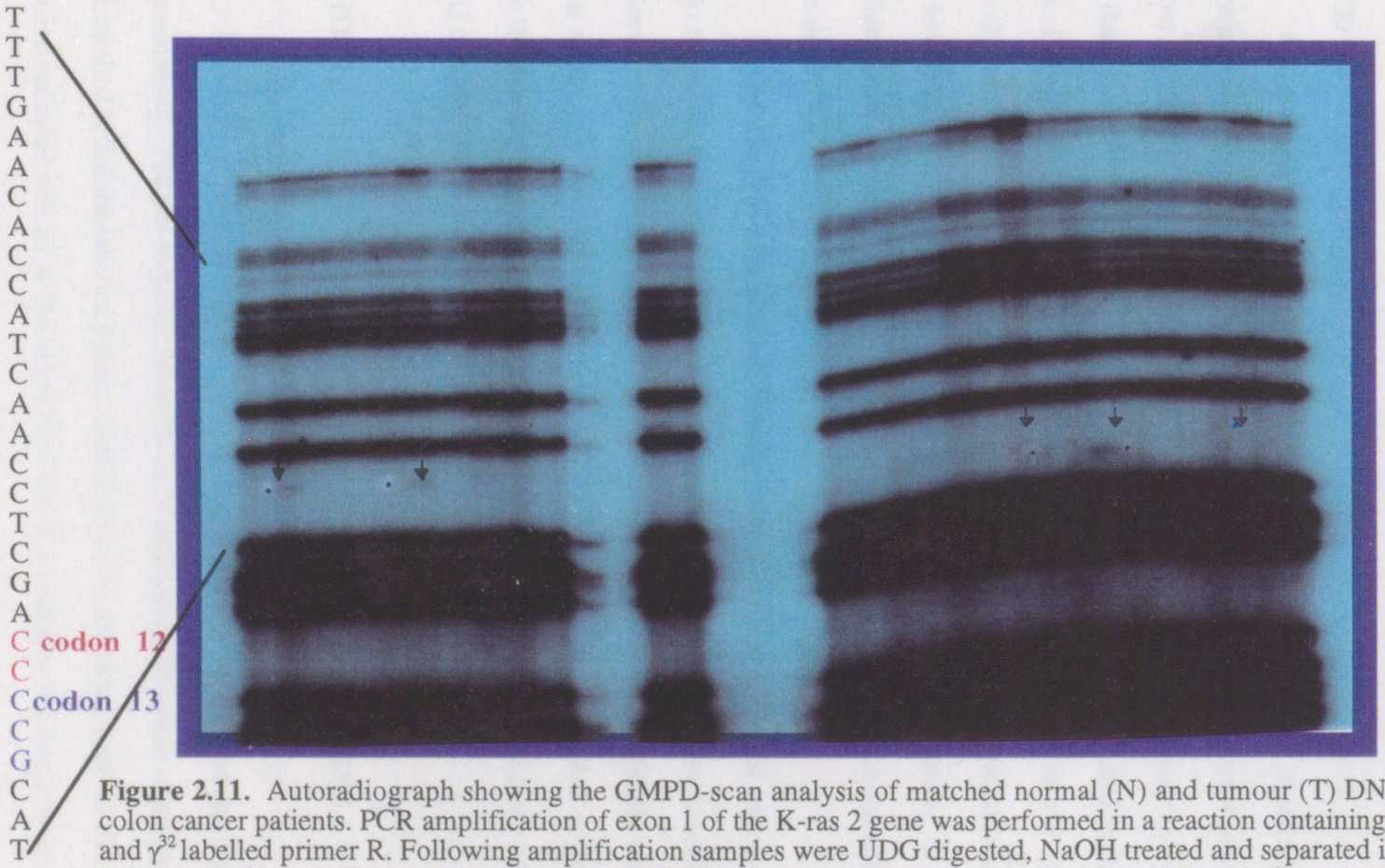


Figure 2.11. Autoradiograph showing the GMPD-scan analysis of matched normal (N) and tumour (T) DNA from 14 sporadic colon cancer patients. PCR amplification of exon 1 of the K-ras 2 gene was performed in a reaction containing unlabelled F primer and γ^{32} labelled primer R. Following amplification samples were UDG digested, NaOH treated and separated in a 10% denaturing polyacrylamide gel. The corresponding normal DNA sequence is shown and codon 12 and 13 are indicated, additional digestion products are present in tumour DNA from patients 2, 10, 14, 40 and 44 (indicated by arrow).

very faint additional bands are present in samples 2, 10, 14, 40 and 44. The extra bands were observed in tumours from 5 sporadic colon cancer patients and are not visible in the DNA samples obtained from the corresponding normal mucosa. This indicates that the additional fragments are most likely due to a mutation in these samples. The mutations which may be detected when the reverse primer is labelled are GTT to GAT or AGT at codon 12 and GGC to GAC or AGC at codon 13.

On analysis of the results obtained using the GMPD-scan method, it was determined that the absence of mutations in the reaction with the F primer labelled and the low signal obtained for additional bands observed when the R primer was labelled was due to the fact that the tumour DNA may have been contaminated with normal DNA. Contamination may have occurred at the DNA extraction stage as it is likely that surgical removal of tumour from the patient may have also included normal tissue. This may have been inadvertently extracted from the sample on DNA isolation. It was not possible to estimate the amount of contamination as this is dependant on the accuracy of tissue dissection.

In order to determine if the additional bands observed represent true mutations, and to re-assess the remaining samples for mutations, the GMPD-check method was employed. The GMPD-check method is more sensitive as it specifically targets the region of interest directly, furthermore no dTTP residues are included in the PCR, hence the resulting signal is stronger as all the labelled DNA is digested at the site of the first U residue.

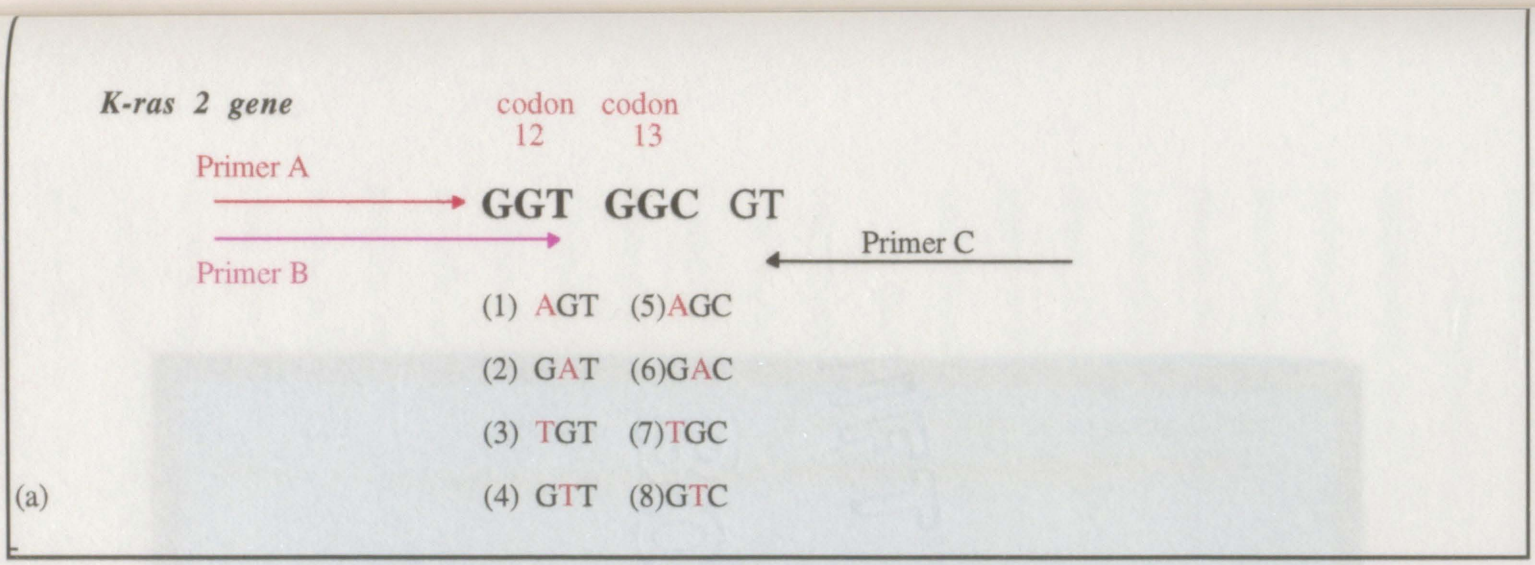
2.3.(III) GMPD- CHECK ANALYSIS OF THE K-RAS GENE

Selection of primers for use in GMPD-check analysis requires careful design and positioning. The pre-requisite for analysis of a particular mutation using the GMPD check method is that the labelled primer must lie close to the target base so that the first uracil incorporated will be at the target base site if a mutation is present or at the next thymine residue if a mutation is absent. Due to the numerous mutations which are

possible at codons 12 and 13 of K-ras, three primers were designed so that depending on the primer combinations and labelling options, all mutations could be detected. Figure 2.12 details the target sequence, primer annealing sites, possible mutations and the primer combinations which may detect them. It should be noted that the template DNA in all cases was the PCR product of a PCR reaction performed with primers U and 2. These primers ensure that only the K-ras 2 gene sequence is analysed.

GMPD analysis was performed on the twenty matched normal and tumour samples using primer combinations of (1) primer A and C with the A primer labelled, (2) primer B and C with B labelled and (3) primer A and C with C labelled. (see fig 2.12). Analysis of the assay design revealed that in the case of detecting mutation number 3 and 7 (fig. 2.12) in codons 12 and 13 (GGT to TGT and GGC to TGC), the resulting products are similar in size to the labelled primer. Therefore, in order to avoid masking the mutation by the presence of the primer, exonuclease I was added to the reaction. Exonuclease I digests all single stranded DNA, thus digesting the labelled primer and leaving the target DNA intact.

PCR amplification of the K-ras gene for GMPD-check analysis was performed using the primer combination of A and C with the A (upper) primer labelled. Twenty matched normal and tumour DNA samples from colorectal cancer affected patients were amplified in an AC PCR reaction, in the presence of dUTP in place of dTTP, resulting products were digested with UDG and DNA cleavage was induced by the addition of NaOH. Digestion products were separated by electrophoresis in a 10% polyacrylamide gel. Figure 2.13 shows the GMPD-check analysis of 10 sporadic colon cancer patients. Lanes 20-40 represent normal and tumour samples respectively from the ten patients. Lane 41 contains labelled primer A only which was treated with exonuclease I (Exo I). The absence of a band in this lane demonstrates that labelled primer is removed in all ExoI treated samples. Lane 42 contains undigested labelled primer A only (20bp product). Normal samples which do not contain a mutation yield products of 22bp after digestion. Full length uncut PCR product of 54bp is visible in all samples, indicating that digestion of amplified products by UDG glycosylase was incomplete. A band of



Primer set	Labelled primer	Primer size	Normal digestion product	Full length uncut product	Mutations which may be detected
AC	A	20bp	22bp	54bp	Mutation (3) codon 12: 20bp Mutation (4) codon 12: 21bp
BC	B	24bp	28bp	55bp	Mutation (7) codon 13: 24bp Mutation (8) codon 13: 25bp
AC	C	27bp	37bp	54bp	Mutation (1) codon 12: 33bp Mutation (2) codon 12: 32bp Mutation (5) codon 13: 30bp Mutation (6) codon 13: 29bp

Figure 2.12.(a) Mutations of codon 12 and 13 of the K-ras gene which may be detected by GMPD-check using various primer combinations and labelling options. (b) Details of the primer pair combinations which were used in the GMPD-check analysis of the K-ras 2 gene. Primer size, full length PCR product size and digestion product sizes are detailed for each primer pair combination.

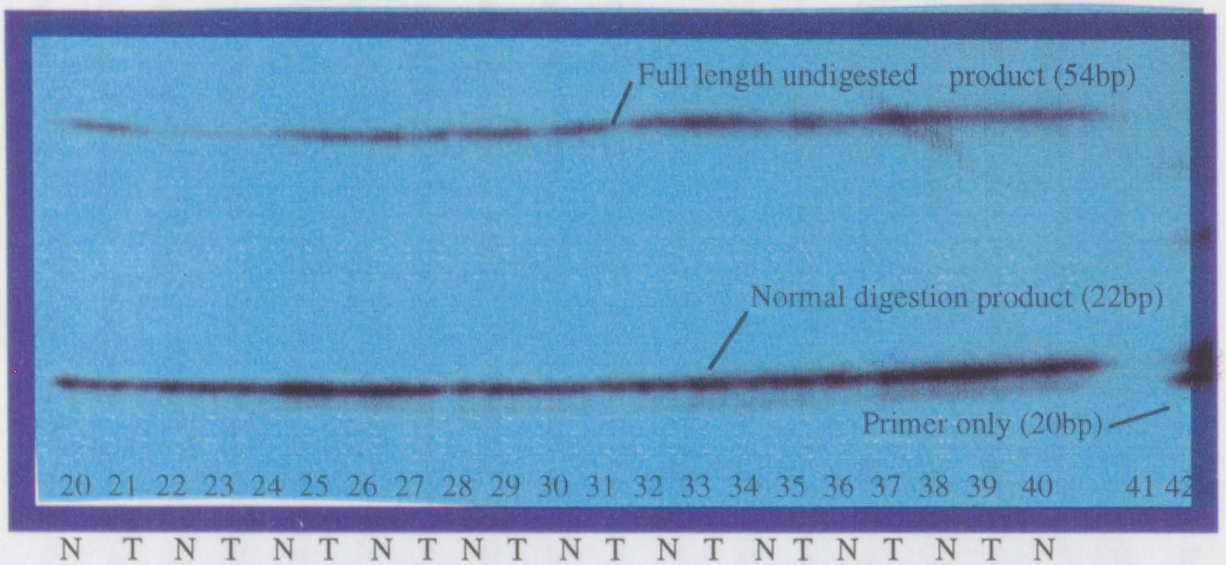


Figure 2.13. Autoradiograph of the GMPD- check analysis of a 54bp PCR product of the K-ras 2 gene. Amplification was performed on matched normal (N) and tumour (T) DNA from 20 sporadic colon cancer patients (not all shown) in a PCR reaction containing unlabelled primer C and γ^{32}P labelled primer A. Following amplification samples were subjected to GMPD analysis, digestion products were separated on a 10% denturing polyacrylamide gel and visualised by autoradiography. Lanes 20 to 40 contain matched N and T DNA digestion products from 10 patients. Full length undigested product (54bp) and normal digestion product (22bp) is present in all samples. Lane 41 contains ExoI treated labelled primer A (digested therefore not visible) and lane 42 contains undigested labelled primer (20bp).

22bp corresponding to normal digestion product is however, present in all samples (normal and tumour), this indicates that GGT to TGT or GGT to GTT codon 12 mutations were not detected in these samples.

GMPD-check analysis was performed on the 20 matched normal and tumour samples using primer set B and C in order to select for codon 13 mutations. PCR was performed with the primers B and C with the B primer labelled. Twenty Normal and tumour DNA samples were analysed as above. The possible mutations which may be detected using this primer combination include GGC to TGC and GGC to GTC of codon 13 (see fig 2.12) As is clear from the autoradiograph (fig 2.14) only labelled primer (24bp), normal digested product (28bp) and full length uncut DNA (55bp) is present in all samples (normal and tumour). No additional bands or loss of fragments were observed indicating the absence of G-T codon 13 mutations in these samples.

