

***The Molecular Genetic
and Epidemiological
Investigation of
Colorectal Cancer
in South Africa***

***Rebecca Felix
BSc(Hons)***

*Thesis presented in fulfilment of the requirements for the
degree of*

Doctor of Philosophy

in the Division of Human Genetics

Faculty of Health Sciences

University of Cape Town

August 2003

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

To Emily-Jane

"The disease of cancer will be banished from life by calm, unhurrying, persistent men and women and the motive that will conquer cancer will not be pity nor horror, it will be curiosity to know how and why."

- H.G. Wells

University of Cape Town

Declaration

This study was performed from 1999 to 2003 under the supervision of Professor R.S. Ramesar from the Division of Human Genetics, University of Cape Town.

I hereby certify that this thesis is my own work and has not been presented for a degree at any other university.

Rebecca Felix

August 2003.

Signed by candidate

TABLE OF CONTENTS

TABLE OF CONTENTS i

ABSTRACT v

ACKNOWLEDGEMENTS vii

LIST OF FIGURES..... x

LIST OF TABLES..... xii

SYMBOLS AND ABBREVIATIONS.....xv

GLOSSARY xvii

PLAN OF THESIS xxiv

CHAPTER 1: REVIEW OF COLON CANCER 1

1.1. INTRODUCTION 1

1.1.1. Cancer Genetics 1

1.1.2. Cancer of the colon and rectum 3

1.2. ANATOMY OF THE COLON..... 5

1.2.1. Embryological development.....7

1.3. MOLECULAR GENETICS OF HEREDITARY COLORECTAL CANCER 8

1.3.1. Hereditary nonpolyposis colorectal cancer (HNPCC).....8

1.3.2. Familial adenomatous polyposis (FAP).....27

1.4. METHYLATION OF *hMLH1*..... 30

1.5. MODIFIERS OF THE HNPCC PHENOTYPE..... 32

1.6. COLORECTAL CANCER IN SOUTH AFRICA	33
1.7. AIMS AND OBJECTIVES.....	35
1.7.1. Aim:	35
1.7.2. Project objectives:	35
 CHAPTER 2: ELUCIDATION OF GERMLINE DEFECTS UNDERLYING COLORECTAL CANCER.	 36
2.1. GENERAL INTRODUCTION.....	36
2.1.1. The Mismatch Repair (MMR) Pathway and CRC.....	36
2.1.2. APC.....	40
2.1.3. <i>β-catenin</i>	40
2.1.4. <i>p53</i>	41
2.1.5. Mutation Detection Methods	44
2.1.6. Genotype/Phenotype correlation in HNPCC.....	46
2.2. MATERIALS AND METHODS	49
2.2.1. Subject recruitment.....	49
2.2.2. Mutation Analysis.....	52
2.2.3. DNA sequence analysis	55
2.2.4. Amplification Refractory Mutation System (ARMS) PCR Analysis.....	55
2.2.5. <i>MvaI</i> Restriction endonuclease analysis.....	56
2.2.6. Haplotype Analysis	56
2.2.7. Microsatellite instability (MSI) studies	59
2.3. RESULTS	61
2.3.1. Subjects	61
2.3.2. Mutation analysis	61
2.3.3. <i>MvaI</i> Restriction Endonuclease Analysis.....	67
2.3.4. Haplotype Analysis	67
2.3.5. Genotype/Phenotype correlation	68
2.3.6. MSI analysis.....	71
2.4. DISCUSSION	74
2.4.1. Mutation and Sequencing Analysis	74
2.4.2. Genotype/phenotype correlation	88
2.4.3. Microsatellite Instability.....	89

2.4.4.	The mutation test procedure	93
CHAPTER 3: MODIFIERS OF THE HNPCC PHENOTYPE		97
3.1.	INTRODUCTION	97
3.1.1.	<i>E-Cadherin (CDH1)</i>	99
3.1.2.	Missense variants in the adenomatous polyposis coli gene (<i>APC</i>).....	102
3.1.3.	The Ataxia telangiectasia gene (<i>ATM</i>).....	104
3.1.4.	The <i>Cyclin D1 gene (CCND1)</i>	106
3.1.5.	The Methylenetetrahydrofolate Reductase gene (<i>MTHFR</i>).....	110
3.1.6.	The <i>Glutathione S-transferases (GSTs)</i>	114
3.1.7.	<i>N-acetyltransferases 2 (NAT2)</i>	116
3.2.	MATERIALS AND METHODS	119
3.2.1.	Subjects	119
3.2.2.	Candidate modifier genes.....	119
3.2.3.	Polymerase chain reaction	120
3.2.4.	Restriction fragment length polymorphism (RFLP)	121
3.2.5.	DNA sequencing analysis.....	122
3.2.6.	Statistical methods	123
3.3.	RESULTS.....	124
3.3.1.	Subjects' characteristics	124
3.3.2.	ARMS PCR Assay analysis	126
3.3.3.	Polymerase chain reaction	132
3.3.4.	RFLP analysis.....	136
3.4.	DISCUSSION AND CONCLUSION	141
3.4.1.	<i>E-cadherin (CDH1)</i>	142
3.4.2.	<i>Adenomatous polyposis coli gene (APC)</i>	143
3.4.3.	<i>Ataxia telangiectasia gene (ATM)</i>	143
3.4.4.	<i>Cyclin D1 (CCND1)</i>	145
3.4.5.	<i>Methylenetetrahydrofolate Reductase (MTHFR)</i>	147
3.4.6.	<i>Glutathione-S-transferases (GSTs)</i>	147
3.4.7.	<i>N-acetyltransferase 2 (NAT2)</i>	148
CHAPTER 4: CONCLUDING REMARKS		153
4.1.	CONCLUDING REMARKS	153

REFERENCES.....	161
APPENDICES.....	192
APPENDIX A.....	193
APPENDIX B.....	197
APPENDIX C.....	209
APPENDIX D.....	211

University of Cape Town

ABSTRACT

Cancer of the colon and rectum is the third-most common cause of death due to neoplasia in Western countries. Colorectal cancer (CRC) can broadly be divided into the hereditary, non-hereditary and sporadic forms. Hereditary cancers account for approximately 10% of the CRCs. Familial cancers can be divided into those types in which polyps are a dominant feature (e.g. familial adenomatous polyposis-FAP) and the non-polypotic forms (e.g. hereditary nonpolyposis colorectal cancers-HNPCC).

The recognition of the familial basis of some cancers has proven to be an effective contributor to improved management through early detection of neoplasia. HNPCC has been recognised as a familial cancer and families have been identified for management on the basis of several clinical criteria *viz.* The Amsterdam criteria I and II and the Bethesda criteria. However, the recognition of familiarity and hereditary risk with presymptomatic screening of those at risk, has not been a part of management of cancers in any meaningful way in developing countries. The current study represents an approach towards identifying the genetic defects in individuals who present with CRC in an attempt to recognise those individuals who are likely to be “tips of icebergs” and would benefit from familial recruitment and management.

HNPCC is dominantly inherited and individuals are usually affected by 45 years of age. Subjects generally present with cancer of the proximal colon (Type I). In addition, in a proportion of families, tumours occur at a variety of extracolonic sites (Type II). HNPCC is characterised by germline mutations in one of the DNA MMR genes i.e. *hMLH1*, *hMSH2*, *hMSH6*, *hPMS2* and *hMLH3*.

A research cohort of 100 individuals diagnosed with CRC at a young age (mostly under 45 years) was chosen for mutation analysis of the *hMSH2*, *hMLH1* and *hMSH6* genes. In addition, this research cohort was screened for mutations in exon 15 of the *adenomatous polyposis* gene, exons 3, 5 and 6 of the β -catenin gene, exons 4 to 9 of the *p53* gene and for the presence of eight previously-

described mutations in the *E-cadherin* gene. All these genes were chosen for their known role in the development and progression of CRC.

Twelve mutations were identified in the research cohort during the first phase of screening of the three MMR genes. Of the 12 mutations identified, five were previously reported. The second phase of mutation screening involved examining the entire CRC cohort (n=260 individuals) for the 13 identified mutations. The C1528T mutation in exon 13 of the *hMLH1* gene was detected in 12 additional individuals, each representing a different family. A disease-associated haplotype was identified in the 12 families, suggesting a common founder. All 13 families belong to the same population group and have a shared geographical origin. In addition, two novel mutations were identified in *p53* in four individuals and four mutations (one novel) in the MCR of *APC*.

The study went on to examine any genotype/phenotype correlation that may exist between the mutations identified in the MMR genes and the clinical phenotype observed in the subjects under investigation. The 13 families who harbour the C1528T mutation were also examined independently for any genotype/phenotype correlation. The investigation of the sites of colonic neoplasia demonstrated that those individuals diagnosed with rectal cancer do not harbour mutations in the MMR genes ($p=0.09$). In addition no significant correlation could be made between the C1528T mutation and age at diagnosis or the site of cancer in the 13 families, suggesting that DNA variations in other genes could be modulating the HNPCC phenotype in this group. A variety of mutations and polymorphisms were subsequently investigated in genes known to modify HNPCC and CRC as a whole (*ATM*, *Cyclin D1*, *MTHFR*, *NAT2*, *GSTM1* and *GSTT1*). Significant correlations were observed between the null allele of *GSTM1*, the C677T polymorphism in *MTHFR*, the D1853N variant in *ATM* and the age at diagnosis of CRC in individuals who harbour the C1528T mutation.

ACKNOWLEDGEMENTS

My project supervisor, Professor Rajkumar Ramesar, for his constructive criticism during the experimental aspects and for his advice during the writing of this thesis.

My co-supervisor Sir Walter Bodmer who supervised this research when I spent a year in his laboratory at Oxford university, and for his guidance and constant enthusiasm during the writing of this thesis.

Dr. Rene Goliath for editing sections of this thesis and for her expert guidance of experimental procedures conducted.

A special thank you to Lord and Lady Sainsbury of Candover for awarding me with the John Sainsbury/Lindbury Fellowship for PhD students at the University of Cape Town that allowed me to spend a valuable year at Oxford University in the United Kingdom. This opportunity made me realise that there are no barriers in life except the ones that I make for myself.

A very special thanks to the doctors and nurses involved in this project, especially Sister Ursula Algar, who was primarily instrumental in meeting patients and supplying us with biological material and clinical information for this study.

Thank you also to the many patients who consented to the research being conducted on their biological material.

I am grateful to Emeritus Professor Peter Beighton for his constructive criticism of parts of this thesis.

I am greatly indebted to friends and colleagues in the Division of Human Genetics.

Professor Ian Tomlinson's laboratory for providing me with primer sequences whilst working at the ICRF laboratory at Oxford University. Dr. Nicola Fearnhead for supervising the work carried out at the ICRF laboratory at Oxford University and supplying me with oligonucleotides for the work carried out there.

My colleagues and friends at the Cancer Immunogenetics/ICRF laboratory at the John Radcliffe Hospital, Oxford University for their friendship and encouragement during the year that I spent in the United Kingdom.

I am gratefully acknowledge various funding bodies for financing my research and travelling expenses when attending overseas conferences:

The Cancer Society of South Africa (CANSA)

The Medical Research Council (MRC)

The National Research Foundation (NRF)

The Guy Elliott Trust

The Division of Human Genetics at the University of Cape Town

The Faculty of Health Sciences at the University of Cape Town

I would especially like to thank certain family members who were co-operative and understanding during the duration of my PhD:

My mother, Lucinda Martin for encouraging me to further my career and for always believing in me and availing herself to help me when I needed it, especially in the final stages of PhD. My father, Daniel Martin, who was always encouraging and concerned and never hesitated to provide me with much needed help at the spur of the moment.

My sister, Edeona Martin, and brother, David Martin, for their constant encouragement. My sister-in-law Tracey Martin for her interest and moral support.

Then I would especially like to acknowledge my husband, Kenneth Felix for his amazing support, love, patience and encouragement during the course of my studies and especially during the writing of this thesis.

This thesis is dedicated to my seven-month old daughter, Emily-Jane Felix, who has brought a new dimension to my life and who has made this journey worthwhile.

Then last, but certainly not least. I would like to thank my Heavenly Father for giving me the privilege to study. I truly am blessed.

LIST OF FIGURES

Figure 1.1.	Components of the large intestine.	6
Figure 1.2.	Development of the human large intestine – structure and position of the hindgut, midgut and foregut in a 20day old embryo.....	7
Figure 1.3.	Putative genetic pathways in colorectal cancer.-.....	23
Figure 2.1.	The enzymes that repair DNA mismatches in <i>Escherichia coli</i> ..	38
Figure 2.2.	Ideogram of chromosome 3 with the relative positioning of hMLH1 and the physical distances of the microsatellite markers as used for linkage/ haplotype analysis.	58
Figure 2.3.	Bar graph of affected individuals in whom mutations were detected vs. those in whom no mutations were detected in the MMR genes investigated.....	69
Figure 2.4.	Numbers of individuals (n = 29) displaying cancer at the various sites of the colon.	70
Figure 2.5.	The distribution of the age at diagnosis for individuals belonging to the C1528T-research cohort (n=29).	71
Figure 2.6.	Illustration of MSI in tumour tissue.	72
Figure 2.7.	Pedigree of NPC 85.	75
Figure 3.1.	Interactions of cyclins and cyclin-dependent kinases (CDK) during the cell cycle.	108
Figure 3.2.	The metabolic role of MTHFR in folate metabolism.	112
Figure 3.3.	Proportion of individuals in the C1528T research cohort unaffected at each age in years.	125
Figure 3.4.	Proportion of individuals in the C1528T-research cohort who are unaffected at each age (years) stratified by gender.....	126
Figure 3.5.	Proportion of unaffected individuals in the C1528-research cohort at each age, stratified by ATM genotype.	128
Figure 3.6.	Proportion of unaffected individuals in the C1528T-research cohort at each age, stratified by CCND1 genotype.....	129

Figure 3.7. Proportion of the C1528T-research cohort unaffected at each age, stratified by MTHFR genotype.	130
Figure 3.8. Proportion of unaffected individuals in the C1528T-research cohort at each age, stratified by GSTT1 genotype.	133
Figure 3.9. Proportion of unaffected individuals in C1528T-research cohort at each age, stratified by GSTM1 genotype.	135
Figure 3.10. Proportion of unaffected individuals in the C1528T-research cohort at each age, stratified by NAT2.	137
Figure 3.11. Proportion unaffected for each age at diagnosis stratified by combined genotype as defined above.	140

LIST OF TABLES

Table 1.1. Details of screening recommendations for extracolonic tumours in HNPCC.....	14
Table 2.1. Distribution of tumours in the colon of patients chosen for mutation analysis (n=100).....	50
Table 2.2. Mutations and polymorphisms detected on analysis of hMLH1, hMSH2 and hMSH6.....	62
Table 2.3. CRC phenotype in additional families in whom mutations have been identified.....	64
Table 2.4. Mutations and polymorphisms detected on screening of exon 15 of the APC gene.	66
Table 2.5. Sequence variants in exon 4 and 9 of p53.....	67
Table 2.6. Mutations detected in DNA extracted from fresh tumour tissue displaying MSI-H status.	73
Table 2.7. Microsatellite instability displayed for formalin-fixed paraffin embedded tissue samples.	73
Table 2.8. Increase in the number of family members of kindreds harbouring the Q510X mutation.....	83
Table 3.1. Summary of germline mutations in CDH1.	102
Table 3.2. Mutations associated with the various NAT2 genotypes.	117
Table 3.3. Patterns obtained for NAT2 alleles after cleavage with KpnI, DdeI, TaqI and BamHI.....	122
Table 3.4. Characteristics of subjects in C1528T research cohort.....	124
Table 3.5. ATM genotype frequencies in the C1528T-research cohort.....	127
Table 3.6. CCND1 genotype frequencies in the C1528T-research cohort.	129
Table 3.7. MTHFR genotype frequencies in the C1528T-research cohort.	130
Table 3.8. Gender-stratified age at diagnosis for males and females who harbour the MTHFR CC and CT genotypes.	131
Table 3.9. GSTT1 genotype frequencies in the C1528T-research cohort.	132
Table 3.10. GSTM1 genotype frequencies in the C1528T-research cohort. .	134

Table 3.11. <i>GSTM1</i> genotypes for males and females who harbour the C1528T mutation.	136
Table 3.12. <i>NAT2</i> genotype frequencies in the C1528T-research cohort.	136
Tables of primer sequences used for PCR amplification.	211
Table D.1. Oligonucleotide sequences for hMSH2 for the purpose of PCR based mutation detection.	212
Table D.2. Oligonucleotide sequences for hMLH1 for the purpose of PCR-based mutation detection.	213
Table D.3. Oligonucleotide sequences for hMSH6 for the purpose of PCR-based mutation detection.	215
Table D.4. Oligonucleotide sequences for PCR amplification of the MCR of APC for dHPLC Wave analysis.	216
Table D.5. Oligonucleotide sequences for PCR amplification of β -catenin for dHPLC Wave analysis.	217
Table D.6. Oligonucleotide sequences for PCR amplification of exons 4 to 9 of p53 for dHPLC Wave analysis.	217
Table D.7. ARMS primers for the amplification of the P60R and S315C mutations in p53.	218
Table D.8. Microsatellite markers used for MSI analysis.	218
Table D.9. ARMS primers for PCR amplification of eight mutations in the CDH1 gene.	219
Table D.10. ARMS primers for PCR amplification of the I1307K and E1317Q missense variants in the APC gene.	219
Table D.11. ARMS primers for PCR amplification of the D1853N variant in the ATM gene.	220
Table D.12. ARMS primers for PCR amplification of the G242A variant in the CCND1 gene.	220
Table D.13. ARMS primers for PCR amplification of the C677T variant in the MTHFR gene.	220
Table D.14. Primer sequences for the PCR amplification of the <i>GSTM1</i> and <i>GSTT1</i> polymorphisms.	220

<i>Table D.15. Primer sequences for generating PCR fragment for NAT2</i>	
<i>polymorphisms.</i>	<i>221</i>

University of Cape Town

SYMBOLS AND ABBREVIATIONS

°C	degrees Celsius
[$\alpha^{32}\text{P}$] dCTP	radio-isotope of phosphorus
μg	microgram
μM	micromolar
μl	microlitre
%	percentage
A	adenine
bp	base pair (s)
C	cytosine
cDNA	complementary deoxyribonucleic acid
dCTP	α -deoxycytosine-triphosphate
CRC	colorectal cancer
dHPLC	denaturing high performance liquid chromatography
DNA	deoxyribonucleic acid
dNTP	deoxynucleoside triphosphate
e.g.	for example
<i>et al.</i>	and others
Fig.	Figure
G	guanine
HCl	hydrochloric acid
HNPCC	hereditary nonpolyposis colorectal cancer
hr	hour
ICRF	Imperial Cancer Research Fund
kb	kilobase
KCl	potassium chloride
l	litre
lod	log of the odds
M	molar

Mg	magnesium
Min	minute (s)
ml	millilitre
mM	millimolar
MMR	mismatch repair
mRNA	messenger ribonucleic acid
NaCl	sodium chloride
NaOH	sodium hydroxide
ng	nanogram
O/N	overnight
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
pH	log H ⁺ ions
pmol	picomoles
SA	South Africa
Sec	second
SNP	single nucleotide polymorphism
SSCP	single strand conformation polymorphism
T	thymine
TBE	Tris-borate/EDTA
UCT	University of Cape Town
UK	United Kingdom
www	world wide web
USA	United States of America
■ □	affected/unaffected male
● ○	affected/unaffected female
⊠ ∅	deceased male/ female

GLOSSARY

Adenocarcinoma:	Cancer that begins in cells that line certain internal organs and that has glandular (secretary) properties.
Adenoma:	A noncancerous tumour.
Allele:	Individual variant versions of a gene. Individuals have two copies of each gene called alleles, one on each chromosome.
Amino acid:	The repeating unit of a protein.
Benign:	Not cancerous; does not invade nearby tissue or spread to other parts of the body. Not cancerous; does not invade nearby tissue or spread to other parts of the body.
CA-125:	CA 125: a protein that can be measured in the blood and an important tumour marker in ovarian cancer. Elevated levels of CA 125 are found in 80% of the patients with epithelial ovarian cancer.
Carcinogen:	Any substance that is known to cause cancer.
Carcinogenesis:	The process by which normal cells are transformed into cancer cells.
Carcinoma:	A malignant new growth that arises from epithelium, found in skin or, more commonly, the lining of body organs. Carcinomas tend to infiltrate

into adjacent tissue and metastasise to distant organs such as the bone, liver, lung or the brain.

Chromosome: A bundle of coiled DNA visible under a microscope. Humans have 22 pairs of chromosomes and two unpaired sex chromosomes.

Codon: A nucleotide triplet that specifies an amino acid or a translational stop signal.

Colectomy: An operation to remove all or part of the colon. In a partial colectomy, the surgeon removes only the cancerous part of the colon and a small amount (called a margin) of surrounding healthy tissue.

Colonoscopy: An examination in which the doctor looks at the colon through a flexible, lighted instrument called a colonoscope.

CpG dinucleotide: The sequence 5' CG 3' within a longer DNA molecule. CpG dinucleotides are targets of a specific DNA methylation system in mammals that is important in the control of gene expression.

CpG island: Short stretch of DNA often less than 1kb containing frequent unmethylated CpG dinucleotides. CpG islands tend to mark the 5' ends of genes.

Cyst: A sac or capsule filled with fluid.

DNA: Deoxyribonucleic acid, the molecules that code for most life-forms.

Dominant:	Describes any trait that is expressed in a heterozygote.
Dysplasia:	Abnormal cells that are not cancer.
Exon:	A segment of a gene that is represented in the mature RNA product. Individual exons may contain coding DNA and/or noncoding DNA.
Frameshift mutation:	A mutation that alters the normal translation reading frame of a mRNA by adding or deleting a number of bases that is not a multiple of three.
Gene:	A Sequence of DNA that codes for a particular protein.
Genetic Counselor:	A professional who is trained to discuss, interpret and advise people about genetic disease and present treatment options.
Genotype:	The genetic constitution of an individual either overall or at a specific locus.
Germ-line:	The germ cells and those cells, which give rise to them. A gene change in the body's reproductive cells (egg or sperm) that becomes incorporated into the DNA of every cell in the body of offspring; hereditary mutations are passed on from parents to offspring.

Heredity:	The biological transmission of features determined by genetic information from one generation to the next.
Hyperplasia:	A precancerous condition in which there is an increase in the number of normal cells lining the uterus.
Incidence:	The number of new cases of a disease diagnosed each year.
Informed consent:	Consent given by a patient allowing research to be undertaken on his/her biological material.
Intron:	Part of the DNA of a gene that does not code for protein.
Kindred:	A group of individuals related to each other either genetically or by marriage.
Lod score:	A measure of the likelihood of genetic linkage between loci. The log (base 10) of the odds that the loci are linked (with recombination fraction θ) rather than unlinked. For a Mendelian character, a lod score of +3 is evidence of linkage, one that is -2 is evidence against linkage.
Malignant:	Cancerous; can invade nearby tissue and spread to other parts of the body.
Metachronous tumour:	Multiple primary cancers that develop at different intervals.

Microsatellite:	A small run of tandem repeats (usually less than 0.1kb) of a very simple DNA sequence, usually 1-4bp e.g. (CA) _n .
Mutation:	Variations in DNA sequence of a gene that damages function and may cause disease.
Neoplasm:	Abnormal new growth of cells.
Nonsense mutation:	A mutation that occurs within a codon and changes it into a termination codon.
Oncogene:	A gene involved in the control of cell proliferation which, when overactive, can help transform a normal cell into a tumour cell.
Penetrance:	The frequency with which a genotype manifests itself in a given phenotype.
Polymorphism:	The existence of two or more variants at significant frequencies in the population; any sequence variant present at a frequency of more than 1% in a population.
Polyp:	A mass of tissue that projects into the colon.
Prevalence:	The proportion of individuals in a population having a disease.
Proband:	The family member whose disease led to diagnosis of the family's mutation.

CHAPTER 1: REVIEW OF COLON CANCER

University of Cape Town

CHAPTER 1: REVIEW OF COLON CANCER

1.1. INTRODUCTION

1.1.1. Cancer Genetics

Carcinogenesis is a multi-step process, which results in the activation of a dominantly-acting oncogene or inactivation of a tumour-suppressor gene. The normal behavioural response of the cell is then altered and this confers a growth advantage to the cell (Ilyas *et al.*, 1999). As these mutations increase in number and accumulate, a change occurs from normal cell growth to malignant behaviour. This process is characterised by unrestrained growth leading to invasion into the surrounding tissues and subsequent metastasis (Ilyas *et al.*, 1999). Since the mutation rate in human tissues is low, it is unlikely that a single mutation will give rise to the malignant phenotype. Mutations occur at different times and it follows then that a multi-step process occurs which leads to the “genetic pathway” for that particular cancer. In theory, this genetic pathway can be constructed for each cancer that occurs.

Tumorigenesis results when the cell escapes its growth-control mechanisms. Different tissues have different growth control mechanisms and thus the mutations involved will differ between tissues. Tumours that are derived from the same tissue will escape the same growth controls and thus may develop along similar genetic pathways.

Most cancers are sporadic i.e. the causative mutation only occurs in cells of the affected tissue. In contrast, a germline mutation is inherited from previous generations and is present in every cell of the body, including the gametes (Lewis, 2001).

Carcinogenic mutations usually affect the genes that control the cell cycle or apoptosis. Two main categories of genes can be distinguished *viz.* oncogenes and tumour suppressor (TS) genes. These encode for a diverse range of proteins whose physiological role involves the maintenance of normal growth.

Proto-oncogenes normally trigger cell division and are active where high rates of cell division are necessary. When proto-oncogenes are mutated, they are termed oncogenes and these have a role in the development of cancer (Lewis, 2001). Oncogenes are usually dominant growth stimulating genes.

Tumour suppressor gene products inhibit events leading to cancer. Mutant TS genes in cancer cells have lost their function and uncontrolled cell proliferation occurs. A proportion of TS genes prevent cell cycle progression, some steer deviant cells into apoptosis and others keep the genome stable and mutation rates low by ensuring accurate replication, repair and segregation of cellular DNA. Inactivation of both alleles of a TS gene is required to alter normal cell behaviour (Lewis, 2001).

Alfred Knudson (1971) developed the "two hit hypothesis", which is a fundamental concept in cancer genetics. Knudson hypothesised that two successive mutations (hits) were required to transform a normal cell into a tumour cell and that in familial forms, one of these hits were inherited. This two-hit hypothesis is demonstrated in the *Rb* gene responsible for retinoblastoma (RB), which has been an ideal model in the elucidation of TS genes. RB is a relatively rare tumour that originates in the retina and has an incidence of about 1 in 20,000. Familial retinoblastoma usually manifests from birth up to the 6th or 7th year of life. Wild-type expression of *Rb* counteracts tumour formation. Because *Rb* is a growth-inhibitory gene, both copies need to be inactivated within a single cell to release that cell from its growth inhibitory effects. Considering that the rate at which

somatic mutations occur in a mammalian cell is extremely rare (10^{-6} to 10^{-7}), the occurrence of two independent somatic mutations at both alleles of the same gene is likely to represent a very rare event (10^{-12} to 10^{-14}).

A further class of genes exist that are mutated in cancer cells. These do not play a direct role in controlling the cell cycle, but have a role as caretakers, ensuring the integrity of the genome (Strachan and Read, 1999). These loss-of-function mutations can be observed as nucleotide excision repair defects or as replication error defects.

- i. Nucleotide excision repair defects: When DNA has been damaged by ionising radiation, ultraviolet light or chemical mutagens, it needs to be repaired before the next round of replication. If loss-of-function mutations occur in repair enzymes, these mutations are not repaired and could lead to cancer-prone syndromes e.g. xeroderma pigmentosum (XP). No somatic second hit occurs in the repair gene as the damaged DNA can be considered as the somatic mutation.
- ii. Replication error repair defects: patients with hereditary nonpolyposis colorectal cancer (HNPCC) are heterozygous for a loss-of-function mutation. Their normal cells have one mutated copy of the mismatch repair (MMR) system that recognises mismatched DNA bases after replication. Normal functioning of the MMR system does, however occur in these cells. The tumour DNA of individuals with HNPCC contains extra alleles when amplified with microsatellite markers, which is a feature of this disorder. The second copy of the MMR system is lost in the tumour, resulting in the incorporation of mismatched bases.

1.1.2. Cancer of the colon and rectum

Cancer of the colon and rectum is one of the leading causes of mortality due to neoplasia in Westernised countries (Nicum *et al.*, 2003) after lung and

breast cancer (Houlston and Tomlinson, 2001). The highest incidence of the disease occurs in the 7th decade of life and is fairly equally distributed between men and women (Ilyas *et al.*, 1999). It has been estimated that at least half a million new cases occur worldwide each year (Finlay and Pickford, 1994). It is the second leading cause of cancer-related deaths in the United States and Canada, following lung cancer (Greenlee *et al.*, 2000) with approximately 160 000 new cases and 60 000 deaths per year (Parker *et al.*, 1999). In England and Wales, approximately 30 000 new cases are diagnosed each year, resulting in 17 000 deaths (HMSO mortality statistics; HMSO Cancer statistics). Colon and rectal cancer are collectively referred to as colorectal cancer (CRC). Prognosis depends on the stage at which the disorder is diagnosed. At present, surgery is the main form of treatment. Progress in understanding the biology of CRC has led to improved management due to early detection and subsequent prevention of the disorder.

CRCs can be divided into the hereditary and non-hereditary forms. In addition, there are the classes that do not appear to be inherited in a Mendelian fashion due to the effects of various susceptibility genes resulting in different penetrances. The inherited forms of CRC are further classified into two groups *viz.* those that do not manifest primarily with polyps (known as 'non polyposis' syndromes) and those that do (known as 'polyposis' syndromes) (Cohen and Winawer, 1995). The nonpolyposis syndromes include hereditary nonpolyposis colorectal cancer (HNPCC). The inherited polyposis syndromes include familial adenomatous polyposis (FAP) and the hamartomatous polyposis conditions of Peutz-Jeghers syndrome and familial juvenile polyposis (Haggitt and Reid, 1986). Gardner's syndrome, which results in colonic polyps with epidermoid skin cysts and benign osteoid tumours of the mandible and long bones is now recognised as a variant of FAP (Fearnhead *et al.*, 2001). Each of these polyposis syndromes is inherited in an autosomal dominant mode with a high penetrance. HNPCC accounts for approximately 2-7% of all cases of CRC (Muller and Fishel,

2002). FAP has an occurrence of approximately 1 in 10 000 births, while Peutz-Jeghers syndrome and familial juvenile polyposis are approximately one tenth as common (Cohen and Winawer, 1995). The hereditary forms together account for approximately 10% of all CRC (Cohen and Winawer, 1995). FAP and HNPCC account for a large proportion of these familial cases. The non-hereditary forms of the condition are referred to as 'sporadic' cases.

Hereditary forms of CRC have facilitated access to large families with many affected individuals leading to gene localisation. Recently, several genes and mutations underlying the hereditary forms of CRC have been identified.

Internationally, molecular genetic research aims to elucidate the precise disease-causing mutations in each family with CRC. Identification of all CRC-causing mutations has been recognised as a valuable means for management and treatment of at-risk and affected individuals respectively.

This chapter provides an overview of pertinent aspects of CRC including anatomy, pathology, molecular genetics and clinical management as a backdrop for the research in the body of the thesis.

1.2. ANATOMY OF THE COLON

A brief review of the structure of the colon is essential towards an understanding of CRC.

The colon extends from the ileocaecal valve to the anus, measures approximately 1.5m in length and has a diameter of 5cm at its midpoint. It is divided into various sections (Fig. 1.1).

- i. The caecum, which leads into the ascending colon, is the most proximal section

- ii. The transverse colon then continues horizontally towards the splenic flexure
- iii. The descending colon extends from the splenic flexure to an S-shaped curve known as the sigmoid colon
- iv. The latter is followed by the rectum, which tapers to form the anal canal.

The colon is made up of smooth muscle, lined with the mucosa (Northover and Kettner, 1992), which has a dual function *viz.* the production of mucus, which lubricates the faeces, and absorption of water from the intestinal contents. The main function of the large bowel is the collection, concentration, transport and temporary storage of intestinal waste products.

The rectum acts as a storage reservoir for waste products before they leave the body. To some extent, the colon and rectum control the concentrations of water and electrolytes in the body.