GMPD-check analysis of 20 matched normal and tumour samples from sporadic colon cancer samples was performed using the primer combination of A and C in the PCR amplification. In this case the lower C primer was labelled which facilitates the analysis of the complementary DNA strand. PCR products were digested with UDG glycosylase and cleaved with NaOH. The resulting fragments were analysed on a 10% denaturing polyacrylamide gel. The possible mutations which could be detected in this analysis include codon 12 GGT to AGT and GGT to GAT and codon 13 GGC to AGC and GGC to GAC. These mutations produce digestion products of 33bp, 32bp, 30bp and 29bp respectively. Figure 2.15 shows the autoradiograph of GMPD-check analysis in these samples, labelled primer C (27bp) and full length uncut product (54bp) are clearly visible in lanes 43 and 41 respectively. Lanes 20-40 contain normal and tumour DNA samples from the ten patients. All samples clearly display a 37 bp fragment which corresponds to the normal digestion product, additional bands are seen in samples 2, 10, 14, 40 and 44, the additional bands in samples 2 and 44 are not clearly visible in the autoradiograph image (fig 2.15) but are visible in the original autoradiograph. Samples 2 and 40 display bands of similar size, whereas samples 10, 14 and 44 exhibit

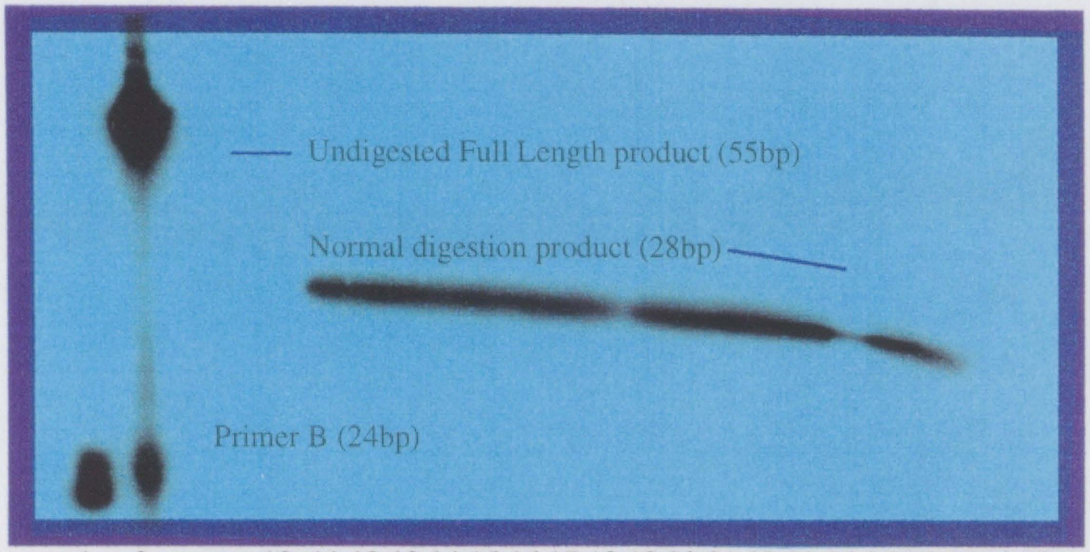
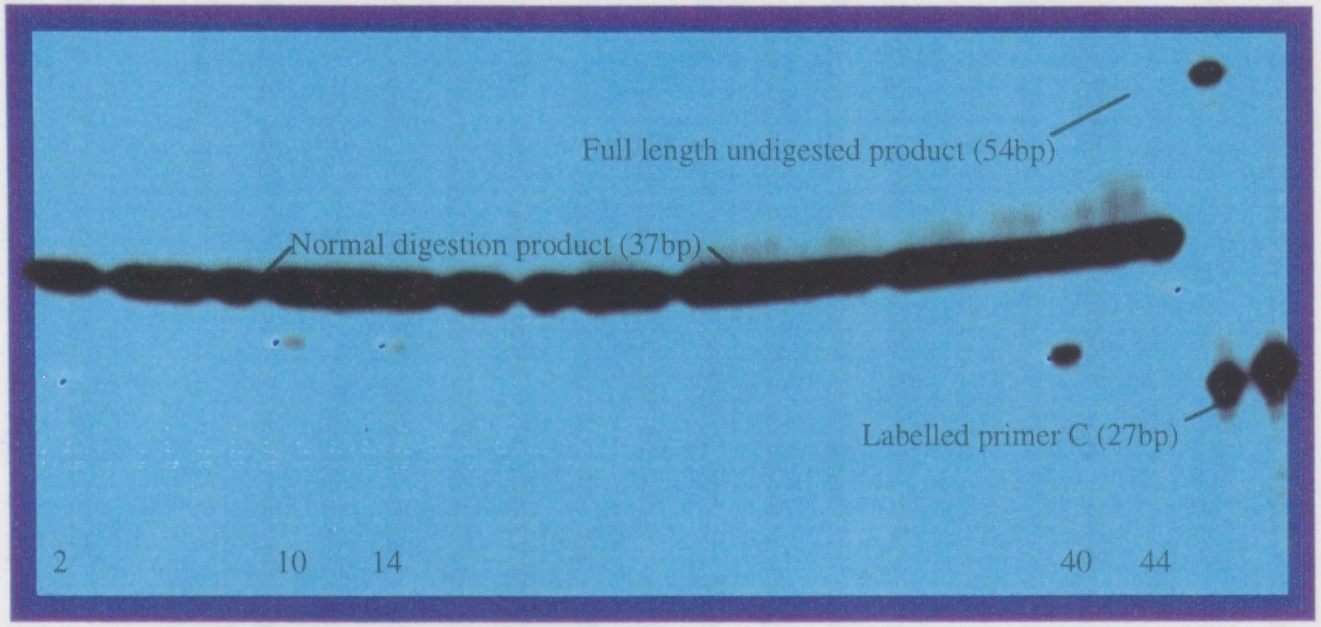


Figure 2.14. Autoradiograph of the GMPD-check analysis of a 55bp PCR product of the K-ras 2 gene. PCR amplification was performed on matched normal (N) and tumour (T) DNA samples from 20 sporadic colon cancer patients (not all shown). The 55bp product was amplified in a PCR reaction containing unlabelled primer C and end labelled primer B. Following amplification samples were UDG digested, NaOH treated and separated by electrophoresis in a 10% denaturing polyacrylamide gel. Lanes 10-29 contain digestion products from 10 patients. Normal digestion product is visible in all samples (28bp). Lane 1 contains labelled primer B only (24bp) and lane 3 contains undigested full length product (55bp).



	N	T	N	T	N	T	N	T	N	T	N	T	N	T	N	T	N	T	N	T	N	T	N	T	N	T	N	T	N	T	N	T		
lanes	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	48

Figure 2.15. Autoradiograph of the GMPD-check analysis of a 54bp PCR product of the K-ras 2 gene. PCR amplification was performed on matched normal (N) and tumour (T) DNA samples from 20 sporadic colon cancer patients (not all shown). The 55bp product was amplified in a PCR reaction containing unlabelled primer A and end-labelled primer C. Following amplification samples were UDG digested, NaOH treated and separated by electrophoresis in a 10% denaturing polyacrylamide gel. Lanes 12-45 contain digestion products from 17 patients. Normal digestion product is visible in all samples (37bp). Lane 48 contains labelled primer C only (27bp) and lane 46 contains undigested full length product (54bp). Additional digestion products are visible in the tumour DNA from patients 2, 10, 14, 40 and 44 (indicated by arrows). The additional bands are not present in the matched normal DNA from these patients.

larger additional fragments. The presence of the additional fragments in these samples indicates the presence of K-ras mutations in these patients. Furthermore the fragments are only present in the tumour DNA from the patients and not in the DNA extracted from the normal mucosa. The presence of the normal digestion product in all samples including those bearing mutations suggests the presence of some normal cells in the tumour tissue.

2.3.(IV) BACTERIAL CLONING OF THE K-RAS GENE PCR PRODUCTS

In order to analyse the mutant tumour DNA without interference from the normal DNA contaminant, PCR products were cloned and subsequently analysed using GMPD. PCR amplification was performed using the original nested primer set U and 2. These primers ensure that exon 1 is amplified from the K-ras 2 gene only and not the pseudogene. PCR products were ligated into the pTag cloning vector (R&D systems), which exploits the T overhangs produced by Taq DNA polymerase during amplification. pTag transformed colonies were selected based on chromogenic assay. Colonies were further analysed using the GMPD technology for the presence or absence of mutations. Cloning of the initial PCR products ensures that only one copy of the DNA is analysed- either a PCR product bearing the mutation or a normal product. All samples in which an additional mutant fragment was observed during GMPD analysis were cloned and re-analysed.

RE-ANALYSIS OF CLONED PCR PRODUCTS

Sixty colonies were selected from each of the five patient samples which displayed a mutation on previous analysis. Colony PCR was performed using the A and C primer combination with the latter primer labelled. Products were UDG digested, NaOH treated and separated on a 10% denaturing polyacrylamide gel.

Figures 2.16-2.20 display the autoradiographs of GMPD analysis of cloned fragments.

Additional (mutant) K-ras fragments were detected in 11/60 (18.3%) of colonies from sample 40, 2/60 (3.3%) of colonies from sample 44, 4/60 (6.6%) of colonies from sample 10, 2/60 (3.3%) from sample 2 and 7/60 (11.6%) of colonies from sample 14. These values provide a basic indication of the amount of contamination by normal DNA in the tumour samples. The number of colonies containing mutant samples is representative of the true tumour DNA and on average 14% of colonies analysed contained tumour DNA. Furthermore it should be noted that samples 2 and 44 exhibited mutant fragments in only 3.3% of colonies, this correlates with the intensity of the additional fragments observed on uncloned GMPD analyses of these samples (fig 2.15) In the previous GMPD scan analysis samples 2 and 44 display additional bands, however the signal intensity is relatively lower than that in samples 10, 14 and 40.

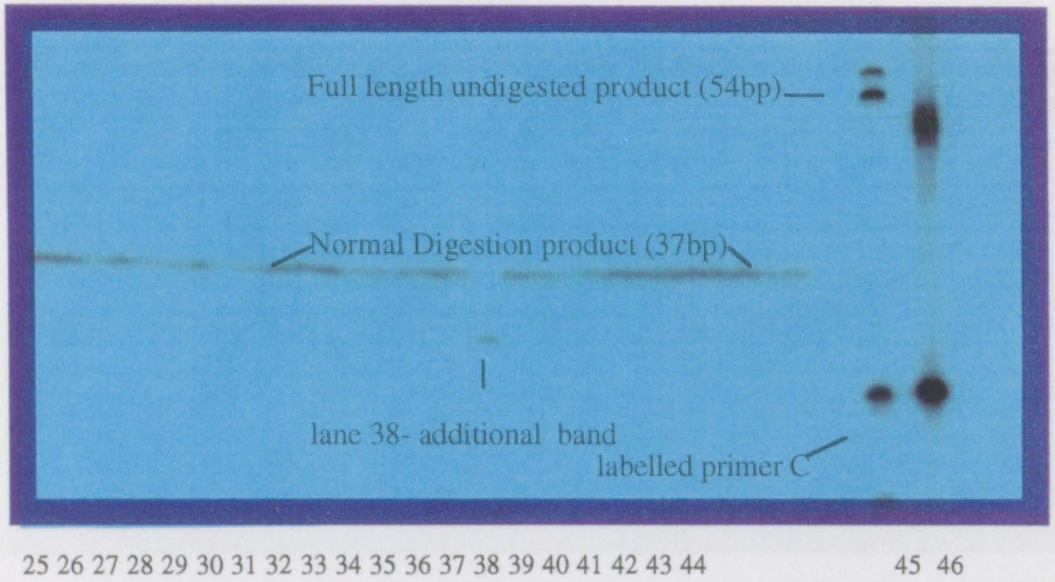


Figure 2.16. Autoradiograph of the GMPD-check analysis of a 54 bp segment of the K-ras 2 gene. PCR amplification of this fragment was performed on 60 *E.coli* colonies transformed with the pTag vector containing exon 1 of the K-ras 2 gene amplified from tumour DNA of patient 2. Colony PCR was performed in a reaction containing unlabelled primer A and end labelled primer C. Following amplification samples were UDG digested, NaOH treated and separated by electrophoresis in a 10% denaturing polyacrylamide gel. The normal GMPD-check profile was observed in 58 of 60 colonies analysed (not all shown). Lane 38 displays a single additional band however no band corresponding to the normal digestion product is observed indicating that a pure colony containing only mutant tumour samples was analysed. Lane 45 contains labelled primer C only and lane 46 contains undigested full length product (54bp) and primer C.

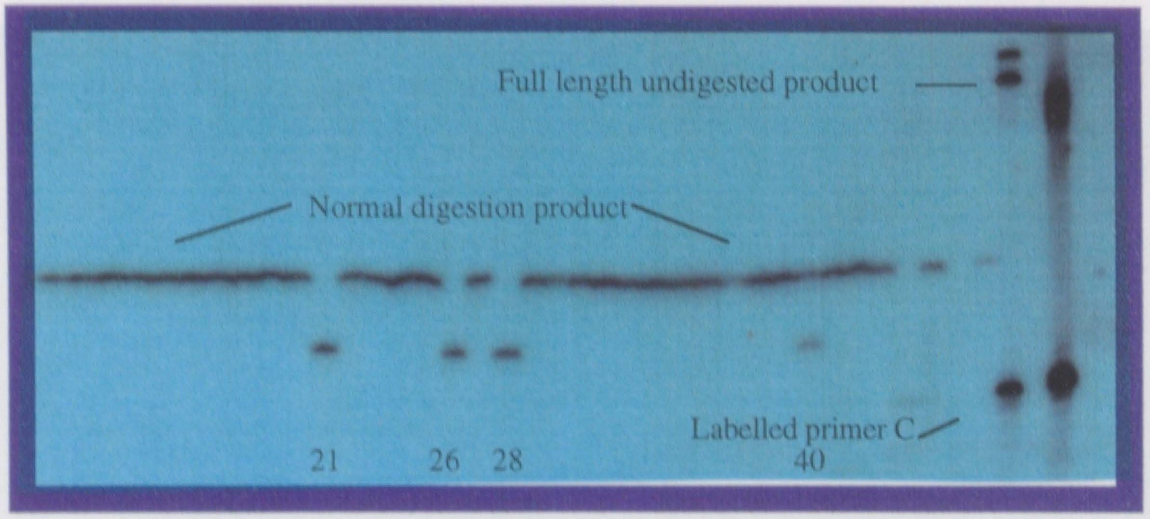
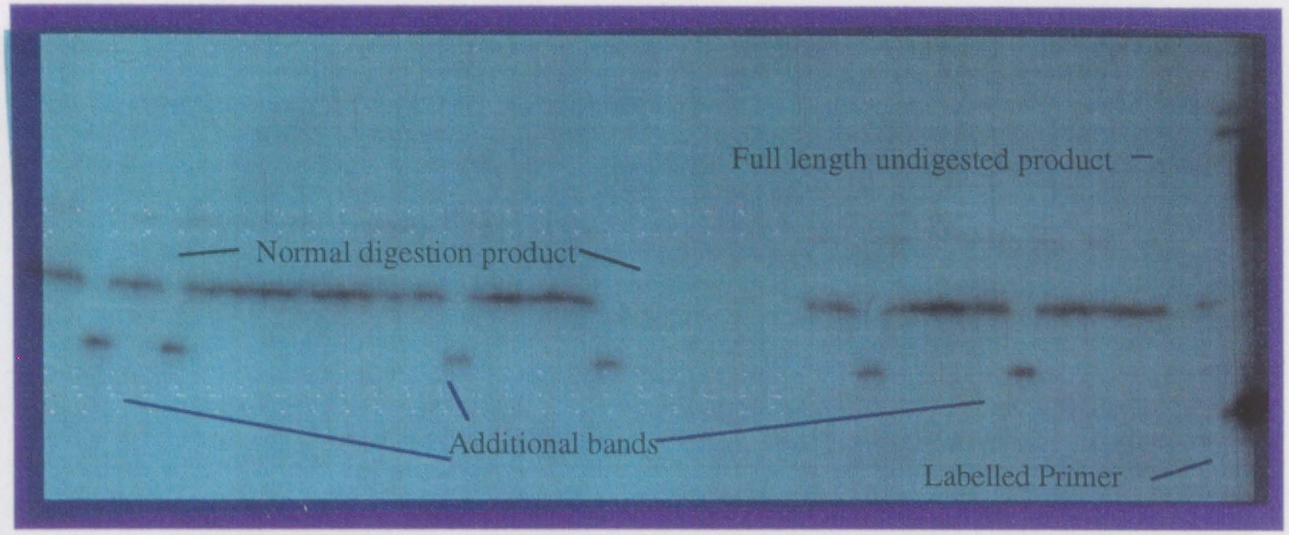
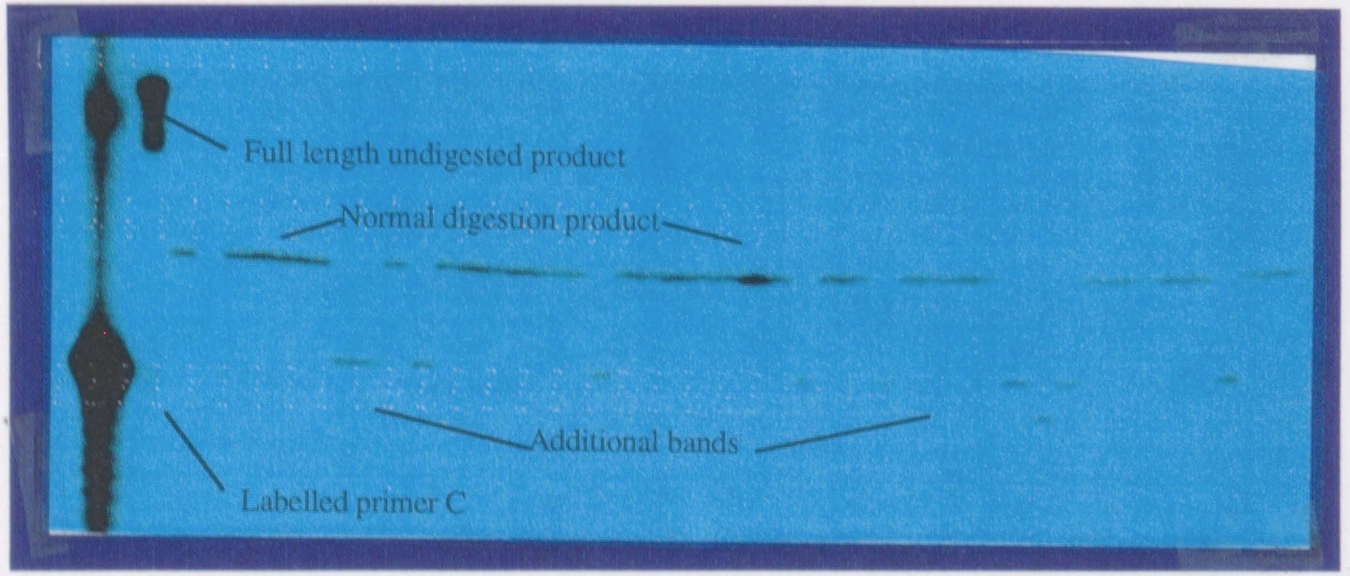