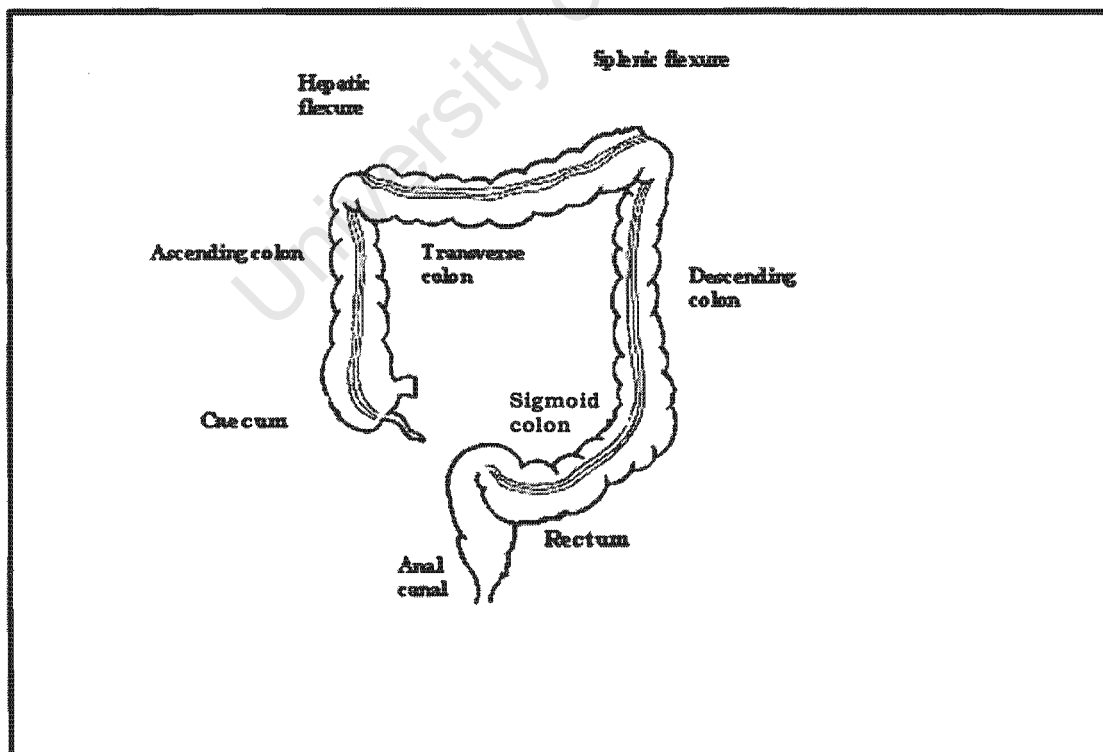


Figure 1.1. Components of the large intestine.
(Adapted from Kirsner and Shorter, 1988)

1.2.1. Embryological development

The endodermal roof of the embryonic yolk sac gives rise to the primitive gut tube (Fig. 1.2). The caudal portion of this tube invaginates into the developing tailfold of the embryo and becomes the hindgut (Kirsner and Shorter, 1988) and the cranial portion, within the headfold, becomes the foregut. The midgut is the central portion of the tube. The embryonic hindgut gives rise to the descending colon, the rectum and the anal canal whereas the midgut gives rise to the terminal ileum, caecum, appendix and the proximal colon (Kirsner and Shorter, 1988).

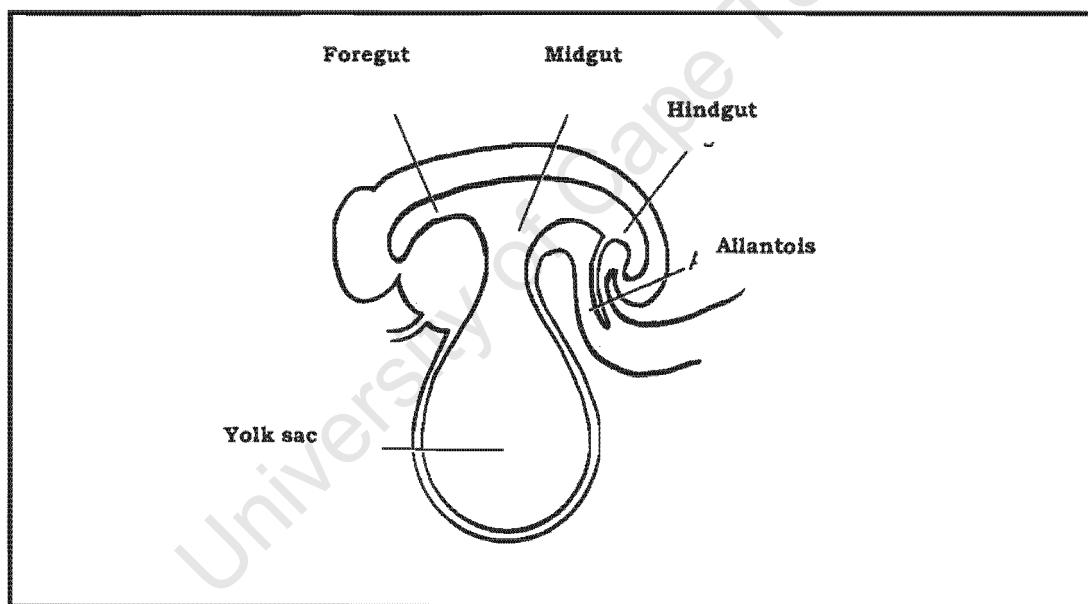


Figure 1.2. *Development of the human large intestine – structure and position of the hindgut, midgut and foregut in a 20day old embryo.*
(Kirsner and Shorter, 1988).

After the 6th week of embryonic development, the elongation of the gut tube leads to a U-shaped prolapse of the midgut into the umbilical cord. This hernia is present until the end of the 10th week of embryological

development, when the abdomen accommodates all the viscera. The midgut twists in an anti-clockwise direction around the axis of the superior mesenteric cavity as it returns to the abdominal cavity. Subsequently, the caecum passes in front of the developing coils of the small intestine to occupy the lower right quadrant of the abdominal cavity. The caecum and the ascending colon on the right and the descending colon on the left then become closely applied to the dorsal body wall. Their mesenteries will eventually get fused to the posterior parietal peritoneum. The transverse colon is the last portion of the midgut to return to the abdominal cavity and remains freely mobile anterior to the small intestine.

1.3. MOLECULAR GENETICS OF HEREDITARY COLORECTAL CANCER

The hereditary CRCs are comprised largely of HNPCC and FAP.

1.3.1. Hereditary nonpolyposis colorectal cancer (HNPCC)

The earliest reports on HNPCC include the first by Warthin (Warthin, 1913) and that of Lynch *et al.* (Lynch *et al.*, 1966) where the syndrome was described in 2 families.

HNPCC is an autosomal dominant disorder (Potter, 1999) with a penetrance of 70-85% (Peltomaki, 2001a). Clinical features manifest, on average, around 41 years of age (Muller and Fishel, 2002). Affected individuals may exhibit a few adenomatous polyps, but not diffuse polyposis. These polyps are endoscopically and histologically indistinguishable from adenomas that occur in the general population (Cohen and Winawer, 1995). Tumours of the colon are more common than rectal tumours occurring in about 70% of patients (Bellacosa *et al.*, 1996). There is increased occurrence of synchronous and metachronous CRCs in the proximal colon in patients affected with HNPCC (Muller and Fishel, 2002).

Two types of HNPCC families can be distinguished. Type I HNPCC tumours are located exclusively in the colon. Individuals who are diagnosed with Type II HNPCC have all the features of Type I but also harbour extracolonic cancers (stomach, endometrium, ovary, and urinary tract) (Muller and Fishel, 2002). A study of mutation-positive HNPCC individuals by Aarnio *et al.* (1999) found that women with HNPCC have a higher chance of developing endometrial cancer (60%) than colon cancer (54%) during their lifetime. Interfamilial phenotypic variation occurs, resulting in certain families with different cancer risks to others (Hampel and Peltomaki, 2000). The extracolonic cancers that do occur complicate the clinical management of this disorder.

1.3.1.1. Diagnostic criteria

The variable clinical characteristics observed in HNPCC patients and families highlighted the need for a specific definition of the disorder, which would facilitate more directed HNPCC research. In addition, no single specific characteristic of the syndrome exists. HNPCC is diagnosed in patients on familial criteria. The classic diagnostic criteria are the so-called “Amsterdam criteria”, which were proposed to standardise clinical studies (Vasen *et al.*, 1991). These criteria include:

- i. At least three relatives should have histologically verified CRC. One of the affected individuals should be a first-degree relative of the remaining two. FAP should be excluded
- ii. Affected individuals should span at least two successive generations
- iii. CRC should be diagnosed before the age of 50 years in at least one of the affected individuals.

An early detection program was also recommended for at-risk individuals in families diagnosed with HNPCC. The Bethesda criteria were proposed by the National Cancer Institute (NCI) workshop in 1997 (Rodriguez-Bigas *et al.*, 1997). These criteria allow for the inclusion of smaller families, as well as

individuals with cancer of organs other than the colon and rectum. The “Bethesda Criteria” include: (Rodriguez-Bigas *et al.*, 1997):

- i. Individuals who meet the Amsterdam criteria
- ii. Individuals with two HNPCC-related cancers:
synchronous/metachronous CRCs, endometrial, ovarian, gastric, hepatobiliary, small intestine, or renal tract transitional cell cancers
- iii. Individuals with CRC who have a first-degree relative with CRC, or any other HNPCC-related cancer or adenoma. One of the cancers should be diagnosed before 45 years of age and the adenoma before 40 years
- iv. Individuals with endometrial or CRC diagnosed before 45 years
- v. Individuals with cancer of the proximal colon diagnosed by age 45 years, with an undifferentiated pattern histopathologically
- vi. Individuals with signet-ring-cell-type CRC diagnosed by 45 years
- vii. Individuals with adenomas diagnosed by age 40 years

The Amsterdam criteria were not meant to formulate a new definition for HNPCC, but rather to serve as a guideline ensuring uniformity in multi-centre studies (Vasen *et al.*, 1991). These criteria continue to promote discussion, as they are overly restrictive when compared to more recent findings based on molecular diagnosis. To resolve this problem, new clinical criteria were introduced, known as the Amsterdam criteria II (Vasen *et al.*, 1999). These take into account the extracolonic manifestations that have been associated with HNPCC:

- i. There should be at least 3 relatives with HNPCC-associated cancer (i.e. cancer of the colon and rectum, endometrium, small bowel, ureter or renal pelvis). FAP should be excluded
- ii. One individual should be a first degree relative of the other two
- iii. At least two successive generations should be affected
- iv. At least one individual should be diagnosed before the age of 50
- v. Tumours should be verified by pathological examination.

There appears to be evidence that the Bethesda criteria are more sensitive and less specific in identifying HNPCC families with disease-causing mutations than either the Amsterdam criteria I or II (Hampel and Peltomaki, 2000).

A classification of tumours/neoplasms into various stages was devised based on the degree to which it has spread. The 'Dukes' classification system comprises the following stages:

- i. Dukes' A: a cancer that affects the innermost lining of the colon or rectum
- ii. Dukes' B: the cancer has grown into the muscle layer of the colon or rectum
- iii. Dukes' C: the cancer has spread to at least one lymph node in the area and
- iv. Dukes' D: the cancer has metastasised to another site in the body e.g. the liver or lung.

Various procedures such as faecal occult blood tests, digital examinations, rigid sigmoidoscopy, flexible colonoscopy and barium enemas could aid in the diagnosis of the disorder. It is not often possible to 'stage' the cancer precisely until after a resection; however, various tests such as chest X-rays, liver ultrasonic scans, abdominal computerised tomography (CT) scans, liver function tests, CT scans of the pelvis, endoluminal ultrasound in rectal cancer and Magnetic Resonance Imaging (MRI) may provide the physician with an indication of the stage, which could guide the management of the disorder.

1.3.1.2. The adenoma-to-carcinoma sequence

Adenomas do form in the large bowel of individuals affected with HNPCC and there is strong evidence that these are the precursors to malignancy in this

disorder (Jass, 2000). There does not seem to be a definitive morphological parameter that distinguishes colorectal adenomas occurring in HNPCC from those occurring in sporadic CRC cases. However, adenomas in individuals with HNPCC may grow more rapidly than those that occur in sporadic cases. There appears to be an accelerated rate of adenoma-to-carcinoma occurring with HNPCC-associated mutations and/or accelerated growth of cancer arising *de novo* (Moslein, 2003). In addition, patients with HNPCC develop no more colonic adenomas than the general population. Jass *et al.* (1994) suggest that patients with HNPCC adenomas differ from those in the general population in their accelerated progression to cancer, but not in their frequency. This acceleration of tumourigenesis is regarded as the hallmark of the “Mutator Phenotype” and the increased mutation rates associated with MMR mutations (Muller and Fishel, 2002).

HNPCC produces poorly differentiated or mucinous carcinoma (Percesepe *et al.*, 2001), which usually results from an adenoma (Jass, 2000). A higher proportion of adenomas are distributed in the proximal colon of HNPCC patients (Muller and Fishel, 2002).

1.3.1.3. Clinical surveillance and management

The ICG-HNPCC has proposed a set of guidelines for the clinical surveillance of HNPCC individuals (Weber, 1996). They suggested that individuals who belong to HNPCC families should receive their first colonoscopy at age 20 to 25, and every 2 years thereafter. They also suggested screening for endometrial cancer in females at age 30 to 35 with a follow-up every 1 to 2 years. Any other cancers included in the HNPCC spectrum should be screened for if there is a family history. In 1994, a Cancer Genetics Studies Consortium (CGSC) was established to assess the social and psychological effects of testing for cancer-predisposing mutations (Burke *et al.*, 1997). A task force was convened by the CGSC to evaluate the appropriate surveillance, treatment options, and follow-up for individuals who have an inherited predisposition to cancer. This task force recommended that

colonoscopies begin at 20 to 25 years, but they suggested a 1 to 3 year interval, as there is insufficient knowledge as to the most appropriate interval. It is generally accepted that colonoscopic screening should be specifically tailored for families when the complete natural history of HNPCC is not known for the particular kindred. In addition, it was suggested that annual screening for endometrial cancer be carried out, beginning between the ages 25 and 35. Evidence now exists that colonoscopy every 1 to 2 years in individuals with HNPCC leads to a 7 to 9 year increase in life expectancy (Hampel and Peltomaki, 2000). Furthermore, the financial costs in countries such as the United States are less to monitor individuals who are carriers of mutated MMR genes, regardless of the frequency of the screening, as opposed to no screening at all (Vasen *et al.*, 1998).

Table 1.1 represents the various suggested screening periods for extracolonic tumours in families with HNPCC (Menko, 1993).

Table 1.1. Details of screening recommendations for extracolonic tumours in HNPCC.

REFERENCE	SCREENING RECOMMENDATION	INTERVAL	RECOMMENDED START AGE
Fitzgibbons, 1987	• Pelvic examination ^a • Endometrial aspiration biopsy ^a • Echogram of the ovaries • Other, depending on pedigree	• Every two years • Annually	• NA
Houlston, 1990	• Mammography, echogram • Mammography • Mammography	• Annual • Annual • As in national screening program	• 25 • 40 • 50
Vasen, 1990	• Echogram of the pelvis • Endometrial aspiration biopsy • Upper GI endoscopy ^c	• Annual • Every 2 years ^b	• 25 • 30
Mecklin and Jarvinen, 1991	• Regular urine cytology for patients with cancer • Consider gastroduodenoscopy • Liver function tests • Abdominal ultrasound • Urine cytology		
Trimble, 1991	• Endometrial biopsy • Echography of the pelvis • Serum CA 125 estimation	• Annual • Every 6 months	• 25 • 30
Lynch, 1992	• Endometrial cytology and histology • Vaginal ultrasound of ovaries, serum CA 125 estimation	• Annual	

a: starting at the same age as the colonoscopy, 5 years earlier than earliest colonic cancer in the family

b: exception: a single case diagnosed at an advanced age

c: if the incidence of stomach cancer is very high in the family.

NA: not applicable

The discovery of the genetic defects underlying HNPCC and the capacity to provide presymptomatic "genetic diagnosis" has allowed clinicians to advise those individuals at risk about the possibility of a prophylactic colectomy. The recommended treatment for an affected HNPCC patient is total abdominal colectomy and ileorectal anastomosis (TAC-IR) (Rodriguez-Bigas,

1996) in order to prevent metachronous colon cancer. The CGSC task force suggested that a subtotal colectomy with ileorectal anastomosis be considered in HNPCC-associated mutation-carriers who have colon cancer. In addition, this procedure is also suggested prophylactically for selected HNPCC-associated mutation-carriers who have adenomas at the time of surveillance (Burke *et al.*, 1997). Individuals with rectal cancer should be offered a proctocolectomy. This recommendation is also an option for mutation-carriers who present with adenomas at the time of examination. Despite these recommendations, surgical treatment options for CRC in HNPCC remain controversial. Effort is currently under way in Germany to initiate a prospective randomized trial on the surgical treatment of HNPCC at the time of colon or rectal cancer (Moslein, personal communication, 06/11/2003). In addition, should prophylactic hysterectomy and oophorectomy be suggested to affected mutation-carriers at the time of CRC surgery (Pistorius, 2001)?

It is recommended that the patient continues regular surveillance due to the lifetime risk of rectal and extracolonic cancers. A study carried out by the ICG-HNPCC showed that the risk of developing rectal cancer was 12% at 12 years after abdominal colectomy (Rodriguez-Bigas *et al.*, 1997b). The median age at diagnosis of rectal cancer after colectomy was 51 years (range 46-73 years).

The age at which the prophylactic TAC-IR should be performed in individuals at high risk for HNPCC has not been established, however, this varies considerably within and between families. The average age at diagnosis of HNPCC is 45 years. Rodriguez-Bigas (1996) showed that 27% of his patients were diagnosed before the age of 39 years and 88% by 69 years. The average age at diagnosis was 48 years. Vasen *et al.* (1994) reported that 92% of CRC had developed in their patients by the age of 60 years.

The diagnosis of HNPCC is largely dependent on a well-defined family history. However, diagnosing HNPCC is now significantly enhanced by molecular genetic studies (Lynch and Lynch, 1998).

Chemotherapy is not usually considered as a treatment for HNPCC since MMR-deficient cells are tolerant to killing by alkylating agents (Branch *et al.*, 1995). These agents are not recommended as it may lead to the enrichment of tumour cells (Fink *et al.*, 1998). MMR proteins can act on DNA damage caused by alkylating agents, including those used in cancer chemotherapy, heterocyclic amines or oxidation (Li *et al.*, 1996; Yamada *et al.*, 1997; deWeese *et al.*, 1998). MMR-deficient cells are more resistant to these alkylating agents than MMR-proficient cells (Karran and Hampson, 1996). In the latter, repeated attempts to remove the mismatched bases incorporated opposite the damaged DNA base may lead to persistent DNA strand breaks, recognised by cellular damage sensors that induce subsequent apoptosis. In MMR-deficient cells, alkylation tolerance occurs as the alkylated base damage remains in the DNA. The MMR system can initiate a damage-signalling cascade that can lead to G2/M-arrest or cell death in response to the alkylation system (Hawn *et al.*, 1995). The presence of mutagenic lesions may enhance tumourigenesis in MMR-deficient cells and may contribute to the "HNPCC tumour spectrum" (de Wind *et al.*, 1998). Management regimes including the in/appropriateness of chemo- and radiotherapy for the Lynch spectrum of cancers for individuals affected with CRC will evolve as knowledge of the biology of the condition increases.

1.3.1.4. Genetic testing and counseling for HNPCC

New discoveries in the genetics of CRC have aided in the diagnosis and management of the disorder. It is important for clinicians to understand factors such as test validity and disease heterogeneity, and that this information is communicated to the patient (Young *et al.*, 1996).

Clinical diagnosis for HNPCC can only be verified by detecting a molecular defect in a patient or family. Individuals should be informed about certain factors prior to any genetic test. These include the natural history of the disorder, the purpose of the genetic test and the limitations thereof; these include possible human and technical error in the laboratory (Lynch *et al.*, 1996). Individuals who are mutation-positive (disease-predisposing) need to be aware of the potentials and limitations of colonoscopic and additional surveillance for the disorder and contextualise this with the option of prophylactic surgery (Anwar *et al.*, 2000).

In addition to dealing with the burden of being diagnosed as mutation-positive and the effect that it has on their own lives, newly-diagnosed patients need to consider how it will impact on other family members. There is likely to be a heavy emotional burden, associated with the knowledge that an individual is the carrier of a disease-predisposing mutation (Lerman *et al.*, 1991). This information could have an adverse psychological effect (fear, denial) on the patient. On the other hand, a mutation-positive result could be beneficial to the individual and allows for informed decision-making regarding life issues about the future and relieves uncertainty. The at-risk individual knows that monitoring for early signs of the disorder is essential (Baba, 1997). Those individuals, whose results indicate that they do not carry a disease-predisposing mutation i.e. negative, could also undergo psychological distress (Huggins *et al.*, 1992; Lynch *et al.*, 1993). If emotional and psychological needs are not addressed, comprehension of the genetic information given to individuals may be reduced.

Presymptomatic testing for late onset genetic disorders became possible initially for Huntington Disease (HD) in the early 1990s (Harper, 1991). The experience gained in testing and setting up guidelines for HD has since been used to establish protocols for a variety of genetic disorders. Thus far, the counseling programs for predictive genetic testing for HNPCC have more or less followed those guidelines set up for HD. These programs include a minimum of 2 pre-test educational sessions and several post-test sessions combined with psychological support (Aktan-Collan *et al.*, 2000).

Counseling should begin with preparatory sessions and individuals that have requested testing should do so voluntarily. They should be given the most updated information concerning the available tests in order to make an informed, voluntary decision (Young *et al.*, 1996). Individuals should realise that they may withdraw from the testing procedure at any stage. In these sessions individuals ought to be educated about the hereditary nature of the disease, the available tests and the potential implications of test results. A disclosure session follows, in which the results of the gene test are disclosed to the individual. If a disease-causing mutation is found in one of the MMR genes, a follow-up session is advised in which the clinician will decide whether this individual requires any additional support.

In a study undertaken by Aktan-Collan (2000) and colleagues, 2% of mutation-positive individuals, who went through a predictive genetic testing counseling session, felt they needed psychological support one year after receiving their test results. They concluded that the most distressing part of the predictive testing procedure was in most instances over one year later and that the individuals had adjusted to their situation. In addition, they perceived that the need for support was greatest at the time of decision-making, when waiting for the results and especially at the moment of receiving the test results. The genetic test result should not be disclosed to a third party without the individual's consent.

A pertinent problem is the issue of medical insurance for coverage for surveillance of high-risk individuals (Lynch *et al.*, 1993). Insurance companies may refuse to cover the cost of these persons, which could lead to poor compliance by the family concerned.

1.3.1.5. The Mismatch repair (MMR) system and HNPCC

In 1993 linkage was shown between HNPCC and a locus on chromosome 2 (Peltomaki *et al.*, 1993). Two large kindreds with HNPCC, one from North

America and one from New Zealand, were investigated. Three hundred and forty five microsatellite markers were analysed before linkage between the disease phenotype and the marker D2S123 was demonstrated. The first family demonstrated a LOD score of 6.39 at a recombination fraction of zero, while the second family showed a LOD score of 1.45. At that stage the disease gene was located to chromosome 2p15-16. This gene was later localised to chromosome 2p22-p21 as *hMSH2* (**h**uman **M**utS **h**omologue 2) (Fishel *et al.*, 1993; Leach *et al.*, 1993). This gene is the human homologue of the bacterial MutS, accounts for approximately 60% of HNPCC cases and functions as a mismatch repair (MMR) gene (Service, 1994). If a mutation exists in the MMR pathway it could lead to an increased rate of spontaneous mutations. This increase in mutation rate and the genomic instability associated with it would then lead to the development of HNPCC (Fishel *et al.*, 1993). A feature of tumour cells in HNPCC is the rearrangement of genetic material, including chromosomal loss. It is postulated that the loss of *hMSH2* function could increase the frequency of homologous recombination, leading to chromosomal loss and/or rearrangement (Fishel *et al.*, 1993). The *hMSH2* gene consists of 16 exons and encodes a 934 amino acid protein reviewed in Jacob and Praz (2002).

It became clear that heterogeneity exists for HNPCC since a number of families did not map to the *hMSH2* locus on chromosome 2p. Lindblom *et al.* (1993) studied 3 families with HNPCC who did not map to chromosome 2p. One of these families showed linkage between HNPCC and a marker on chromosome 3p21. This prompted them to use additional microsatellite markers in this region, and they obtained a LOD score of 3.01 at a recombination fraction of zero with marker D3S1029, strongly suggestive of linkage. In the second family the LOD score was below the accepted significant level of linkage and they could not conclude definite linkage between HNPCC in this family and a locus on chromosome 3. The third family showed no evidence of linkage to chromosome 3p or 2p. The defective gene causing the disorder in a subsequent subset of families with HNPCC

was later localised to chromosome 3p21-23 (Bronner *et al.*, 1994; Papadopoulos *et al.*, 1994) and shown to be the *hMLH1* (**h**uman **M**utL **h**omologue 1) gene, which is the human homolog of the bacterial and yeast MutL gene. The *hMLH1* gene consists of 19 exons and encodes a 756 amino acid protein reviewed in Jacob and Praz (2002). The MLH1 protein shares 41% identity with its yeast homolog (Bronner *et al.*, 1994).

Mutation rates in tumour cells with a deficient MMR system are 100 to 1000 fold higher when compared with normal cells (Baba *et al.*, 1997; Lynch *et al.*, 1997; Peltomaki, 2001).

Despite intensive investigation of *hMSH2* and *hMLH1* genes, some HNPCC families remained unmapped. A search for additional genes was undertaken and subsequently, mutations in other MMR genes associated with HNPCC have been identified. The *hPMS1* (**h**uman **p**ost**m**eiotic **s**egregation 1) gene has been localised to chromosome 2q31-33 and the *hPMS2* (**h**uman **p**ost**m**eiotic **s**egregation 2) gene to chromosome 7p21 (Nicolaidis *et al.*, 1994). Two kindreds with mutations in *hPMS2* have been reported to date. A germline mutation in PMS1 was reported for an HNPCC-like family, but re-examination of this family revealed an additional mutation in MSH2 that co-segregated with the disease phenotype whereas the mutation in PMS1 did not. There is thus no evidence presently that mutations in PMS1 cause HNPCC (Peltomaki, 2003). Miyaki *et al.* (1997) reported a family with Type II HNPCC who displayed a mutation in the *GTBP* (**G**uanosine/**T**hymidine [G/T] **b**inding **p**rotein) gene, and this gene was subsequently termed *hMSH6* (Baba, 1997). The International Collaborative Group on HNPCC (ICG-HNPCC) have characterised more than 400 mutations in MMR genes in over 500 families to date and this can be found on their web-site at <http://www.nfdht.nl> (Peltomaki, 2001; Quaresima *et al.*, 2003).

The MMR genes function in a manner similar to tumour suppressor genes. Thus, loss of both alleles (one due to inheritance of a mutated copy and the

other because of a somatic mutation) causes accumulation of mutation and altered expression of corresponding MMR proteins. This leads to an increased mutation rate with the subsequent accumulation of mutations in tumour suppressor genes and proto-oncogenes required for malignant transformation and progression.

Recent data suggests that mutations in the various MMR genes give rise to different clinical scenarios (Anwar *et al.*, 2000). It appears that *hMSH2* mutations result in a higher proportion of patients suffering from ovarian, urinary tract and gastric cancer than *hMLH1* mutations (Vasen *et al.*, 1996). In addition, mutations in *hMSH6* show a predilection to endometrial and ovarian cancer (Anwar *et al.*, 2000). Research is now focussing on possible genotype/phenotype correlations with these mutator genes.

Perucho (1996) has proposed a model based on a 'mutator that mutates the other mutator' theory. This author suggests that primary mutators such as *hMLH1* and *hMSH2* cause frameshift mutations in secondary mutators such as *hMSH6* and *hMSH3*, which ultimately result in the MMR process.

Recent studies show that there might be additional genes that have a role in the development of HNPCC (Fig. 1.3). In a study by Yamamoto *et al.* (1998) 32% of the families fulfilling the Amsterdam criteria had no detectable microsatellite instability (see section 1.3.1.8), leading the authors to conclude that genes other than those involved with MMR may play a role in the tumourigenesis of HNPCC. Evidence suggests that apoptotic-response genes such as *BAX* may contribute to the adenoma-to-carcinoma transition in HNPCC tumourigenesis (Yagi *et al.*, 1998). Parsons *et al.* (1995) found that 90% of HNPCC tumours had mutations in the transforming growth factor Beta II receptor (RII) gene (*TGFBR2*). Lu *et al.* (1998) found a mutation in the coding region of the *TGFBR2* gene in an HNPCC family. The tumourigenic pathway in the reported family is suggested to be different from the typical mutator phenotype pathway found in HNPCC. TGF-B

inhibits the growth of colon epithelial cells. If the receptor is inactivated, then the inhibition of growth is not accomplished. This particular receptor may be susceptible to mutations in tumours from persons with defective MMR function since the gene contains a sequence of 10 consecutive adenine nucleotides, which could be prone to mismatch errors.

University of Cape Town

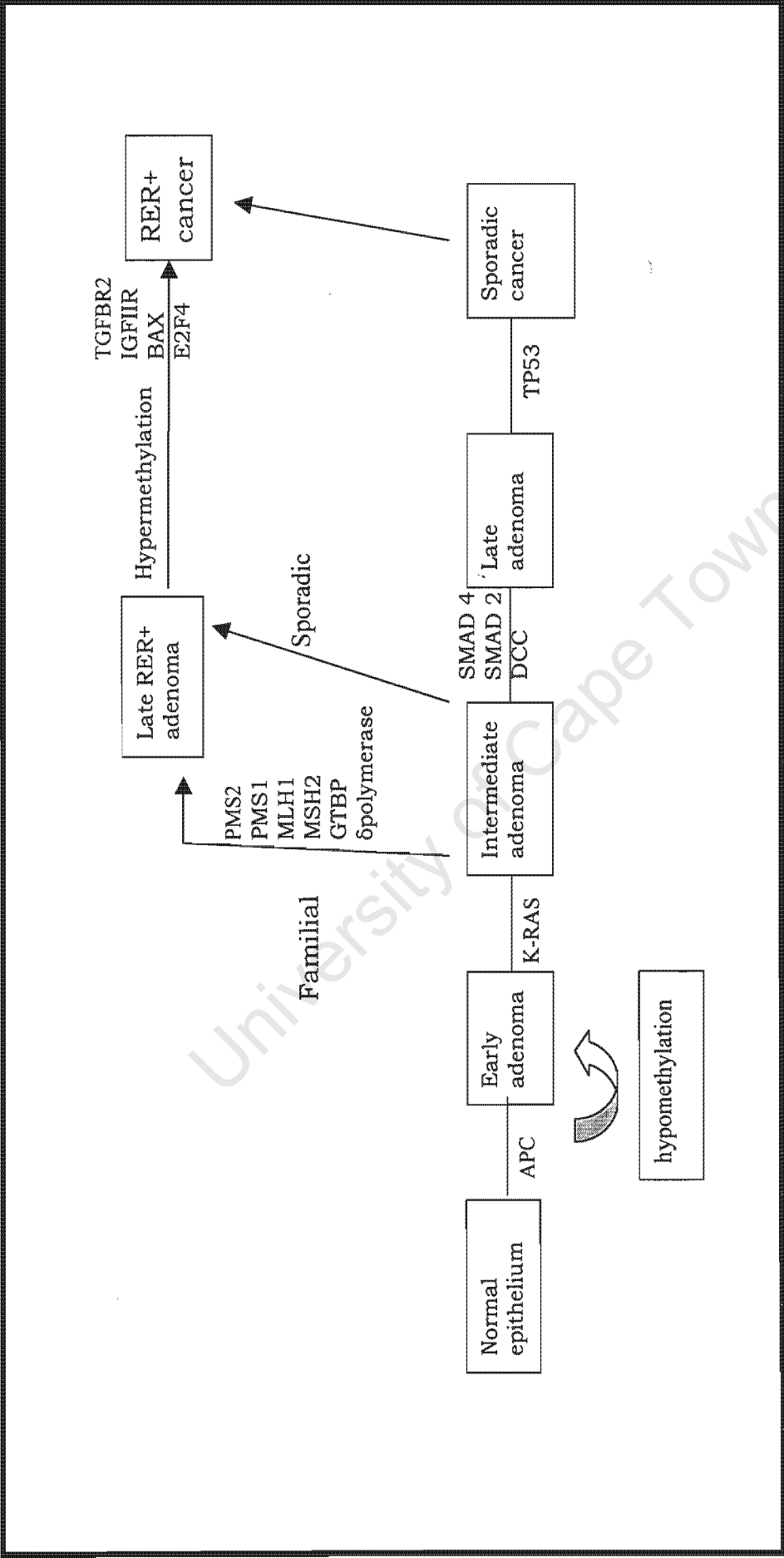


Figure 1.3. Putative genetic pathways in colorectal cancer.
(Ilyas et al., 1999)

1.3.1.6. Germline Mutations of *hMSH2* and *hMLH1*

Over 150 mutations have been described in *hMSH2* accounting for approximately 40% of disease-causing mutations detected in HNPCC kindreds (Jacob and Praz, 2002). Approximately 80% of these mutations lead to a truncation of the protein product.

Nearly 200 different mutations have been described in *hMLH1* in individuals with HNPCC, which represents approximately 55% of mutations detected to date. A cluster of mutations has been found in the region spanning exon 15 and 16 (Wijnen *et al.*, 1996). Most of the *hMLH1* mutations that have been identified to date result in a truncated gene product either due to frameshift or splice-site mutations, leading to exon deletion, which occurs at an unusually high frequency (Jacob and Praz, 2002).

The interaction domains of the three MutL homologues (*hPMS2*, *hMLH3* and *hPMS1*) have recently been determined. The C-terminal domain of MLH1 interacts with 36 amino acid residues that are conserved within PMS2, MLH3 and PMS1 (Guerrette *et al.*, 1999; Kondo *et al.*, 2001). A significant loss of MLH1 interaction with PMS2 has been observed for the HNPCC missense mutations located in the C-terminal region of MLH1 (Guerrette *et al.*, 1999; Kondo *et al.*, 2001). The majority of mutations reported are either into the ATPase domain or the MutL interacting domain reviewed in Jacob and Praz (2002). Based on the crystal structure of the 40kDa ATPase fragment of the *E.coli* MutL, it has been proposed that the missense mutations that occur at the N-terminal conserved region of MLH1 either affect ATP-binding and ATPase activity or alter the overall folding of the protein (Ban and Yang, 1998).

1.3.1.7. Microsatellite Instability (MSI)

Microsatellite sequences are comprised of short repetitive units of DNA, 1-6 base pairs in length, which occur approximately every 50 000-100 000 base

pairs (Saletti *et al.*, 2001) and have a high polymorphism content in the eukaryote genome (Litt and Luyt, 1989; Weber and May, 1989). These stretches of DNA should be the same length at the same locus in the cells of one person, but vary between individuals. MSI -formerly known as "replication error" (RER) phenotype (Aaltonen *et al.*, 1993; Peltomaki *et al.*, 1993)-is defined as a relatively frequent change in length of these loci due to either an insertion or a deletion of the repeated units (Wheeler *et al.*, 2000). This instability has been reported in an increasingly large subset of tumours reviewed by Hussein and Wood (2002). According to the extent of MSI, tumours have been classified into three groups:

- i. MSI-L: low frequency MSI; 1-30% unstable loci
- ii. MSI-H: tumours with high frequency MSI; more than 30% unstable loci
- iii. MSS: microsatellite stable: 0% unstable loci (Hussein and Wood, 2002).

MSI is usually the result of slippage of the DNA polymerase on the template strand, resulting in the formation of a small loop in either the template or the nascent strand (Levison and Gutman, 1987). This loop is detected by the MMR system in normal cells, but this may not occur optimally in the tumour, which is replicating at an accelerated rate (Marx, 1994). These loops may become permanent and alleles of different sizes will be formed at the next round of replication (Wheeler *et al.*, 2000). The mutation rate in microsatellites in the germline is estimated to be in the order of 10^{-3} to 10^{-4} (Weber and Wong, 1993; Talbot *et al.*, 1995).

MSI-H tumours have inactivating mutations of the MMR genes and wide destabilisation of microsatellite repeats in the genome (Hussein and Wood, 2002). In addition, these tumours have altered lengths of the microsatellite sequences in some cancer-related genes containing mononucleotide repeats in their coding region. These genes include *TGF β RII*, insulin-like growth

factor II receptor (*IGF1R*), the proapoptotic gene *BAX*, and the MMR genes *hMSH6* and *hMSH3* (Boland *et al.*, 1998). These genes are particular targets in the hypermutability pathway driven by DNA MMR deficiency. In addition, these mutations affect growth-regulatory genes including the genes that play a role in the multi-step pathway of colon tumourigenesis such as *APC*, *KRAS* and *p53* (Peltomaki, 2001a). Mismatch repair deficiency leads to the selection of these and other genes resulting in the extracolonic malignancies observed in HNPCC. The underlying mechanisms of MSI-L tumours are not clearly understood, but the presence of this instability appears to be of diagnostic value in tumour tissue from urothelial, colorectal and bronchial cancers. Lu *et al.* (1998) examined tumours from five MSS HNPCC kindreds and screened the coding region of the *TGFBR2* gene. Their data suggests that mutations in *TGFBR2* and other as yet unknown genes could be causative of MSS tumours in HNPCC. These MSS tumours generally have a better prognosis than MSI-L tumours (Hussein and Wood, 2002).

An increase in the somatic mutation rate of microsatellites in tumour tissue has been observed in patients with HNPCC-related CRC with mutations in the MMR genes (Sturzeneker *et al.*, 1998). The majority of tumours (approximately 95%) occurring in individuals with HNPCC display MSI due to defective MMR (Hampel and Peltomaki, 2000). Identification of tumours for MSI prospectively could allow for the identification of individuals who are likely to have mutations in DNA repair genes.

As a proportion of tumours from verified HNPCC cases do not display MSI, this test should not completely rule out the diagnosis of HNPCC (Hampel and Peltomaki, 2000). Furthermore, 12-15% of all colorectal tumours displays MSI-H, while 1-6% of these is associated with HNPCC (Aaltonen *et al.*, 1998). The larger proportions of individuals whose tumours display instability are not from HNPCC families and the MSI test should therefore be considered as a preliminary molecular screening procedure rather than a diagnostic test (Hampel and Peltomaki, 2000; Gryfe and Gallinger, 2001).

Hussein and Wood (2002) state that MMR gene mutations and their corresponding MMR protein defects are associated with specific changes in the microsatellite status of tumour cell populations. The interaction between MMR genes, MMR proteins and MSI alterations leads to tumourigenesis. They add that their speculation is supported by the fact that MMR protein expression is absent or markedly reduced in most MSI-H tumours, whereas the alteration is not seen in MSI-L or MSS tumours (Hussein and Wood, 2002).

HNPCC provides researchers with a model to investigate the process of tumourigenesis as driven by MMR deficiency (Peltomaki, 2001a). Not much is known about the mechanisms that translate the HNPCC genotype into the observed clinical phenotype. Why do individuals who harbour an identical mutation manifest the disorder at different ages and some individuals develop cancer of the colorectum while others develop extracolonic cancers? These questions all need to be answered. In addition, the MMR genes known to cause susceptibility to HNPCC account for only two thirds of HNPCC cases that meet the Amsterdam I criteria and show MSI in tumour tissue (Peltomaki, 2001). The investigation of the unexplained one third of families that remain will result in the identification of additional MMR genes or other genes participating in MMR (Peltomaki, 2001a).

1.3.2. Familial adenomatous polyposis (FAP)

FAP is an autosomal dominant disorder with an incidence of approximately 1 in 10 000 births (Jarvinen, 1992; Rustgi, 1994) and a penetrance of approximately 100%. Adenomatous colonic polyps have been observed in persons with the nonpolyposis syndromes, but the diagnosis of FAP requires the presence of 100 or more polyps (Young *et al.*, 1996). Polyps, a hundred to several thousand initially, usually develop during adolescence; a proportion of these will, if left untreated, inevitably give rise to one or more colorectal carcinomas in the third or fourth decade of life (Bodmer, 1999; Heinimann *et al.*, 2001). The average age of onset of FAP is in the early to

mid-teens. In the absence of treatment tumours arise in the 2nd or 3rd decade. FAP displays phenotypic heterogeneity. Within so-called 'typical' FAP families, polyp number and age of onset might differ considerably (Sieber *et al.*, 2000).

A variant of FAP, "attenuated FAP" (AFAP) has been described, characterized by the presence of less than 100 adenomatous polyps, most of which are located in the proximal colon (Lynch *et al.*, 1995). Colon cancer will occur in most patients with AFAP, but with a later age of onset – average age 55 years (Spirio *et al.*, 1993; Friedl *et al.*, 1996).

The classical form of FAP results in colorectal adenomas often accompanied by adenomas in the upper gastrointestinal tract, congenital hypertrophy of the retinal pigment epithelium (CHRPE). CHRPE occurs in approximately 60% of families with FAP, can be detected with ophthalmoscopy at any age and can thus be used to identify at-risk individuals well before polyps appear in the colon (Blair and Trempe, 1980). Abdominal desmoid tumours and epidermoid cysts also occur, most notably of the scalp, osteomas of the mandible, carcinomas of the duodenum, stomach, thyroid and brain (Burn *et al.*, 1991; Giardiello *et al.*, 1991; Giardiello *et al.*, 1991a; Giardiello *et al.*, 1993; Frayling *et al.*, 1998).

1.3.2.1. The Adenomatous Polyposis Coli Gene (*APC*)

A patient presenting with mental retardation, colorectal polyposis and other abnormalities was the first clue to localising the position of the *APC* gene (Fearnhead *et al.*, 2001). A deletion on chromosome 5q was observed in this patient and subsequent linkage analysis of families with FAP led to the mapping of the *APC* gene to 5q21 (Bodmer *et al.*, 1987). Positional cloning was used to identify and characterise the *APC* gene (Grodén *et al.*, 1991; Kinzler *et al.*, 1991). Subsequently, germline mutations, shown to be causative of FAP were identified in *APC*. These mutations account for less than 1% of all CRC (Nakamura, 1993). Turcot syndrome, which results in

multiple colorectal polyps and cerebellar medulloblastoma is also a consequence of a germline defect in *APC* (Hamilton *et al.*, 1995).

The *APC* gene consists of 8535bp spanning 15 exons (Grodin *et al.*, 1991; Thliveris *et al.*, 1996). Approximately 95% of germline mutations that do occur in *APC* result in nonsense or frameshift mutations, producing a truncated protein with abnormal function (Fearnhead *et al.*, 2001). In addition, 70-80% of sporadic CRCs have such truncating mutations, supporting Knudson's hypothesis (Bodmer, 1999).

Mutations have been detected between codon 78 and 2839 in *APC* (Frayling *et al.*, 1998). Exon 15 comprises more than 75% of the coding sequence of *APC* and is the common target for germline and sporadic mutations (Beroud and Soussi, 1996). In sporadic, non-familial tumours, over 60% of mutations are found in a hotspot region within exon 15, termed the mutation cluster region (MCR). The MCR is situated between codons 1286 and 1513, which comprises less than 10% of the coding region of *APC* (Miyoshi *et al.*, 1992). This region is located near the β -catenin binding site; mutations that occur here result in a truncated APC protein that binds to but does not degrade β -catenin (Peifer, 1996; Munemitsu *et al.*, 1995), resulting in its loss of regulation.

The APC protein consists of 2843 amino acids in its commonest form and is organised into several domains with distinct functions (Nakamura, 1993; Ilyas and Tomlinson, 1997). The 300-kDa APC protein mediates growth regulatory signals by its association with the microtubule cytoskeleton and β -catenin. Exon 10A, which is located downstream of exon 10, is subject to alternative splicing and 18 amino acids are added onto the APC protein when transcribed (Sulekova *et al.*, 1995; Xia *et al.*, 1995). The wild-type APC protein is active as a homodimer and the first 45 amino acids are necessary and sufficient for dimerisation (Joslyn *et al.*, 1993; Su *et al.*, 1993). The binding of the catenins occurs between amino acids 1013 and 1174, a region

that contains 15 amino-acid repeats (Rubinfeld *et al.*, 1993). The N-terminus contains oligomerisation sites, which indicate that the molecule exerts some of its effects as a polymer. The central part of the molecule contains heptad repeats and Arm repeats as well as sites for phosphorylation and complexing with β -catenin. The C-terminal region contains sites which can bind the micro-tubule cytoskeleton (Munemitsu *et al.*, 1994) as well as sites that degrade β -catenin (Rubinfeld *et al.*, 1996). The distal 3' end of the APC protein contains 284 amino acid residues for binding EB1 protein, which is 30-kDa and of unknown function. The 3' end also contains the DLG tumour suppressor protein (Fearnhead *et al.*, 2001).

More than 400 disease-causing germ-line mutations have been reported most of which lead to a truncated APC protein (Laurent-Puig *et al.*, 1998). Despite extensive investigation, no APC germline mutations have been identified in a subset of approximately 20% of classical FAP patients with multiple adenomas (Beck *et al.*, 1997). Their phenotype could be due to a variation at a locus or at loci elsewhere in the genome (Beck *et al.*, 1997). Attenuated FAP (AFAP) is observed with germline mutations at the 5' or the 3' end of the APC gene, but could also be associated with missense variants (van der Luijt *et al.*, 1996).

1.4. METHYLATION OF *hMLH1*

Epigenetic silencing of tumour-related genes, due to hypermethylation of their CpG sites in the 5' promoter regions, has emerged as one of the pivotal genetic alterations in cancer development (Baylin *et al.*, 1998; Jones and Laird, 1999). When hypermethylation occurs within the promoter of a normally unmethylated tumour-suppressor gene, it inhibits gene transcription and translation and possibly provides the cell with a growth advantage in a manner similar to mutations (Leung *et al.*, 2001).

Cytosine residues can acquire a methyl group in the C-5 position. This acquisition occurs on the opposite DNA strand of the palindromic CpG sequence. Many genes are rich in CpG doublets, referred to as “CpG islands”. In widely expressed genes, CpG islands are typically, but not exclusively, located in the promoter region (Lynch and de la Chapelle, 1999) and are usually unmethylated. When methylation does however occur, it inhibits the binding of transcription factors, which impedes transcription initiation and this ultimately leads to the silencing of the gene. *De novo* changes in methylation are one mechanism by which genes are regulated during normal development.

Fifteen to twenty percent of sporadic CRC from patients with no obvious family history display microsatellite instability. Hypermethylation of the *hMLH1* promoter occurs frequently in MSI-H sporadic CRCs and is associated with transcriptional loss, which results in lack of protein expression, and this leads to the MSI phenotype in MSI-H sporadic CRC (Wheeler *et al.*, 2000). The hypermethylation can be reversed *in vitro* by treatment with 5-aza-2'-deoxycytidine (5-aza-C), a DNA methyltransferase inhibitor, which restores MLH1 function (Herman *et al.*, 1998). Herman and colleagues demonstrated that no corresponding findings could be shown for MSH2, and that approximately 85% of all MSI-positive sporadic tumours could be accounted for by the silencing of MLH1 in this manner.

It is estimated that 15 to 20% of all CRCs are MSI-H, but only 20 to 25% of this represents HNPCC (Lynch and de la Chapelle, 1999). Individuals affected with HNPCC usually have a primary mutation in *hMLH1* or *hMSH2* as well as a subsequent somatic change, which inactivates the second allele. Somatic hypermethylation of the *hMLH1* promoter causes bi-allelic knockout in 80% of cases. (Lynch and de la Chapelle, 1999). The exact mechanism/s remain to be ascertained, but it has been shown that the ‘second hit’ in *hMLH1* can either be a mutation or a deletion resulting in loss of heterozygosity (LOH). It now appears that hypermethylation can also

produce this 'second hit', resulting in tumour formation, but the frequency with which this occurs is not known. While in some instances, bi-allelic hypermethylation of *hMLH1* has been documented, it is possible that in a subset one allele may be knocked out by promoter methylation and the other by LOH or mutation (Lynch and de la Chapelle, 1999).

Since most MSI-positive sporadic CRCs result from the hypermethylation of *hMLH1* it prompts questions about the role of epigenetic factors:

- i. is hypermethylation gene-specific, organ-specific or even age-specific? (Ahuja *et al.*, 1998)
- ii. could reversal or inhibition of hypermethylation be a worthwhile preventative or therapeutic strategy?

The answers to these questions could have an impact on the understanding of the molecular processes in a subset of CRC (Lynch and de la Chapelle, 1999).

1.5. MODIFIERS OF THE HNPCC PHENOTYPE

Modifiers to the primary genetic defect are defined as polymorphic genes which have alleles of differing levels of activity, distinct from the disease locus, resulting in the phenotype (Ilyas *et al.*, 1999; Muller *et al.*, 2000). These genes are known to modulate the penetrance and expressivity of the disease in individuals known to harbour a disease-associated mutation (Nadeau, 2003). It is generally difficult to identify genetic modifiers, but once detected, they provide clues to both the molecular and functional interactions of the genes under investigation (Nadeau, 2003). Several genes, shown to play a modulating role in CRCs, have been studied intensively.

Environmental factors are known to play a role in carcinogenesis. The colon is continuously exposed to xenobiotic substances, resulting in

genotoxic stress on the colonic epithelium (Muller *et al.*, 2000). There are a host of enzyme systems whose role it is to detoxify and metabolise foreign compounds. These families include the N-acetyltransferases (NATs) and the glutathione-S-transferases (GSTs), which are polymorphic at the genetic level. If any alteration occurs in the enzyme activity, it could result in susceptibility to cancer. These enzyme systems thus represent good candidate modifiers for CRC.

Individuals who have mutations in *hMLH1* and *hMSH2* show variability in the phenotypic expression of the disease with regard to the age at onset of the disease, the cancer site, and the cancer occurrence. These modifier genes could partially account for the variation observed in affected individuals and families. Further investigation of these genes in persons that harbour mutations in the MMR genes is required to make firm correlations between the polymorphisms that exist and the phenotypic variation observed.

1.6. COLORECTAL CANCER IN SOUTH AFRICA

Various studies have been undertaken to determine the frequency of CRC in South Africa. The incidence rates for CRC amongst the ethnic Black and White population group is 3 and 23 per 100 000 respectively, according to the National Cancer Registry of South Africa (Sitas *et al.*, 1998). These figures are taken from a pathology-based registry but give an indication of the cancers observed in these populations.

A retrospective study over a period of seven years, from 1986 to 1992 was carried out at Pelonomi Hospital, in Bloemfontein, South Africa (Bonnet and Grobler, 1997). The aim of this study was to compare patients younger than 40 years affected with CRC to a group older than 40 years. A total of 121 patients were included in the study. The Dukes staging was used to classify colorectal tumours. The mean age at diagnosis of patients was 30.9 years on

the whole, while the mean age at diagnosis in patients older than 40 years (n=77) was 56.3 years. A high proportion of cancer of the proximal colon was noted and adenomucinous carcinomas were found in 16.7% of patients below 40 years and 26% in patients over 40 years. Bonnet and Grobler (1997) concluded that there were no significant differences in the stage of the disease between the under-40 year group and the over-40 group. There was, however, a large proportion of CRC in the younger age group. The prognosis of the disorder in both groups was comparable. No distinction was made between tumours derived from individuals from various ethnic groups.

The incidence of large bowel cancer in the Caucasian population in South Africa typically mimic those of Westernised countries (Oettle and Strijdom, 1997). It also appeared in the late 1970s that the incidence in the ethnic Black population was very low. These low rates for ethnic Blacks remained unchanged from 1950 to 1965, with typical values ranging from 1.6/100 000 for colon cancer and 1.2/100 000 for rectal cancer. Between 1954 and 1956, the frequency at Baragwanath Hospital was 0.9 cases/100 000 admissions. These figures were consistent until 1980. In fact, colon cancer is not listed in a publication discussing the histologically diagnosed cancers in this population group, published in 1988 (Sitas, 1994). It was then reported in the late 1980s that the incidence of CRC in the ethnic Black population had increased 2 to 3 fold, but accurate rates could not be quoted due to the lack of proper incidence rates.

Mortality rates for the Indian and Mixed Ancestry populations are intermediate (Walker and Segal, 1988). In 1981, the reported the incidence of CRC in the Mixed Ancestry population of Johannesburg was 1.33/100 000 for colon cancer and 1.08/100 000 for rectal cancer (Maisto and Bremner, 1981). These figures are even lower than that found in the rural Black population. The lower incidence in the Mixed Ancestry population group could not be explained; it is possibly peculiar to this population group in Johannesburg and remains to be confirmed in subsequent studies.

Differences exist in the risk of developing colon cancer amongst the various population groups in South Africa. In addition, the incidence rates observed historically in the Indigenous Black population is starting to mimic that of Westernised countries. The results obtained to date indicate that a focussed approach is needed to ascertain incidence and mortality rates due to CRC in the different population groups. In addition, the incidence of inherited CRC (in particular HNPCC) in the South African population has not been ascertained as yet.

1.7. AIMS AND OBJECTIVES

1.7.1. Aim:

To identify the molecular genetic basis of hereditary colorectal cancers in South African patients.

1.7.2. Project objectives:

1. To investigate the frequency of germ line mutations in the mismatch repair genes, the mutation cluster region of the *APC* gene, exons 3, 5 and 6 of the *β -catenin* gene and exons 4 to 9 of the *p53* gene, in South African colorectal cancer patients.
2. To estimate the extent of HNPCC due to mutations in the DNA mismatch repair genes, *hMLH1*, *hMSH2* and *hMSH6*.
3. To identify what proportion of CRC families (that do not meet the Amsterdam criteria for HNPCC) and sporadic cases of CRC is due to defects in mismatch repair genes.
4. To investigate the modifier genes of the HNPCC phenotype, which in turn may provide improved estimation of cancer risk.

**CHAPTER 2: ELUCIDATION
OF GERMLINE DEFECTS
UNDERLYING COLORECTAL
CANCER.**

CHAPTER 2: ELUCIDATION OF GERMLINE DEFECTS UNDERLYING COLORECTAL CANCER.

2.1. GENERAL INTRODUCTION

Familial CRCs occur as a result of germline mutations in various genes. Changes in the adenomatous polyposis coli (*APC*) gene give rise to familial adenomatous polyposis (FAP) and mutations in the mismatch repair (MMR) genes result in hereditary nonpolyposis colorectal cancer (HNPCC). In addition to the aforementioned genes, mutations in *p53* and β -catenin are also known to play a role in the etiology of CRC. The current study included the investigation of the MMR genes (*hMLH1*, *hMSH2* and *hMSH6*), *APC*, β -catenin and *p53* due to their reported involvement in the pathology of CRC.

A brief overview of the above-mentioned genes and their involvement in CRC follows with a description of the experimental work carried out in the current investigation.

2.1.1. The Mismatch Repair (MMR) Pathway and CRC

HNPCC, which is one of the most common CRC syndromes in humans, is usually the result of germline mutations in the MMR genes. To date, changes have been reported in six human MMR genes predisposing to HNPCC. Inheritance of a germline mutation in any one of these genes defines the syndrome at the molecular level.

Individuals, who inherit a mutation through the germline, do not acquire widespread developmental abnormalities, as one would expect, although these genes have a vital role and are expressed ubiquitously in the body (Chung and Rustgi, 2003). A possible explanation is that the

second wild-type allele provides sufficient DNA MMR function. Early-onset tumours occur primarily in the colon and endometrium in subjects affected with HNPCC; the precise reason for this degree of organ specificity is not known. It has been observed that the second allele of the MMR gene is inactivated, leading to tumour formation (Chung and Rustgi, 2003).

The MutHLS pathway in *Escherichia coli* is the best-studied MMR system (Muller and Fishel, 2002). The proteins in bacteria that form a complex to repair the mismatched bases are products of the *mutS*, *mutH*, *mutL* and *mutU* (*uvrD*) genes (Modrich, 1997). This MMR system, which has been reconstructed *in vitro*, requires the MutS, MutH, MutL, MutU (helicase II) proteins, as well as DNA polymerase III holoenzyme, DNA ligase, single-stranded DNA-binding protein and one of the four single-stranded DNA exonucleases (ExoI, ExoVII, Exo X or RecJ) (Modrich, 1997) (Fig. 2.1). Newly synthesised DNA strands are usually hemi-methylated. MutS recognises this newly-synthesised strand and binds to mismatched base pairs in the nascent strand, which is then followed by the interaction with MutL (Su and Modrich, 1986; Grilley *et al.*, 1989). The MutS-MutL-complex then interacts with MutH bound to the adenine at –GATC– sites on the hemi-methylated DNA strand (Welsh *et al.*, 1987). The MutH protein is also responsible for incision of the unmethylated strand (Fishel *et al.*, 1993). MutL then facilitates interaction between the MutS protein bound to the DNA mismatch and the MutH bound to the hemi-methylated site. UvrD interacts with one of the single-stranded DNA-specific exonucleases to excise the unmethylated strand. The gap that is left is repaired by the conglomerate action of DNA polymerase III holoenzyme, single-stranded DNA-binding protein and DNA ligase.

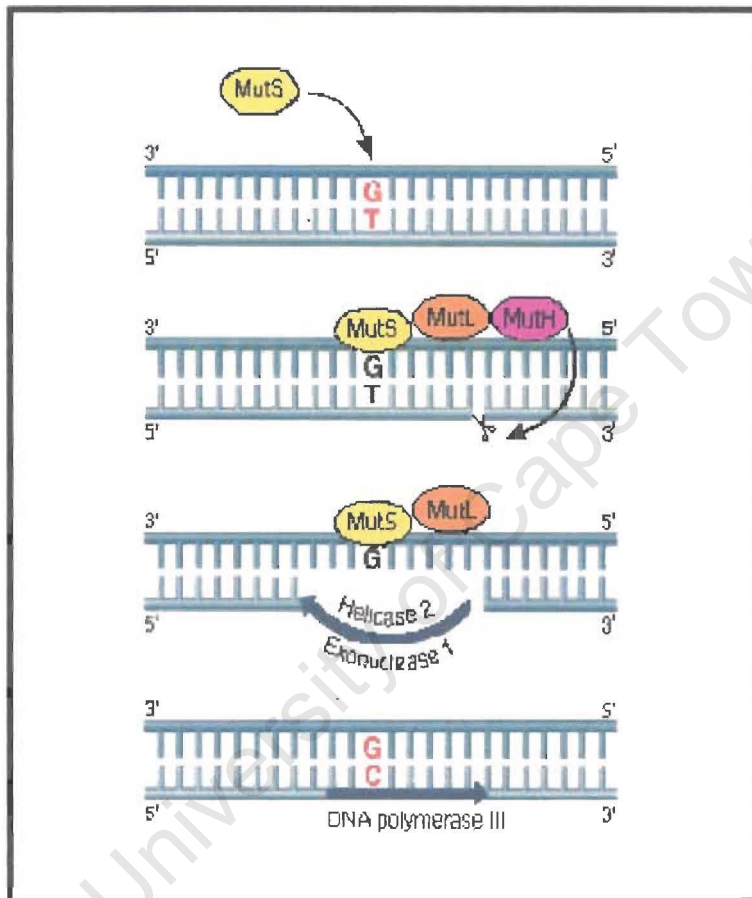


Figure 2.1. *The enzymes that repair DNA mismatches in Escherichia coli.*
(Service, 1994).

The principal function of the MMR system is to remove single-base mismatches and insertion-deletion loops that arise during the DNA-replication process (Peltomaki, 2003). Insertion-deletion loops arise when short DNA repeats are lost or gained within microsatellite repeat

sequences. The latter process is also referred to as microsatellite instability.

A minimum of six MMR proteins is required for the efficient removal of these mismatches and/or insertion-deletion loops in humans (Peltomaki, 2003). These include MSH2, MSH6, MSH3, PMS2, MLH1 and PMS1. The MSH2 protein forms a mismatch-binding heterodimer with either MSH6 or MSH3, depending on the type of lesion that needs to be repaired. The MSH6 protein corrects single-base mismatches, while both the MSH6 and the MSH3 proteins contribute to the correction of insertion-deletion loops. A heterodimer of MLH1 and PMS2 functions as a facilitator for the mismatch-recognition complex and the additional proteins necessary for MMR (Quaresima *et al.*, 2003). The additional proteins may include exonuclease 1 (EXO1), helicases/s, proliferating cell nuclear antigen (PCNA), replication protein A (RPA), replication factor C (RFC), flap structure-specific endonuclease1 (FEN1) and DNA polymerases δ and ϵ (Peltomaki, 2003; Quaresima *et al.*, 2003). The MLH1/PMS2 complex may heterodimerise with two additional proteins, *viz.* MLH3 and PMS1. PMS2 is required for the correction of single-base mismatches and PMS2 and MLH3 function together to correct insertion-deletion loops. The role of PMS1 in MMR still has to be elucidated. There are additional homologues of the human MMR system, which include MSH4 and MSH5. These homologues are necessary for meiotic and possibly mitotic recombination, but have no role in the MMR process (Peltomaki, 2003b).

The vast majority of germline mutations have been identified in MLH1 and MSH2 (Fishel, 2001). This could be explained by the overlapping and redundant MMR activities in the human MLH and MSH heterodimers, since the common denominators are *hMLH1* and *hMSH2* (Fishel, 2001; Peltomaki, 2003b).

2.1.2. **APC**

Mutations in *APC* are associated with the initiation of colorectal adenomas (Bodmer, 1999). Germline mutations in this gene are known to result in FAP; however, 60-80% of sporadic colorectal carcinomas show evidence of mutations in *APC* (Fearnhead *et al.*, 2001).

In sporadic CRCs, more than 60% of all mutations detected in *APC* occur between codons 1286 and 1513, known as the mutation cluster region (MCR), which makes up less than 10% of the entire coding region of the *APC* gene. Two hotspots for somatic mutations occur within this MCR *viz.* codons 1309 and codons 1450.

A correlation exists between the phenotype observed and the site of the mutation within *APC*. Mutations in the MCR are associated with allelic loss (LOH), while tumours with non-MCR mutations result in truncation of the protein product. The region in *APC* where somatic mutations are clustered corresponds to a region of the protein that interacts with β -catenin and other relevant proteins. Mutations in this part of the molecule may have relatively strong dominant-negative effects with more extreme functional consequences (Bodmer, 1999). Missense mutations occurring in the MCR resulting in charge differences, may be more significant here than mutations elsewhere in *APC* (Bodmer, 1999).

A more extensive review of *APC* is provided in section 1.3.2.1.

2.1.3. **β -catenin**

The β -catenin gene (*CTNNB1*) maps to chromosome 3p21 and is comprised of 16 exons spanning 23.2 kb (Nollet *et al.*, 1996). The gene product is a 92 kDa protein.

The β -catenin protein is important in the functional activities of both E-cadherin and APC (Ilyas and Tomlinson, 1997) and it is not surprising that changes in this gene have been associated with CRC. β -catenin is a component of both the cadherin cell-adhesion system and the Wnt-signalling pathway (Behrens *et al.*, 1998). It shares almost 80% homology with the Armadillo protein in *Drosophila* and contains 13 central Arm repeat sequences. These Arm repeats consist of a highly conserved sequence of 42 amino acids. β -catenin can complex with α -catenin, APC and E-cadherin. The binding sites for APC and E-cadherin are different, but they compete for mutually exclusive associations with β -catenin and both molecules cannot be bound to β -catenin simultaneously (Hulsken *et al.*, 1994).

One of the functions of the APC protein is to control the turnover of β -catenin. Mutations in *β -catenin* have been detected in up to 25% of colorectal carcinomas (Bodmer, 1999). Activating mutations in *β -catenin* can occur in tumours without the occurrence of APC gene mutations. Some of these missense mutations in *β -catenin* have been found in colorectal carcinomas, which do not have mutations in APC.

Since germline mutations in exons 3, 5 and 6 of *β -catenin* have been associated with CRC, these exons were chosen for mutation analysis in the current study.

2.1.4. p53

p53 is located on chromosome 17p13.1 and encodes a 53 kDa protein, which is expressed in virtually all cell types (Harlow *et al.*, 1985). The p53 protein resides in the nucleus of the cell (Steele *et al.*, 1998).

p53 is made up of four highly conserved domains, each contributing to the cellular response to DNA damage:

- i. the N-terminal domain-required for transcriptional transactivation (Fields and Jang, 1990)
- ii. a sequence-specific DNA-binding domain (Cho *et al.*, 1994; Pietenpol *et al.*, 1994)
- iii. a tetramerisation domain near the C-terminal end (Sturzbecher *et al.*, 1992)
- iv. the C-terminal domain, which interacts directly with single-stranded DNA (Selivanova *et al.*, 1996).