Figure 2.17. Autoradiograph of the GMPD-check analysis of a 54 bp segment of the K-ras 2 gene. PCR amplification of this fragment was performed on 60 *E.coli* colonies transformed with the pTag vector containing exon 1 of the K-ras 2 gene amplified from tumour DNA from patient 10. Colony PCR was performed in a reaction containing unlabelled primer A and end labelled primer C. Following amplification samples were UDG digested, NaOH treated and separated by electrophoresis in a 10% denaturing polyacrylamide gel. The normal GMPD-check profile was observed in 56 of 60 colonies analysed (not all shown). Lanes 21, 26, 28 and 40 display additional bands however no band corresponding to the normal digestion product is observed in these samples indicating that a pure colony containing only mutant tumour samples was analysed. Lane 48 contains labelled primer C only and undigested full length product (54bp).



lanes 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35

Figure 2.18. Autoradiograph of the GMPD-check analysis of a 54 bp segment of the K-ras 2 gene. PCR amplification of this fragment was performed on 60 *E.coli* colonies transformed with the pTag vector containing exon 1 of the K-ras 2 gene amplified from tumour DNA from patient 14. Colony PCR was performed in a reaction containing unlabelled primer A and end labelled primer C. Following amplification samples were UDG digested, NaOH treated and separated by electrophoresis in a 10% denaturing polyacrylamide gel. The normal GMPD-check profile was observed in 53 of 60 colonies analysed (not all shown). Additional bands were present in 7 samples however normal digestion product was absent from these samples indicating that a pure colony containing only mutant tumour samples was analysed. Lane 35 contains labelled primer C only and undigested full length product (54bp).



lanes 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35

Figure 2.19. Autoradiograph of the GMPD-check analysis of a 54 bp segment of the K-ras 2 gene. PCR amplification of this fragment was performed on 60 *E.coli* colonies transformed with the pTag vector containing exon 1 of the K-ras 2 gene amplified from tumour DNA from patient 40. Colony PCR was performed in a reaction containing unlabelled primer A and end labelled primer C. Following amplification samples were UDG digested, NaOH treated and separated by electrophoresis in a 10% denaturing polyacrylamide gel. The normal GMPD-check profile was observed in 49 of 60 colonies analysed (not all shown). Additional bands were present in 11 samples however normal digestion product was absent from these samples, indicating that a pure colony containing only mutant tumour samples was analysed. Lane 3 contains labelled primer C and undigested full length product (54bp).

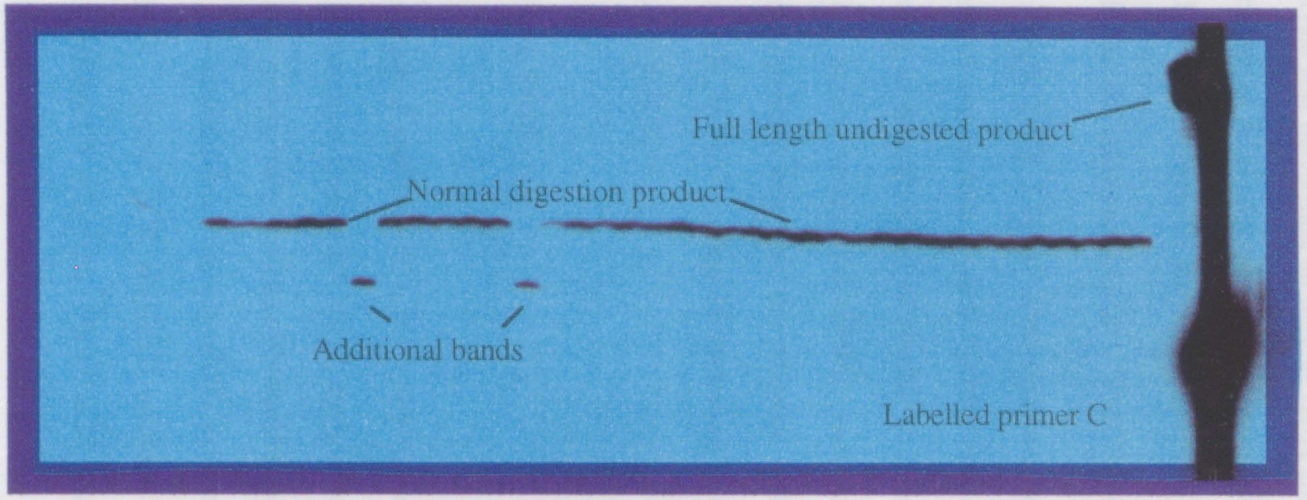


Figure 2.20. Autoradiograph of the GMPD-check analysis of a 54 bp segment of the K-ras 2 gene. PCR amplification of this fragment was performed on 60 *E.coli* colonies transformed with the pTag vector containing exon 1 of the K-ras 2 gene amplified from tumour DNA from patient 44. Colony PCR was performed in a reaction containing unlabelled primer A and end labelled primer C. Following amplification samples were UDG digested, NaOH treated and separated by electrophoresis in a 10% denaturing polyacrylamide gel. The normal GMPD-check profile was observed in 58 of 60 colonies analysed (not all shown). Additional bands were present in 2 samples (lanes 16 and 22) however normal digestion product was absent from these samples, indicating that a pure colony containing only mutant tumour samples was analysed. Lane 39 contains labelled primer C only and undigested full length product (54bp).

2.3.(V) SEQUENCING OF MUTATION BEARING COLONIES

In order to verify that the presence of an additional fragment in the GMPD analysis profile of a sample is due to a mutation, nucleotide sequencing was performed. Cycle sequencing analysis was performed on tumour DNA from sample 10 and sample 2. Analysis performed on an ALF-express semi-automated DNA sequencer identified the presence of a GGT to GAT codon 12 mutation in sample 10 (Fig 2.21a) and cycle sequencing performed with the incorporation of α^{33}P-dCTP identified a GGC to GAC codon 13 mutation in sample 2 (fig 2.21b). Two other samples (14 and 44) displayed additional mutant bands of similar size in their GMPD profiles as sample 10 and the additional band observed in sample 40 is of similar size to that of sample 2. The cloning and sequencing results verify that GMPD analysis of tumour DNA from sporadic colorectal cancer patients identified the presence of three GGT to GAT codon 12 mutations and two GGC to GAC codon 13 mutations in the twenty sporadic colon samples analysed (Figure 2.21).

2.3.(VI) SSCP ANALYSIS

The GMPD method of analysis did not identify mutations in fifteen out of the twenty sporadic tumour samples analysed. SSCP analysis was performed in order to determine if the technique had failed to detect any mutations which may be present in the samples. Fifteen samples which did not display additional bands in their GMPD profile were analysed by SSCP. Nested PCR amplified products were separated in a 6% polyacrylamide gel containing 5% glycerol. Figure 2.22 shows the autoradiograph of the SSCP analysis of K-ras exon 1 in 15 normal and tumour samples. SSCP analysis under these conditions failed to identify any aberrant banding patterns, this is consistent with the GMPD findings and suggests that the technique is accurate in its ability to detect mutations.

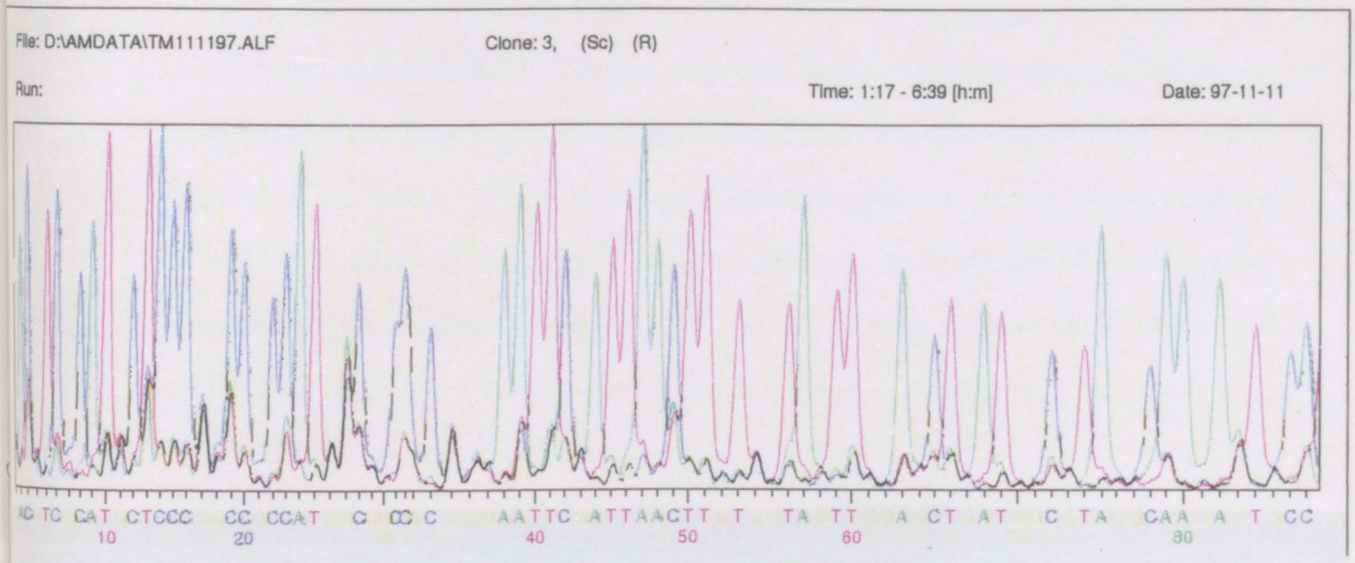


Figure 2.21a.

GGT GGC
|
A

Codon 12 13

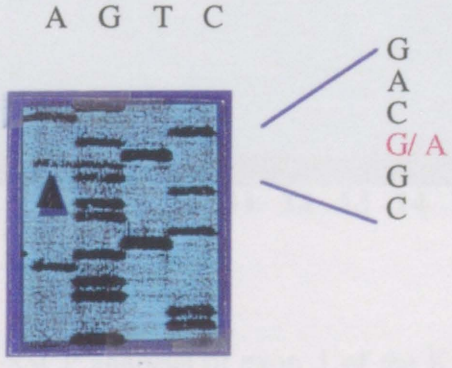


Figure 2.21b.

Figure 2.21 (a) Nucleotide sequence analysis of PCR products of exon 1 of the K-ras 2 gene, from tumour DNA from patient 10. Fluorescent sequence analysis was performed using cycle sequencing employing Cy5 labelled primer A and unlabelled primer C. The sequence analysis was performed by electrophoresis in an ALF-express fluorescent sequencer. A GTT to GAT mutation of codon 12 of exon 1 was identified in the tumour DNA sample from patient 10.

Figure 2.21 (b). Digital scanned image of the autoradiograph of the sequence analysis of a PCR product of exon 1 of the K-ras 2 gene from tumour DNA of patient 40. Cycle sequencing was performed in a reaction containing primers A and C and α^{33} labelled dideoxy terminators. Sequence products were separated in a 6% denaturing polyacrylamide gel and visualised by autoradiography. A GGC to GAC mutation of codon 13 of the K-ras 2 gene was detected in the tumour DNA of patient 40.