The naturally-occurring mutations in *p53* reduce the ability of the protein to activate transcription. The genes controlled by *p53* fall into one of four categories, *viz.* cell cycle inhibition, apoptosis, maintaining genetic stability or inhibition of blood vessel formation (Vogelstein *et al.*, 2000).

When the cell is exposed to stressful stimuli, resulting in DNA damage, metabolic changes, heat shock and hypoxia, the *p53* protein is activated, eliciting a series of events, which either results in apoptosis (programmed cell death) or cell cycle delay (Levine, 1997). The final outcome of *p53* activation is mediated through the action of downstream effector genes transactivated by *p53* (Amundsen *et al.*, 1998).

p53 binds both single- and double-stranded DNA in a sequence-specific manner dependant on the *p53* COOH terminal domain (Steinmeyer and Deppert, 1988; Foord *et al.*, 1991). This property of *p53* is lost in naturally-occurring mutants. The amino terminal of the protein has transcriptional activity (Fields and Jang, 1990; Raycroft *et al.*, 1990). Mutant *p53* protein has a conformational change and oligomerises with both wild-type and mutant proteins into inactive complexes with a prolonged half-life (Finlay *et al.*, 1988; Gannon *et al.*, 1990). Altered proteins are able to drive wild-type protein into the mutant

conformation (Milner and Medcalf, 1991) by accumulating in the cytoplasm, by altering its conformation, and possibly by competing for target sites or regulatory molecules.

Lane (1992) proposed a model for *p53* function, likening it to a 'guardian of the genome'. DNA damaging agents lead to accumulation of *p53*, which switches off replication. *p53* encodes a protein transcription product, which arrests cells at G1 to allow time for repair. If repair fails to occur, *p53* may trigger an apoptotic process. Mutant *p53* cannot carry out this protective cellular arrest. These cells, which are genetically less stable, accumulate mutations and chromosomal aberrations at an increased rate and are more likely to develop into malignant clones.

2.1.4.1. Mutations in *p53*

Approximately 50% of all human cancers may exhibit an abnormal or missing *p53* (Bodmer, 1999). To date, investigation of cancerous tissue of the bladder, blood, brain, breast, colon, liver, lung, oesophagus and skin show distinct *p53* mutations (Lewis, 2001). The most common genetic alterations are missense mutations in one allele; however, deletions and chain-termination mutations occur as well (Steele *et al.*, 1998). Mutations in the DNA-binding domain of *p53* have the greatest effect on function. Most of the missense mutations observed in *p53* are located in this domain, leading to the production of *p53* protein that fails to bind in the normal sequence-specific manner.

p53 mutations occur late in the adenoma-to-carcinoma sequence in CRC (Vogelstein and Kinzler, 1993). Colorectal tumours that show loss at chromosome 17p have been shown to be aggressive (Kern *et al.*, 1991). In addition, a high mutation rate in the remaining non-deleted *p53* in CRC has been observed. Point mutations in the non-deleted *p53* have been identified in adenomas and carcinomas from FAP patients

(Shirasawa *et al.*, 1991). Replacement of a p53 protein in CRC cell lines results in rescue of the wild-type cell proliferative capacity (Baker *et al.*, 1990).

Germ-line mutations of p53 have been associated with the Li-Fraumeni syndrome (Malkin *et al.*, 1990; Srivastava *et al.*, 1990), indicating that p53 is not only associated with sporadic tumours. Varley (2003) reported approximately 250 germline *p53* mutations, most of which were associated with Li-Fraumeni syndrome or Li-Fraumeni-like syndrome.

2.1.5. Mutation Detection Methods

In any screening process aimed at identifying genetic changes, the mutation detection technique should be simple and effective and should be able to work on DNA from a variety of sources of biological material (blood, fresh and archived tumour tissue). A number of different techniques may be employed to identify germline mutations. These include the protein truncation test (PTT) (Roest *et al.*, 1993), chemical mismatch cleavage (CCM) (Grompe, 1993), denaturing gradient gel electrophoresis (DGGE) (Grompe, 1993), single-stranded conformational polymorphism (SSCP) (Orita *et al.*, 1989) and heteroduplex analysis (HA) (Grompe, 1993). Each of these methods has limitations. PTT requires biological samples from living family members, which is not always possible. In addition, the method will only detect mutations that result in the generation of a truncated protein product; not all mutations will therefore be detected. CCM involves the handling of highly toxic chemicals and radioactivity but its sensitivity approximates 100%. DGGE is effective, but one of the oligonucleotide primers used must have a GC clamp attached, which can be costly and these primers cannot be used for sequencing analysis due to the presence of the GC

clamp. The generation of additional primers for sequencing adds to the cost burden.

The SSCP method was the one method of choice for mutation analysis. This method has a sensitivity of 70-95% when the PCR products are less than 200bp, however, the sensitivity reduces to less than 50% when fragments of more than 400bp are analysed (Grompe, 1993). The exons of the MMR genes are suited for screening by the SSCP technique. In addition, the SSCP method can be performed on fresh and archived tissue. HA is a rapid technique and can be used when the mismatch in the DNA strand causes a conformational change in the DNA heteroduplex. The mutation detection method of choice was a combination of SSCP and heteroduplex analysis (HA), which allowed for the detection of SSCP and HA banding patterns, which increased the efficiency of detecting mutations.

Denaturing High-Performance Liquid Chromatography Analysis (dHPLC) was the second method of choice for mutation analysis. This method is based on temperature-dependant separation of DNA containing mismatched bases. Although the initial cost of acquiring a dHPLC device is high, longer-term running costs are low. The method for screening is automated with a sensitivity comparable to DGGE but without the laborious gel-casting requirement and need for specialised primers. It has been suggested that the profiles for certain mutations are so specific that dHPLC could eventually replace direct sequencing, which would ultimately reduce the cost and time of analysis as well (Gross *et al.*, 1999). This may well rely on confidence being established through screening of large numbers of samples to address the issue of reproducibility.

The mutation detection rate in this study could underestimate the actual number of mutations for the following reasons:

- (i) SSCP and HA analysis has a sensitivity of 80-90% to identify single base substitutions (Grompe, 1993; Peeters and Kotze, 1994)
- (ii) SSCP and dHPLC WAVE analysis may not reflect deletions of whole exons.
- (iii) The promoter regions of the various genes were not screened for mutations

The use of Southern blot analysis to identify chromosomal deletions in the MMR genes is likely to increase the reliability of the DNA-based methods.

SSCP and dHPLC WAVE analysis were the methods of choice for mutation analysis since they are quick and reproducible. In addition, due to the nature of the available equipment, many samples could be processed at once.

2.1.6. Genotype/Phenotype correlation in HNPCC

Over the last decade, much progress has been made in the field of human genetics with regard to the identification of disease-causing genes. Researchers now have access to sequence information that was unavailable ten years ago. Identifying the disease-causing genes for Mendelian disorders has been considerably successful to date using techniques such as linkage analysis and positional cloning. Researchers are now in the era of the molecular genetic elucidation of complex disorders. The next aim of genetics is to use the emerging information (i.e. the thousands of sequences and expression profiles) made available from the Human Genome Project to link the phenotype observed in an affected individual with what is detected at the DNA level i.e. the genotype.

Insights into genotype/phenotype correlations have been obtained from research into diseases such as cystic fibrosis (Kulczycki *et al.*, 2003), pseudoachondroplasia and multiple epiphyseal dysplasia (Briggs and Chapman, 2002), cleidocranial dysplasia (Yoshida *et al.*, 2003), Gaucher's disease (Zhao *et al.*, 2003) and Parkinsons disease (Mori *et al.*, 2003), to mention but a few. Evidence for genotype-phenotype correlations have also been shown in certain cancer syndromes. These include familial adenomatous polyposis (Groves *et al.*, 2002), von Hippel-Lindau syndrome (Friedrich, 2001), polycystic kidney disease (Deltas, 2001) and Hirschprung disease (Decker *et al.*, 1998). The relation between the observed phenotype and the inherited genotype is complex, with factors such as diet, smoking, bacterial infection, socio-economic status and therapeutic agents all having an additional influence on the expression of the disease.

HNPCC is one of the most common forms of hereditary CRC, accounting for between 2-5% of all colorectal cancers (Aaltonen *et al.*, 1998). The diagnosis of this disorder is based on a family history and/or a mutation in one of the MMR genes. The relationship between the nature and site of mutation, the age of onset, the severity of the cancer and the penetrance of the disorder have yet to be elucidated. Studies have been conducted to draw correlations between genotype and phenotype. The investigation by Vasen *et al.* (1996) compared two groups of individuals (i.e. one group with mutations in *hMLH1* and the other with mutations in *hMSH2*) and demonstrated that the risk for CRC between the two groups was the same (80%). Endometrial cancer was more common than CRC in the *hMSH2* mutation-carriers (61% vs. 42%). Extracolonic cancers were observed more often in *hMSH2* mutation-carriers compared to *hMLH1* mutation carriers in the study conducted by Lin *et al.* (1998). Dunlop *et al.* (1997) investigated 64 subjects affected with HNPCC and stratified the cohort by gender. The risk of developing CRC was 74% for males and 30% for females and

endometrial cancer was more common than CRC for the female subjects. Miyaki *et al.* (1997) identified a germline mutation in the *hMSH6* gene in a family who did not meet the Amsterdam criteria. The subjects in this family were more prone to cancer of the endometrium and ovary. More studies are required for definitive and conclusive results of the effect that the genotype has on the observed phenotype. Until conclusive results can be obtained, the development of genetic or biochemical systems to test the mutated MMR genes will be useful in the management and counseling of individuals who harbour mutations in the MMR genes (Quaresima *et al.*, 2003).

A cohort of 100 individuals diagnosed with CRC (mostly before 45 years of age) who do not meet the Amsterdam criteria (Vasen, *et al.*, 1990) were chosen for mutation analysis in this phase of the current study. The genes that were screened included the MMR genes (*hMSH2*, *hMLH1* and *hMSH6*), the MCR of *APC*, exons 3, 5 and 6 of *β -catenin* and exons 4 to 9 of *p53*. Genotype/phenotype assessments were carried out on those individuals in whom mutations were identified in an attempt to investigate any relationship between the inherited mutations, the age of onset of disease and the site of cancer.

Various methods (i.e. mutation analysis of genomic DNA and microsatellite instability studies) were employed for the optimal detection of mutations in a variety of tissues from individuals with CRC.

2.2. MATERIALS AND METHODS

2.2.1. Subject recruitment

The research subjects were recruited from the Surgical Gastroenterology Unit at the Groote Schuur Hospital in Cape Town, South Africa. All subjects and parents of subjects, who were younger than 18 years of age, gave informed written consent prior to investigations. A copy of the consent form completed by patients is included in the Appendix (section A.1.). Biological material was obtained from these subjects as whole blood, from which DNA was isolated. In addition, fresh tumour tissue was obtained if a subject underwent surgery. Tumour tissue in the form of 5 μ M sections of formalin-fixed paraffin-embedded tissue was also available for a subset of individuals.

DNA from a total of 260 families (n=1072) diagnosed with CRC, was banked in the biological materials archive in the Division of Human Genetics at the University of Cape Town. The patient database was interrogated for information on the subjects under investigation. A research subject cohort was then extracted for mutation analysis of the MMR genes. Towards identifying an optimal research subject cohort, the records of the biological materials archive of 260 subjects were interrogated for signs of the Amsterdam and other HNPCC-identifying criteria. One hundred unrelated probands, diagnosed with CRC were chosen as the primary research cohort for mutation analysis. These subjects were all diagnosed with CRC before 45 years of age (males and females), preferably with a family history of CRC or any of the cancers that form part of the spectrum of HNPCC-associated cancers. Familial adenomatous polyposis was excluded. Biological material (venous blood) from seemingly sporadic individuals were archived if the subjects

were diagnosed with CRC at a young age (under 45 years). The cohort of 260 subjects is representative of all the population groups in South Africa.

The anatomical localisation of colorectal cancer in the primary research cohort is shown in Table 2.1.

Table 2.1. Distribution of tumours in the colon of patients chosen for mutation analysis (n=100).

REGION OF COLON	BREAKDOWN OF SAMPLES
Proximal	30
Transverse	17
Descending	23
Rectum	30
TOTAL	100

2.2.1.1. Cataloguing and processing of biological material

Following the collection procedure, the blood samples were stored at 4°C and transported within 24 hours, at room temperature, to the laboratory where they were catalogued according to the standard sample-coding procedure. Biological material from each disorder and each subject had a unique identifier (e.g. NPC 1.12DAV). The laboratory code for blood specimens from patients with HNPCC was ‘NPC’, followed by the numerical family identifier (family 1) and the “serial recruitment” number of the particular subject (the 12th subject/recruitment in this instance). A unique confirmatory 3-letter code followed (DAV in this case), which was usually the first 3 letters of the individual’s first name. If not processed within 24 hours of receipt, blood samples were stored at -20°C until the DNA isolation procedure.

The DNA was extracted from whole blood using the Genomix Kit (*Talent*, Italy), with modifications as outlined in the Appendix (section B.1.).

This procedure involved separation of lymphocytes, lysis of cells, coagulation and separation of proteins and precipitation of DNA.

2.2.1.2. Access to and processing of paraffin-embedded tissue for PCR

DNA from tumour tissue was used for microsatellite-instability analysis. Tumour tissue was obtained for 19 subjects from the Division of Anatomical Pathology at the University of Cape Town as 5 μ M sections of paraffin-embedded tumour tissue (PETT). The method for DNA extraction from formalin-fixed paraffin sections is described in the Appendix (section B.2.), but mainly involved the de-paraffination process, followed by removal of protein with proteinase K and finally isolation and purification of the extracted DNA (Frank *et al.*, 1996).

2.2.1.3. DNA extraction from fresh tumour tissue

The DNA from fresh tumour tissue was used for microsatellite instability analysis. Following surgical resection, 20 samples of fresh tumour tissue were received from the Surgical Gastroenterology Unit at Groote Schuur Hospital, Cape Town. The tissue was immediately placed in liquid nitrogen and stored at -70°C until the DNA isolation procedure. The *High Pure PCR Template Preparation Kit* (Boehringer Mannheim, Germany) was used to isolate DNA from fresh tumor tissue, as outlined in the Appendix (section B.3.).

The available biological material for the current study allowed for the detection of mutations by means of SSCP and dHPLC analysis. In addition, 20 fresh tumour tissue samples and 19 preserved tumour tissue sections were available. The ideal approach for searching for mutations underlying HNPCC in MMR genes would start with linkage analysis in suitable pedigrees. This would direct the researcher to the gene associated with the disorder. In addition, MSI analysis of DNA

derived from tumour tissue would direct one to mutations in MMR genes. These methods would lead to mutation analysis of the gene of interest. The available biological material for the research cohort in the current study did not allow for linkage analysis or large-scale MSI; thus SSCP and dHPLC were carried out to identify the disease-causing mutations.

2.2.2. Mutation Analysis

The research described here was geographically divided between two different institutions, the Division of Human Genetics at the University of Cape Town and the Weatherhall Institute for Molecular Medicine at Oxford University. The methods used for mutation analysis reflect the facilities available at the separate laboratories.

Essentially, two methods were used for mutation analysis:

- i. Single stranded conformational polymorphism (SSCP) for all the exons of *hMLH1* and *hMSH2* and
- ii. dHPLC analysis for all the exons of *hMSH6*, the MCR of *APC*, exons 3, 5 and 6 of *β -catenin* and exons 4-9 of *p53*.

2.2.2.1 Single stranded conformational polymorphism (SSCP) analysis

The method used for SSCP analysis was based on the technique developed by Orita *et al.* (1989) with modifications to allow for non-radioactive detection of the single-stranded DNA i.e. through silver staining (Lohmann *et al.*, 1992). The method is outlined in Appendix (section B.4.).

The coding regions of candidate genes were targeted through analysis of their exons via PCR-based amplification of these elements.

Oligonucleotide primers were designed to include the investigation of the intron/exon boundaries of the various candidate genes screened.

Amplification was performed in a total volume of 25 µl containing the following:

- 100 ng genomic DNA
- 20 pmol of each primer (Appendix D, Table D.1. and D.2.)
- PCR buffer containing 50 mM KCl, 10 mM Tris-HCl (pH 8.3), 1.5 mM MgCl₂
- 200 µM (dATP, dCTP, dGTP, dTTP) (Gibco, UK)
- 1 unit of *Taq Polymerase* (Gibco, UK)
- dH₂O to a volume of 25 µl

The reactions were overlayed with 20 µl of liquid paraffin and subjected to a denaturation step at 95°C and then processed through 30 cycles of 95°C for 30 seconds, 50°C, 55°C or 60°C for 30 seconds and 72°C for 40 seconds, followed by a final extension period at 72°C for 5 mins.

2.2.2.2. Screening using dHPLC

With the availability of a dHPLC WAVE DNA fragment-analysis system (*Transgenomic*, UK), a more efficient means of detecting DNA variants was available. After the hybridisation procedure, PCR products were automatically loaded on a DNasep® column (*Transgenomic*) and eluted with a linear acetonitrile (ACN) gradient (*BDH Merck*, UK) consisting of buffer A (0.1 M triethylammonium acetate; TEAA)/buffer B (0.1 M TEAA, 25% acetonitrile) at a constant flow rate of 0.9 ml/min. Analysis time per amplified sample was 6.8 mins. The temperature for optimal heteroduplex analysis was determined using WAVEMaker™ software (*Transgenomic*, UK). The temperature at which heteroduplex detection occurred was deduced from the melting profile of the specific DNA fragment. A melting curve was constructed and the specific

temperature point showing 50% denaturation of the fragment was defined. The linear acetonitrile gradient was then adjusted so that the peaks were eluted between 3.5 and 4.5min.

PCR amplification was performed in a total volume of 50 μ l containing the following:

- 30 ng of genomic DNA
- 200 nM F-primer (10 pmol/ μ l) (Appendix D, Tables D.3. to D.6.)
- 200 nM R-primer (10 pmol/ μ l) (Appendix D, Tables D.3. to D.6.)
- 1X KCl buffer
- 200 μ M dNTPs (*Gibco*, UK)
- 1.5 mM $MgCl_2$ (*Gibco*, UK)
- Sterile MilliQ dH_2O to 50 μ l
- 0.2 μ l Amplitaq Gold (*PE Biosystems*, Germany)

The reactions were subjected to the following PCR conditions:

95°C –10min

95°C –30s

Annealing temp. +7.5°C, - 0.5°C per cycle-30s

72°C -30s

} 14 cycles

95°C –30s

Annealing temp-30s

72°C –30s

} 20 cycles

72°C -7min

end

A final hybridisation procedure to optimise the formation of heteroduplexes and homoduplexes was conducted:

95°C - 4min

95°C - minus 1.6°C per cycle-1min-42 cycles

15°C - hold

After the hybridisation step products were stored at -20°C until dHPLC analysis.

2.2.3. DNA sequence analysis

PCR (DNA) products displaying either variant SSCP or dHPLC melting profiles were sequenced in both 5' to 3' and 3' to 5' directions on an Applied Biosystems 377 automated sequencer using the BigDye Terminator cycle sequencing kit [PE Biosystems, Germany] according to the manufacturer's protocol (Appendix section B.5.).

The resolution of the sequencing process was improved through the use of Centri-Sep Spin Columns (Applied Biosystems, USA), which effectively and reliably removed excess DyeDeoxy terminators from the completed sequencing reactions. The procedure is outlined in the Appendix (section B.6.).

2.2.4. Amplification Refractory Mutation System (ARMS) PCR Analysis

The unrelated, unaffected control cohort was screened by using an ARMS PCR assay for the presence of the mutations detected in exons 4 and 9 of *p53*. The oligonucleotide primers were designed using the OLIGO software (version 4.1, Medprobe, Norway). Two forward primers were used for each PCR, one specific for the wild type allele and the second specific for the mutated allele (Appendix D, Table D.7.). These

primers were used in separate PCR reactions, each utilising the same common reverse primer. The assay was validated by analysis of samples known to be heterozygous for both the mutant and the wild-type allele.

2.2.5. *MvaI* Restriction endonuclease analysis

The entire cohort of 260 subjects was screened for the presence of the C1528T mutation in the *hMLH1* gene. This transversion in exon 13 resulted in loss of an *MvaI* restriction site and replaced an arginine with a termination codon at position 510 in the translated amino acid sequence (Q510X). The screening of the cohort of 260 subjects involved a PCR of exon 13 as described in section 2.2.2.1, followed by a digestion of eight microlitres of the PCR product with ten units of *MvaI* and resolution by agarose gel electrophoresis.

2.2.6. Haplotype Analysis

Haplotype analysis was facilitated with microsatellite markers in the region of *hMLH1*. Figure 2.2 illustrates a physical map of these microsatellite markers and their positions in relation to each other and the *hMLH1* gene. Subjects included affected and unaffected individuals from family NPC 1, in whom the C1528T mutation was originally detected, as well as the affected individuals from other families in which the C1528T mutation was subsequently identified.

The microsatellite markers were amplified by the PCR with the following constituents:

- 100 ng of genomic DNA
- 20 μ M of each primer/microsatellite marker (Fig. 2.2.)(Research Genetics, USA)
- 50 mM KCl, 10 mM Tris-HCl (pH 8.3), 1.5 mM MgCl₂ (Gibco, UK)
- 200 μ M (dATP, dGTP, dCTP and dTTP) (Gibco, UK)

- 1 μ Ci 32-P dCTP (Amersham, USA)
- 1 unit of *Taq Polymerase* (Gibco, UK)
- dH₂O to 25 μ l

The reactions were overlayed with 20 μ l of liquid paraffin and subjected to a denaturation step at 95°C and then processed through 30 cycles of 95°C for 30 seconds, 50°C, 55°C or 60°C for 30 seconds and 72°C for 40 seconds, followed by a final extension period at 72°C for 5 mins. Ten μ l of formamide loading dye (Appendix, section C.2.) was added to 25 μ l PCR product. Five microlitres of the denatured sample was loaded into each well of a polyacrylamide gel (Appendix, section C.3.) and electrophoresed in a 1XTBE (Appendix, section C.1.) running buffer of a sequencing gel (333 x 422 x 0.35mm) (OMEG Scientific, South Africa) at 50°C and at 65W constant power for 2-3hrs. Gels were dried onto 3MM paper (*Schleicher and Schuell*, Germany), on a horizontal slabgel drier for 1-2hrs at 80°C. The gels were exposed to Agfa X-ray film (*Agfa Curix*, Germany) in cassettes containing intensifying screens, at -70°C for a period of 24-56hrs. The X-ray film was developed in an Agfa Curix 60 automated developing machine (*Agfa Curix*, Germany) and the genotypes were read.

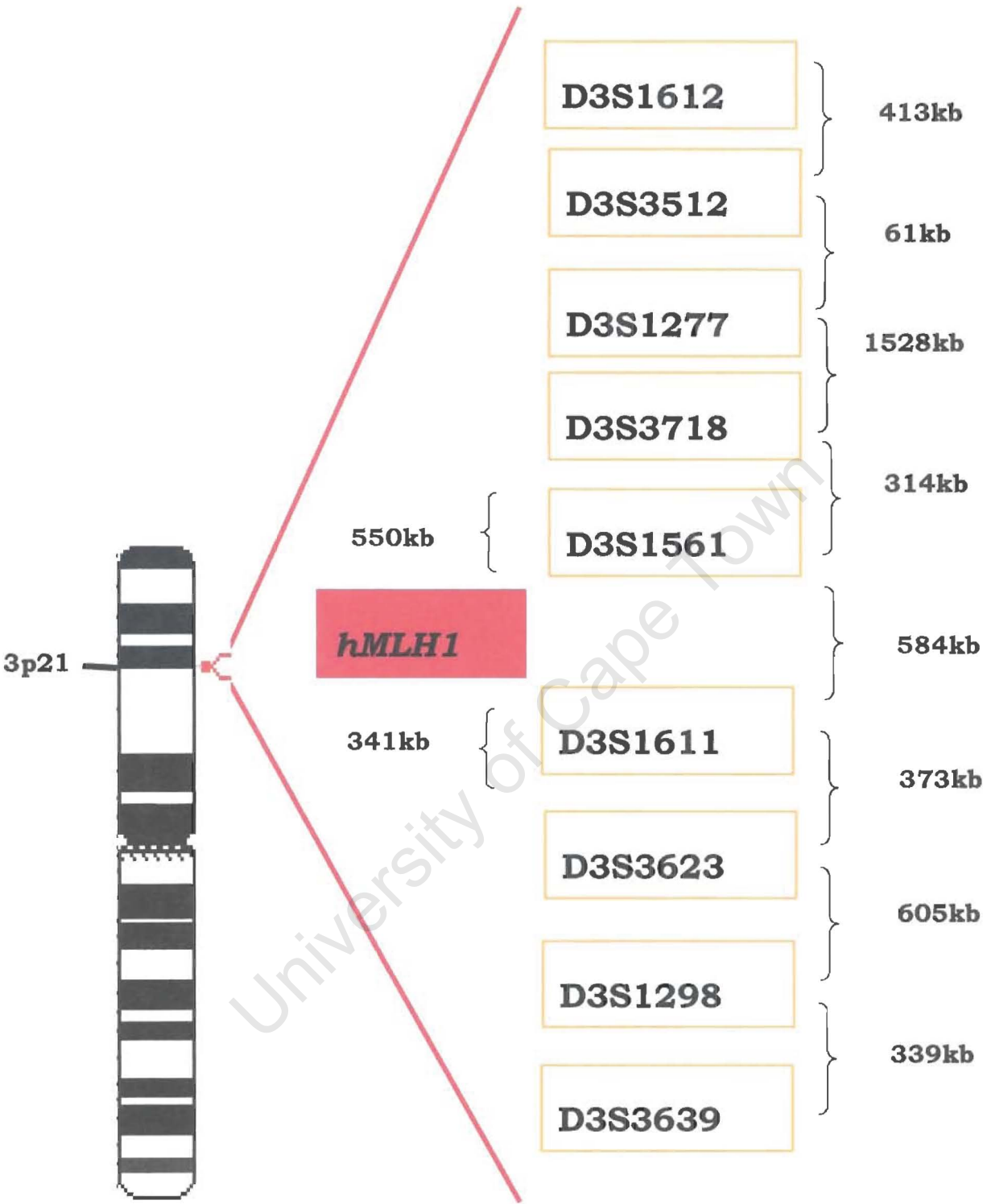


Figure 2.2. Ideogram of chromosome 3 with the relative positioning of *hMLH1* and the physical distances of the microsatellite markers as used for linkage/haplotype analysis.

2.2.7. Microsatellite instability (MSI) studies

DNA extracted from 20 fresh tumour tissue samples and 19 paraffin-embedded tissue samples was used as a template for PCR amplification. DNA extracted from peripheral blood samples was also used in parallel PCRs to serve as controls against which the alleles in the PETT/tumour tissue could be compared.

DNA was amplified as follows:

- 5 µl of template DNA
- 20-25 µM of each microsatellite marker (Appendix D, Table D.8.)
- 50 mM KCl, 10 mM Tris-HCl, pH 8.3, 1.5 mM MgCl₂
- 200 µm (dATP, dGTP, dCTP and dTTP) (Gibco, UK)
- 1 µCi [P32] dCTP (Amersham, USA)
- 1 unit of *Taq* polymerase (Gibco, UK)
- dH₂O to 25 µl

The reactions were overlayed with 20 µl of liquid paraffin and subjected to a denaturation step at 95°C and processed through 30 cycles of (95°C for 30 seconds, 50°C, 55°C or 60°C for 30 seconds and 72°C for 40 seconds), followed by a final extension period of 72°C for 5 mins. Samples were denatured at 95°C for 5 mins and chilled on ice for 5-10 mins before loading onto a 6% polyacrylamide gel (Appendix, section C.3.). Five microlitres of the denatured sample were loaded into each well and electrophoresed in a 1XTBE running buffer at 50°C and at 65W constant power for 2-3hrs (depending on the expected fragment size), using a standard sequencing apparatus (333 x 422 x 0.35mm) (OMEG Scientific, South Africa). The marker dyes in the loading buffer were

used to indicate the length of time of electrophoresis for optimal resolution of the repeat fragments. Gels were dried onto 3MM paper (*Schleicher and Schuell*, Germany), on a horizontal slabgel drier for 1-2hrs at 80°C. The gels were exposed to Agfa X-ray film (*Agfa Curix*, Germany) in cassettes containing intensifying screens, at -70°C for a period of 24-56hrs. The X-ray film was developed in an Agfa Curix 60 automated developing machine (*Agfa Curix*, Germany) and the genotypes were ascertained.

University of Cape Town

2.3. RESULTS

2.3.1. Subjects

A cohort of 100 subjects was investigated for mutations in *hMLH1*, *hMSH2*, *hMSH6*, the MCR of *APC*, β -*catenin* and *p53*. MSI analysis was conducted on 20 fresh tumour tissue sections and 19 paraffin-embedded tissue sections. To prove that the single base substitutions were causative of disease, the presence of the mutation was checked for in 100 unaffected, unrelated individuals from the corresponding population group. These individuals were referred to as 'controls'

2.3.2. Mutation analysis

The MMR genes, *APC*, β -*catenin* and *p53* were investigated for genetic changes using a combination of SSCP and dHPLC WAVE analysis.

2.3.2.1. MMR genes (*hMLH1*, *hMSH2* and *hMSH6*)

Analysis of the MMR genes (SSCP and dHPLC WAVE) revealed 13 mutations in 15 probands (Table 2.2). In addition to the primary research cohort, 100 controls were tested for the presence of the mutations listed in Table 2.2. The accession numbers used for analysis are as follows: *hMLH1* (NM_000249), *hMSH2* (NM_000251) and *hMSH6* (NM_000179).

The mutation analysis aspect of the project allowed the identification of novel and previously reported mutations and polymorphisms in the genes investigated. The following numbers of mutations were detected:

- i. five in *hMLH1*,
- ii. six in *hMSH2* and
- iii. one in *hMSH6*

Table 2.2. Mutations and polymorphisms detected on analysis of hMLH1, hMSH2 and hMSH6.

SAMPLE	SITE	NUCLEOTIDE POSITION	AMINO ACID POSITION	PATHOGENICITY OF MUTATION
<u>MUTATIONS</u>				
<u>hMSH2</u>				
NPC 85.1, 107.1	Exon 6	G965A**	G322D	Pathogenic
NPC 45.1	Exon 6	C1046G	P349R	Not known
NPC 120.1	Exon 7	1220delCT	Frameshift	Pathogenic
NPC 5.6	Exon 8	1340insGG	Frameshift	Pathogenic
NPC 42.1	Exon 9	A1489G	I497V	Not known
NPC 145.2	Exon 15	2501delCTAATTT	Frameshift	Pathogenic
<u>hMLH1</u>				
NPC 85.1	Exon 4	C350T**	T117M	Pathogenic
NPC 43.1	Exon 6	G->A 5' donor site + 3**	Total skipping of exon 6	Pathogenic
NPC 15.1	Exon 9	731delGTTA	Frameshift	Pathogenic
NPC 94.1	Exon 13	1521insT	Frameshift	Pathogenic
NPC 1.2	Exon 13	C1528T	Nonsense	Pathogenic
<u>hMSH6</u>				
NPC 101.1	Exon 4	T2633C**	V878A	Not known
<u>POLYMORPHISMS</u>				
<u>hMSH2</u>				
1 individual	Intron 12	IVS12-6T->C	-	Not Pathogenic
2 individuals	Exon 3	C402T	I150I	Not Pathogenic
<u>hMLH1</u>				
1 individual	Intron 4	IVS4-41A->G	-	Not Pathogenic
9 individuals	Exon 8	A655G**	I219V	Not Pathogenic
22 individuals	Intron 14	IVS14-19A->G **	-	Not Pathogenic
NPC 26.2, 121.1	Exon 19	C2152T\$\$	H718Y	Not Pathogenic
1 individual	Exon 19	T2184C	Y721Y	Not Pathogenic
<u>hMSH6</u>				
17 individuals	Intron 2	IVS2+52T->A**	-	Not Pathogenic
20 individuals	Intron 4	IVS4-56C->T	-	Not Pathogenic
3 individuals	Intron 4	C1164T**	H388H	Not Pathogenic
4 individuals	Intron 4	C1869T	P623P	Not Pathogenic
14 individuals	Intron 7	IVS7-51del 3647-3651	-	Not Pathogenic

** : Mutation previously described in the literature (<http://www.nfdht.nl>)**:

Mutation previously described in the literature (<http://www.nfdht.nl>)

\$\$: (Kowalski et al., 1997; Farrington et al., 1998; Weber et al., 1999;)

The DNA variants detected in the 100 subjects of the primary research cohort were used to screen the entire archived research cohort of 260 probands to determine if any additional index patients harboured any

one of the mutations listed in Table 2.2. The results of the second screening phase are shown in Table 2.3.