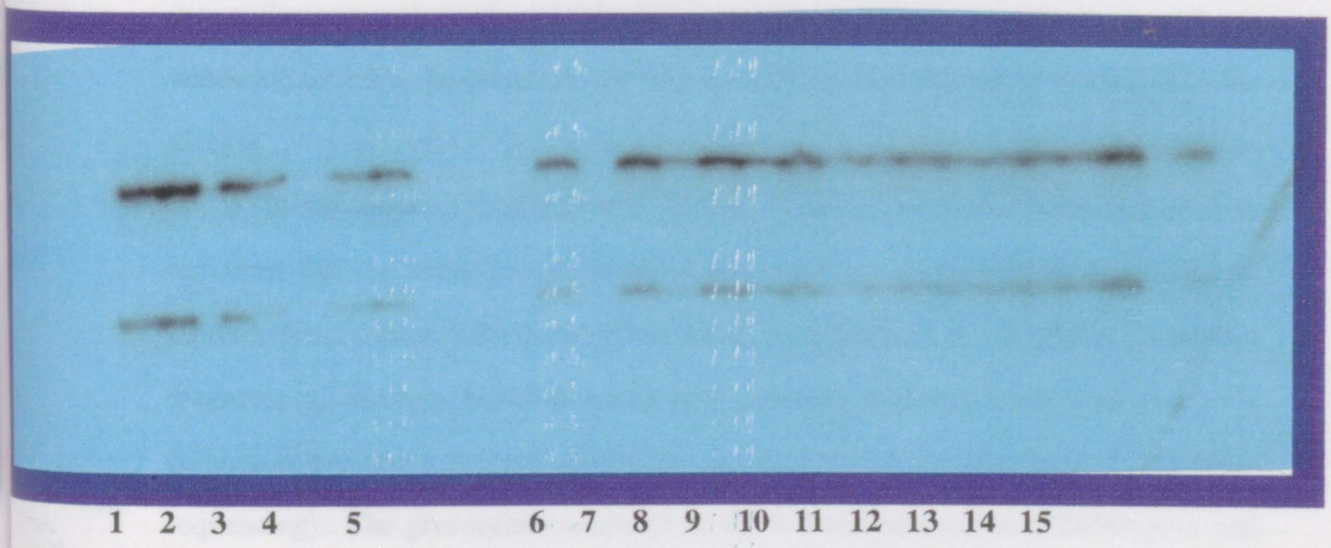


Figure 2.22. Autoradiograph of the SSCP analysis of exon 1 of the K-ras 2 gene in fifteen tumour DNA samples from sporadic colon cancer patients which did not display mutations on GMPD analysis. PCR amplification was performed on samples in a reaction containing primers U and 2, products were labelled by α^{32}PdCTP incorporation. Amplification products were separated on a 6% polyacrylamide gel containing 5% glycerol, and visualised by autoradiography.

2.4. DISCUSSION:

The ability to detect mutations in colorectal cancer associated genes has major implications for diagnosis, prognosis and treatment of colorectal cancer patients. More than 90% of patients with carcinomas confined to the mucosa can be successfully treated with surgical intervention, whereas less than 10% of patients with distant metastases are successfully treated (Rich, *et al.*, 1983). Mutations of K-ras are associated with early stages of development in colon carcinogenesis, with more than 50% of adenomas displaying K-ras defects (Vogelstein *et al.*, 1988). The ability to detect K-ras mutations in asymptomatic patients may have significant impact on widescale screening programmes and may help to significantly reduce death due to the disease.

In the routine identification of genetic mutations associated with disease it is important that the technique employed is efficient, easy to implement, sensitive and reproducible. Many different mutation detection techniques are employed in genetic diagnosis of diseases however many are laborious, time-consuming and require a combination of multiple steps (usually a mutation scanning reaction coupled with DNA sequencing). The glycosylase mediated mutation detection systems (GMPD scan and check) were employed in the analysis of the K-ras 2 gene in matched normal and tumour tissue DNA samples obtained from twenty Irish sporadic colon cancer patients. This analysis was carried out in order to develop and validate a GMPD based mutation detection protocol specifically designed for K-ras mutations and to assess the efficacy and suitability of this technique in the detection of mutation activated oncogenes in colorectal cancer samples.

2.4.(II) GMPD-SCAN

The initial mutation screening of exon 1 of the K-ras gene employed the GMPD technology to scan the entire exon for mutations. This method is less selective than the GMPD-check which selects a small target sequence of DNA. GMPD scan analysis was

carried out using nested PCR amplification products. The primary PCR reaction selectively amplified DNA from the functional gene and not for the K-ras 1 pseudogene. The oligonucleotide primers employed in the second stage analytical PCR were designed so that by labelling the forward (F) primer, mutations of GGT to TGT and to GTT of codon 12 could be detected and mutations of GGC to TGC and to GTC of codon 13 could be detected. The resulting GMPD profile was consistent with normal DNA which does not harbour mutations. It was concluded that GMPD scanning analysis of the twenty samples did not detect any mutations in an assay where the F primer is the labelled diagnostic primer.

A GMPD scan method where the reverse primer is the labelled diagnostic primer was designed to detect GGT to GAT and to AGT codon 12 mutations and GGC to GAC and to AGC codon 13 mutations. The resulting GMPD scan profile identified five samples with additional bands. These additional fragments were only identified in the tumour DNA samples and were not present in the DNA from normal mucosa DNA in the same patients. The results suggested that the additional bands observed in the tumour samples arose from mutations in these samples.

The GMPD -scan method of analysis was successful in identifying G-A codon 12 and 13 mutations in five colorectal cancer tumour samples. The method is highly sensitive to contamination by normal DNA which results in a weak signal for the mutant fragment when a large population of normal DNA is present in the sample. The ability of the technique to detect mutations in a mixed DNA population may prove useful in the detection of mutations in clinically "normal" mucosa which is at the early stage of neoplasia. The technique is highly efficient, does not require multiple manipulations and can be completed rapidly. The GMPD scanning technique is applicable to the large scale analysis of multiple samples where the exact region of the mutation is unknown or where a variety of mutations may exist in a cluster region. The technique has a distinct advantage over conventional scanning methods such as SSCP, DGGE or heteroduplex analysis in that it does not require subsequent sequence analysis to verify the presence of the mutation. The presence or absence of a band on the GMPD scan analysis profile

indicates that either a T residue has been inserted or lost from the sequence. By comparing the banding pattern obtained for a sample with the pattern of a normal sequence the position of the change can be determined and the mutation identified.

2.4 (III) GMPD-CHECK

The GMPD-check method was employed to determine the sensitivity of the GMPD scan technique. The GMPD-check method, unlike the scanning technique, targets the exact mutation position, furthermore only uracil residues are incorporated at the T positions during the amplification reaction, this results in a stronger signal as all fragments are digested at the first U position. A GMPD-check assay was designed to specifically detect mutations of codon 12 and 13 of the K-ras 2 gene. Using combinations of primers and various labelling options the mutations outlined in Figure 2.12 could be detected. No mutations were identified in reactions with either the A or the B primer labelled thus GTT to TGT or to GTT mutations of codon 12 and GGC to GTC or to TGC mutations of codon 13 were not detected. Comparison of the results for GMPD-check with the GMPD-scan method revealed that in both methods mutations of this type were undetected.

GMPD-check analysis of the samples in reaction with the C (lower) primer labelled identified five mutations, two of a higher molecular weight than the other three samples. The mutations which could be detected with the C diagnostic primer are outlined in table 2.12. Comparison of the GMPD scanning technique results with those obtained on GMPD- check analysis revealed that mutant fragments were observed in the GMPD profiles of the same samples using both techniques. The GMPD-check method provides a stronger signal than GMPD-scan as U residues are incorporated at all the T positions, thus if a mutation exists at the first T position, all fragments will be digested by UDG at this position.

In order to determine if the GMPD technology had failed to identify any mutant samples SSCP analysis was performed. SSCP analysis of exon 1 was performed on the fifteen samples in which a mutation had not been previously identified by GMPD.

No aberrant SSCP profiles were observed which is consistent with the results of the GMPD analyses.

Sequencing of cloned PCR products from the tumour samples of the five patients identified the higher molecular weight fragment as a GGT to GAT mutation of codon 12. Analysis of the lower molecular weight fragment identified the mutation as a GGC to GAC codon 13 change.

The combined GMPD technology identified 5 K-ras mutations in twenty Irish colon cancer patients (40%). This result is consistent with previous studies which report K-ras gene mutations in about 50% of colon tumours with 84% of these confined to codon 12 or 13 (Vogelstein *et al.*, 1988).

Two GGC to GAC codon 13 mutations and three GGT to GAT codon 12 mutations were identified in tumour samples only. Analysis of previously reported mutations reveals that the G-A mutations are the most common K-ras mutations and that codon 12 G-A mutations occur more frequently than any other K-ras mutations (Moerkerk *et al.*, 1994).

2.4 (IV) GMPD ANALYSIS OF HETEROGENEOUS DNA SAMPLES.

The presence of wild type DNA in tumour samples poses problems for mutation detection in cancer cells. A large amount of wild type DNA can often mask the mutant sequence with the result that the disease causing mutation may remain undetected. The ability to detect mutant DNA in a heterogeneous DNA sample depends on the sensitivity of the mutation detection technique employed. It has been demonstrated, in DNA mixing experiments, that direct sequencing is only successful when there is at least 10% mutant sequence in a mixed DNA sample, whereas SSCP is sensitive at detecting mutations in a sample containing 5% mutant sequence (Burmer *et al.*, 1989). The GMPD analysis of the K-ras gene in tumour samples from Irish colon cancer patients shows that the technique can detect mutations in samples where the mutant to normal allele is as low as 3.3%. Exon 1 of the K-ras 2 gene was amplified from tumour

samples and subsequently cloned and transformed into *E.coli* cells. Subsequent analysis of the resulting colonies identified only two mutant inserts in 60 colonies containing DNA from patients 44 and 2. This is representative of the proportion of mutant sequence in the original DNA samples. GMPD analysis of tumour DNA from these patients was successful in detecting the mutation in heterogeneous DNA obtained from tumour biopsies. Precise DNA mixing experiments may demonstrate that the technique is effective at mutation detection in samples containing even less than 3% mutant sequence.

Analysis of GMPD-scan and GMPD-check demonstrated that the latter technique is more successful in detecting mutations in a heterogeneous DNA sample. The GMPD-check technique involves the incorporation of U residues at every T position in a PCR product of a target region, thus even in mixed DNA sequences, digestion occurs at the first U position, therefore both tumour and normal DNA digestion products will be observed on gel analysis. The GMPD-scan technique was successful in identifying mutations in DNA samples with a low percentage of mutant sequence, however because a ratio of U:T residues is incorporated into the PCR products, digestion does not occur at the first U in every PCR product, thus the signal from the UDG digestion products is much weaker for the mutant sequence than it is with GMPD-check.

The ability of GMPD to detect mutations in heterogeneous DNA samples is particularly significant in the analysis of mutations which predispose to cancer. K-ras mutations have been detected in histologically normal mucosa in regions adjacent to carcinoma, suggesting the presence of a field of genetically abnormal tissue which may be predisposed to the formation of further neoplasms (Burmer *et al.*, 1989). The ability of GMPD to detect K-ras mutations in samples containing predominantly wild type DNA may facilitate the identification of potential tumour sites at a stage preceding histological evidence of neoplasia.

2.4.(V) ADVANTAGES OF GMPD-SCAN AND GMPD-CHECK OVER EXISTING MUTATION DETECTION TECHNOLOGIES.

Current technologies for the detection of known and unknown sequence changes are either cumbersome, difficult to optimise, poorly informative with respect to the location and nature of unknown mutations or inefficient. The GMPD-scan and GMPD check techniques offer a number of advantages over current technologies:
GMPD technology offers a high probability of detecting a mutation/polymorphism as it can detect all deletions, insertions and 83% of point mutations. With the inclusion of a second glycosylase specific for G-C mutations the technique will detect 100% of base changes.

The GMPD methodology is faster, and more cost effective than dideoxy or chemical DNA sequencing methods, the processes can be performed directly on amplified genomic DNA, require significantly fewer manipulations and utilise less hazardous reagents. Moreover the techniques are not sequence dependant and are not dependent on heteroduplex formation. Most significantly in terms of cancer mutation detection, unlike heteroduplex and mismatch cleavage methods, the technology allows the detection of known and unknown mutations in homozygous or heterozygous samples. Furthermore, the intensity of the digestion product signal is likely to reflect the proportion of mutant to normal DNA in the sample. The methodology is superior to SSCP, cleavage fragment length polymorphism (CFLP) or DGGE as it not only detects but also determines the sequence of the mutation. It is superior to the protein truncation test (PTT) since mutations including those leading to the formation of an in-frame stop codon are detectable and the position of the mutation is determined whereas in PTT it is not.

In conclusion, it has been demonstrated that the GMPD analysis protocol established provides a practical basis for a new and rapid approach for detecting the presence of mutated colorectal cancer cells. Furthermore, as K-ras mutations have been

detected in both pre-malignant tissues and carcinomas it may be employed as a method to detect surgically treatable tumours in the preneoplastic stages of tumour progression and may have a role in reducing deaths from the disease. Moreover, the type and position of the mutation may provide information on tumour stage and prognosis, with some reports suggesting that the metastatic potential of tumours is related to the nature of the genetic alteration. The glycosylase mediated mutation detection method can accurately detect all nucleotide base changes except G-C or C-G mutations. However the inclusion of a new glycosylase will overcome this problem. This additional glycosylase, coupled with the K-ras GMPD analysis protocols established in this investigation may provide an accurate simple and diagnostically applicable method for detecting K-ras gene mutations.

In addition to K-ras gene mutations other mutant genes are involved in the tumorigenic pathway in colon cancer. The GMPD method could be manipulated to scan for and detect mutations in other colorectal cancer genes involved in the tumorigenic pathway. The method also has potential application in the detection of mutations in a non-invasive manner via the analysis of stool samples and colonic washings.

Further research is required to establish the exact association between K-ras mutation type with tumour stage and potential, however if a strong link is defined the GMPD analyses methods may prove an important tool in the diagnosis and prognosis of disease and may play a significant role in reducing deaths due to colorectal carcinoma.

2.5. BIBLIOGRAPHY

Baeg. G.H., Matsumine, A., Kuroda, T., Bhattacharjee, R.N., Miyashiro, I., Toyoshima, K., Akiyama, T. *The tumour suppressor gene product APC blocks cell cycle progression from Go-G1 to S phase.* (1995). EMBO J. 14, 5618-5625.

Barbacid, M. (1987). *Ras Genes.* Annu Rev Biochem. 56, 779-827.

Bodmer, W.F., Bailey, C.J., Bodmer, J., Bussey, H.J., Ellis, A., Gorman, P., Lucibello, F.C., Murday, V.A., Rider, S.H., Scambler, P.et al,. (1987). *Localization of the gene for familial adenomatous polyposis on chromosome 5.* Nature. 328, 614-6.

Bos, J.L., Fearon, E.R., Hamilton, S.R., Verlaan-de Vries, M., van Boom, J.H., van der Eb, A.J., Vogelstein, B. (1987). *Prevalence of ras Gene Mutations in Human Cancers.* Nature 327, 293-297.

Burmer, G.C and Loeb, L. (1989). *Mutations in the KRAS2 oncogene during progressive stages of human colon carcinoma.* Proc. Natl. Acad. Sci. USA. 88, 2403-2407.

Chang, E.H., Gonda, M.A., Ellis, R.W., Scolnick, E.M., Lowy, D.R. (1982). *Human genome contains four genes homologous to transforming genes of Harvey and Kirsten murine sarcoma viruses.* Proc Natl Acad Sci USA. 79,4848-4852.

Cooper, D. L., Lahue, R.S and Modrich, P. (1982). *Methyl-directed Mismatch repair is Bidirectional.* J Biol Chem. 268, 11823-11829.

de la Chapelle, A and Peltomaki, P. (1995). *Genetics of Hereditary Colon Cancer.* Annu. Rev. Genet. 29, 329-348.

Fearon, E.R. and Vogelstein, B. (1990). *A Genetic Model for Colorectal Tumorigenesis.* Cell, 61, 759-767.

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

Forrester, K., Almoguera, C., Han, K., Grizzle, W.E., Perucho, M. (1987). *Detection of High Incidence of K-ras Oncogenes During Human Colon Tumorigenesis.* Nature, 327, 298-303.

Kamb, A. (1995). *Cell-cycle regulators and cancer.* Trends Genet. 11, 136-40.

Keino-Masu, K., Masu, M., Hinck, L., Leonardo, E.D., Chan, S.S., Culotti, J.G., Tessier-Lavigne, M. (1996). *Deleted in Colorectal Cancer (DCC) encodes a netrin receptor.* Cell. 87,175-85.

Knudson, A.G. (1985). *Hereditary Cancer, Oncogenes, and Antioncogenes.* Cancer Res. 45, 1437-1443.

Land, H., Parada, L.F., Weinberg, R.A. (1983). *Cellular oncogenes and multistep carcinogenesis.* Science. 222, 771-8.