University of Cape Town

Table 2.3. CRC phenotype in additional families in whom mutations have been identified.

Family	Mutation	Site of Cancer	Average Age of Onset	Extracolonic cancer present in the family	Ethnic Group
<u>hMSH2</u>					
NPC 85.1	G965A/C350T	Transverse	26	Skin, lung, pituitary, breast, endometrial,	C
NPC 107.1*	G965A	Descending	30	No	MA
NPC 36.1*	1220delCT	Ascending & descending	42	Ovarian	MA
NPC 36.3*	1220delCT	Transverse	39	Ovarian	MA
NPC 120.1	1220delCT	Ascending & transverse	40	No	MA
NPC 37.1*	1220delCT	Ascending	44	No	C
NPC 209.1*	1220delCT	Descending	34	No	
NPC 5.6, 5.28, 5.31, 5.37	1340insGG	Transverse, Ascending	49	Throat, ovarian, breast	C
NPC 158.1*	1340insGG	Ascending	52	No	MA
NPC 145.2	2501delCTAATTT	Ascending	53	Endometrial	MA
<u>hMLH1</u>					
NPC 43.1	G->A 5' donor site + 3	Rectum	38	No	B
NPC 15.1	731delGTTA	Descending	25	No	C

Family	Mutation	Site of Cancer	Average Age of Onset	Extracolonic cancer present in the family	Ethnic Group
NPC 94.1	1521insT	Transverse	30	No	MA
NPC 1.2	C1528T	Ascending	38	No	MA
NPC 16.1*	C1528T	Transverse	44	No	MA
NPC 49.1*	C1528T	Ascending	38	No	MA
NPC 54.1*	C1528T	Ascending	39	No	MA
NPC 66.1*	C1528T	Rectum	31	No	MA
NPC 68.1*	C1528T	Unknown	21	No	MA
NPC 77.2*	C1528T	Rectum	27	No	MA
NPC 82.1**	C1528T	Ascending	33	No	MA
NPC 97.1*	C1528T	Transverse colon	39	No	MA
NPC 109.4*	C1528T	Ascending	31	No	MA
NPC 134.1*	C1528T	Ascending	39	No	MA
NPC 157.1*	C1528T	Not affected	-	No	MA
NPC 166.1*	C1528T	Descending	40	No	MA
	C2152T	Descending & rectal	59	No	MA
NPC 26.2					
NPC 121.1*	C2152T	Ascending	24	Endometrial, Throat, Testes	B

*: Additional families identified with mutations previously detected in the primary research cohort of 100 individuals (MA: Mixed Ancestry, B: Indigenous black population, C: Caucasian)

2.3.2.2. APC

The results of mutation analysis of the MCR of APC are shown in Table 2.4.

Table 2.4. Mutations and polymorphisms detected on screening of exon 15 of the APC gene.

SAMPLE	EXON	NUCLEOTIDE POSITION	AMINO ACID POSITION
MUTATIONS			
NPC 17.1	15	T3920A	I1307K
NPC 24.1	15	G3949C	E1317Q
NPC 10.1, 80.1	15	A5392G	N1798D
10 samples	15	A5495T	D1822V
POLYMORPHISMS			
15 individuals	15	T4326A	P1443P

The changes listed in Table 2.4 were not detected in the 100 controls tested.

2.3.2.3 β -catenin

No variant dHPLC melting profiles were detected in the lymphocytic DNA of both the primary research cohort and matched control DNA samples for exons 3, 5 and 6 of β -catenin.

2.3.2.4. p53

Four samples within the primary research cohort displayed variant dHPLC melting profiles in the screen of p53. Two of the samples, *viz.* NPC 45.1 and NPC 101.1 displayed variants in exon 4 and NPC 23.2 and NPC 119.1 displayed shifts in exon 9 of p53. The sequencing results are shown in Table 2.5.

Table 2.5. Sequence variants in exon 4 and 9 of *p53*.

SAMPLE	EXON	NUCLEOTIDE POSITION	AMINO ACID POSITION
NPC45.1, NPC101.1	4	G423C	P60R
NPC23.1, NPC1191.	9	C2989G	S315C

The P60R and S315C variants in exon 4 and 9 of *p53* were not detected in the 100 controls tested.

2.3.3. *MvaI* Restriction Endonuclease Analysis

The individuals who harbour the C1528T mutation were identified through PCR of exon 13, restriction endonuclease digestion with *MvaI* and subsequent agarose gel electrophoresis. Both alleles were restricted in those individuals negative for the mutation. Archived biological material was available for 219 individuals from family NPC 1. All of the individuals in this family have been tested for the C1528T mutation and the mutation co-segregated with every affected family member. One hundred population-matched controls were screened and found negative for this mutation.

In addition, the C1528T mutation was screened for in the cohort of 260 probands with *MvaI* restriction endonuclease analysis. Twelve additional apparently unrelated individuals displayed the identical band shift as seen in NPC 1 on 3% agarose gels. These families included NPC 16, NPC 49, NPC 54, NPC 66, NPC 68, NPC 77, NPC 82, NPC 97, NPC 109, NPC 134, NPC 157 and NPC 166 (Table 2.3).

2.3.4. Haplotype Analysis

Further analysis of these 13 individuals/families with the C1528T mutation in *hMLH1* (referred to as the C1528T-research cohort) with microsatellite markers around the *hMLH1* gene was carried out. A

disease-associated haplotype (1-4-2-1) for the markers D3S1612, D3S3512, D3S1611 and D3S1277 respectively, identified in family NPC 1 was observed in the 12 additional families, but the phase was unknown for the 6 'sporadic' individuals (i.e. NPC 16.1, 49.1, 54.1, 68.1, 82.1, and 166.1). All those individuals who were positive for the C1528T mutation inherited a specific allele, denoted allele 1, for marker D3S1611. This was an intragenic microsatellite marker, situated in intron 1 of the *hMLH1* gene (Fig. 2.2).

2.3.5. Genotype/Phenotype correlation

2.3.5.1. The primary research cohort

The primary research cohort (n=100) was divided into two groups in order to assess whether any correlation existed between the mutation that an individual harboured and the age at diagnosis of disease, as well as the site of the cancer. The two groups were:

- i. those affected subjects who harboured a disease-associated mutation in the MMR genes and
- ii. those affected subjects in whom no mutation was detected in the aforementioned genes.

No correlations could be made between any particular mutation and the age at diagnosis in the primary research cohort.

Incidence of cancers at various sites in each of the two groups was calculated and compared and are graphically displayed in Fig. 2.3.

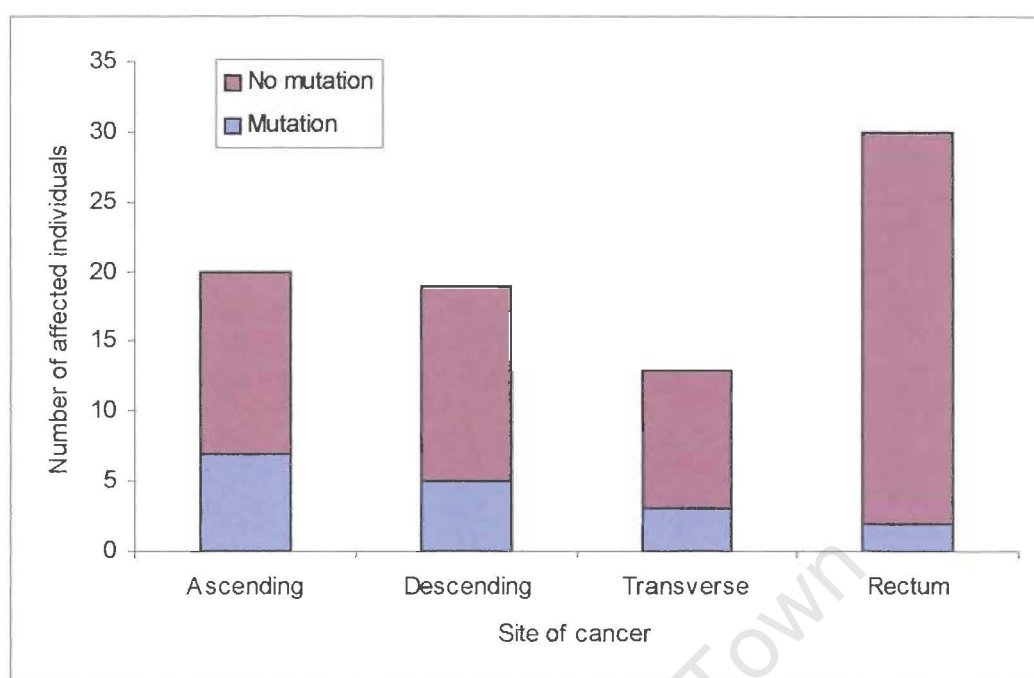


Figure 2.3. Bar graph of affected individuals in whom mutations were detected vs. those in whom no mutations were detected in the MMR genes investigated.

There was a weak association between the site of the cancer and a mutation in the MMR genes investigated ($p=0.09$). Those individuals who were diagnosed with rectal cancer had the least mutations detected in the MMR genes. Conversely, mutations in the MMR genes were mostly associated with cancer of the ascending colon.

2.3.5.2. The C1528T research cohort

The group of affected individuals (29/86) who harbour the C1528T mutation was further analysed to determine whether any correlation existed between this mutation and the site of cancer or the age at diagnosis in the colon. The results are shown in Fig. 2.4 and Fig. 2.5.

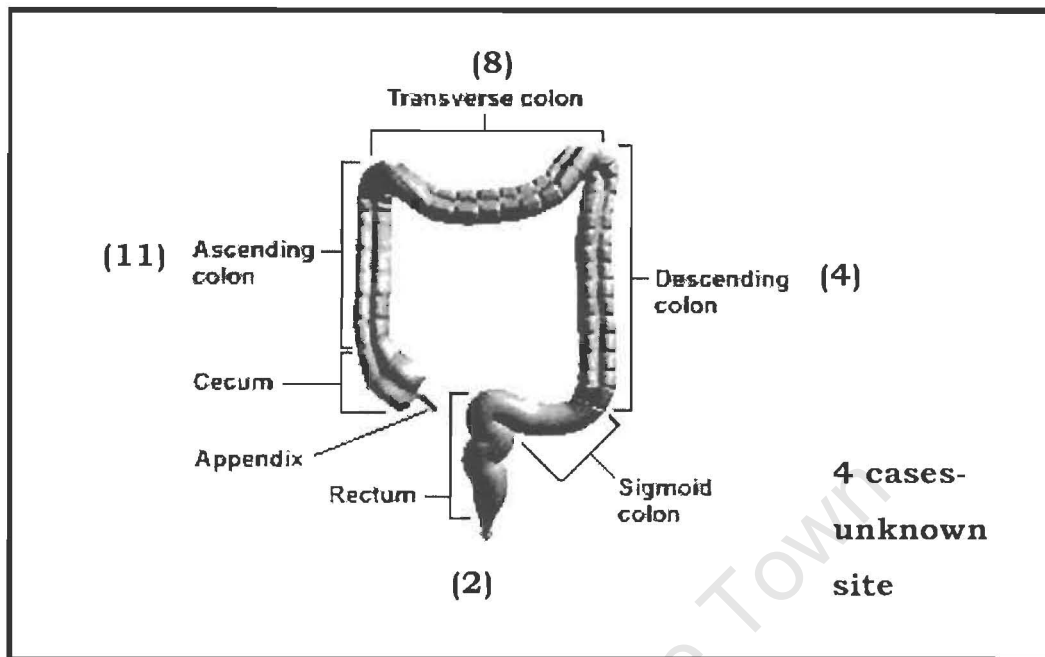


Figure 2.4. *Numbers of individuals (n = 29) displaying cancer at the various sites of the colon.*

Eleven of the 29 symptomatic subjects (38%) presented with cancer of the ascending colon ($p=0.059$) (Fig. 2.4), while 8 (28%) were diagnosed with cancer of the transverse colon. Cancer of the ascending colon was thus mostly associated with the C1528T mutation in this cohort of subjects, followed by cancer of the transverse colon. Individuals with cancer of the ascending and descending colon were diagnosed, on average, at 43 years of age, whereas those with cancer of the transverse colon were diagnosed, on average, at 33 years of age. Individuals with rectal cancer were symptomatic by age 30 years, on average.

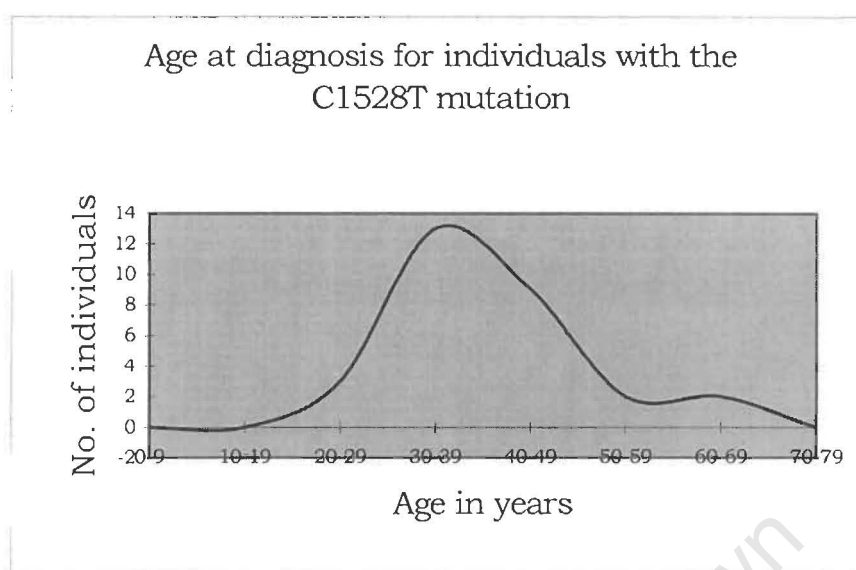


Figure 2.5. *The distribution of the age at diagnosis for individuals belonging to the C1528T-research cohort (n=29).*

2.3.6. MSI analysis

2.3.6.1. Fresh tumour tissue

Four of the 20 fresh tumour tissue samples were classified as MSI-H i.e. more than three of the five MSI markers listed in Table D.8. (Appendix D) displayed instability for one particular sample. Figure 2.6 illustrates the alleles in normal tissue (lane N) (which does not display any instability) and that in the tumour tissue (lane T), which does exhibit instability.

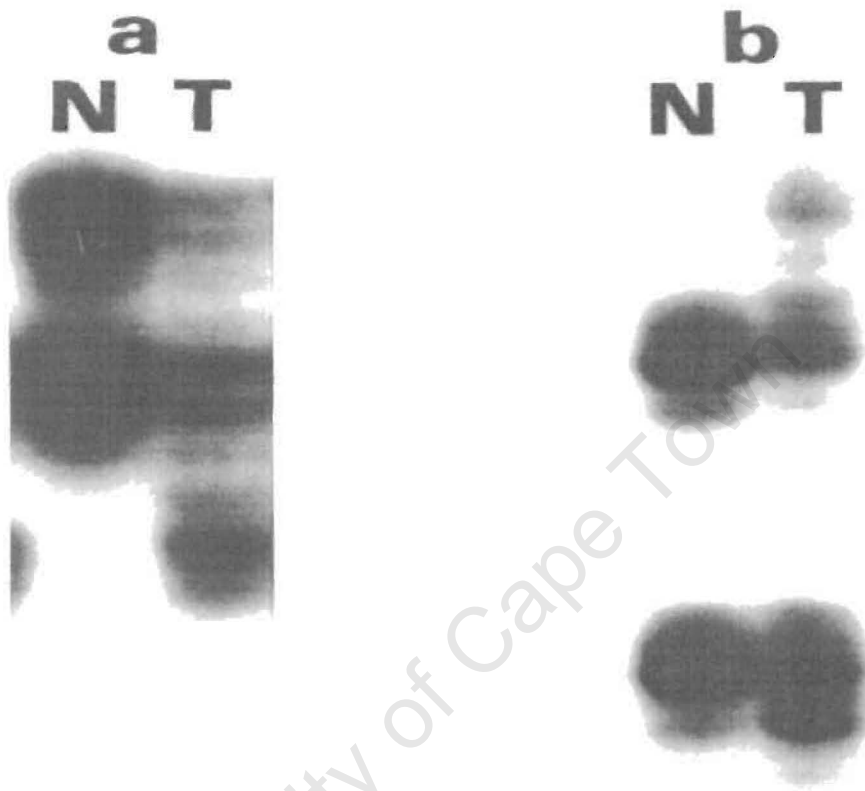


Figure 2.6. *Illustration of MSI in tumour tissue.*
(a): BAT25 (b): D2S123
[N: DNA from peripheral blood; T: DNA from tumour tissue]

MSI-H status is strongly suggestive of mutations in the MMR genes. Mutation analysis of the coding regions of *hMLH1*, *hMSH2* and *hMSH6* was performed for samples that were classified as MSI-H and sequencing analysis of the variant banding patterns revealed changes in *hMLH1* and *hMSH2* in NPC 36.3, NPC 63.1, NPC 1.175 and NPC 121.1 (Table 2.6). NPC 1.175 was known to harbour the C1528T mutation in *hMLH1* prior to MSI studies.

Table 2.6. Mutations detected in DNA extracted from fresh tumour tissue displaying MSI-H status.

SAMPLE	MSI STATUS	GENE/EXON	MUTATION
NPC 36.3	MSI-H	<i>hMSH2</i> exon7	1220delCT
NPC 63.1	MSI-H	<i>hMSH2</i> exon 8	1340insGG
NPC 1.175	MSI-H	<i>hMLH1</i> exon 13	C1528T
NPC 121.1	MSI-H	<i>hMLH1</i> exon19	C2152T

2.3.6.2. Paraffin-embedded tissue

Two of the 19 PETT samples displayed MSI. Mutation analysis of the coding regions of *hMLH1*, *hMSH2* and *hMSH6* was performed and sequencing analysis of the variant banding patterns revealed mutations in *hMLH1* and *hMSH6* in NPC 94.1 and NPC 101.1, respectively (Table 2.7).

Table 2.7. Microsatellite instability displayed for formalin-fixed paraffin embedded tissue samples.

SAMPLE	MSI STATUS	GENE/EXON	MUTATION
NPC 94.1	MSI-H	<i>hMLH1</i> exon 13	1521insC
NPC 101.1	MSI-H	<i>hMSH6</i> exon 4	T2633C

2.4. DISCUSSION

2.4.1. Mutation and Sequencing Analysis

As part of a large study investigating the genetics of CRCs, this investigation was meant to ascertain the contribution that HNPCC makes to the burden of this disease. One hundred individuals affected with CRC were screened for the presence of mutations that predispose to disease in a range of candidate genes (*hMLH1*, *hMSH2*, *hMSH6*, the MCR of *APC*, exons 3, 5 and 6 of *β-catenin* and exons 4 to 9 of *p53*) in South African individuals affected with CRC.

The mutation analysis aspect of the project led to the identification of both novel and previously-reported mutations and polymorphisms in the genes investigated. Five mutations were detected in *hMLH1*, six in *hMSH2*, one in *hMSH6*, four in *APC*, two in *p53* and none in *β-catenin*. The following sections discuss the mutations identified in this study with respect to the available clinical information on the affected individuals. In addition, the novel mutations and/or interesting scenarios are discussed in this section.

2.4.1.1. *hMSH2* :disease-associated mutations

Exon 6: The G965A mutation was detected in 2 unrelated affected individuals, NPC 85.1 and NPC 107.1. Twenty-four additional family members were available for analysis for family NPC 85. This family was classified as having Type II HNPCC as various family members have been diagnosed with endometrial, gastric, uterine, skin, pituitary and breast cancer.

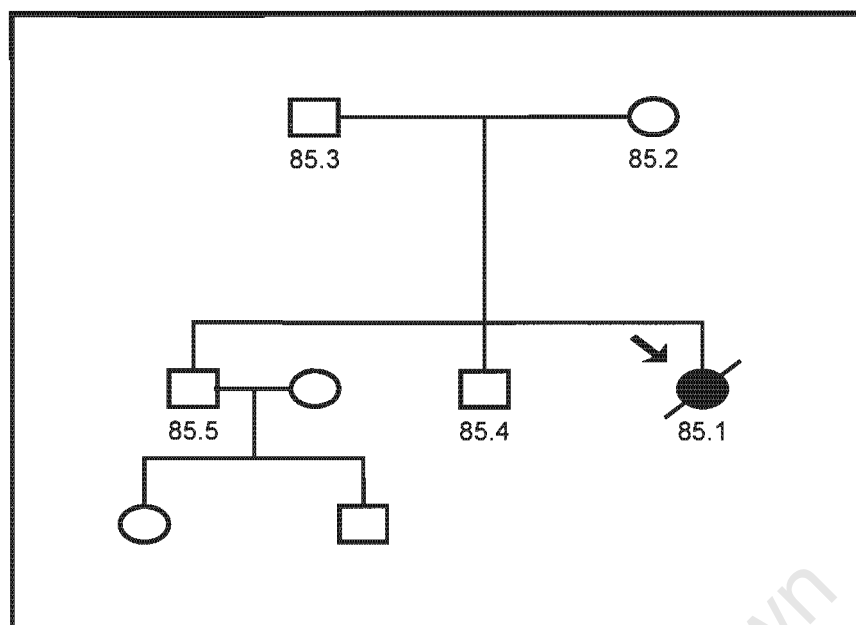


Figure 2.7. Pedigree of NPC 85.

Analysis of the pedigree (Fig. 2.7) indicated that NPC 85.1 and NPC 85.4 each inherited 2 germline mutations, the G965A in *hMSH2* (G322D) from their father NPC 85.3, and the C350T (T117M) mutation in *hMLH1* from their mother NPC 85.2 (Fig. 2.7). NPC 85.1 died at age 27 years from CRC. Her older brother, NPC 85.4 is alive (age 38 years), but experiences 'stomach problems'. He is currently enrolled in a colonoscopy-screening program and no abnormality has been detected to date.

Ten separate international groups reported the T117M mutation in *hMLH1* (<http://www.nfdht.nl>). This mutation maps to a region in the gene that is conserved and involved in mediating binding and/or hydrolysis of ATP (Jacob and Praz, 2002). The G965A transition has also been reported previously, but the pathogenicity of this missense mutation has not been established as yet (Maliaka *et al.*, 1996, <http://www.nfdht.nl>). This mutation was not detected in 100 controls tested.

The principal players in DNA mammalian mismatch repair function together as heterodimers. The key initiators of MMR are the mismatch-binding heterodimers MSH2/MSH6 and MSH2/MSH3. In addition, MLH1/PMS2 form a heterodimer and acts as a facilitator between the MSH2/MSH6 and the MSH2/MLH3 heterodimer and other players in MMR. The mutations that arise in subjects NPC 85.1 and NPC 85.4, if pathogenic, would have an effect on the role of both the MSH2/MSH6 and the MLH1/PMS2 heterodimers since mutations occur in both the *hMLH1* and *hMSH2* genes. This would lead to ineffective mismatch repair, which would result in the phenotype observed in NPC85.1. Since this theory does not account for the phenotype observed in her brother, NPC 85.4, it could be postulated that modifying factors may be playing a role in the variable phenotype observed in one or both individuals. If the mutated copies of these MMR genes were silenced due to methylation, it would also explain the less severe phenotype observed in individual NPC 85.4 as opposed to subject NPC 85.1. These possibilities all need to be investigated to determine the phenotypic difference observed.

Kariola *et al.* (2003) reported two families with CRC in which the index patients carry missense mutations in both *hMSH2* and *hMSH6*. These families do not meet the Amsterdam criteria for the clinical classification as HNPCC, but the tumours from the index patients displayed MSI. Previous reports of patients who harbour two MMR gene variations either in the same or in two different MMR genes indicate that this is a relatively rare event (Hackman *et al.*, 1997; Ricciardone *et al.*, 1999; Wang *et al.*, 1999; Wu *et al.*, 1999; De Rosa *et al.*, 2000; Samowitz *et al.*, 2001; Vilkki *et al.*, 2001; Wu *et al.*, 2001). Those individuals who harbour homozygous mutations in *hMLH1* and *hMSH2* display clinical features of neurofibromatosis type I (Wang *et al.*, 1999; Ricciardone *et al.*, 1999). Those subjects who harbour two mutations in different MMR genes do not display a more severe HNPCC phenotype

than the heterozygous carriers in general, which could imply that only one of the mutations is pathogenic (Kariola *et al.*, 2003). If allele segregation studies are not possible, because of a limited family size, the pathogenicity of these missense mutations may not be defined (Kariola *et al.*, 2003). Functional studies then need to be carried out to determine which of the mutations cause MMR deficiency in the tumour cells. Ideally, functional studies need to be carried out to determine the compound pathogenicity of the two mutations. The mutations identified by Kariola *et al.* (2003) did not appear to be pathogenic individually; the effect of the compound pathogenicity of these mutations within the cell could not be ruled out.

The C1046G alteration identified in NPC 45.1 has not been previously reported. Proline has an aliphatic hydrocarbon with a nonpolar R group. Arginine contains a positively-charged guanidinium group. The amino acid alteration is thus from a nonpolar group to a positively charged one. It was difficult to postulate the pathogenicity of this missense mutation since segregation analysis with additional affected individuals could not be carried out. This mutation was, however, not detected in the 100 controls tested. Individual NPC 45.1 was diagnosed with cancer of the descending colon at 54 years of age. The subject did not have recurrence of the cancer since her left hemi-colectomy surgical procedure in 1965; she recently died at 78 years of age of unrelated causes.

Exon 7: The 1220delCT identified in NPC 120.1 led to a stop nine codons downstream in the same exon. This DNA alteration has not been previously reported. No additional family members were available for analysis. This individual was diagnosed with synchronous cancer (the caecum and splenic flexure) at 40 years of age. She has not had a

recurrence of the cancer since her surgical procedure (right hemicolectomy) in 1988.

The 1220delCT mutation was also detected in a second family, *viz.* NPC 36. This family originally comprised of 6 individuals. Additional family members were recruited when the disease-causing mutation was identified and the biological materials archive contained material from 20 individuals (1 affected and 19 unaffected members). Of these individuals, 5 members (1 affected and 4 unaffected) harboured this mutation. Family NPC 36 is enrolled in the genetic counseling program and results have been given to 5 of the 16 family members.

Individual NPC 209.1 was the only member for whom biological material is archived in the laboratory. In addition, no history of colorectal cancer was documented.

The 1220delCT was detected in NPC 37.1. Biological material was archived from 2 other individuals in this family, who have as yet not been tested for the presence of this mutation. The index patients' paternal aunt was affected with colorectal cancer. No other family history was documented.

Exon 8: The 1340insGG detected in family NPC 5 results in a frameshift mutation with a termination signal being introduced seven codons downstream within exon 8. This mutation has not been previously reported.

Fifteen of the 43 recruited family members harboured this mutation, including five affected and ten unaffected individuals (age range 29-75 years). This mutation was associated with Type II syndrome as extra-colonic cancers (throat, ovarian and breast) were present in NPC 5. Mutation results have been given to 22 family members to date and

carriers of this mutation were made aware of the importance of monitoring the Type II spectrum of cancers.

Exon 9: The A1489G alteration in proband NPC 42.1 has not been previously reported. Both amino acids (isoleucine and valine) contain nonpolar R-groups. It was difficult to postulate the pathogenicity of this missense mutation since allele segregation analysis with additional affected individuals could not be verified. This was the only individual for whom DNA is available in this family. This alteration was not present in the 100 controls tested.

Exon 15: Genotyping of NPC 145 was carried out with microsatellite markers close to the *hMSH2* gene. Visual examination of the inheritance pattern for markers around *hMSH2* showed that a disease-associated haplotype for D2S119, D2S2378 and D2S123 (5-2-1) co-segregated with the disease phenotype (Appendix, section A.3.). In addition, two unaffected individuals (NPC 145.7 and NPC 145.14) harboured this disease-associated haplotype. This led to the preliminary conclusion that *hMSH2* could be the disease-causing gene. SSCP analysis of all of the exons of *hMSH2* revealed a band shift in exon 15. In addition to the SSCP co-segregating with the condition in family NPC 145, the variant was observed in the 2 unaffected individuals, NPC 145.7 and NPC 145.14. DNA sequencing of exon 15 of *hMSH2* revealed a 7bp deletion (CTAATTT) at nucleotide position 2501 (2501delCTAATTT) in the three affected individuals and in NPC 145.7 and NPC145.14. This mutation led to a frameshift and has been reported previously on the ICG-HNPCC web site (<http://www.nfdht.nl>). This family had individuals affected with endometrial cancer in addition to CRC.

2.4.1.2. *hMLH1*: disease-associated mutations.

Exon 6: An A->G substitution in the third position of intron 6 was observed in NPC 43.1. This change is in the region of the 5' donor-site consensus sequence (G->A 5' donor site + 3). While investigating RNA splicing errors, Carstens *et al.* (1991) reported a patient who displayed the identical base alteration (A->G) in the same position (+3) as the mutation in individual NPC 43.1. Although this alteration did not involve the GT of the splice donor, it did affect the established

consensus sequence for 5' donors. This mutation caused the preceding exon to be skipped in the patient studied by Carstens *et al.* (1991), even though the 3' splice acceptor was not affected. Individual NPC 43.1 was diagnosed with rectal cancer (Dukes' B carcinoma) at 38 years of age. The patient underwent a course of chemotherapy (5FU) and is currently being managed through regular colonoscopy. There has been no recurrence of the cancer since 1994. This mutation, which was detected in an Italian patient with HNPCC, caused total skipping of exon 6 in this individual (<http://www.nfdht.nl>). The individual in the current study originates from the indigenous black population group of South Africa.

Exon 9: Individual NPC 15.1 harboured the 731delGTGA, which resulted in a frameshift in exon 9, resulting in a truncated protein product. This mutation has not been reported previously.

Investigation of the proband's pedigree indicated that her deceased father was affected with CRC while she was diagnosed with cancer of the sigmoid colon and had a colectomy at age 25 years. She underwent a course of chemotherapy (Leucovorin and 5FU) and had no recurrence of the cancer. She died at 49 years of age of causes unrelated to cancer.

Exon 13: The 1521insT was detected in subject NPC 94.1, which resulted in a frameshift, introducing a termination signal eight codons downstream within this exon. This mutation has not been reported previously. Individual NPC 94.1, who was diagnosed with cancer of the transverse colon at age 30 years, had no family history of cancer.

The C1528T mutation was detected in family NPC 1 (C1528T) resulting in replacement of a glutamine by a termination codon at position 510 in the amino acid sequence (Q510X). The individuals from NPC 1, who harboured this mutation, have Type I HNPCC i.e. no extracolonic cancers were present in individuals who were positive for this mutation (Goldberg *et al.*, 1998).

Members of family NPC1 lived mostly in a remote village in the Northern Cape province of South Africa, approximately 50km from the nearest tarred road and about 800km North of Cape Town, the nearest city. The extended family members resided in coastal villages between Port Nolloth in the north and Cape Town in the south. A "clinical management team" which included geneticists (for genealogical tracing, risk assessment and genetic counseling), gastroenterologists (to perform colonoscopies) and an endoscopy nurse, visited these coastal areas annually. Since the advent of a mutation detection test, genetic counseling was offered to those individuals who requested it. Through the detection of this mutation, the family could be separated into those at high-risk (i.e. positive for this mutation) and those at negligible-risk for HNPCC.

Before the availability of the mutation test, the surgeons performed colonoscopies on every consenting high-risk family member (deemed to be at high risk i.e. more than 12.5%) who attended the clinic. This resulted in approximately 89 individuals being eligible for colonoscopic screening. Of the 219 family members on whom genetic tests were

done, 54 (25%) were positive for the C1528T mutation and therefore were most eligible for surveillance colonoscopies. The remaining 165 (75%) individuals who were negative for this mutation need not undergo routine surveillance for CRC. In addition to decreasing the number of warranted colonoscopies from 89 to 54 per year (and in some cases every second year), the mutation screening identified family members who could be released from this relatively burdensome follow-up and emotional concern.

Five of the 13 families that harboured the C1528T mutation *viz.* NPC 1, NPC 66, NPC 77, NPC 97 and NPC 157 are currently participating in the predictive testing program for HNPCC. A proportion of family members of NPC 77 were undergoing genetic counseling and mutation results were being offered to those individuals who requested it. Mutation test results have been given to 35 of 64 individuals in NPC 77 and the program has been successful, as subjects have been forthcoming and enthusiastic about receiving their mutation results. The rest of the family members are yet to receive results.

The proband in NPC 157 was diagnosed with CRC at 13 years of age. The disease progressed at an accelerated rate and the patient died at the age of 16 years. This is the youngest recorded individual in this cohort to die of CRC. In addition, an individual belonging to NPC 1 is positive for the C1528T mutation and remains clinically unaffected at 82 years of age. She has one mutation-positive son who was diagnosed with CRC at 40 years of age. Furthermore, this son has four children, who are all mutation positive, but asymptomatic at present.

DNA was available for more than one family member in 7 of the 13 kindreds harbouring the C1528T mutation. One subject was available for testing in the remaining 6 families.

Mutation results were given to 103 individuals to date who are 18 years of age and older. These five families are well informed about the implications of CRC and are currently participating in clinical surveillance programs in an attempt to monitor the disorder.

Table 2.8 shows the initial and final number of members per family of the kindreds currently enrolled in the predictive testing program. NPC 1 has grown from 55 members to 219 members over the last 12 years. The other four families have all grown in number over the past five years to the numbers depicted in Table 2.8. These five families are actively enrolled in a genetic counseling and predictive testing program.

Table 2.8. Increase in the number of family members of kindreds harbouring the Q510X mutation.

Family	Initial Number of members	Family size in 2003
NPC 1	55	219
NPC 66	6	21
NPC 77	4	108
NPC 97	1	25
NPC 157	3	63

The initial examination of NPC 66, 77, 97 and 157 led to the conclusion that these kindreds did not meet the Amsterdam criteria. Yet on further genealogical and clinical investigation, prompted by the detection of the C1528T mutation in the proband, it became clear that these families do indeed meet these criteria and may be labelled as HNPCC for research and management. The additional 7 families have yet to be investigated in this way, and could thus also fulfil the Amsterdam criteria on further analysis of their pedigrees, illustrating the power of identifying a mutation in a MMR gene, even in a single individual.

2.4.1.3. *hMSH6*: disease-associated mutations.

Exon 4: The T2633C substitution occurred in subject NPC 101.1. It is the only individual for whom biological material was available in this family.

Individual NPC 101.1 belonged to the Indigenous Black population group of South Africa and displayed an additional mutation in exon 4 of *p53* as well. This individual was diagnosed with cancer of the sigmoid and hepatic flexure (Dukes' B) at 33 years of age. In addition, paraffin-embedded tissue was available for this individual and microsatellite instability analysis with markers BAT26, BAT40, D2S123, D5S346 and D17S250 led to the tumour being classified as MSI-H. The T2633C mutation has previously been described in the literature by Wijnen *et al.* (1999) in subjects originating from the Netherlands and by Figer *et al.* (2002) in a family diagnosed with early-onset colorectal and endometrial cancer. Although this mutation has been logged onto the ICG-HNPCC mutation database (<http://www.nfdht.nl>) as a pathogenic mutation, its pathogenicity still has to be elucidated.

2.4.1.4. Silent polymorphisms in *hMSH2*, *hMLH1* and *hMSH6*.

The silent polymorphisms detected in *hMLH1*, *hMSH2* and *hMSH6* are shown in Table 2.2.

Of the silent polymorphisms, two were detected in *hMSH2*, five in *hMLH1* and five in *hMSH6*. Two of the seven polymorphisms in *hMLH1* and two of the six polymorphisms in *hMSH6* were previously described in the literature (Table 2.2). These polymorphisms are either in the non-coding regions of these MMR genes or they are positioned in exons but do not alter the sequence of the amino acids at the sites where they

occur, or they do alter the amino acid sequence but has been shown by other researchers to be non-disease causing..

There is a tendency to dismiss silent polymorphisms when detected in a large-scale screening of genes, but these polymorphisms should be used as tools by geneticists as they are important in the follow-up of probands, since these individuals have a history of bowel cancer. The identification of mutations in areas outside of the coding regions need not be dismissed as they might influence the expression levels of the genes under investigation and would assist international efforts to address the influence of these variations on disease. In fact, the first description of the identification of a disease-causing mutation in the 5' promoter region of *hMLH1* was reported in a Newfoundland HNPCC kindred (Green *et al.*, 2003).

2.4.1.5. APC missense variants

Exon 15: The T3920A mutation was identified in individual NPC 17.1. This was the only individual for whom DNA was available in this family. The subject was diagnosed with cancer of the ascending colon at 69 years of age and died two years later.

This mutation was not detected in the control cohort. Laken *et al.* (1997) identified the T3920A (I1307K) mutation in approximately 6% of the Ashkenazi Jewish population and postulated that it is associated with an elevated risk of colon cancer. Individual 17.1 was of Jewish descent.

The G3949C was identified in NPC 24.1, who is the only individual for whom DNA was available in this family. The subject was diagnosed with non-polypotic CRC (Dukes' D) at 41 years of age and died three years later. This variant was previously reported by Frayling *et al.* (1998)

The A5392G was detected in two individuals, NPC 10.1 and NPC 80.1, both of whom have rectal cancer. Individual NPC 10.1 was diagnosed at 69 years of age. A family history of non-polypotic CRC existed for NPC 10.1 i.e. her mother, three paternal uncles and a nephew died of CRC; no biological material was available for any of these individuals to test for the presence of the variation. Subject NPC 80.1 was diagnosed at 35 years of age, with no family history of cancer. This change has not been reported previously.

The D1822V alteration was identified in 15 unrelated individuals affected with the non-polypotic form of CRC in the heterozygous form and in the homozygous form in one individual. These individuals were from the Mixed Ancestry, Indigenous Black and Caucasian population groups. The significance of this variant is unknown but it has been postulated that the D1822V change may result in disruption of the putative cell signalling function of the APC protein as it is situated within the β -catenin down-regulation domain (Wijnen *et al.* 1999).

2.4.1.6. β -catenin

The absence of mutations in exons 3, 5 and 6 of β -catenin in the cohort and controls under investigation may reflect that this group of cancers is not influenced by germline mutations in these exons.

2.4.1.7. *p53*

Two different mutations were detected in *p53* in four of the 100 individuals in the primary research cohort. The G423C mutation in exon 4 of *p53* was detected in NPC 45.1 and NPC 101.1. The change resulted in a proline to arginine (P60R) substitution at position 60 of the amino acid sequence.

Both of these individuals had additional mutations in the MMR genes. A mutation, C1046G, was detected in exon 6 of the *hMSH2* gene in subject NPC 45.1 who was diagnosed with cancer of the descending colon at 54 years of age. A second mutation in *hMSH6* (V878A) was also detected in NPC 101.1. This individual was diagnosed with cancer of the sigmoid and hepatic flexure at 33 years of age.

The second mutation C2989G within exon 9 of the *p53* gene was detected in subjects NPC 23.1 and NPC 119.1. Individual NPC 23.1 was diagnosed with cancer of the descending colon (age at diagnosis not known) with no history of cancer. NPC 119.1 was diagnosed with cancer of the caecum at 44 years of age, with no family history of cancer. Neither of these mutations has been previously reported in *p53*. These mutations were not detected in the control cohort.

p53 is a tumour suppressor gene, which is mutated in 50-60 % of tumours. Germline mutations in this gene are associated with Li-Fraumeni and Li-Fraumeni-like syndromes. If a mutation in *p53* is observed in the lymphocytic tissue, there is a possibility that the individual inherited the mutation from a parent. Alternatively, the mutation could have developed early in the embryological stages of development, giving rise to its ubiquitous presence. Since biological material was not available for any of their parents, it was not clear whether individuals NPC 23.1, NPC 119.1, NPC 45.1 and NPC 101.1 inherited the *p53* mutations from their parents or if these represented new mutations in these subjects. These mutations may now be passed on to the offspring of these individuals. Since subjects NPC 101.1 and NPC 45.1 both harboured a mutation in the MMR system in addition to the *p53* mutations, there was a chance that these mutations could be passed on to their offspring as well. A mutation in the MMR system could reduce the cell's ability to correct errors introduced during DNA replication. A mutation in an oncogene such as *p53* could result in

uncontrolled proliferation of tissue. The consequences of inheriting a combination of these two mutations could result in the CRC phenotype observed in subjects NPC 101.1 and NPC 45.1.

2.4.2. Genotype/phenotype correlation

This phase of the study was an attempt to identify if any of the aspects of the disease phenotype correlated with the genotypes identified in the research cohort.

Those individuals who harboured the C1528T mutation were affected mostly with cancer of the ascending or proximal colon i.e. 38% of these individuals had cancer of the proximal colon, while 7% were diagnosed with rectal cancer. In addition, no extracolonic cancers were observed in this group. Furthermore, the age at diagnosis was variable for the individuals in this cohort i.e. 76% were diagnosed with CRC between 30 and 49 years of age. This information will be of value for the management of the disorder in this group of individuals. Since this mutation was associated with Type I syndrome in all 13 families, the at risk subjects could be advised to have regular colonoscopies to monitor the colon and rectum, but need not have screening of the other organs that form part of the HNPCC disease spectrum. It would be difficult however, to accurately predict the site of cancer within the colon, or the precise age at diagnosis, since these were not specifically defined by the C1528T mutation alone. This “varied phenotype” suggested that DNA variations other than the C1528T mutation modulate the HNPCC phenotype in this group of subjects. The mutation-positive individuals were subsequently screened for a range of variants in additional genes known to have a modifying effect on the phenotypic expression of CRC (Chapter 3).

A total of 41% of the mutation-positive subjects in the primary research cohort of 100 individuals were diagnosed with cancer of the ascending (or proximal colon) (Fig. 2.6). In comparison, individuals in the primary research cohort in whom mutations were not detected, had a higher incidence (43%) of rectal cancer ($p=0.09$), indicating that there was a weak association between the site of cancer and whether a mutation is detected in the MMR genes or not. These results for the primary research cohort relating to the site of cancer and age at diagnosis for HNPCC is in accordance with published data that demonstrate that mutations in the MMR genes are predominantly associated with cancer of the proximal colon, with an average age at diagnosis of 45 years (Bisgaard *et al.*, 2002). In addition, individuals affected with rectal cancer do not usually have mutations in MMR genes. Furthermore, it allows for a more directed mutation screening approach i.e. if individuals are clinically diagnosed with cancer of the proximal colon at a young age (before 45 years), then the MMR genes should be screened for mutations as a first-line screen. Screening for mutations is expensive and time-consuming and thus a directed genetic screening approach for CRC samples that are catalogued in the biological materials archive would reduce both the cost of the screening procedure as well as the time taken to do so. The results from studies such as the present one will enable researchers to make correlations between the genotype and clinical phenotypes observed within individuals and utilise this information in the genetic and clinical management of the disorder in these families.

2.4.3. Microsatellite Instability

Screening for mutations in the MMR genes in individuals who appear to be affected with HNPCC can be expensive and time-consuming, often with no definitive outcome. MSI analysis can be used as a pre-screen test prior to germline mutation analysis. The ICG-HNPCC and the NCI has recommended the use of a panel of five microsatellite markers to

identify MSI in tumour tissue from individuals affected with CRC. MSI has been identified in approximately 90% of patients diagnosed with HNPCC and in 12-15% of individuals with sporadic CRC. In addition, MSI analysis is simple and inexpensive and this makes it a useful tool for HNPCC screening. Once MSI-H has been detected in tumours the MMR genes can be investigated for the causative mutation.

MSI was detected in four tumour samples *viz.* NPC 36.3, NPC 121.1, NPC 63.1 and NPC 1.95 (Table 2.6), which were subsequently shown to harbour disease-causing mutations in the MMR genes. All 4 samples were MSI-H.

Tumour tissue was available for individual NPC 36.3 and microsatellite instability studies with markers BAT26, BAT40, D2S123, D5S346 and D17S250 led to the tumour being classified as MSI-H. Subsequent SSCP analysis of the affected individual, NPC 36.3, revealed a band shift in exon 7 of *hMSH2* and DNA sequencing revealed the del1220CT. The blood relatives of NPC 36.3 are enrolled in the predictive testing program in the Division of Human Genetics. Five of the 16 members in this family have received mutation test results to date. Those individuals who were positive for the del1220CT change underwent regular colonoscopic screening to ensure early detection and thus improved management of the disorder.

Subject NPC 121.1 harboured the C2152T missense mutation in exon 19 of *hMLH1*. The disease in individual 121.1 has progressed rapidly, with a period of approximately 24 months recorded between the onset of the disorder and death due to CRC. This individual belongs to the Indigenous Black population group of South Africa. This missense mutation was identified in the homozygous form in another subject, NPC 26.2, with a less severe prognosis. This individual was diagnosed at 59 years of age with rectal cancer at the Duke's D stage. This DNA

alteration was previously reported by three other groups, two studied colon cancer and the third was a study of MSI-H endometrial carcinoma (Farrington *et al.*, 1998, Kowalski *et al.*, 1997 and Weber *et al.*, 1999). This polymorphism is unique to the African American population group, with an allele frequency of 14% in a control group of individuals.

Analysis of 100 chromosomes from unrelated, unaffected population-matched individuals revealed that the heterozygous pattern was present in 2% of the population, with a clear absence of homozygotes for the C2152T transition.

Two of the 19 paraffin-embedded tumour tissue samples displayed instability (Table 2.7). The disease-associated mutations were identified upon mutation analysis of *hMLH1* and *hMSH2* and direct DNA sequencing.

DNA derived from tumour tissue was not available for all the subjects in the primary research cohort of 100 affected individuals. It was thus not possible to carry out microsatellite instability analysis on all these samples. Twenty fresh tumour tissue samples and 19 paraffin-embedded samples were available, which allowed for the implementation of MSI analysis in the research laboratory. The results obtained has indicated, as confirmed by the literature reports on MSI, that the approach to the search for mutations in MMR genes should start with MSI analysis, which would indicate the presence of a MMR defect. The work outlined in this thesis is part of the larger project conducted by the Colorectal Cancer Research Consortium (CCRC) of South Africa. The consortium comprises researchers from the Division of Human Genetics, the Division of Anatomical Pathology as well as the Surgical Gastroenterology Unit at Groote Schuur hospital. The identification of families with HNPCC requires the utilisation of a range of methods including microsatellite instability (MSI),

immunohistochemistry (IHC), which will direct the researcher to the mutation screening of the appropriate MMR gene. The findings presented in this thesis was carried out in the Division of Human Genetics, which included the large-scale screening of genes that are known to contribute to the aetiology of CRC. Microsatellite instability could not be conducted on all samples included in the primary research cohort of 100 individuals as tumour tissue was not available for the majority of samples. Thus IHC could not be carried out on these samples either. Researchers in the Division of Anatomical Pathology have recently set up the method for IHC, aimed at testing the methods and the detection rates in the environment. A separate research cohort of 94 tumour samples is a pilot study and the results of this work are part of another investigation. At the post-thesis level, MSI and IHC analysis will be conducted on each of the samples that enter the Division of Human Genetics/CCRC for further investigation. This protocol has been expanded to include the Lynch II spectrum of tumours. In addition, an attempt is being made to obtain tumour tissue from every patient with CRC who undergoes surgery at the Surgical Gastroenterology Unit at Groote Schuur Hospital.

The importance of correlating the MSI and IHC result has been recognised in order to identify mutations in the MMR genes in these subjects affected with CRC before 45 years of age.

Detection of mutations in the MMR genes allows for the identification of families with HNPCC by genetic means. Any at-risk individuals can then be monitored for premonitory signs of the disease. In the current study, mutation analysis of the MMR genes in subjects who were diagnosed with CRC (mostly under 45 years of age) allowed 12 disease-causing mutations to be detected in 13 index cases/families, a frequency of 13% of HNPCC in these CRC cases. Additional genes known to be causative of CRC were screened in these 100 individuals in

an effort to detect any mutations. Four mutations (4%) were detected in the MCR of *APC*, two mutations (2%) were detected in *p53* and none in *β-catenin*. This data set of 100 families still needs to be screened for mutations in the other mismatch repair genes associated with HNPCC (*hPMS2* and *hMLH3*). These individuals can now join the genetic screening and counseling program that is conducted by the Division of Human Genetics, as well as the clinical screening program that is conducted by the Surgical Gastroenterology Unit at Groote Schuur Hospital. Experience has shown that these index cases are representative of larger families in which the disease is mostly familial. Knowledge of mutation status could prevent the disease before it manifests, improving the quality of life for patients and lowering the cost relating to the burden of disease.

2.4.4. The mutation test procedure

The biological material under investigation was collected purely for research purposes. Upon the identification of a disease-causing mutation, the individual/family is then contacted again and invited to enroll in the genetic counselling program and mutation test procedure. Those individuals who belong to a family in whom the disease-causing mutation is known attend an information session where the mutation test procedure is explained and they are handed an information pamphlet (Appendix, section A.2.). The individual/family is then counseled with regard to their risk and asked whether they would like to be informed of their mutation status. No one is compelled to know his or her mutation status. A genetic counselor and a nurse are present when results are given. Thus far, results have been given to five of the thirteen families who harbour the C1528T mutation and patients usually arrive in family groups. The family is addressed and any questions are then answered. The patients then meet individually with the genetic counselor who administers the results. An additional

family member is allowed to be present when receiving results, if requested by the patient.

The relevant mutation tests in the current study were carried out on at least two separate DNA samples, on different occasions. This enables verification of test results. The utmost care is taken at every step during the test to maintain 100% accuracy and reproducibility.

The policy for individuals at high risk of developing HNPCC is to offer an initial colonoscopy at age 16 years. When the subject reaches the age of 18 years, a genetic counselor explains the mutation identification test procedure, the implications of a positive or negative result and the voluntary basis of the test. Those who refuse the test are counseled regarding the need for colonoscopic follow-up and offered this form of surveillance once every two years before the age of 30 and once every year thereafter. Those who submit to testing and who are shown to carry the mutation remain in a high-risk group and the importance of continued colonoscopic screening is stressed. In addition, it is also emphasised that their children are at 50% risk of inheriting the mutated gene and would require colonoscopic screening from the age of 16 years onwards. DNA testing could be provided to the offspring, if requested by them, once they were over the age of 18 years. Individuals who test negative for the C1528T mutation were removed from the at-risk group requiring regular colonoscopic surveillance. Care was taken to inform these individuals that their risk of developing CRC was not 0% but the same as that for the general population. These individuals no longer have to be concerned about the possibility that their children could inherit the disease-causing mutation. Genetic counseling is provided and available both before and after the results are given to the subject.

In conclusion, the current study suggests that individuals with a history suggestive of HNPCC be referred to a geneticist even if the Amsterdam criteria are not fulfilled. None of the 100 individuals in the primary research cohort diagnosed with CRC and chosen for mutation analysis met the Amsterdam criteria, yet 13 disease-causing mutations were found in these individuals.

Although searches for mutations in the MMR genes are costly and time-consuming, it is likely that more sensitive and cost-effective methods for assessing these genes will be developed. In addition, the mutation analysis procedures such as SSCP and dHPLC used in the current study will underestimate the number of mutations that are present since these methods cannot detect large genomic deletions, which makes up approximately 27% of the mutations identified in *hMSH2* and *hMLH1* (Wagner *et al.*, 2003) and 10-15% in *APC* (Sieber *et al.*, 2002). Methods such as the multiplex ligation-dependent probe amplification (MPLA) is a quantitative multiplex PCR approach to determine the relative copy number of each MLH1 and MSH2 exon (Gille *et al.*, 2002). This is a relatively new method and requires relatively low concentrations of DNA (20 ng) and is able to discriminate sequences that differ in only a single nucleotide. The next phase of the project is to identify large genomic deletions in the primary research cohort by using the MPLA and/or alternative methods.

The ability to identify the different types of DNA defects causative of HNPCC and confirmation that the MMR genes are defective in the majority of HNPCC patients have important practical implications.

Health care workers who are the first to establish contact with a patient in the hospital setting need to enquire about and record a detailed accurate family history of the patient. This will enable researchers to identify those families who are most likely to be affected with hereditary

forms of CRC and then to screen the genes that underly these disorders.

Identifying the disease-causing genes is merely one aspect of a successful screening programme. Educating patients on how to successfully manage the disease is a vital step in successfully managing this disease. For individuals diagnosed with HNPCC it would involve colonoscopy and screening for extracolonic cancers.

The recording of data to construct an accurate family pedigree as well as direct genetic screening methodology represent the best approach for finding mutations in MMR genes in those individuals and families diagnosed with early-onset CRCs.

CHAPTER 3: MODIFIERS OF THE HNPCC PHENOTYPE

University of Cape Town

CHAPTER 3: MODIFIERS OF THE HNPCC PHENOTYPE

3.1. INTRODUCTION

It has become increasingly apparent, in the field of medical genetics that few human diseases are caused entirely by a mutation in a single gene (Nebert, 2002). Certain diseases are monogenic (caused by a single gene), but most are polygenic (caused by two or more genes) or multifactorial (affected by environmental, genetic and epistatic factors) (Nebert, 2002). The lack of observed genotype/phenotype correlation in the single-gene disorders suggests that additional modifier genes or loci could be influencing the clinical phenotypes of various diseases (Dipple and McCabe, 2000). A “modifier” is an inherited genetic variation that leads to a qualitative or quantitative difference in the disease phenotype (Muller *et al.*, 2000). These variations are distinct from the disease locus and may play a role as low penetrance risk factors (Houlston *et al.*, 2001). Dipple and McCabe (2000) describe 5 disorders *viz.* cystic fibrosis, Hirschprung disease, neurofibromatosis type 2, craniosynostosis and adrenal hypoplasia congenita in which modifying genes play a role in the pathogenesis of the disorder. These modifiers range from chromosomal loci, to environmental factors (including diet), to known genes. Dipple and McCabe (2000) argue that there is one gene responsible for the pathogenesis of single-gene disorders with one or more independently inherited modifier genes. Independently inherited alleles together with additional modifier genes could result in the development of a complex disease.

Cancer development and progression is a multiplex process involving numerous primary and modifier genes. As the knowledge of the influence and involvement of primary and modifying cancer genes

increases, it should be possible to more accurately predict the severity and the age at diagnosis of an individual's risk of developing the disorder. Understanding the role of genetic modifiers could contribute to the knowledge of the process of carcinogenesis in humans (Ilyas *et al.*, 1999).

Modifying genes are thought to play a role in the pathogenesis of FAP since identical mutations display phenotypic differences in the manifestation of disease (Houlston *et al.*, 2001). Experiments conducted on the *Min* mouse model of polyposis has shown the existence of a modifier locus, termed *Mom1*, on mouse chromosome 4, syntenic to human chromosome 1p35-p36. The secretory phospholipase A2 gene was mapped to this locus, but no variants were identified in subjects affected with FAP. Candidate modifier genes for FAP include the human homologue of *Mom1* (1p35-p36), members of the *N-acetyltransferase* family, *SMAD3*, and *COX2* (Houlston *et al.*, 2001).

Individuals from families diagnosed with HNPCC, carrying identical mutations in the DNA mismatch repair genes (*hMLH1* and *hMSH2*), exhibit phenotypic variation with respect to the age at diagnosis of disease symptoms, the site of the cancer and the recurrence of cancer after treatment. The phenotypic variation observed between individuals who harbour the same predisposing mutation in *hMLH1* and *hMSH2* is possibly indicative of the influence of modifying genes. DNA variants known to modulate the HNPCC phenotype have been reported in several genes. These genes include the *ataxia telangiectasia* gene, *cyclin D1* and those involved in carcinogen metabolism (*N-acetyltransferases*, *glutathione-S-transferases*) (Kong *et al.*, 2000; Moisio *et al.*, 1998; Maillet *et al.*, 2000; Bala and Peltomaki, 2001). In addition, a polymorphism in *MTHFR* is known

to influence CRC with microsatellite instability (Shannon *et al.*, 2002).

A research cohort of 86 subjects was investigated in this phase of the study. These individuals, who belong to 13 different families, harbour the same nonsense mutation (C1528T) in exon 13 of *hMLH1*. This cohort was screened for the presence of known mutations and polymorphisms in various candidate modifier genes to determine if these variants had any modifying effect on the age at diagnosis of the disease. All the variants described in these genes are believed to play a role in the initiation and progression of CRC. These genes, and the relevant variants chosen for analysis, are described below.

3.1.1. *E-Cadherin (CDH1)*

Mutations in the genes whose protein products form part of the *Wnt* signalling pathway have been observed in approximately 90% of CRC investigated (Giles *et al.*, 2003). Activation of this pathway is initiated when Wnt binds to a membrane receptor, called Frizzled. This binding signals a host of intracellular proteins, including a protein called β -catenin, which binds to the cytoplasmic domain of E-cadherin, a cell-cell adhesion molecule (Rubinfeld *et al.*, 1993).

In normal cells, GSK-3 and APC proteins promote degradation of free β -catenin by phosphorylation of the -NH₂ terminal sequence. GSK-3 activity and β -catenin degradation are inhibited by the activation of the Wnt pathway as a result of the action of the Frizzled receptor and Dishevelled (DSH) signalling protein. In the absence of Wnt signaling, APC and GSK3 cooperate to degrade β -catenin. In response to Wnt signaling, via the Frizzled receptor, GSK3 is inhibited and β -catenin is stabilized (Moon and Miller, 1997). Inactivation of APC in colorectal tumours allows β -catenin to accumulate and complex with the Tcf-4

transcription factor in the nucleus, activating the expression of Tcf-4-regulated genes. Tcf-4 is expressed in the nuclei of intestinal epithelial cells (Korinek *et al.*, 1997).

E-cadherin expression is reduced in individuals with CRC and germline mutations in *CDH1* have been described in subjects affected with colorectal and gastric cancer (Guilford *et al.*, 1998; Richards *et al.*, 1999; Shinmura *et al.*, 1999; Kim *et al.*, 2000). *CDH1* has been suggested as a candidate gene for familial or early-onset CRCs (Richards *et al.*, 1999). Since mutations in *E-cadherin* have been associated with CRC, the disease cohort was investigated for the presence of eight reported germline mutations, previously shown to cause colorectal and/or gastric cancer (Table 3.1). E-cadherin will be discussed in this chapter. An introduction to *APC* and β -catenin is provided in chapter 1 (1.3.2.1) and chapter 2 (2.1.3) respectively.

E-cadherin (*CDH1*) belongs to the cadherin family of cell surface glycoproteins (Kim *et al.*, 2000) and is a Ca^{2+} dependent cell-cell adhesion molecule, located primarily in epithelial tissues (Kim *et al.*, 2000). The *E-cadherin* gene located on chromosome 16q23, consists of 16 exons and encodes a transmembrane protein, which exists as a homodimer (Shapiro *et al.*, 1995; Nagar *et al.*, 1996). *CDH1* has five extracellular repeat domains and each contains a histidine-arginine-valine (HAV) motif (Overduin *et al.*, 1995). The cytoplasmic tail of the protein binds to β -catenin (or γ -catenin) which, in turn, complexes with α -catenin. E-cadherin works with α and β -catenin as a functional unit termed the E-cadherin-catenin unit (ECCU) (Ilyas and Tomlinson, 1997). Through β -catenin, the ECCU is thought to bind to the actin cytoskeleton and a mutation in any part of the ECCU results in loss of function of E-cadherin (Ilyas and Tomlinson, 1997). The cadherin-catenin complex is essential for normal epithelial cell-cell adhesion (Kim *et al.*, 2000).

transcription factor in the nucleus, activating the expression of Tcf-4-regulated genes. Tcf-4 is expressed in the nuclei of intestinal epithelial cells (Korinek *et al.*, 1997).

E-cadherin expression is reduced in individuals with CRC and germline mutations in *CDH1* have been described in subjects affected with colorectal and gastric cancer (Guilford *et al.*, 1998; Richards *et al.*, 1999; Shinmura *et al.*, 1999; Kim *et al.*, 2000). *CDH1* has been suggested as a candidate gene for familial or early-onset CRCs (Richards *et al.*, 1999). Since mutations in *E-cadherin* have been associated with CRC, the disease cohort was investigated for the presence of eight reported germline mutations, previously shown to cause colorectal and/or gastric cancer (Table 3.1). E-cadherin will be discussed in this chapter. An introduction to *APC* and β -catenin is provided in chapter 1 (1.3.2.1) and chapter 2 (2.1.3) respectively.

E-cadherin (*CDH1*) belongs to the cadherin family of cell surface glycoproteins (Kim *et al.*, 2000) and is a Ca^{2+} dependent cell-cell adhesion molecule, located primarily in epithelial tissues (Kim *et al.*, 2000). The *E-cadherin* gene located on chromosome 16q23, consists of 16 exons and encodes a transmembrane protein, which exists as a homodimer (Shapiro *et al.*, 1995; Nagar *et al.*, 1996). *CDH1* has five extracellular repeat domains and each contains a histidine-arginine-valine (HAV) motif (Overduin *et al.*, 1995). The cytoplasmic tail of the protein binds to β -catenin (or γ -catenin) which, in turn, complexes with α -catenin. E-cadherin works with α and β -catenin as a functional unit termed the E-cadherin-catenin unit (ECCU) (Ilyas and Tomlinson, 1997). Through β -catenin, the ECCU is thought to bind to the actin cytoskeleton and a mutation in any part of the ECCU results in loss of function of E-cadherin (Ilyas and Tomlinson, 1997). The cadherin-catenin complex is essential for normal epithelial cell-cell adhesion (Kim *et al.*, 2000).

E-cadherin is important in the formation and maintenance of *adheren* junctions when epithelial cells come into contact with each other (Ilyas and Tomlinson, 1997).

3.1.1.1 Mutations in *CDH1*:

Mutations in *CDH1* have been detected along the whole length of the gene with no clear hotspots. Loss of E-cadherin function is not only the result of a mutation in *CDH1*. Mutations in other components of the ECCU can also abolish E-cadherin-mediated adhesion (Morton *et al.*, 1993; Oyama *et al.*, 1994). In addition, methylation of the promoter region of *CDH1* causes transcriptional inactivation, which leads to down-regulation of E-cadherin expression in some tumours (Graff *et al.*, 1995).

CDH1 germline mutations were first observed in familial forms of gastric cancer (Guilford *et al.*, 1998). Prior to this, somatic mutations were identified in this gene in histologically diffuse gastric carcinoma, lobular breast cancers and carcinomas of the endometrium and ovary (Guilford *et al.*, 1998). The inactivation of both alleles of *CDH1* by loss of heterozygosity and a separate mutation on the altered allele is common in sporadic cancers (Guilford *et al.*, 1998). Epigenetic silencing of *CDH1* has also been observed in tumours (Graff *et al.*, 1995).

Mutations in *CDH1* result in the disruption of normal cell-cell adhesion (Guilford *et al.*, 1998). This could play a role in the initiation of cell proliferation due to the escape of growth-control signals (Guilford *et al.*, 1998). Alternatively, the cytoplasmic domain of E-cadherin could alter Wnt signalling by reducing the availability of free β -catenin.

Table 3.1. Summary of germline mutations in CDH1.