Leppert, M., Dobbs, M., Scambler, P., O'Connell, P., Nakamura, Y., Stauffer, D., Woodward, S., Burt, R., Hughes, J., Gardner, E et al., (1987). *The gene for familial polyposis coli maps to the long arm of chromosome 5.* Science. 238, 1411-1413.

Lindahl, T., Ljungquist, S., Siegert, W., Nyberg, B., Sperens, B. (1977). *DNA N-glycosidases: properties of uracil-DNA glycosidase from Escherichia coli.* J Biol Chem. 252, 3286-3294.

Lindahl, T., Nyberg, B. (1972). *Rate of depurination of native deoxyribonucleic acid.* Biochemistry. 11, 3610-3618.

Loeb, L. (1991). *Mutator Phenotype May Be Required for Multistage Carcinogenesis.* Cancer Res. 51, 3075-3079.

Loeb, L.A. (1994). *Microsatellite Instability: Marker of a Mutator Phenotype in Cancer.* Cancer Res. 54, 5059-5063.

Lynch, H.T., Smyrk, T.C., Watson, P., Lanspa, S.J., Lynch, J.F., Lynch, P.M., Cavalieri, J.R., Boland, C.R. (1993). *Genetics, Natural History, Tumour Spectrum, and Pathology of hereditary Nonpolyposis Colorectal Cancer: An Updated Review.* Gasteroenterology, 104, 1535-1549.

McGrath, J.P., Capon, D.J., Smith, D.H., Chen, E.Y., Seeburg, P.H., Goeddel, D.V., Leninson, A.D. (1983). *Structure and organisation of the human Ki-ras proto-oncogene and a related processed pseudogene.* Nature, 304, 501-506.

Mehlen, P., Rabizadeh, S., Snipas, S.J., Assa-Munt, N., Salvesen, G.S., Bredesen, D.E. (1998). *The DCC gene product induces apoptosis by a mechanism requiring receptor proteolysis.* Nature. 395, 801-804.

Moerkerk, P., Arends, J.W., van driel, M., de Bruine, A., de Goeij, A., ten Kate, J. (1994). *Type and Number of ki-ras Point Mutations relate to Stage of Human Colorectal Cancer.* Cancer Res, 54, 3376-3378.

Muller, R. *Transcriptional regulation during the mammalian cell cycle.* (1995). Trends Genet. 5, 173-178.

Nowell, P.C. (1976). *The clonal evolution of tumor cell populations.* Science.194, 23-28.

O'Sullivan, M.J., Mulcahy, T.M., Cambell, J., O'Suilleabhain, C.B., Kirwan, W.O., Doyle, C.T., McCarthy, T.V. (1989). *Detection of five novel germline mutations of the APC gene in Irish familial adenomatous polyposis families.* Hum Mutat. S 1(6), S251-S253.

Orita, M., Iwahana, H., Kanazawa, H., Hayashi, K., Sekiya, T. (1989(a)). *Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms.* Proc Natl Acad Sci USA 6(8), 2766-2770.

Orita, M., Suzuki, Y., Sekiya, T., Hayashi, K. (1989(b)). *Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction.* Genomics 5 (4), 874-879.

Peifer. M. (1993). *Cancer , catenins and cuticle pattern: a complex connection.* Science. 262, 1731- 1743.

Pounder , R.E, Allison, M.C., Dhillon, A.P. (1989). *A Color Atlas of the Digestive System.* Wolfe Medical Publications Ltd. Pp 136-151.

Price, A., Lindahl, T. (1991). *Enzymatic release of 5'-terminal deoxyribose phosphate residues from damaged DNA in human cells.* Biochemistry. 30, 8631-8637.

Rich, T., Gunderson, L.L., Lew, R., Galdibini, J.J., Cohen, A.M., Donaldson, G. (1983). *Patterns of recurrence of rectal cancer after potentially curative surgery.* Cancer. 52, 1317-1329.

Sambrook, J., Fritsch,E.F., Maniatis, T.(1989). *Molecular Cloning a Laboratory Manual.* 2nd. Edition. Cold Spring Harbour Laboratory Press.

Shivapurkar, N., Huang, L., Ruggeri, B., Swalsky, P.A., Bakker, A., Finkelstein, S., Frost, A, Silverberg, S. (1997). *K-ras and p53 mutations in aberrant crypt foci and colonic tumours from colon cancer patients.* Cancer Lett. 115, 39-46.

Smith, K.J., Johnson, K.A., Bryan, T.M., Hill, D.E., Markowitz, S., Willson, J.K., Paraskeva, C., Petersen, G.M., Hamilton, S.R. (1993). *The APC gene product in normal and tumor cells.* Proc Natl Acad Sci USA.90(7), 2846-2850.

Spandidos, R.A., Glarakis, I.S., Kotinas, A., Ergazaki, M., Kiaris, H. (1995). Ras Oncogene activation in Benign and Malignant Colorectal Tumours. Tumori 81(S) 7-11.

Vaughan, P and Mccarthy, T. (1998a). *A novel process for mutation detection using uracil DNA glycosylase.* Nucleic Acids Res. 26, 810-815.

Vaughan, P., McCarthy, T. *Glycosylase mediated polymorphism detection (GMPD), a novel process for genetic analysis.* Genetic Analysis: Biomolecular Engineering. *In press.*

Vogelstein, B., fearon, E., Hamilton, S.R., Kern, S.E., Preisinger, A.C., Leppert, M., Nakamura, Y., White, R., Smits, A.M.M., Bos, J.L. (1988). *Genetic Alterations During Colorectal Tumour Development.* N Engl J Med. 319, 525-532.

Yonish-Rouach, E., Resnitzky, D., Lotem, J., Sachs, L., Kimchi, A., Oren, M. (1991). *Wild-type p53 induces apoptosis of myeloid leukaemic cells that is inhibited by interleukin-6.* Nature. 352, 345-347.

Zhu, D., Keohavong, P., Fikelstein, S.D., Swalsky, P., Bakker, A., Weissfeld, J., Srivastava, S, and Whiteside, T.L. (1997). *K-ras Gene Mutations in Normal Colorectal Tissues from K-ras Mutation Positive Colorectal Cancer Patients.* Cancer Res, 57, 2485-2492.

CHAPTER 3

ANALYSIS OF INSTABILITIES AT REPEAT DNA SEQUENCES IN IRISH COLON CANCER PATIENTS

3.1. INTRODUCTION

3.1.(I). MICROSATELLITE INSTABILITY AND COLON CANCER.

Instability in short tandem repeat sequences (microsatellites) has been identified in a variety of different tumours including gastric cancer (Han *et al.*, 1993), endometrial carcinoma (Risinger *et al.*, 1993), Ovarian cancer (Orth *et al.*, 1994), and colon cancers including HNPCC (Peltomaki *et al*, 1993) and sporadic colorectal cancers (Lui *et al.*, 1995).

Microsatellites are regions of highly repetitive DNA sequences ((CA)n,(GT)n, and polyA tracts) found interspersed randomly throughout the genome. There are between 50,000-100,000 (CAn) repeats within the human genome, the majority are located within introns or between genes. The length of many of these repeats vary naturally from one individual to another and have proved very useful as a tool for following genetic transmission and in localisation of disease causing regions within a gene (Weber *et al.*, 1989).

Slipped strand mispairing (SSM) is widely believed to be the mechanism by which alterations in the sizes of repeats occur and plays a vital role in the evolution of repetitive DNA sequences by generating large numbers of short frameshift mutations within these regions. The tandem nature of simple repeat sequences is thought to render them prone to slipped strand mispairing and hence insertion or deletion mutagenesis during replication. The process of slipped strand mispairing involves mispairing of two strands of duplex DNA within a tandem repeat area, the mispaired duplexes may serve as primers for subsequent DNA synthesis, thus duplication occurs (Kunkel, 1993).

Examination of the mutation frequency within a Poly CA sequence inserted into bacteriophage M13 in *E. coli*, showed an increase in frameshift mutation frequency of greater than 1% in a 40 base pair tract, when compared to random non-repeat DNA. This change in frequency increases more than 13 fold in cells deficient in methyl-

directed mismatch repair (MutS$^-$, MutL$^-$ hosts) suggesting that mismatch repair plays a vital role in correcting frameshift mutations (Levinson and Gutman, 1987).

Microsatellite instability (MI) is observed as an increase or decrease in the electrophoretic mobility of a repeat sequence of a tumour cell which is not consistent with the similar sequence in a normal cell (Figure 3.1). This instability is believed to be a manifestation of DNA repair defects in the cell and is commonly known as the replication error phenomenon (RER) (Parsons *et al.*, 1993; Loeb, 1994).

The RER phenotype has been detected in more than 70% of HNPCC tumours (Aaltonen *et al.*, 1993) and is associated with mutations in the mismatch repair genes hMSH2, hMLH1, hPMS1 and hPMS2 (Leach *et al.*, 1993; Bronner *et al.*, 1994; Papadopoulos *et al.*, 1994). In HNPCC, mutations in the mismatch repair (MMR) genes are more common in families which satisfy the strict Amsterdam selection criteria, (true HNPCC) (Vasen *et al.*, 1991; Wijnen, *et al.*, 1997), concordantly MI is observed more frequently in typical HNPCC pedigrees than in pedigrees which fail to satisfy all the criteria. The analysis of 16 colorectal tumours identified MI in 3 of 11 atypical HNPCC pedigrees compared to 4 of 6 HNPCC pedigrees which satisfy all the Amsterdam criteria (Wijnen *et al.*, 1997).

The association of MI with mismatch repair gene defects in HNPCC led to the proposal of the 'mutator phenotype ' hypothesis (Loeb *et al.*, 1991). This hypothesis suggests that some cells in cancer patients undergo an early mutation event (mismatch repair gene defect) which renders them susceptible to the accumulation of multiple mutations throughout the genome. If these mutations occur in essential regulatory genes such as tumour suppressor genes or oncogenes the cells may become tumorigenic. A mechanism for microsatellite instability involves replication induced errors in repeat DNA sequences as a result of MMR defects. Replication induced errors occur frequently in repeat sequences however the normal mismatch repair system detects and corrects these mismatched nucleotides (Kunkel, 1993). If these mismatched nucleotides are not corrected (due to defective mismatch repair), they become fixed into the genome by subsequent rounds of replication (Kolodner *et al.*, 1994). Replication errors by

DNA polymerase are most likely to occur at microsatellite sequences due to their repeat nature and are the most obvious manifestation of the RER+ phenotype (Schlotterer *et al.*, 1992; Loeb, 1994).

Microsatellite Instability has also been detected in sporadic colorectal cancers, however the incidence is much lower (15%) than in HNPCC tumours (70%) (Aaltonen *et al.*, 1993; Ionov *et al.*, 1993; Thibodeau *et al.*, 1993). The RER phenomenon in sporadic colon cancer tumours, as in HNPCC is associated with mutations in mismatch repair genes (Lui *et al.*, 1995; Borresen, 1995). A notable feature of RER + sporadic colorectal tumours is that they display HNPCC like characteristics including localisation to the proximal colon, diploid DNA content and a more favourable prognosis than RER- tumours (Aaltonen *et al.*, 1993; Ionov *et al.*, 1993; Lothe *et al.*, 1993), this suggests that a subset of sporadic colorectal tumours may develop through a tumorigenic pathway similar to that of HNPCC tumours. Instability in DNA has also been observed in mono, tri and tetra repeat sequences (Parsons *et al.*, 1993; Loeb, 1994).

Genetic models to explain the pathological differences between RER+ and RER- tumours have been developed. The most plausible hypothesis proposes that two pathways for tumorigenesis exist, one based on the RER+, MMR defect route and the other is the more classical tumour suppressor, oncogene activation pathway. The first model proposes that mutations at repeat sequences remain unrepaired due to defects in the mismatch repair genes, if the increased mutation rate results in mutations of sequences in essential regulatory genes such as the transforming growth factor β gene receptor, a selective growth advantage for the cell results in rapid tumour advancement. The second pathway is not characterised by MI as it is believed that mismatch repair gene defects may not have a major causative role as in RER+ tumours. This model is based on the classical K-ras, APC, p53 tumorigenic pathway in which the accumulation of mutations in a number of specific genes is required for tumour development (Fearon *et al.*, 1990).

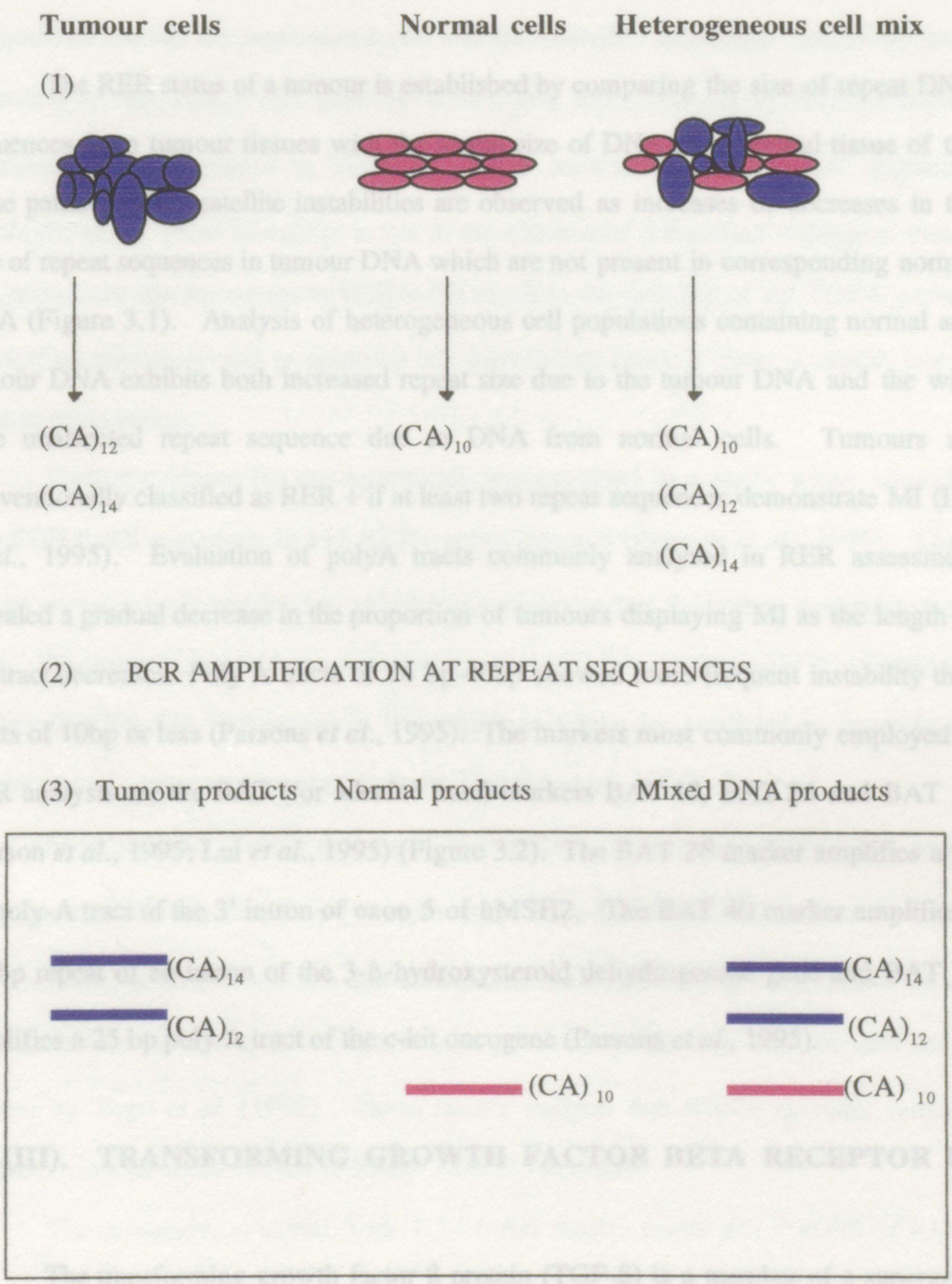


Figure 3.1. Schematic representation of the procedure for the analysis of the replication error status (RER) of normal, tumour and heterogeneous cell samples. (1) DNA isolation from cell population. (2) PCR amplification of repeat region of DNA. (3) Electrophorectic analysis of PCR products: products are radioactively labelled and separated on a 6% denaturing polyacrylamide gel.