PRIMER	EXON	NUCLEOTIDE CHANGE	CONSEQUENCE	REFERENCE
CDH(49-2)FM1 CDH(49-2)FW1 CDH(49-2)REV	2bp before exon 2	A(49-2)G	Premature truncation	Richards <i>et al.</i> , 1999
G59AFW1 G59AFW1 G59AREV	2	G59A	Premature truncation	Richards <i>et al.</i> , 1999
G62V WTF G62VMUTF G62VREV	3	G185T	Missense	Shinmura <i>et al.</i> , 1999
CDH7FW1 CDH7FM1 CDH7REV	7	G1008T at splice site	Splice site	Guilford <i>et al.</i> , 1998
T340AFW1 T340AFM1 T340AREV	8	A1142G	Missense	Kim <i>et al.</i> , 2000
L649VFW1 L649VFM1 L649VREV	12	C2012G	Missense	Kim <i>et al.</i> , 2000
CDH13FW1 CDH13FM1 CDH13REV	13	C2095T	Premature termination	Guilford <i>et al.</i> , 1998
CDH15FW1 CDH15FM1 CDH15REV	15	2382-2386 ins C	Frameshift	Guilford <i>et al.</i> , 1998

3.1.2. Missense variants in the adenomatous polyposis coli gene (APC)

Missense germ-line variants of APC have been described in individuals who develop multiple adenomas or carcinomas at a young age but who have not been diagnosed with FAP (Fearnhead *et al.*, 2001). A T3920A missense mutation was identified in the APC gene, which results in an isoleucine to lysine variant at codon 1307 (I1307K) (Laken *et al.*, 1997). This mutation substitutes an (A)₈ tract for the (A)₃T(A)₄, which may cause the DNA polymerase to incorporate

errors during DNA replication (Laken *et al.*, 1997). It is postulated that the (A)₈ tract creates an unusual secondary structure that is processed incorrectly by DNA polymerase in human cells (Laken *et al.*, 1997), which could indirectly predispose to cancer. Further investigations of this mutation have shown that it occurs because of a selective effect, which could be due to a dominant negative influence of mutations in this critical region of *APC* (Fearnhead *et al.*, 2001). The I1307K mutation was reported to confer susceptibility to CRC in individuals with the sporadic form of the disease. This mutation has only been detected in individuals from Ashkenazi Jewish descent (Laken *et al.*, 1997).

Another germ-line variant in *APC* is the G3949C transversion, resulting in a glutamic acid to glutamine substitution at position 1317 (E1317Q) (Frayling *et al.*, 1998). This variant predisposes to colorectal adenomas and carcinomas in populations studied to date, but with low and variable penetrance. It has been hypothesised that the E1317Q missense mutation either exerts a pathogenic effect directly, by a yet undetermined mechanism, or is linked to a cryptic *APC* mutation close-by, which predisposes to the development of multiple adenomas and carcinomas (Frayling *et al.*, 1998). The E1317Q variant substitutes an uncharged hydrophilic amino acid for an acidic hydrophilic one. This substitution may itself be sufficient to affect the function of the *APC* protein (Frayling *et al.*, 1998).

Tumours develop in the colon and rectum of carriers of both the I1307K and the E1317Q mutations and surveillance by colonoscopy is recommended. A study by Frayling and colleagues (1998) suggests that some carriers of E1317Q may have a tendency to develop multiple metaplastic polyps of the colorectum. Since the I1307K and the E1317Q variants have a low penetrance, they will not usually be present in families containing cases of colorectal carcinomas.

Individuals who harbour these variants would manifest with sporadic CRC or possibly have one or two family members affected with colorectal cancer (Frayling *et al.*, 1998). The existence of I1307K and E1317Q raises the possibilities that 'other variants exist at subpolymorphic frequencies that could have a significant, but not entirely penetrant effect on the incidence of colorectal adenomas and carcinomas' (Frayling *et al.*, 1998). Since these missense variants are present in sub-polymorphic frequencies (i.e. significantly higher than rare mutations maintained by mutation selection balance) they are associated with little or no selective disadvantage and the frequencies of these mutations can thus easily be increased in the population due to random genetic drift (Bodmer, 1999; Fearnhead *et al.*, 2001). In certain populations, however, the variants may reach polymorphic frequencies of between 6-8%, such as the I1307K mutation in the Ashkenazi Jewish populations. These variants may confer a growth advantage and thus individuals harbouring these mutations are at a significant risk of developing colorectal tumours (Fearnhead *et al.*, 2001). Missense variants have also been identified in other candidate genes that form part of the adenoma to carcinoma sequence i.e. *CDH1*, *hMLH1* and *hMSH2*. These missense variants may contribute to multifactorial disease inheritance, possibly at a low penetrance (Bodmer, 1999).

These missense variants have been investigated by other researchers in subjects affected with familial CRC (including HNPCC) to determine if these variants are predisposing factors to the onset of disease in these individuals (Evertsson *et al.*, 2001).

3.1.3. The Ataxia telangiectasia gene (*ATM*)

Ataxia-telangiectasia (A-T) is an autosomal recessive disorder characterised by progressive cerebellar ataxia, chromosomal

instability, immunodeficiency, cancer susceptibility, and radiation sensitivity (Gatti *et al.*, 1991). After considerable work in families with the A-T phenotype, defects in a gene, subsequently termed *ATM*, were detected. The *ATM* protein belongs to the phosphatidylinositol 3-kinase (PI-3-kinase)-like family, which comprises cell cycle regulators in different species (Savitsky *et al.*, 1995).

ATM is largely found in the nucleus of proliferating cells, but 10-20% of this protein is present in cytoplasmic vesicles. Single-strand and double-strand breaks occur in the DNA from cells that are exposed to ionising radiation. This activates *ATM* by an unknown mechanism. *ATM* then phosphorylates a number of substrates, including p53, chk-2, RP, PHAS1 and c-Abl. *ATM* phosphorylates p53 on the serine-15 residue in response to DNA damage. This destabilises and/or activates the transcriptional activity of p53, which, in turn, induces a number of different genes. The activation of the G1/S and the G2/M checkpoint of the cell are defective in the mutated *ATM* protein.

3.1.3.1. Mutations in *ATM*

The *ATM* gene is located on chromosome 11q23.1. *ATM* is 150 kb in length and consists of 66 exons, which translate to a 350 kDa nuclear phosphoprotein with a transcript of 13kb. Individuals heterozygous for germ-line *ATM* mutations do not exhibit the typical phenotype of the disease but have been suggested to be at high-risk of developing epithelial malignancies such as breast and digestive tract cancers (Swift *et al.*, 1991). This led Mailett *et al.* (2000) to regard *ATM* as a good candidate and modifier gene for HNPCC.

A G-> A polymorphism at nucleotide position 5557 in exon 39 of *ATM* substitutes an asparagine for an aspartic acid at position 1853 of the protein (D1853N). The frequency of the (A) and the (G) alleles was estimated to be 0.15 and 0.85 in the general population, respectively

(Maillett *et al.*, 1999). Sandoval *et al.* (1999) estimated the relative frequencies of the (A) and (G) alleles at 0.18 and 0.82 respectively. Maillett *et al.* (2000) evaluated the frequency of the *ATM* D1853N variant and its correlation with phenotypic expression. The variant was associated with a significantly higher incidence of CRC and other HNPCC-related cancers when compared to individuals carrying the *ATM* D1853D variant (92% *vs.* 57.5%, $p=0.02$). *MLH1* and *MSH2* mutation-positive individuals who carried the *ATM* D1853N variant had an eight times increased risk of developing CRC and other HNPCC-related cancers (OR: 8.9; $p=0.02$). It was suggested that the *ATM* D1853N polymorphism modulates the penetrance of *hMLH1* and *hMSH2* germ-line mutations.

3.1.4. The Cyclin D1 gene (CCND1)

The cyclin-dependent kinases (Cdks) are protein kinases that are central to the cell cycle process (Clurman and Roberts, 1995). These kinases aid in the progression of cell proliferation by phosphorylating specific substrates in a cell-cycle dependant manner. An active protein kinase consists of a Cdk subunit that binds with a cyclin subunit to form a heterodimer. Unique combinations of cyclins and Cdks assemble during each phase of the cell cycle and the specific activities of these protein complexes are essential for progression through the various phases of the cell cycle (Clurman and Roberts, 1995; Donnelin and Chetty, 1998).

Classes of cyclins: Eight major classes of mammalian cyclins have been identified (Donnelin and Chetty, 1998). At least 11 major mammalian cyclins have been identified *viz.* A, B1, B2, C, D1, D2, D3, E, F, G and H (Hunter and Pines, 1994; Sherr, 1994). The D-type cyclins (D1, D2 and D3) are primarily involved in the G1 phase

in mammalian cells and assemble into holoenzymes with the kinase catalytic subunits Cdk4 and Cdk6.

The role of cyclin D1 in the cell cycle: Proteins that drive the cell cycle, such as the cyclins, are frequently over-expressed in primary tumours. Over-expression of cyclin D1 in malignant cells shortens the G1 phase of the cell cycle, decreases cell size and reduces the dependency of the cell on mitogens in both animal models and cell lines (Ohtsubo and Roberts, 1993; Resnitzky *et al.*, 1994). Transfection of cyclin D1 into normal fibroblasts stimulates proliferation (Quelle *et al.*, 1993). Conversely, microinjection of cyclin D1 antibodies into dividing cells causes them to remain in the G1 phase (Quelle *et al.*, 1993).

Cyclin D1 is one of the important components driving cells through the restriction START, early in the G1 phase (Fig. 3.1) as cells are committed to complete the cell cycle when reaching this point. *CCND1* reaches maximal activity during the G1 phase where it plays an important role in the transition of the G1 to S phase of the cell cycle (Morgan, 1995; Sherr *et al.*, 1996). The fundamental task of a cell during this process is to ensure that DNA is replicated once during the S phase and that the identical chromosomal copies are distributed equally to two daughter cells during the M phase (Heichman and Roberts, 1994). Oncogenic processes exert their greatest effect by targeting the regulators of the G1 phase of the cell cycle (Hall and Peters, 1996). During this phase, cells respond to extracellular signals by either advancing toward another division, or withdrawing from the cell cycle and entering the resting state (Go) (Pardee, 1989; Sherr, 1994). Unlike the transit through the S, G2 and M phases, the G1 progression normally relies on stimulation by mitogens and can be blocked by anti-proliferative cytokines. Cancer cells are capable of overcoming these controls and continue cycling.

Cell cycle exit facilitates maturation and terminal differentiation and cancer cells by-pass these processes as they remain in a continuous cycle.

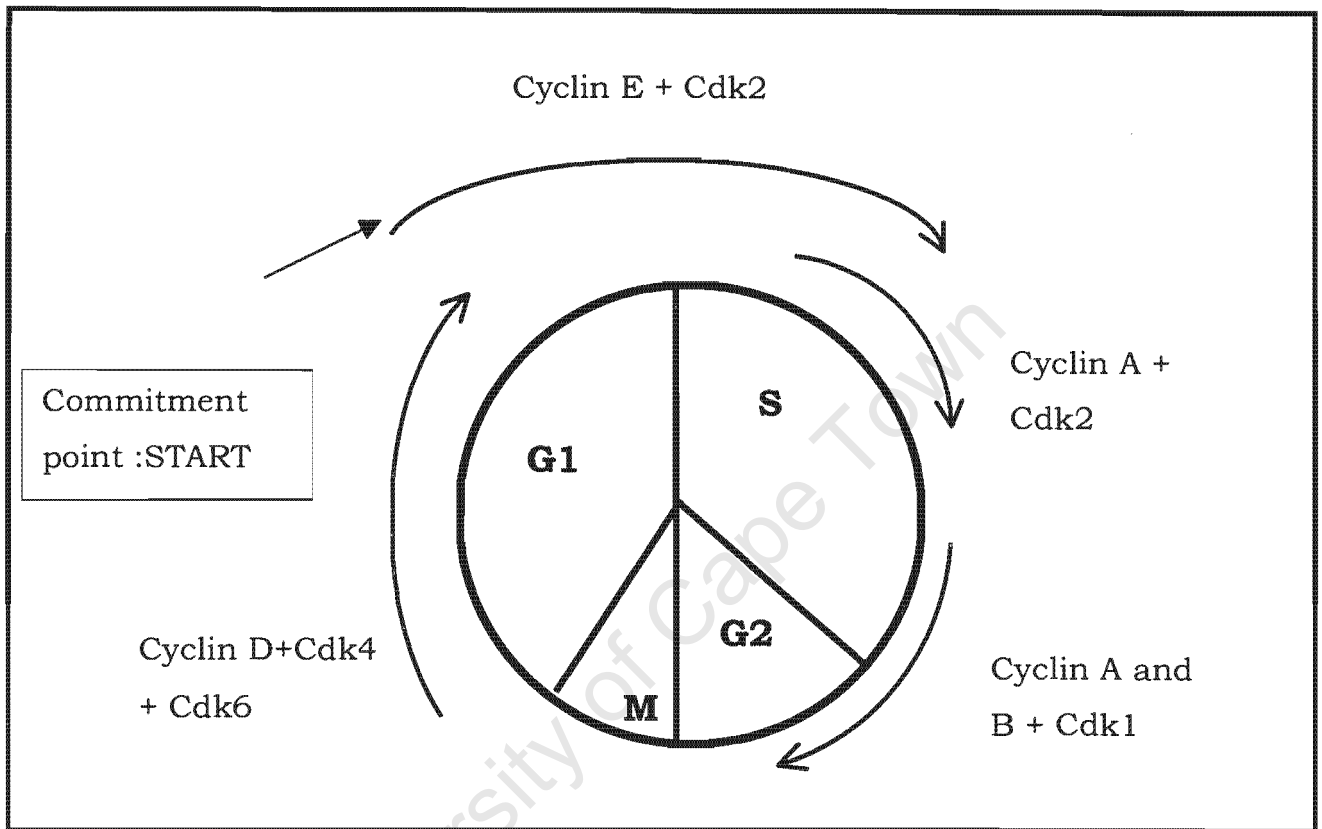


Figure 3.1. Interactions of cyclins and cyclin-dependent kinases (CDK) during the cell cycle.
(Betticher, 1996)

3.1.4.1. Mutations in *CCND1*

The gene coding Cyclin D1 viz. *CCND1* is located on chromosome 11q13 and has 5 exons and 4 introns spanning approximately 15 kb of genomic DNA (Donnellan and Chetty, 1998).

Cyclin D1 mRNA exhibits alternative splicing and the translation of the different transcripts (transcripts a and b) result in protein

products with different COOH-terminal domains, with possible functional differences (Betticher *et al.*, 1995; Sawa *et al.*, 1998). Transcript (a) is normally spliced while transcript (b) is the variant transcript, which reads into intron 4 and skips intron 5. In addition, a G to A polymorphism at the splice donor site of exon 4 at codon 242 has been postulated to increase alternate splicing (Betticher *et al.*, 1995). However, both the (G) and (A) alleles at codon 242 can produce the alternate transcript arising from the altered splicing. The (AG) and (AA) genotypes are associated with transcript (b).

The normal and the altered transcripts encode a protein that contains the amino acids (55-161) thought to be responsible for cyclin D1 function (the cyclin D1 box), but the protein coded by the alternate transcript (b) may have a longer half-life (Kong *et al.*, 2000). The COOH-terminal of the *CCND1* protein (a) has a PEST region (the destruction box) responsible for a rapid turnover (Rogers *et al.*, 1986). The cyclin D1 variant has no PEST sequence, which could account for the longer half-life for transcript (b).

The aforementioned observations suggest that *CCND1* may act as a cancer modifier (Bala and Peltomaki, 2001). Patients with resected small cell lung carcinoma and the (AG) or (AA) genotype showed a shorter relapse-free survival compared to those with the (GG) genotype. The (A) allele is associated with a poor clinical outcome in non-small cell lung carcinoma (Betticher *et al.*, 1995; Matthias *et al.*, 1998). Kong *et al.* (2000) reported an earlier age of onset by 10 years in HNPCC mutation-positive patients who were homozygous or heterozygous for the mutant (A) allele as compared to patients who were homozygous for the normal (G).

3.1.5. The Methylenetetrahydrofolate Reductase gene (*MTHFR*)

Clinical and epidemiological evidence suggests that folate deficiency may play a role in the predisposition to colorectal carcinoma (Shannon *et al.*, 2002). Findings indicate that the development of colorectal adenomas and carcinomas is associated with diets high in alcohol but low in folate and methionine (Benito *et al.*, 1991; Giovannucci *et al.*, 1995). Figure. 3.2 illustrates the biological process involved in folate metabolism.

Methionine is a precursor for S-adenosylmethionine, the methyl donor for most biological transmethylation reactions in the body, including that of DNA (Selhub and Miller, 1992). The enzyme Methylenetetrahydrofolate Reductase (*MTHFR*) regulates the metabolism of folate and indirectly, methionine, which are both important factors in DNA methylation and synthesis (Chen *et al.*, 1996). *MTHFR* irreversibly converts 5,10-methylenetetrahydrofolate (5,10-methyleneTHF) to 5-methylenetetrahydrofolate (5-methyleneTHF). The latter provides the methyl group for de novo methionine synthesis and DNA methylation. A decreased 5-methyleneTHF pool may affect DNA methylation, resulting in imbalanced DNA methylation and methylation of usually unmethylated CpG islands, thereby contributing to carcinogenesis. Imbalanced DNA methylation is observed consistently in colonic neoplasias (Laird and Jaenisch, 1994). Alternately, 5,10-methyleneTHF is required for conversion of deoxyuridylate (dUMP) to thymidylate (dUTP), and depletion of this form of folate may interfere with the thymidylate biosynthesis and result in deoxynucleotide pool imbalances (Blount and Ames, 1994). Folate in the form of methyltetrahydrofolate and vitamin B12 methylase is involved in these pathways. *MTHFR* is thus involved in regulation of the

homocysteine levels and maintaining an adequate methionine pool. (Frosst *et al.*, 1995).

University of Cape Town

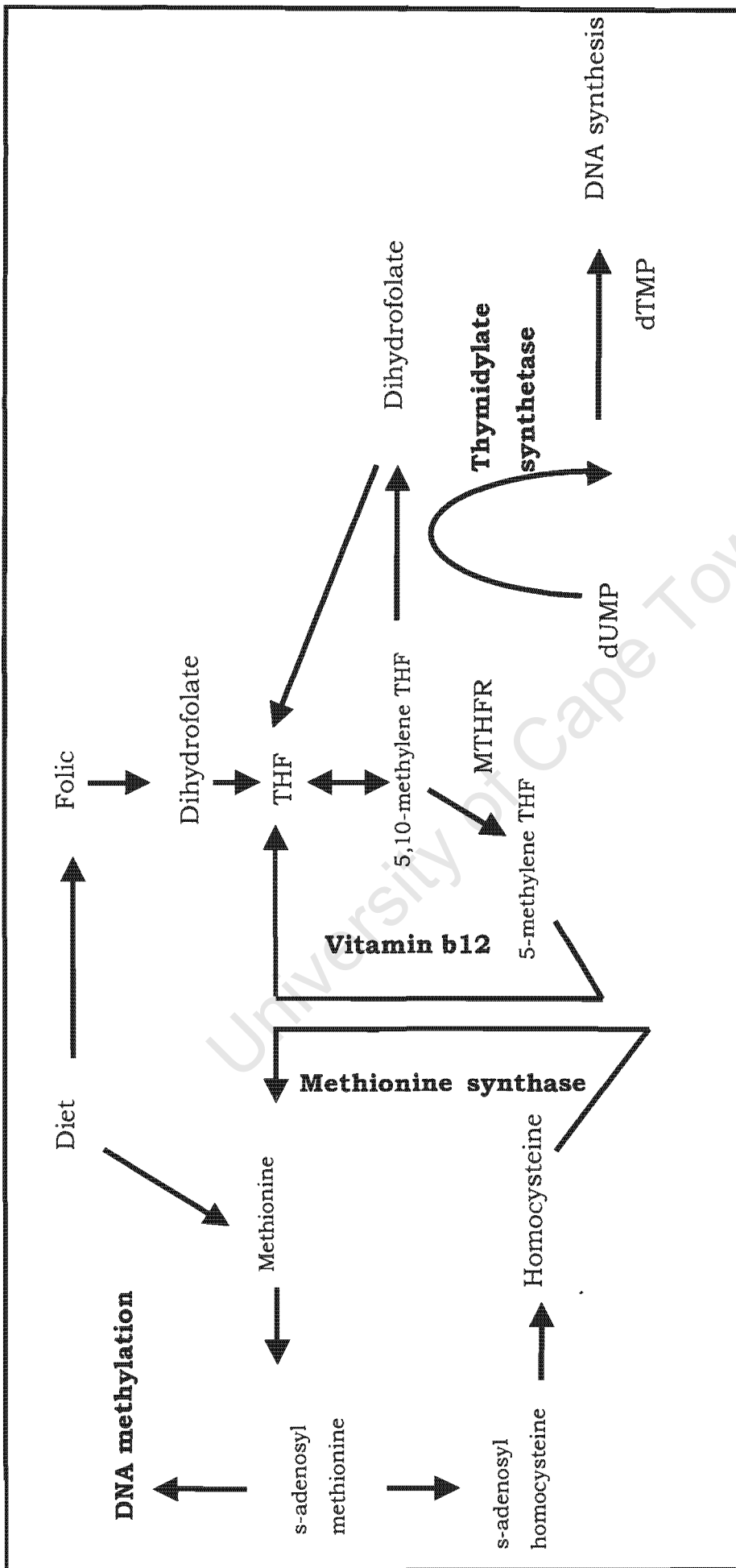


Figure 3.2. The metabolic role of MTHFR in folate metabolism.

MTHFR: methylenetetrahydrofolate reductase; (dUMP) deoxyuridylate ;(dUTP) thymidylate.
(Adapted from Chen *et al.*, 1996; Ma *et al.*, 1997)

3.1.5.1. Mutations in *MTHFR*

MTHFR is located on chromosome 1p36.3. A C->T substitution at nucleotide 667 (C677T), converting an alanine to valine has been identified (Frosst *et al.*, 1995) and is associated with decreased enzyme activity, which in turn leads to a reduction in folate levels, and increased circulating homocysteine levels (Frosst *et al.*, 1995). Individuals homozygous for the TT genotype have approximately 30% normal enzyme activity and this variant was shown to be associated with lower levels of circulating folate (5-methyleneTHF), leading to an accumulation of 5,10-methyleneTHF. The relationship between the *MTHFR* polymorphism and the development of CRC is not clear. A number of case-control studies have been carried out effectively showing a significant reduction in CRC risk in those individuals harbouring the TT variant (Chen *et al.*, 1996; Ma *et al.*, 1997; Slattery *et al.*, 1999). This protective effect appeared to be dependent on an adequate intake of folate in the diet. Studies have also suggested that the TT genotype was associated with an increased risk of CRC in older men with tumours in the proximal colon as well as in men who consume high amounts of alcohol and have poor dietary folate intake (Chen *et al.*, 1996; Ma *et al.*, 1997).

Chen *et al.* (1996) suggested that the relationship of the *MTHFR* genotype with CRC might depend on dietary methyl content. Individuals harbouring the mutated TT genotype may be at a reduced risk of developing CRC due to an imbalance in the nucleotide pool during DNA synthesis when the methyl supply was low. When there is a shortage of methyl groups, abnormal methylation may become the primary mechanism of colorectal tumourigenesis. Multiple pathways are involved in the relation between methyl group metabolism and the development of CRC, which depends on the dietary supply of methyl as well as the activity of *MTHFR*.

3.1.6. The Glutathione S-transferases (GSTs)

Cells contain a variety of enzyme systems that protect them from injury by toxic chemicals and carcinogens. One such system is the family of glutathione S-transferases (GSTs). GSTs play a role in the detoxification of a wide variety of potentially toxic carcinogenic electrophiles by using reduced glutathione to conjugate or to reduce these compounds. This usually inactivates the electrophiles and facilitates their excretion into urine or bile (Walen and Boyer, 1998). Most GST substrates are xenobiotics or products of oxidative stress, including some environmental carcinogens (Hayes and Pulford, 1995). In addition to their catalytic activity, GSTs engage in the intracellular transport of endogenous metabolites and steroid hormones (Gsur *et al.*, 2001).

Of the range of potential carcinogenic metabolites, heterocyclic amines are implicated in causing CRC in humans and are detoxified by GSTs (Lin *et al.*, 1994). In addition, foods capable of inducing the expression of GSTs are thought to be protective against CRC (Boone *et al.*, 1990). Confirmation of a role of GSTs in CRC will aid in the identification of individuals at risk for the disease. GST expression could then be manipulated to prevent CRC either by altering the diet or by pharmacological means (Welfare *et al.*, 1999).

There are studies that did not find a significant association between *GSTT1* and *GSTM1* null genotypes and CRC when comparing patients and controls (Chevenix-Trench *et al.*, 1995; Katoh *et al.*, 1996; Heinimann *et al.*, 1999). Other studies, however, found a significant excess of the *GSTM1* null genotype in patients with CRC when compared to controls (Strange *et al.*, 1991; Zhong *et al.*, 1993). The study undertaken by Deakin *et al.* (1996) and Zhang *et al.* (1999) found a statistically higher frequency of the *GSTT1* null genotype in patients with CRC when compared to controls.

The human cytosolic GSTs comprise six main classes based on sequence homology and substrate specificity. These are α (GSTA), μ (GSTM), π (GSTP), θ (GSTT), kappa (GSTK) and microsomal GST (Zhang *et al.*, 1999). Functional polymorphisms have been identified for 3 classes, *viz.* GSTP, GSTM and GSTT. This section will focus on *GSTM1* and *GSTT1* due to their association with colorectal cancer.

3.1.6.1. Mutations in *GSTM1* and *GSTT1*

Polymorphisms in the glutathione-S-transferases and its role in the susceptibility to CRC have been investigated to determine any protective function that may exist. The variant allele for both *GSTM1* and *GSTT1* is a deletion of the gene. Those individuals who are homozygous for this deleted allele are said to possess the 'null' genotype (0/0) and therefore are unable to produce the enzyme.

There are 3 alleles at the *GSTM1* locus on chromosome 1p13.3 *viz.* *GSTM1* null (a deletion), *GSTM1a* and *GSTM1b* (Rebbeck, 1997). *GSTM1a* and *GSTM1b* differ by a substitution in one base pair with no evidence of a functional difference (Rebbeck, 1997).

Various studies have examined the role of GSTs in CRC. The study by Zhong *et al.* (1993) found an association with the null genotype in *GSTM1* and CRC, with a relative risk of 1.8. Gene frequencies for the *GSTM1* null genotype range from 23-100%, varying by ethnic group (Cotton *et al.*, 2000). *GSTM1* has primarily been studied in bladder, breast, lung and colon cancers (Seidergard *et al.*, 1986; Zhong *et al.*, 1993).

The *GSTT1* gene is on chromosome 22q11.2. The study by Deakin *et al.* (1996) found a positive association between the *GSTT1* null genotype and CRC, with an odds ratio of 1.88, Katoh *et al.* (1996) found an odds ratio of 1.2 and Butler *et al.* (1997) reported an odds ratio of 3.4. Zhang

et al. (1999) found that the frequency of the *GSTT1* null genotype was increased in CRC patients compared to healthy controls. Zhang *et al.* (1999) suggested that the absence of *GSTT1* might be used as a susceptibility marker for CRC. Chenevix-Trench *et al.* (1995) and Katoh *et al.* (1996) did not find any significant changes in the frequencies of *GSTT1* or *GSTM1* in patients with CRC when compared with control samples.

3.1.7. N-acetyltransferases 2 (NAT2)

Germ-line mutations of proto-oncogenes and tumour suppressors account for less than 15% of CRC (Cannon-Albright *et al.*, 1988). The great majority of cancers are due to non-inherited factors. Diet is estimated to account for up to 90% of the international variation in rates of CRC (Doll and Peto, 1981; Smith *et al.*, 1995). Altered enzyme activity in xenobiotic metabolising enzymes such as *N*-acetyltransferase (NATs) and the glutathione S-transferases (GSTs) affects cancer susceptibility and could partially explain variation in disease expression in conditions such as HNPCC (Heinimann *et al.*, 1999). Three distinct NAT genes have been localised to chromosome 8p22 and consist of two active genes (*NAT1* and *NAT2*) and an inactive pseudogene (*NATP1*).

NAT2, is found in a variety of tissues including the colonic mucosa (Ilett *et al.*, 1994; Windmill *et al.*, 2000). It catalyses the metabolism of xenobiotics and carcinogens by transferring an acetyl group to these agents. Phenotypic variation in the rate at which acetylation occurs is partially attributable to the polymorphic nature of the *NAT2* gene.

3.1.7.1. Mutations in NAT2

Mutations in *NAT2* correlate with the rate at which drugs and chemicals that contain primary aromatic amine or hydrazine groups are metabolised. At least seven different alleles, resulting from single base

changes, are known to cause decreased expression of the enzyme. The *NAT2* gene is responsible for the acetylation polymorphism, which determines if individuals are rapid, intermediate or slow acetylators. Possession of two mutant alleles results in the slow acetylator phenotype and inheritance of at least one wild-type allele results in a fast acetylator (Vatsis *et al.*, 1991). The latter is seen in approximately 10 to 90% of individuals, depending on their ethnic origin. The various mutant alleles for *NAT 2* are shown in Table 3.2.

Table 3.2. Mutations associated with the various *NAT2* genotypes.

ALLELES	MUTATIONS
NAT2*4	None
NAT2*5A	C481T
NAT2*5B	C481T; A803G
NAT2*5C	A803G
NAT2*6	G590A
NAT2*7	G857A

Taken from Smith et al., 1997

Studies investigating the association between the *NAT2* genotype and CRC risk have been conflicting. Lang *et al.* (1986) and Ilett *et al.* (1987) reported a positive association between the fast acetylator phenotype and risk of CRC, whereas Kirlin *et al.* (1991) did not. Heinimann *et al.* (1999) suggested an increased risk of developing CRC among slow acetylators with HNPCC. The various conclusions regarding the association of slow acetylation phenotypes in different populations might be due to different genetic and environmental influences. Individuals with HNPCC may be more sensitive to certain environmental influences than those with no MMR defects (Frazier *et al.*, 2001). Frazier *et al.* (2001) observed that the *NAT2**7 mutant allele increased the risk of colorectal and all other cancers in HNPCC carriers. They also found that the *NAT2**5 mutant allele increased the risk for CRC in heterozygotes, but the level of significance was not as high as that for *NAT2**7. The detrimental effects of environmental factors might be

enhanced by the *NAT2*7* mutant allele in HNPCC while having little or no effect in other individuals.

This phase of the study involved the screen of specific variants of candidate modifier genes to determine if they had any effect on the age at diagnosis of the disease in the cohort under investigation.

University of Cape Town

3.2. MATERIALS AND METHODS

From a comprehensive screen of the biological materials archive with PCR primers for the exons of *hMLH1*, *hMSH2* and *hMLH6*, several disease-causing mutations were detected. The largest cohort of individuals carrying a single disease-causing mutation was used as the “research cohort” for this phase of the study aimed at determining the effect of potential modifier genetic factors on the disease phenotype.

3.2.1. Subjects

Each subject harbours the C1528T mutation in exon 13 of *hMLH1*, which causes truncation of the MLH1 protein product at this site. The study cohort will be referred to as the “C1528T-research cohort”. This cohort (n=86) was analysed according to affection status, age at diagnosis (with relevant medians, means and ranges) and gender.

Twenty ml of venous blood samples was collected from each of 86 subjects under sterile conditions in tubes containing the anti-coagulant EDTA. DNA was prepared using the *Genomix* blood scale-up kit [Talent, Italy] (Appendix, section B.1.).

3.2.2. Candidate modifier genes

This process was facilitated through the following two methods:

- i. the PCR amplification of target regions of the various genes, followed by gel electrophoresis or
- ii. alkaline-mediated differential interaction (AMDI) analysis - a fluorescence technique for the detection of SNPs that have been amplified by using the amplification refractory mutation system (ARMS)-PCR procedure (Bartlett *et al.*, 2001)

3.2.2.1. Amplification Refractory Mutation System (ARMS) PCR Analysis

An amplification refractory mutation system (ARMS) PCR assay was employed to investigate whether the mutations in the modifier genes described play a role in the age of expression of the disease in the C1528T-research cohort.

The C1528T-research cohort was screened by using an ARMS PCR assay for the variants described in *CDH1*, *ATM*, *APC*, *CCND1* and *MTHFR*. Oligonucleotide primers were designed using the OLIGO software (version 4.1, *Medprobe*, Norway). Two forward primers were used for each PCR, one specific for the wild type allele and the second specific for the mutated allele (except for *ATM*-see Table D.11.). These primers were used in separate PCR reactions, each utilising the same common reverse primer. Where possible, the assay was validated by analysis of control samples known to be heterozygous for the mutant allele. AMDI fluorescent detection was carried out using a BMG Fluorostar Galaxy (*BMG Labtechnologies*, UK) according to the method described by Bartlett *et al.* (2001) (Appendix, section B.8.).

The sequences for the ARMS primers used to amplify the DNA sequence variants in the various genes mentioned previously are displayed in Appendix D (Tables D.9. to D.13.)

Conditions for PCR and the AMDI fluorescent detection are outlined in Appendix (section B.8.).