3.1.(II). ANALYSIS OF MICROSATELLITE INSTABILITY

The RER status of a tumour is established by comparing the size of repeat DNA sequences from tumour tissues with the repeat size of DNA from normal tissue of the same patient. Microsatellite instabilities are observed as increases or decreases in the size of repeat sequences in tumour DNA which are not present in corresponding normal DNA (Figure 3.1). Analysis of heterogeneous cell populations containing normal and tumour DNA exhibits both increased repeat size due to the tumour DNA and the wild type unaffected repeat sequence due to DNA from normal cells. Tumours are conventionally classified as RER + if at least two repeat sequences demonstrate MI (Lui *et al.*, 1995). Evaluation of polyA tracts commonly analysed in RER assessment revealed a gradual decrease in the proportion of tumours displaying MI as the length of the tract decreases. Poly A tracts of 26 bp-40bp showed more frequent instability than tracts of 10bp or less (Parsons *et al.*, 1995). The markers most commonly employed in RER analysis are the BAT (for adenine tract) markers BAT 40, BAT 26 and BAT 25 (Parson *et al.*, 1995; Lui *et al.*, 1995) (Figure 3.2). The BAT 26 marker amplifies a 26 bp poly-A tract of the 3' intron of exon 5 of hMSH2. The BAT 40 marker amplifies a 40 bp repeat of an intron of the 3-ß-hydroxysteroid dehydrogenase gene and BAT 25 amplifies a 25 bp poly A tract of the c-kit oncogene (Parsons *et al.*, 1995).

3.1.(III). TRANSFORMING GROWTH FACTOR BETA RECEPTOR II.

The transforming growth factor ß protein (TGF-ß) is a member of a supergene family of factors which regulate cellular growth. TGF-ß inhibits epithelial cell growth and as such it is believed that ablation of the normal activity of this protein may have a role in tumorigenesis (Roberts *et al.*, 1990).

The activity of TGF-ß results due to it's interaction with two receptors- type I (RI) and type II (RII) which function as a heteromeric complex in the cell (Chen RH *et al.*, 1993). In an attempt to determine whether the inactivation of TGF-ß receptors is a

mechanism by which TGF-ß exerts its tumorigenic activity Markowitz *et al.* (1995) analysed RI and RII for expression levels and microsatellite instability. Mutations were identified in the TGF-ß-RII gene in RER+ cell lines but not in RER- samples. Furthermore the mutations in TGF-ß-RII were confined to small repeat sequences within the gene. These mutations result in the absence of cell surface receptors, thus it was postulated that mutations in TGF-ß-RII result in the inability of the TGF-ß protein to exert its regulatory role in epithelial cell growth and hence it plays a central role in tumour progression.

Further evidence for this hypothesis was provided in a study which identified 100 TGF-ß-RII mutations in 111 RER+ colon tumours (Parsons *et al.*, 1995). These mutations were again confined to polyAdenine tracts within the coding sequence of the gene. The most common frameshift mutation occurred in a 10bp poly-A sequence at nucleotides 709-718. Mutations in this repeat sequence are predicted to result in the production of a truncated receptor protein (Parsons *et al.*, 1995).

In an attempt to clarify the role of TGF-ß-RII mutations in sporadic colon cancer Akiyama *et al.* (1996) identified 7 samples with TGF-ß-RII mutations, these mutations were confined to RER+ tumours and all tumours were localised to the proximal colon. The association of TGF-ß-RII mutations with proximal localisation was corroborated in studies by Togo *et al.* (1996). These results suggest that RER+ sporadic tumours progress via a pathway similar to RER+ HNPCC tumours.

The cumulative evidence from TGF-ß-RII studies reveal that 70-90% of RER+ tumours (sporadic or HNPCC) harbour TGF-ß-RII mutations. TGF-ß-RII mutations in RER- tumours are very rare and are most likely chance somatic mutations rather than effects of mismatch repair defects within the cell. A common feature of TGF-ß-RII mutations is that tumours are located for the most part in the proximal colon and prognosis is more favourable in proximal sporadic tumours, as in HNPCC (Akiyama *et al.*, 1996). It has been proposed that due to the nature of the intragenic repeat sequences

of TGF-ß-RII, the gene is a target for mismatch repair defects and is therefore associated with an RER+ phenotype. The loss of tumour suppressor activity of TGF-ß (through the inactivation of RII) provides a selective growth advantage to the cells thus advancing the tumorigenic pathway (Markowitz *et al.*, 1995; Akiyama *et al.*, 1996). The association of TGF-ß-RII mutations with MI and proximal colon location further corroborates the hypothesis that two tumourigenesis pathways exist- the proximal tumour pathway involves early mismatch repair gene defects which result in an RER+ phenotype and a selective growth advantage due to TGF-ß-RII defects. The second is based on the Fearon and Vogelstein model of tumorigenesis which involves the loss of tumour suppressor activity coupled to the mutational activation of oncogenes through APC, K-ras and p53 gene defects. The latter pathway is not associated with the RER+ phenotype and shows no increased association with the proximal colon (Togo *et al* 1996; Akiyama *et al.*, 1996; Tannergard *et al.*, 1997).

3. 1. (IV). OBJECTIVES

In order to investigate the relationship between instability at repeat DNA sequences and TGF-ß-RII mutations in the Irish population, 38 Irish colorectal cancer patients were analysed for alterations in the 10bp polyAtract at nucleotides 708-719 of TGF-ß-RII and for instability at two poly-A tract repeat loci. The 38 samples included 18 Irish HNPCC patients and 20 sporadic colorectal cancer patients. Due to the fact that the HNPCC patients had been previously screened in this study for mutations in the hMSH2 and hMLH1 mismatch repair genes (chapter one) and the sporadic colon tumours had been analysed for K-ras mutations (chapter two), it was envisaged that correlation of the results of the various DNA analyses (MMR- HNPCC genes, K-ras, MI and TGF-ß-RII) might provide better insight into the tumourigenesis pathways in the Irish samples analysed.

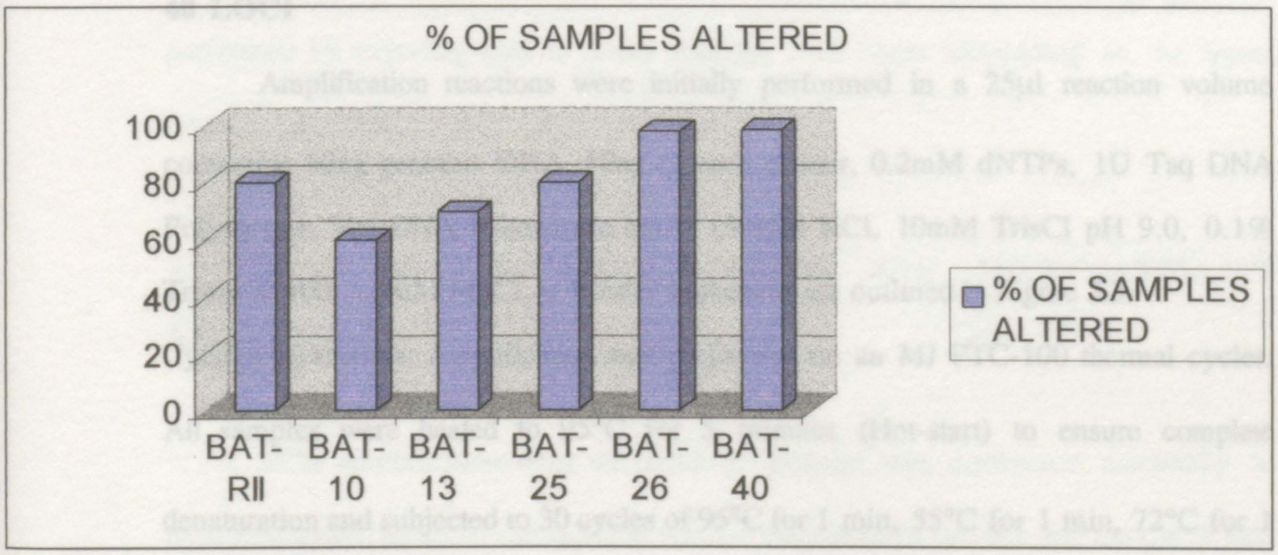


Figure 3.2. Graph of the frequency of alterations in polyA tracts of different lengths. BAT 10, 13, 25, 26 and 40 amplify polyA repeats of 10, 13, 25, 26 and 40 base pairs respectively. BAT RII amplifies a 10bp intragenic polyA tract of the TGF-β RII gene. Adapted from Parsons et al., 1995.

3.2 MATERIALS AND METHODS

3.2 (I). DNA ISOLATION:

DNA samples used in the analysis of the BAT-IIR, BAT-26 and BAT-40 loci had been previously isolated for the analysis of the k-Ras gene and the HNPCC mismatch repair genes (chapter 1 and 2) and were isolated as outlined in materials and methods sections of chapter 1 and 2.

3.2.(II) PCR AMPLIFICATION AT THE BAT-IIR, BAT-26 AND BAT-40 LOCI

Amplification reactions were initially performed in a 25µl reaction volume containing 60ng genomic DNA, 50ng of each primer, 0.2mM dNTPs, 1U Taq DNA Polymerase, Taq DNA Polymerase buffer (50mM KCl, 10mM TrisCl pH 9.0, 0.1% Triton X-100, 1.5mM $MgCl_2$). Primer sequences are outlined in Figure 3.3.
Cycling parameters: Amplification was performed on an MJ PTC-100 thermal cycler. All samples were heated to 95°C for 5 minutes (Hot-start) to ensure complete denaturation and subjected to 30 cycles of 95°C for 1 min, 55°C for 1 min, 72°C for 1 min followed by a final extension step of 72°C for 10 minutes. PCR optimisation was performed on primer pairs which did not give good PCR yield under these conditions. Optimisation involved performing the PCR reactions according to the above conditions however, either the concentration of $MgCl_2$ or the annealing temperature was altered.

3.2 (III) BAT-IIR AMPLIFICATION FOR REPLICATION ERROR ANALYSIS:

PCR amplification with the BAT-IIR primers was performed essentially as described above, however the reaction volume was reduced to 10µl and the dNTP mix

concentration was reduced to 0.02mM. In addition 2μCi of [α^{32}P]-dCTP (3000Ci/mmol) was included in the reaction. The optimum cycling parameters included 30 cycles of 95°C for 1 minute, 50°C for 30 seconds, 72°C for 30 seconds and a final extension step of 72°C for 10 minutes. Samples were diluted 1:100 with 95% formamide loading dye (95% formamide, 0.025% Bromophenol blue, 0.025% Xylene cyanol) and denatured by heating at 95°C for 5 minutes. Samples were subsequently loaded onto a 6% denaturing (7M urea) polyacrylamide gel and electrophoresis was carried out for 3-4 hours at 60 Watts. Following electrophoresis autoradiography was performed by exposing gels to X-ray film for 5-12 hours (depending on the signal intensity) at -70°C.

3.2 (IV). BAT-26 AMPLIFICATION FOR MICROSATELLITE ANALYSIS:

PCR amplification with the BAT-26 primers was performed essentially as described above however the reaction volume was reduced to 10μl and the dNTP mix concentration was reduced to 0.02mM. In addition 2μCi of [α^{32}P]-dCTP (3000Ci/mmol) was included in the reaction. The optimum cycling parameters included 30 cycles of 95°C for 1 minute, 52°C for 30 seconds, 72°C for 30 seconds and a final extension step of 72°C for 10 minutes. samples were diluted 1:100 with 95% formamide loading dye (95% formamide, 0.025% Bromophenol blue, 0.025% Xylene cyanol) and denatured by heating at 95°C for 5 minutes. Samples were subsequently loaded onto a 6% denaturing (7M urea) polyacrylamide gel and electrophoresis was carried out for 3-4 hours at 60 Watts. Following electrophoresis autoradiography was

performed by exposing gels to X-ray film for 5-12 hours (depending on the signal intensity) at -70°C.

3.2 (V) PCR AMPLIFICATION OF THE BAT-40 LOCUS:

PCR amplification of the Bat-40 locus for microsatellite analysis was performed as for Bat -26 however an annealing temperature of 52°C was required. Samples were loaded onto a 6% denaturing polyacrylamide gel (7M urea) and analysed as for the Bat-26 locus. The optimum cycling parameters included 30 cycles of 95°C for 1 minute, 50°C for 30 seconds, 72°C for 30 seconds and a final extension step of 72°C for 10 minutes. Samples were diluted 1:100 with 95% formamide loading dye (95% formamide, 0.025% Bromophenol blue, 0.025% Xylene cyanol) and denatured by heating at 95°C for 5 minutes. Samples were subsequently loaded onto a 6% denaturing (7M urea) polyacrylamide gel and electrophoresis was carried out for 3-4 hours at 60 Watts. Following electrophoresis autoradiography was performed by exposing gels to X-ray film for 5-12 hours (depending on the signal intensity) at -70°C.

3.2.(VI). MICROSATELLITE INSTABILITY ANALYSIS:

Samples were loaded onto gels in pairs of normal mucosa DNA products adjacent to matched tumour DNA products from the same patient. The size of the resulting products were compared and any discrepancies between pairs of samples were scored as positive for microsatellite instability.

Primer	Sequence	Size
BAT 26 (f)	5' TGA CTA CTT TTG ACT TCA GCC 3'	21mer
BAT 26 (r)	5' AAT CAT TCA ACA TTT TTA ACC 3'	21mer
BAT 40 (f)	5' ACA ACC CTG CTT TTG TTC CT 3'	20mer
BAT 40 (r)	5' GTA GAG CAA GAC CAC CTT G 3'	19mer
BAT RII (f)	5' CTT TAT TCT GGA AGA TGC TGC 3'	21mer
BAT RII (r)	5' GAA GAA AGT CTC ACC AGG C 3'	19mer

Figure 3.3 Nucleotide sequences of the primers employed to amplify the poly A tract repeat loci at the BAT 26 , BAT 40 and the TGF-β-RII Loci.

3.3. RESULTS

MICROSATELLITE INSTABILITY ANALYSIS.

The replication error phenotype (RER) status of 20 sporadic colon cancer patients and 18 HNPCC patients were analysed by comparing the mobilities of PCR products from two repeat loci in matched normal and tumour DNA from these individuals. The markers used to assess the RER status include BAT 26 which amplifies a 26bp polyA tract of the 3' intron of exon 5 of the hMSH2 gene (Lui *et al.*, 1995; Parsons *et al.*, 1995) and BAT 40 which amplifies a 40bp repeat of an intron of the 3-ß-hydroxysteroid dehydrogenase gene (Parsons *et al.*, 1995). These markers were selected as it has been demonstrated that larger repeat sequences display higher frequencies of instabilities than smaller repeat sequences (Parson *et al.*, 1995).

3.3 (I). INSTABILITY ANALYSIS AT THE BAT 26 LOCUS

Matched normal and tumour tissue DNA from 20 sporadic colon cancer patients and 18 HNPCC patients were analysed for instability at the BAT 26 locus. Samples were amplified by PCR with [α-^{32}P]-dCTP incorporation and products were separated by electrophoresis in a 6% denaturing polyacrylamide gel.

Figure 3.4 displays the autoradiograph of the polyacrylamide gel analysis of PCR products amplified from the BAT 26 locus. Lanes were loaded in pairs with normal and tumour DNA respectively, the size of the normal product was compared to the tumour DNA to determine instability. Of the 38 samples analysed (not all shown) instability was observed in 8 samples at this locus - the instabilities were observed in 4 HNPCC samples (IV:I-BI, II:4-H1, II:4-F and II:I-W3) and 4 sporadic colon cancer samples (3, 11, 16 and 19).

3.3 (II). ANALYSIS OF INSTABILITY AT THE BAT 40 LOCUS

Matched normal and tumour DNA from the 20 sporadic colon cancer patients and 18 HNPCC samples was analysed for instability at the BAT 40 locus. The autoradiograph of polyacrylamide gel electrophoresis of the amplified samples is shown in **Figure 3.5**. DNA instability was observed in 7 samples (only samples with instabilities are shown) these include three HNPCC samples (IV:I-B1, II:4-H1, II:1-W3) and 4 sporadic colon cancer samples (3, 16, 17 and 19).

3.3. (III).MICROSATELLITE INSTABILITY AT THE BAT 2R LOCUS

The BAT- IIR oligonucleotide primers amplify a 10bp polyA tract of the TGF-ß-RII gene. This intergenic repeat locus was analysed for instability in matched normal and tumour DNA from 20 sporadic colon cancer samples and 18 HNPCC samples. Radiolabelled [α-^{32}P]-dCTP PCR products were separated by electrophoresis in a 6% polyacrylamide gel. Analysis of the autoradiograph (**Figure 3.6**) reveals no differences in the size of the polyA tract between normal and tumour DNA in any sample. This suggests that the 10bp polyA tract of the TGF-ß-RII gene is stable in the Irish colon cancer samples analysed and is not implicated in the tumourigenic pathway (Table 3.1).

3.3.(IV). DIFFICULTIES WITH THE ANALYSIS OF THE RER PHENOTYPE.

The ability to detect repeat sequence instability is based on the comparison of normal DNA with tumour DNA. DNA instability at repeat sequences is caused by either the expansion or contraction of a repeat sequence in tumour DNA which is not present in DNA from the matched normal tissue. In order to accurately compare the tumour

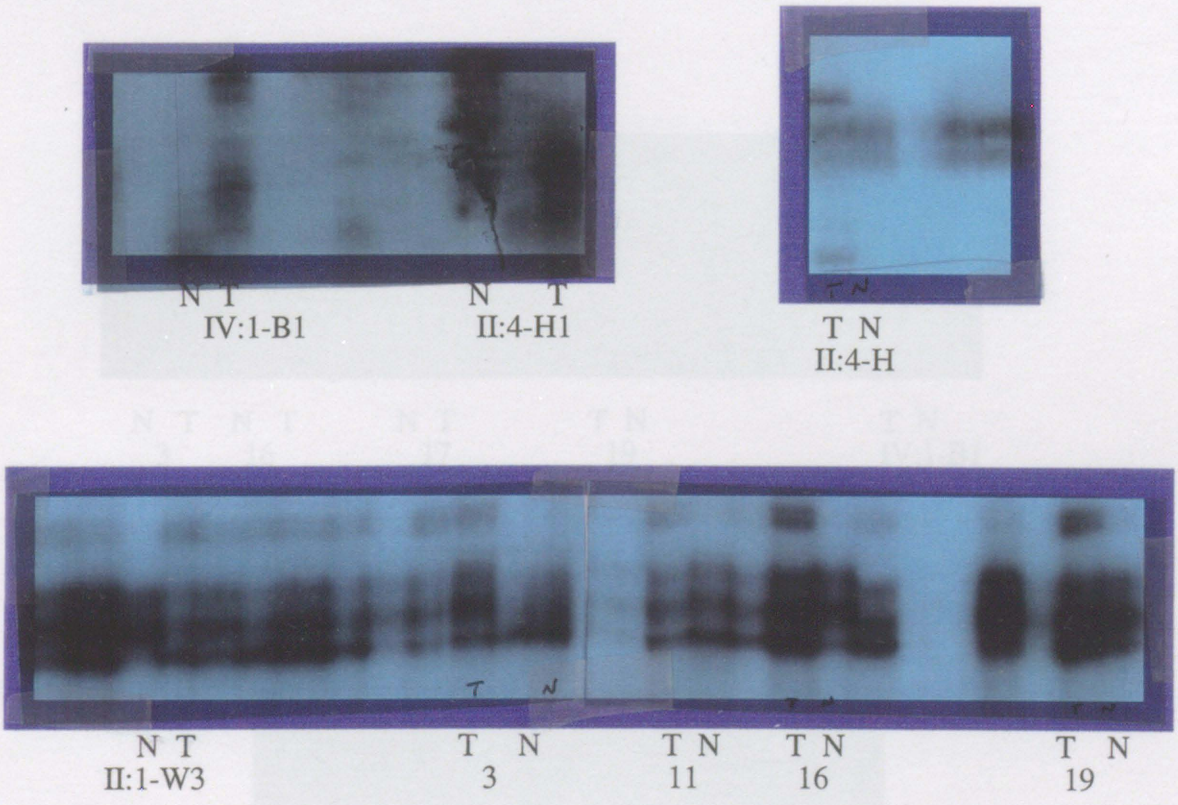
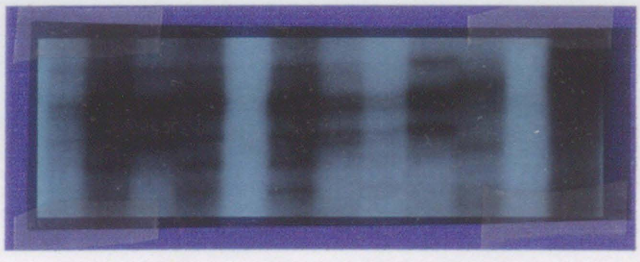


Figure 3.4. Analysis of DNA instability at the 26bp polyA tract at the BAT 26 locus. The Bat 26 locus was amplified from normal (N) and tumour (T) DNA in a PCR reaction containing the BAT 26 forward and reverse primers. α-^{32}P dCTP was included in the reaction to label the PCR products. Products were separated in a 6% denaturing polyacrylamide gel and visualised by autoradiography. Matched normal and tumour DNA from 38 samples were analysed (not all shown). Instabilities were observed in HNPCC samples: IV:1-B1, II:4-H1, II:4-H and II:1-W3 and in sporadic colon cancer samples 3, 11, 16 and 19. Only matched samples displaying differences are numbered.

N T N T N T T N T N
3 16 17 19 IV:1-B1

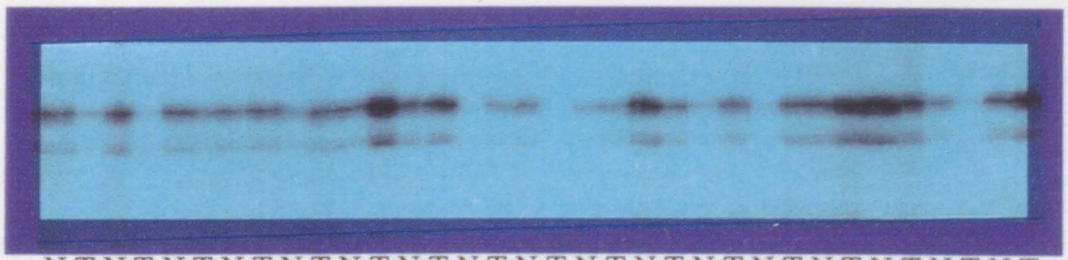


T N T N
II:4-H1 II:1-W3

Figure 3.5 Analysis of DNA instability at the 40bp polyA tract at the BAT 40 locus. The Bat 40 locus was amplified from normal (N) and tumour (T) DNA in a PCR reaction containing the BAT 40 forward and reverse primers. α-^{32}P dCTP was included in the reaction to label the PCR products. Products were separated in a 6% denaturing polyacrylamide gel and visualised by autoradiography. Matched normal and tumour DNA from 38 samples were analysed (not all shown). Instability was observed in 7 of 38 samples including 3 HNPCC samples IV:1-B1, II:4-H1 and II:1-W3 and 4 sporadic colon cancer samples- 3, 16, 17 and 19. Only matched samples displaying differences are numbered.

N T N T N T N T N T N T N T N T N T N T N T N T N T N T N T N T N T N T
1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18

SAMPLE NO



N T N T N T N T N T N T N T N T N T N T N T N T N T N T N T N T N T N T
19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35

SAMPLE NO

Figure 3.6. Analysis of DNA instability at the 10bp polyA tract at the BAT IIR locus. The Bat IIR locus was amplified from normal (N) and tumour (T) DNA in a PCR reaction containing the BAT IIR forward and reverse primers. α-^{32}P dCTP was included in the reaction to label the PCR products. Products were separated in a 6% denaturing polyacrylamide gel and visualised by autoradiography. Matched normal and tumour DNA from 38 samples were analysed (not all shown). No differences in the size of the 10bp polyA tract between normal and tumour DNA were observed.

DNA with normal DNA it is essential that both samples are independent and do not contain any contaminating DNA. The tumour DNA employed in the analysis was obtained from tumour biopsy samples or in some cases from archival paraffin embedded tissues (PET). Accurate removal of purely tumour DNA requires sensitive identification and microdissection of the tumour tissue away from the normal tissue. The samples used in this study were carefully excised either during surgery or in the histology laboratory at Cork university hospital. However it is difficult to assess the purity of the sample and in some cases contamination of the tumour DNA with normal DNA may have occurred and DNA obtained from dissected tumour samples is likely to contain DNA from normal mucosa. Analysis of repeat sequences in heterogeneous DNA samples is observed as the presence of repeat sequences of increased size (instability) in addition to the normal repeat sequence. This phenomenon was observed in the analysis of instability at the BAT 26 and Bat 40 loci performed in the Irish samples investigated, which suggests that the tumour samples do contain cells from the normal colonic mucosa.

3.3.(V). REPEAT SEQUENCE INSTABILITY AND K-RAS GENE DEFECTS

In order to establish the relationship between instability and K-ras gene mutations, results of the mutation analysis of the K-ras gene from chapter 2 were compared with the instability results obtained. Instability at the BAT 26 locus was observed in 4 samples (3, 11, 16 and 19) and at the BAT 40 locus in three similar samples (3, 16 and 19) and one additional sample (17) which did not display microsatellite instability at the BAT 26 locus. K-ras gene analysis identified mutations in 5 sporadic colon cancer samples (2, 10, 14, 40 and 44) however there was no correlation between samples exhibiting instability and K-ras gene mutations (Table 3.3).

3.3.(VI). REPEAT SEQUENCE INSTABILITY AND HNPCC GENE MUTATIONS

In order to determine the relationship between instability and mismatch repair gene defects in HNPCC patients the results of the analysis of the BAT 26 and BAT 40 loci were compared with the results of the mutation analysis of hMSH2 and hMLH1 in 18 Irish HNPCC families (chapter 1). Analysis of the BAT 26 locus identified four samples with instability- individual IV:I of pedigree B1, II:4 of pedigree H1, II:I of pedigree W3 and II:4 from pedigree F. Analysis of the mismatch repair gene mutation data reveals that mutations in either hMSH2 or hMLH1 were identified in the four samples. An exon 5 splice site mutation of hMSH2 was detected in pedigree B1, an exon 17 mutation of hMLH1 was present in pedigree H1, pedigree W3 exhibited an exon 8 hMLH1 mutation and an exon 5 mutation was identified in pedigree F.

Further comparisons with samples which displayed instability at the BAT 40 locus revealed that the three samples (IV:I-B1, II:4-H1 and II:1-W3) correspond to mutation positive samples and displayed BAT 26 instability also. Further evaluation of samples which display instability reveals that they all satisfy the Amsterdam criteria for HNPCC pedigrees and no instability was observed in atypical HNPCC pedigrees suggesting that this subset of HNPCC pedigrees may have an alternative tumourigenic pathway.

Repeat Locus	Sporadic Patients	HNPCC Patients	Total
BAT-26	4/20 (20%)	4/18 (22%)	8/38 (21%)
BAT-40	4/20 (20%)	3/18 (16%)	7/38 (18%)
BAT-RII	0	0	0

Table 3.1. Results of DNA instability assessment at three polyA tract repeat loci. in the Irish sporadic and HNPCC tumour samples

Samples	*MMR Mutations*	*MI*		*BAT-RII*
		BAT-26	*BAT-40*	
IV:1-B1	Exon 5-hMSH2	Yes	Yes	No
II:4-H1	Exon 17-hMLH1	Yes	Yes	No
II:4-F	Exon 5-hMSH2	Yes	No	No
II:1-W3	Exon 8-hMLH1	Yes	Yes	No

Table 3.2. Results of the analysis of DNA instability at three polyA tract repeat loci and of the mismatch repair gene mutation status of the HNPCC cancer samples.

Sample	BAT-26 MI	BAT-40 MI	BAT-RII MI	K-RAS Mutation
3	Yes	Yes	No	No
11	Yes	No	No	No
16	Yes	Yes	No	No
19	Yes	Yes	No	No
17	No	Yes	No	No
2	No	No	No	Yes
10	No	No	No	Yes
14	No	No	No	Yes
40	No	No	No	Yes
44	No	No	No	Yes

Table 3.3. Correlation of DNA instability results with the K-ras mutation status of the sporadic colon cancer samples.

3.4. DISCUSSION

The phenomenon of microsatellite instability has been associated with mismatch repair gene defects in sporadic and hereditary colon cancer (Aaltonen *et al.*, 1995; Lui *et al.*, 1995). Genetic models to explain the pathogenic differences between RER+ and RER- tumours have been postulated (Fearon *et al.*, 1990; Tannergard *et al.*,1997).

It has been proposed that RER+ tumours develop through the mismatch repair pathway in that mutations in mismatch repair genes which occur at a very early stage result in a mutator phenotype within the cell, the mutator phenotype results in multiple mutations in other genes which if essential in the regulation of cellular growth, result in the progression of tumourigenesis. The most obvious manifestation of the RER phenotype is observed as instability at repeat sequences as these regions are the most frequent sites of DNA Polymerase slippage (Aaltonen *et al.*, 1993; Parsons *et al.*, 1995; Loeb 1994; Thibodeau *et al.*, 1993).

The RER - pathway is not characterised by MI as it is believed that mismatch repair gene defects may not have a major causative role as in RER+ tumours. This model is based on the classical K-ras, APC, p53 tumorigenic pathway in which the accumulation of mutations in a number of specific genes is required for tumour development (Fearon *et al.*, 1990). The accumulation of mutations in such genes provide a selective growth advantage to the cell with progression to carcinoma (Fearon and Vogelstein, 1990).

Due to the association of the RER+ phenotype with alterations at repeat sequences, TGF-ß-RII which contains an intragenic repeat sequence has been implicated in the disease and it has been suggested that it may play a vital role in the RER+ tumourigenic pathway (Markowitz *et al.*, 1995).

In order to investigate the RER+/RER- tumourigenesis pathways in the Irish colon cancer population, instability was analysed at two loci and at the TGFB-RII

intergenic repeat locus in the Irish HNPCC and sporadic colon cancer samples. The HNPCC samples had been previously screened for mutations in the hMLH1 and hMSH2 mismatch repair genes (chapter 1) and the sporadic samples had been analysed previously for K-ras gene mutations (chapter 2). The analysis of instability at repeat sequences and the various cancer genes in these samples facilitated the comparison of the factors implicated in the two tumorigenic pathways.

3.4. (II). ANALYSIS OF THE HNPCC POPULATION

Results of the mutation screening of hMSH2 and hMLH1 in 18 Irish HNPCC pedigrees are presented in chapter 1, briefly, mutations were identified in 5 families 3 in the hMSH2 gene and 2 in the hMLH1 gene. Analysis of the BAT 26 locus identified 4 HNPCC patients with instabilities, these samples correspond to samples in which a mismatch repair gene mutation was identified (I:V-B1, II:4-H1, II:4-F and II:1-W3). Analysis of the BAT 40 locus identified 3 samples with instabilities and they correspond to patients with mismatch repair gene mutations and overlap with samples displaying BAT 26 instabilities (I:V-B1, II:4-H1 and II:1-W3).

In summary, analysis of the HNPCC patients identified mismatch repair gene mutations in 5 HNPCC pedigrees (see chapter 1), assessment of the RER phenotype in the same HNPCC patients identified abnormalities at two loci in three samples and one locus in one patient, however there was no evidence for instability at either locus in a single sample which did exhibit a mismatch repair mutation.

Further assessment of RER data reveals a correlation between samples exhibiting instabilities and families which satisfy the Amsterdam selection criteria for HNPCC pedigrees. Of the 18 HNPCC pedigrees analysed, 8 are typical in that they satisfy all the Amsterdam criteria and 10 are atypical in that they fail to satisfy all the criteria. Due to the fact that mismatch repair gene mutations in Irish HNPCC patients were confined to typical HNPCC pedigrees it follows that instability at repeat sequences should be associated more frequently with these families. The correlation serves to

substantiate the belief that the RER phenotype observed in tumour samples is a manifestation of mismatch repair gene defects and that in families which strictly satisfy the Amsterdam criteria- mismatch repair gene defects are most likely the major factors in tumourigenesis.

Based on the premise that the RER+ phenotype is designated when a minimum of two loci display instability in a sample, 3 of 18 HNPCC (16%) samples exhibited this phenotype. However, when only the typical HNPCC pedigrees are analysed 3 of 8 (37.5%) display the RER+ phenotype. These results are somewhat lower than the reported data for HNPCC pedigrees (70%) (Aaltonen *et al.*, 1993), however this may be due to the fact that only two loci were analysed, and the sample size was small. The results do however suggest that instability at repeat sequences is a common feature of tumours which harbour mutations in either hMSH2 or hMLH1 and are not common in atypical HNPCC pedigrees in the Irish samples studied here.

3.4.(II). MICROSATELLITE INSTABILITY IN SPORADIC COLON CANCER.

In order to determine the incidence of instability in the Irish sporadic cancer patients, 20 normal and tumour DNA samples were analysed for instability at the BAT 26 and the BAT 40 loci. Four tumour samples displayed altered migration at the BAT 26 repeat sequence when compared to matched normal DNA and three of these samples also displayed instability at the BAT 40 locus. In addition 1 sample which did not display instability at the BAT 26 locus did display abnormalities at the BAT 40 locus. It has been reported that 15% of sporadic colon cancer samples exhibit the RER phenomenon (Aaltonen *et al.*, 1993; Ionov *et al.*, 1993), this compares favourably with the results obtained in the Irish samples studied here 4/20 (20%). However as the sample size is rather small and only two loci were examined it would be necessary to perform further analysis to verify the incidence of instability in sporadic colon cancer samples.

Sporadic tumours with microsatellite instability have characteristics in common with HNPCC tumour in that they are often located in the proximal colon (Ionov *et al.*, 1993; Thibodeau *et al.*, 1993; Aaltonen *et al.*, 1993), are poorly differentiated (Ionov *et al.*, 1993) and display diploidy (Aaltonen 1993), furthermore they show far lower frequencies of mutations in K-ras (Ionov *et al.*, 1993) and p53 (Kim *et al.*, 1994) and infrequent loss at 5q, 17p and 18q (Thibodeau *et al.*, 1993). It has been suggested that in the absence of mismatch repair gene defects and repeat sequence instability, tumours develop through an alternative pathways which involves the accumulation of mutations in a number of essential genes including K-ras, APC and p53 (Tannergard *et al.*, 1997). In order to determine if there is a negative correlation between MI and K-ras gene mutations the results of the K-ras mutation analysis presented in chapter two were compared with the MI data on the same patients. K-ras gene mutations were identified in five individuals (2, 10, 14, 40 and 44), however microsatellite instability observed in samples 3, 11, 16, 17 and 19 only. These results suggest that there is a negative correlation between K-ras gene defects and MI and provides evidence to corroborate the hypothesis that a distinct MI pathway for tumourigenesis exists and that the K-ras, APC, p53 pathway does not exhibit MI as a characteristic phenomenon.

3.4.(IV).THE TRANSFORMING GROWTH FACTOR BETA GENE AND COLON CANCER.

The TGF-ß-RII gene has been implicated in the RER+ mismatch repair defect pathway of tumourigenesis (Markowitz *et al.*, 1995). Mutations in a 10bp polyA tract of the gene have been identified in 70% of RER+ colon cancer tumours (sporadic and HNPCC) and are associated with mutations of the mismatch repair genes (Parsons *et al.*, 1995; Togo *et al.*, 1996; Markowitz *et al..*, 1995). Analysis of the 10bp polyA tract at nucleotides 709-718 (BAT -RII) of the TGF-ß-RII gene failed to identify instabilities in the 20 sporadic colon cancer samples and in 18 Irish HNPCC samples. These results suggest that although mutations in TGF-ß-RII may play a role in advancing tumourigenesis in some forms of colon cancers, mutations in this gene are not

associated with RER or MMR gene mutations in any of the sporadic or HNPCC samples analysed here.

3.5. BIBLIOGRAPHY

Aaltonen, L.A., Peltomaki, P., Mecklin, J.P., Jarvinen, H., Jass, J.R., Green, J.S., Lynch, H.T., Watson, P., Tallqvist, G., Juhola, M., et al., (1994). *Replication errors in benign and malignant tumors from hereditary nonpolyposis colorectal cancer patients.* Cancer Res 54, 1645-1648.

Aaltonen, L. A., Peltomaki, P., Leach, F. S., Sistonen, P., Pylkkanen, L., Mecklin, J-P., Jarvinen, H., Powell, S. M., Jen, J., hamilton, S.R., Petersen, G.M., Kinzler, K.W., Vogelstein, B., de la Chapelle, A. (1993). *Clues to the Pathogenesis of Familial Colorectal Cancer.* Science 260, 812-819.

Akiyama, Y., Iwanaga, R., Ishikawa, T., Sakamoto, K., Nishi, N., Nihei, Z., Iwama, T., Saitoh, K., Yuasa, Y. (1996). *Mutations of the transforming growth factor-beta type II receptor gene are strongly related to sporadic proximal colon carcinomas with microsatellite instability.* Cancer. 78(12), 2478-2484.

Borresen, A.L., Lothe, R.A., Meling, G.I., Lystad, S., Morrison, P., Lipford, J., Kane, M.F., Rognum, T.O., Kolodner, R.D. (1995). *Somatic mutations in the hMSH2 gene in microsatellite unstable colorectal carcinomas.* Hum Mol Genet. 4(11), 2065-2072.

Bronner, C.E., Baker, S.M., Morrison, P.T., Warren, G., Smith, L.G., Lescoe, M.K., Kane, M., Earabino, C., Lipford, J., Lindblom, A., (1994). *Mutation in the DNA mismatch repair gene homologue hMLH1 is associated with hereditary non-polyposis colon cancer.* Nature 368 , 258-261.

Chen, R.H., Ebner, R., Derynck, R. (1993). *Inactivation of the type II receptor reveals two receptor pathways for the diverse TGF-beta activities.* Science. 260, 1335-1338.

Fearon, E.R., Vogelstein, B. (1990a). *A genetic model for colorectal tumorigenesis.* Cell, 61(5), 759-767.

Han, H. J., Akio, Y., Yo, K., Jae-Gahb, P., Nakamura, Y. (1993). *Genomic Instability in Pancreatic Cancer and Poorly Differentiated Type of Gastric Cancer.* Cancer Res. 53, 5087-5089.

Ionov, Y., Peinado, M.A., Malkhosyan, S., Shibata, D., Perucho, M. (1993). *Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis.* Nature 363,(6429), 558-561.

Kim H , Jen J Vogelstein B and Hamilton SR. 1994, Am. J. Pathology, 145, 148-156.

Kolodner, R.D., Hall, N.R., Lipford, J., Kane, M.F., Morrison, P.T., Finan, P.J., Burn, J., Chapman, P., Earabino, C., Merchant, E., Bishop, T., Garber, J., Bronner, C.E., Baker,

S.M., Warren, G., Smith, L.G., Lindblom, A., Tannergard, P., Bollag, R.J., Godwin, A.R., Ward, D.C., Liskay, R.M., Copeland, N., Jenkins, N., Lescoe, M.K., Ewel, A., Lee,S., Griffith, J., Fishel, R. (1994). *Human Mismatch Repair Genes and their Association with HNPCC.* Cold Spring Harb Symp Quant Biol. 59:331-338.

Kunkel, T.A. (1993). *Nucleotide repeats. Slippery DNA and diseases.* Nature 365(6443), 207-208.

Leach, F.S., Nicolaides, N.C., Papadopoulos, N., Liu, B., Jen, J., Parsons, R., Peltomaki, P., Sistonen, P., Aaltonen, L.A., Nystrom-Lahti, M., et al. (1993). *Mutations of a mutS homolog in hereditary nonpolyposis colorectal cancer.* Cell 75(6), 1215-1225.

Levinson, G., Gutman, G.A. (1987). *High frequencies of short frameshifts in poly-CA/TG tandem repeats borne by bacteriophage M13 in Escherichia coli K-12.* Nucleic Acids Res. 15(13), 5323-5338.

Loeb, L.A. (1991). *Mutator phenotype may be required for multistage carcinogenesis.* Cancer Res. 51(12), 3075-3079.

Loeb, L.A. (1994). *Microsatellite instability: marker of a mutator phenotype in cancer.* Cancer Res. 54 (19), 5059-5063.

Lui, B., Parsons, R., Papadopoulos, N., Nicolaides, N.C., Lynch, H.T., Watson, P., Jass, J.R., Dunlop, M., Wyllie, A., Peltomaki, P., de la Chapelle, A., Hamilton, S.R., Vogelstein, B., Kinzler, K.W. (1996). *Analysis of Mismatch repair genes in hereditary Nonpolyposis Colorectal cancer patients.* Nat. Med. 2, 169-174.

Lui, B., Nicolaides, N.C., Markowitz, S., Wilson, J.K., Parsons, R.E., Jen, J., Papadopoulos, N., Peltomaki, P., de la Chapelle, A., Hamilton, S.R., Kinzler, K.W., Vogelstein, B. (1995). *Mismatch Repair Gene Defects in Sporadic Colorectal cancers with Microsatellite instability.* Nat. Genet. 1, 48-55.

Lothe, R.A., Peltomaki, P., Meling, G.I., Aaltonen, L.A., Nystrom-Lahti, M., Pylkkanen, L., Heimdal, K., Andersen, T.I., Moller, P., Rognum, T.O., et al., (1993). *Genomic instability in colorectal cancer: relationship to clinicopathological variables and family history.* Cancer Res, 53, 5849-5852

Liu, B., Farrington, S.M., Petersen, G.M., Hamilton, S.R., Parsons, R., Papadopoulos, N., Fujiwara, T., Jen, J., Kinzler, K.W., Wyllie, A.H., et al., (1995). *Genetic instability occurs in the majority of young patients with colorectal cancer.* Nat Med. 4, 348- 352

Markowitz, S., Wang, J., Myeroff, L., Parsons, R., Sun, L., Lutterbaugh, J., Fan, R.S., Zborowska, E., Kinzler, K.W., Vogelstein, B., et al., (1995). *Inactivation of the type II TGF-beta receptor in colon cancer cells with microsatellite instability.* Science. 268, 1336-1338.

Orth, K., Hung, J., Gazdar, A., Bowcock, A., Mathis, J.M., Sambrook, J. (1994). *Genetic instability in human ovarian cancer cell lines.* Proc Natl Acad Sci U S A. 91(20), 9495-9499.

Papadopoulos, N., Nicolaides, N.C., Wei, Y.F., Ruben, S.M., Carter, K.C., Rosen, C.A., Haseltine, W.A., Fleischmann, R.D., Fraser, C.M., Adams, M.D, Venter, C., Hamilton, S., Petersen, G., Watson, P., Lynch, H.T., peltomaki, P., Mecklin, J-P., de la Chapelle, A., Kinzler, K., Vogelstein, B. (1994). *Mutation of a mutL homolog in hereditary colon cancer.* Science 263, 1625-1629.

Parsons, R., Myeroff, L.L., Liu, B., Willson, J.K., Markowitz, S.D., Kinzler, K.W., Vogelstein ,B. (1995). *Microsatellite instability and mutations of the transforming growth factor beta type II receptor gene in colorectal cancer.* Cancer Res 55(23), 5548-5550.

Parsons, R., Li, G.M., Longley, M.J., Fang, W.H., Papadopoulos, N., Jen ,J., de la Chapelle, A., Kinzler, K.W., Vogelstein. B.,,Modrich, P. (1993). *Hypermutability and mismatch repair deficiency in RER+ tumor cells.* Cell 75(6), 1227-1236.

Peltomaki, P., Lothe, R.A., Aaltonen, L.A., Pylkkanen, L., Nystrom-Lahti, M., Seruca, R., David, L., Holm, R., Ryberg, D., Haugen, A. (1993). *Microsatellite instability is associated with tumors that characterize the hereditary non-polyposis colorectal carcinoma syndrome.* Cancer Res 53(24), 5853-5855.

Peltomaki, P., Vasen, H.F.A., and the International Collaborative Group on Hereditary Nonpolyposis Colorectal Cancer: *Database and Results of a Collaborative Study.* (1997). Gastroenterology, 113, 1146-1158.

Risinger, J.I., Berchuck, A., Kohler, M.F., Watson, P., Lynch, H.T., Boyd, J. (1993). *Genetic instability of microsatellites in endometrial carcinoma.* Cancer Res. 53(21), 5100-5103.

Roberts, A.B and Sporn, M.B. (1990). *Peptide growth factors and their receptors.* Springer-Verlag. Berlin. pp 420-427.

Schlotterer, C and Tautz, D. (1992). *Slippage synthesis of simple sequence DNA.* Nucleic Acids Res. 20(2), 211-215.

Tannergard, P., Liu, T., Weger, A., Nordenskjold, M., Lindblom, A. (1997). *Tumorigenesis in colorectal tumors from patients with hereditary non-polyposis colorectal cancer.* Hum Genet. 101(1), 51-55.

Thibodeau, S.N., Bren, G., Schaid , D. (1993). Microsatellite instability in cancer of the proximal colon. Science 260 (5109), 816-819.

Togo, G., Toda, N., Kanai, F., Kato, N., Shiratori, Y., Kishi, K., Imazeki, F., Makuuchi, M., Omata, M. (1996). *A transforming growth factor beta type II receptor gene mutation common in sporadic cecum cancer with microsatellite instability*. Cancer Res, 56, 5620-5623.

Vasen, H.F., Mecklin, J.P., Khan, P.M., Lynch, H.T. (1991). *The International Collaborative Group on Hereditary Non-Polyposis Colorectal Cancer (ICG-HNPCC).* Dis Colon Rectum. 34(5), 424-425.

Weber, J.L., May, P.E. (1989). *Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction.* Am J Hum Genet. 44(3), 388-396.

Wijnen, J., Khan, P.M., Vasen, H., van der Klift, H., Mulder, A., van Leeuwen-Cornelisse, I., Bakker, B., Losekoot, M., Moller, P, Fodde, R. (1997). *Hereditary nonpolyposis colorectal cancer families not complying with the Amsterdam criteria show extremely low frequency of mismatch-repair-gene mutations.* Am J Hum Genet. 61(2), 329-335.