3.2.3. Polymerase chain reaction

The C1528T-research cohort was amplified by PCR and fragments were electrophoresed on agarose gels to determine if any of these individuals

carried the null genotype for *GSTM1* and *GSTT1* and if this genotype had any modifying effect on the disease phenotype.

GSTM1 and *GSTT1* genotyping was performed using oligonucleotide sequences as described by Bell *et al.* (1993) and Pemble *et al.* (1994) (Appendix D, Table D.14.). The presence or absence of the *GSTT1* and *GSTM1* alleles was determined by co-amplifying with β -globin gene primers (Bell *et al.* 1993).

Conditions for the PCR amplification of the *GSTM1* and *GSTT1* fragments are outlined in Appendix (section B.9.).

3.2.4. Restriction fragment length polymorphism (RFLP)

The C1528T-research cohort was investigated for acetylation status by PCR amplification and RFLP analysis of the *NAT2* gene.

3.2.4.1. PCR for *NAT2* genotyping

NAT2 genotyping was performed according to the method described by Smith *et al.* (1997). A 547 bp fragment was generated using NAT2F and NAT2R primers (Appendix D, Table D.15.).

Conditions for the PCR amplification of the *NAT2* fragment is outlined in Appendix (section B.10.).

The 547bp fragment was digested with four restriction endonucleases *KpnI*, *DdeI*, *TaqI* and *BamHI* in separate reactions to distinguish the mutations at positions 481, 590, 803 and 857 (Table 3.3) (Appendix, section B.11). Digested PCR products were electrophoresed on 3% agarose gels.

Table 3.3. Patterns obtained for NAT2 alleles after cleavage with KpnI, DdeI, TaqI and BamHI.

Alleles	KpnI	DdeI	TaqI	BamHI	Acetylation status
NAT2*4	+	-	+	+	Fast (no mutation)
NAT2*5A	-	-	+	+	Slow
NAT2*5B	-	+	+	+	Slow
NAT2*5C	+	+	+	+	Slow
NAT2*6	+	-	-	+	Slow
NAT2*7	+	-	+	-	Slow

(Adapted from Smith et al., 1997)

+ Restriction site present
-Restriction site absent

The resulting cleavage pattern allowed for the distinction between the NAT2*4 rapid acetylator phenotype and the NAT2*5, NAT2*6 and NAT2*7 slow acetylators. Individuals who were homo- or heterozygous for the NAT2*4 allele were classified as rapid acetylators.

The C1528T-research cohort was subdivided into two groups i.e. the affected and the unaffected individuals. This was carried out to determine whether any correlation existed between the genotype that an individual inherited in the above-mentioned genes, and the age at diagnosis of disease.

3.2.5. DNA sequencing analysis

PCR products displaying the mutations screened for on fluorescent reading or gel electrophoresis were sequenced in both 5' to 3' and 3' to 5' directions on an Applied Biosystems 377 automated sequencer using the BigDye Terminator cycle sequencing kit [PE Biosystems, Germany] according to manufacturer's protocol, described in Appendix (section B.5.).

3.2.6. Statistical methods

Contingency tables were generated to determine the numbers of individuals that were affected and unaffected, for each of the potentially modifying genotypes under investigation. Chi-squared tests were used to determine the frequency of specific alleles in the affected and unaffected individuals who harbour the C1528T mutation.

Kaplan-Meier survival analysis was used to model the age at diagnosis of CRC. Unaffected individuals at the time of the study were censored observations. Age at diagnosis rates rather than the usual survival rates, were calculated. Graphs of the proportion unaffected (by CRC) at each age (in years) are presented for various subsets of the study group. The median age at diagnosis, with 95% confidence intervals, are also presented. Log-rank tests were used to determine whether the observed difference in age at diagnosis rates was statistically significant.

Cox regression was used to calculate an individuals' relative risk (RR, with 95% confidence interval) of becoming affected with CRC with a specific property/genotype compared to individuals with a reference property. Cox regression was also used to determine whether an interaction between different combinations of genotypes affect the age at diagnosis of CRC. Only the significant combination is discussed.

The analyses were done with Microsoft Excel XP, Stata and NCSS 2001 available from <http://www.ncss.com>.

3.3. RESULTS

3.3.1. Subjects' characteristics

Of the 86 subjects with the C1528T mutation in *hMLH1*, 45 were female and 41 were male (Table 3.4). The gender distribution for diseased state was notable *viz.* 9 of the 45 females (20%) and 20 of the 41 males (49%) were diagnosed with CRC. The mean age at diagnosis was 38.7 years.

Table 3.4. Characteristics of subjects in C1528T research cohort.

	Affected subjects			Unaffected subjects			Total		
	Male	Female	Both	Male	Female	Both	Male	Female	Both
Number	20	9	29	21	36	57	41	45	86
%	23%	10%	34%	24%	42%	66%	48%	52%	100%
<u>Present age</u> <u>(years):</u>									
Mean	49.6	50.8	50.0	35.6	35.4	35.5	42.1	38.5	40.2
Range	35-68	28-75	28-75	21-63	14-82	14-82	21-68	14-82	14-82
<u>Age at</u> <u>diagnosis:</u>									
Mean	37.1	42.2	38.7						
Range	15-63	21-63	15-63						

The age at diagnosis of disease for the C1528T-research cohort is represented in Fig. 3.3. The curve shows the proportion of individuals not yet affected by CRC (on the y-axis) at each age (on the x-axis). For example, the median age at diagnosis of the disorder in the research cohort was 49 years (95%CI=44-54) i.e. at 49 years of age, half the individuals were affected with CRC and the other half were not.

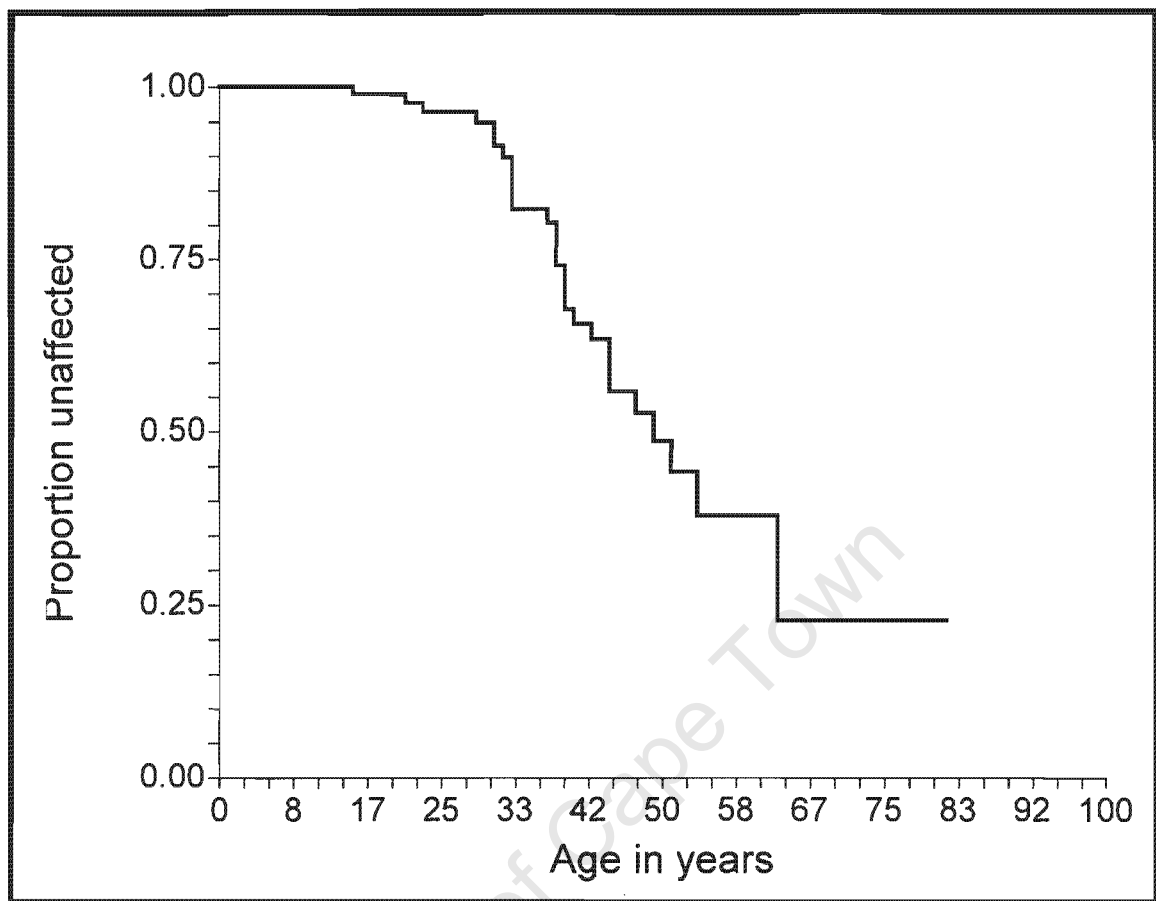


Figure 3.3. Proportion of individuals in the C1528T research cohort unaffected at each age in years.

The C1528T-research cohort was further stratified by gender (Fig. 3.4). There was a significant difference between the two curves ($p=0.0101$) i.e. females have a much later age at diagnosis than males. The median age at diagnosis of CRC for females was 63 years (95%CI=49-63), and 39 years (95%CI=38-44) for males. Thus 50% of females and males were symptomatic by the age of 63 and 39 years respectively. The RR for males (vs. females) developing disease is 2.93 (95%CI=1.24-6.03).

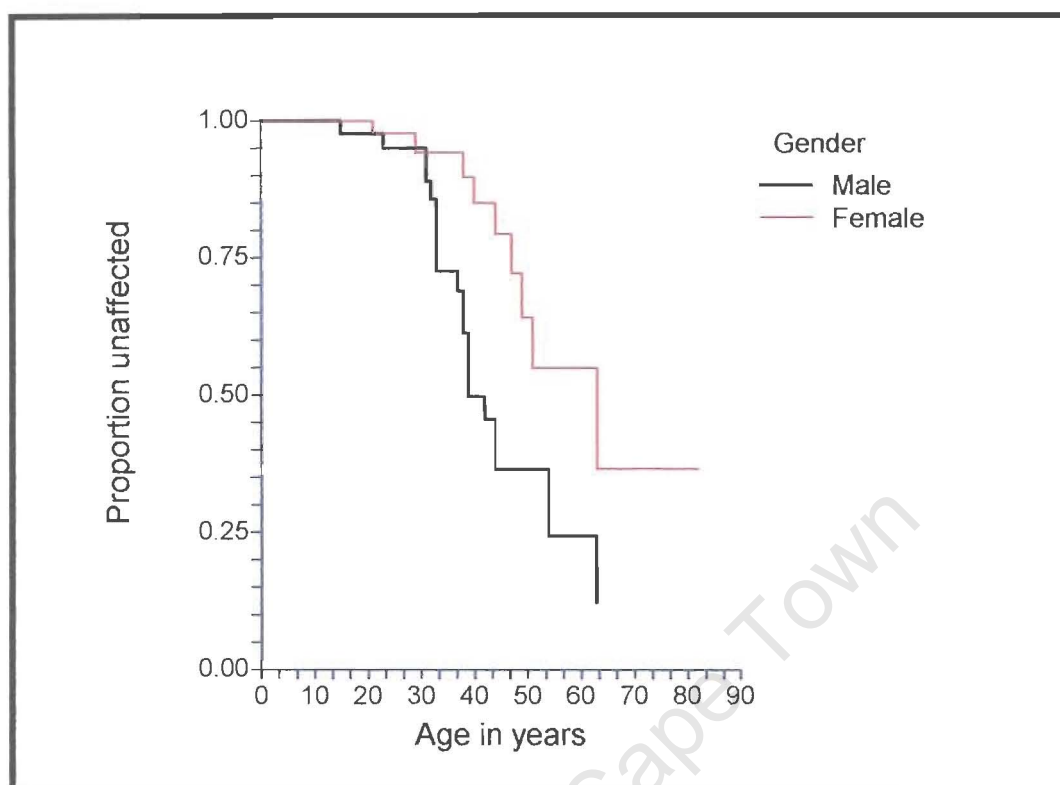


Figure 3.4. *Proportion of individuals in the C1528T-research cohort who are unaffected at each age (years) stratified by gender.*

3.3.2. ARMS PCR Assay analysis

The C1528T-research cohort (n=86) was genotyped for specific variants in *CDH1*, *ATM*, *APC*, *CCND1*, *MTHFR*, *GSTM1*, *GSTT1* and *NAT2*. The phenotype relating to age at diagnosis of cancer was examined as a function of the variations in the genotype at each of the loci. The cohort was stratified on the basis of gender for the analysis of each variant. Only the significant result for gender stratification is presented.

3.3.2.1. *E-cadherin*(*CDH1*)

Neither the subjects nor the 100 unrelated unaffected control individuals displayed any of the eight mutations in the *CDH1* gene listed in Table 3.1.

3.3.2.2. *Adenomatous polyposis coli* gene *APC*

The ARMS PCR test for the presence of the I1307K and the E1317Q variant in *APC* showed that neither of these mutations was present in the research or control cohorts.

3.3.2.3. *Ataxia telangiectasia* gene (*ATM*)

Homozygous (GG) and heterozygous (GA) individuals for the D1853D variant in *ATM* were found in both groups (Table 3.5). Neither the affected nor the unaffected individuals were homozygous for the AA genotype.

Table 3.5. *ATM* genotype frequencies in the C1528T-research cohort.

Disease status	AG	GG	Total
Unaffected	1	56	57
Affected	3	26	29
Total	4	82	86

The age at diagnosis of CRC for the different *ATM* genotypes (AG and GG) is graphically represented in Fig. 3.5. The survival rates of patients with the AG genotype was different to those who harboured the GG genotype at any specific age ($p=0.0559$). The median age at diagnosis was 33 years (95%CI=31-33) for the AG genotype and 51 years (95%CI=44-54) for the GG genotype, an 18-year difference in the median age at diagnosis. The RR of developing CRC for the AG genotype (vs. GG) was 3.2 (95%CI=0.95-10.82).

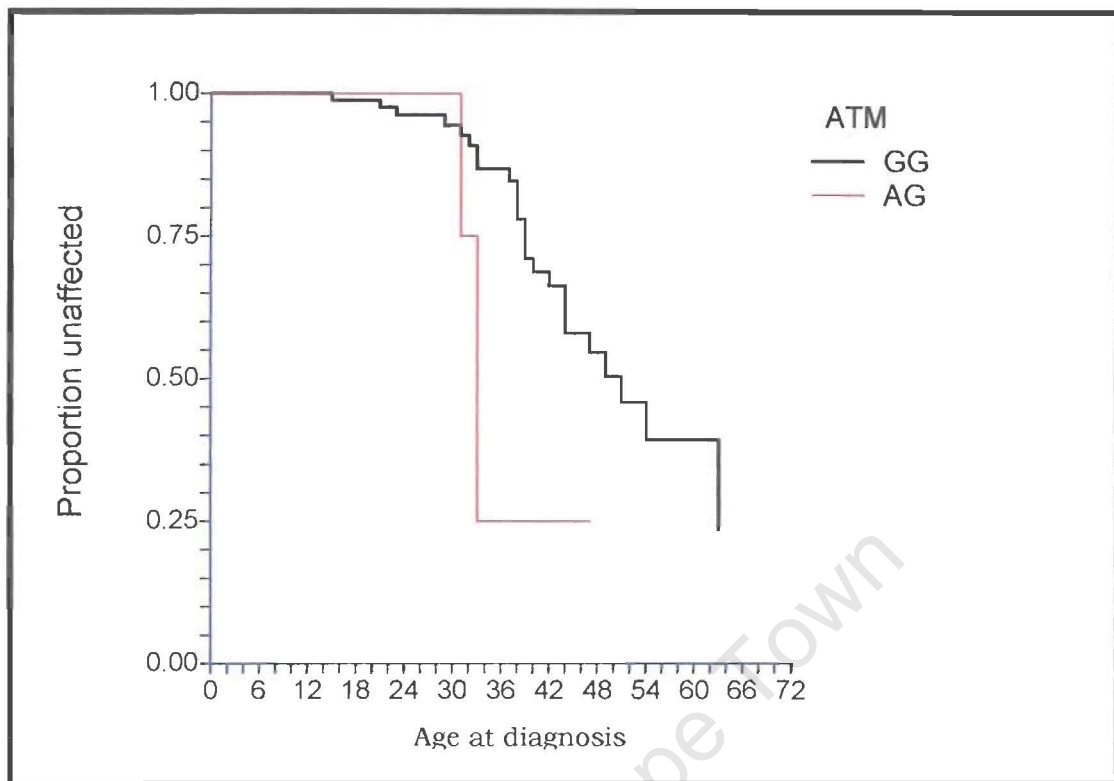


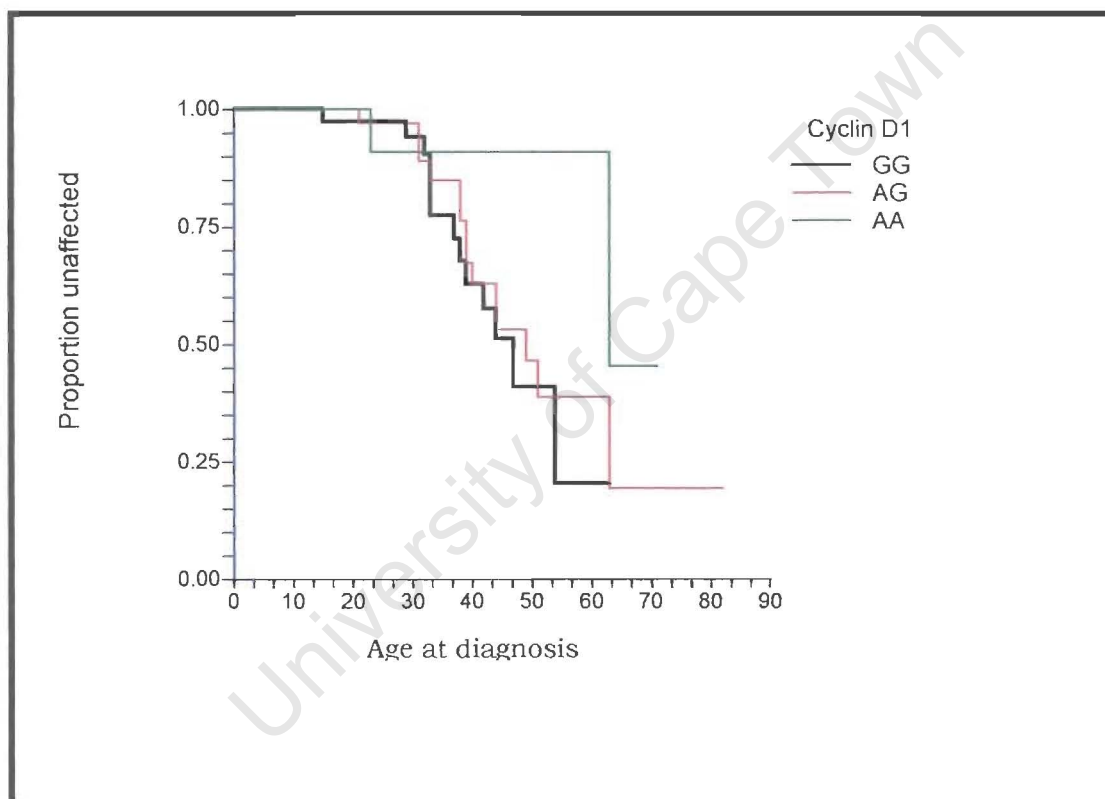
Figure 3.5. Proportion of unaffected individuals in the C1528-research cohort at each age, stratified by ATM genotype.

3.3.2.4. Cyclin D1

The age at diagnosis of CRC for the *CCND1* genotypes, GG, AG and AA is graphically represented in Fig. 3.6. The number of affected individuals who harboured the AA genotype was very small ($n=2$) (Table 3.6). There was no difference between the curves for the AG and GG genotypes ($p=0.4325$). Neither the RR nor the difference in median age of diagnoses were significant.

Table 3.6. CCND1 genotype frequencies in the C1528T-research cohort.

Disease status	AA	AG	GG	Total
Unaffected	10	21	26	57
Affected	2	14	13	29
Total	12	35	39	86

**Figure 3.6. Proportion of unaffected individuals in the C1528T-research cohort at each age, stratified by CCND1 genotype.**

3.3.2.5. MTHFR

The age at diagnosis of CRC for the *MTHFR* genotypes (CC and CT) is represented in Fig. 3.7. An unaffected individual who displayed the TT genotype was excluded from the age at diagnosis analyses (Table 3.7).

Table 3.7. *MTHFR* genotype frequencies in the C1528T-research cohort.

Disease status	CC	CT	TT	TOTAL
Unaffected	43	14	0	57
Affected	23	5	1	29
Total	66	19	1	86

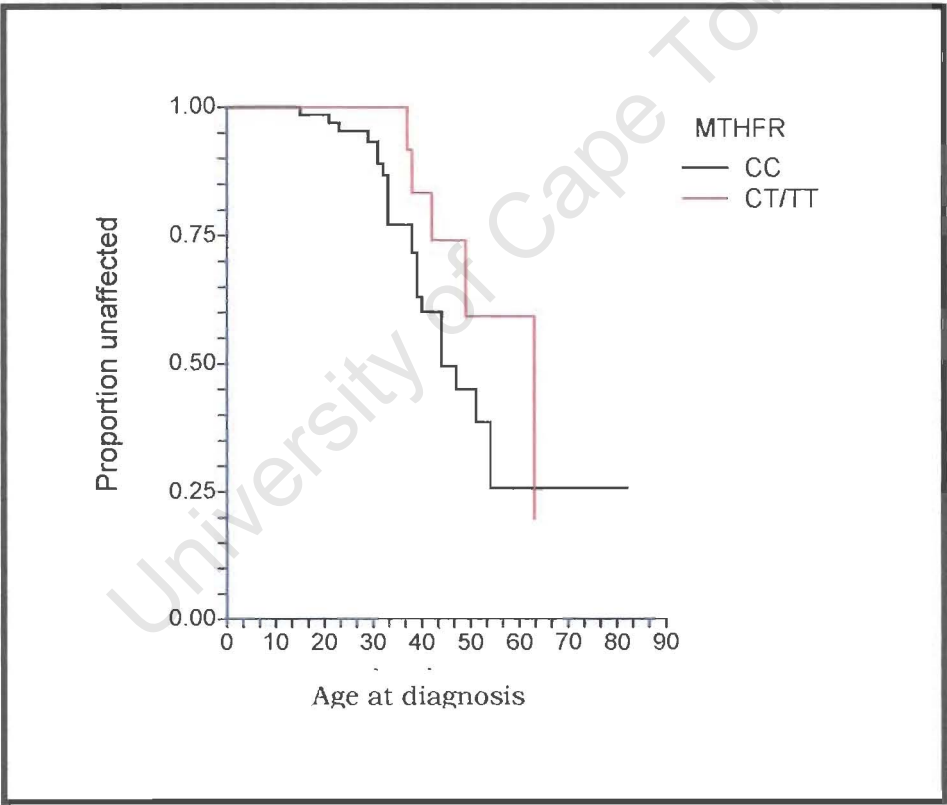


Figure 3.7. *Proportion of the C1528T-research cohort unaffected at each age, stratified by MTHFR genotype.*

The proportion of unaffected subjects at each age was always more for the CT genotype than for the CC genotype, but the difference was not

statistically significant ($p=0.12$). Interestingly, the median age at diagnosis for the CT genotype was 19 years older than that for the CC genotype (63 vs. 44 years), but this difference was not statistically significant.

The joint effect of gender and the CC or CT genotype on age at diagnosis of CRC was subsequently investigated with significant results ($p=0.0014$). The males who harbour the CC genotype had the highest proportion of affected individuals at every age, with a median age of 39 years (95%CI=33-39). The age at diagnosis was identical for the females with the CC genotype and males with CT genotype, each with a median of 63 years. The females who carry the CT genotype exhibited the oldest age at diagnosis i.e. 56 years. Table 3.8 shows the average age at diagnosis for the affected individuals and the average age of those who are unaffected at present.

The gender-adjusted RR was 3.14 (95%CI=1.11-8.88) i.e. when the known effect of gender was adjusted for, an individual with the CC genotype was 3.14 times more likely to develop cancer at any age than those with the CT genotype.

Table 3.8. Gender-stratified age at diagnosis for males and females who harbour the MTHFR CC and CT genotypes.

	Female		Male	
	CT	CC	CT	CC
MTHFR				
Average age at diagnosis (years)	56.0	38.6	47.3	35.1
Average age of unaffected individuals (years)	30.0	36.8	41.3	32.8

3.3.3. Polymerase chain reaction

3.3.3.1. *GSTT1*

The age at diagnosis of CRC for the null and positive genotype for *GSTT1* is shown in Fig. 3.8. A higher proportion of the unaffected individuals in the C1528T-research cohort were null for *GSTT1* (Table 3.9) when compared to the affected individuals. There was no significant difference between the curves ($p=0.5890$). This was further confirmed by the fact that the median age at diagnosis was 51 years for the null allele and 49 years for individuals positive for *GSTT1*.

Table 3.9. *GSTT1* genotype frequencies in the C1528T-research cohort.

Disease status	Null	Positive	TOTAL
Unaffected	41	16	57
Affected	21	8	29
Total	62	24	86

Positive i.e. harbours the *GSTM1* allele
Null i.e. the *GSTM1* allele is deleted

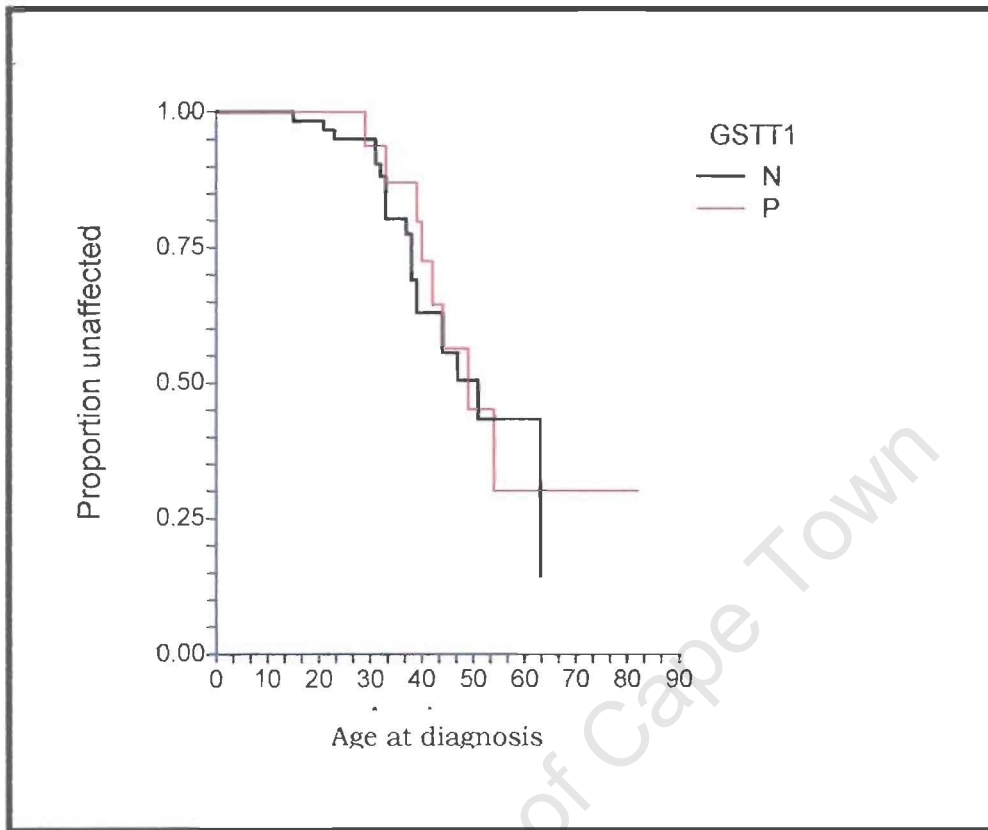


Figure 3.8. *Proportion of unaffected individuals in the C1528T-research cohort at each age, stratified by GSTT1 genotype.*

3.3.3.2. *GSTM1*

A significant difference was observed between the curves for the null and positive genotypes of *GSTM1* ($p=0.0018$) as shown in Fig. 3.9. Fifty percent of the individuals who were null for *GSTM1* were affected by 33 years of age, whereas those who carried the *GSTM1* alleles were affected at a much later age (median age of 51 years). The RR for individuals who are null for *GSTM1* (vs. harbouring the *GSTM1* allele) was 4.1 (95%CI=1.59-10.56). This indicated that individuals who were null for *GSTM1* were 4.1 times more likely to develop cancer at any age than those individuals who harboured the *GSTM1* allele (i.e. *GSTM1* positive). Six of the affected and four of the unaffected mutation-positive individuals harboured the null genotype for *GSTM1* (Table 3.10).

Table 3.10. *GSTM1* genotype frequencies in the C1528T-research cohort.

Disease status	Null	Positive	TOTAL
Unaffected	4	52	56
Affected	6	23	29
Total	10	75	85

Positive i.e. harbours the *GSTM1* allele
Null i.e. the *GSTM1* allele is deleted

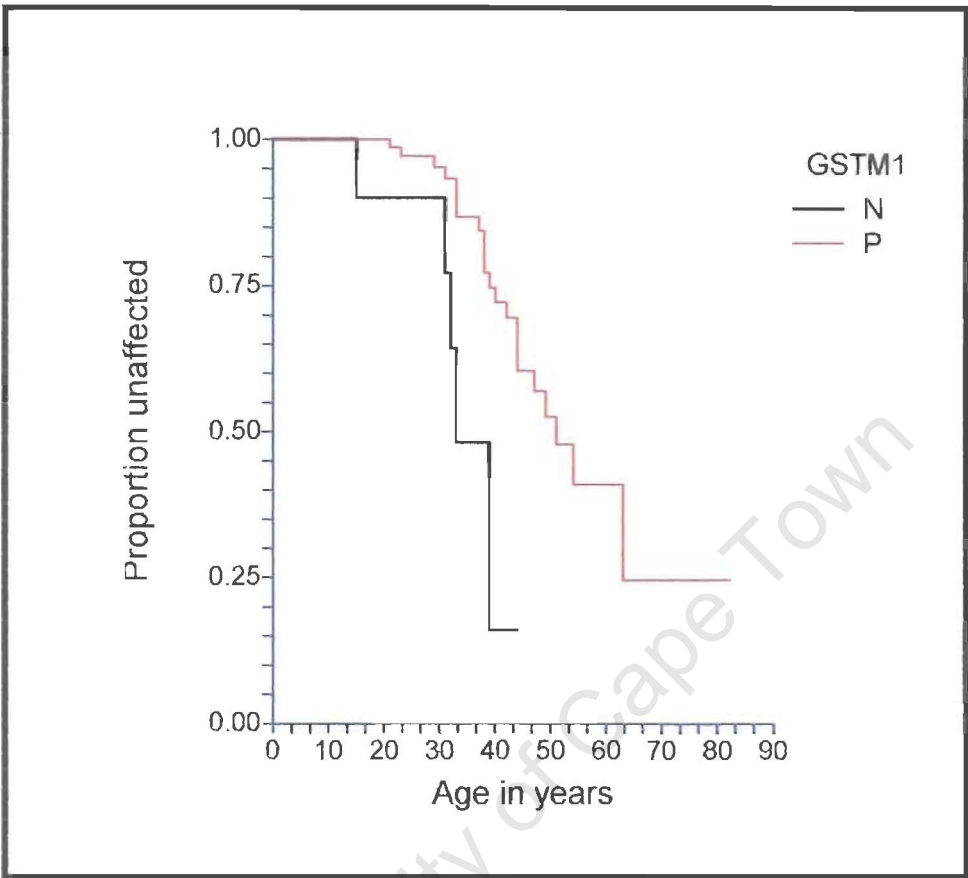


Figure 3.9. Proportion of unaffected individuals in C1528T-research cohort at each age, stratified by GSTM1 genotype.

The C1528T-research cohort was stratified by gender (Table 3.11) to determine if a difference in age at diagnosis occurred for the *GSTM1* genotypes. When the known effect of gender was adjusted for, the RR was reduced to 3.71 (95%CI=1.42-9.7). The affected females who were positive for *GSTM1* had an average age at diagnosis of 43.4 years. There were no females who were null for *GSTM1* in the C1528T-research cohort (Table 3.11). Interestingly, the males who were null for *GSTM1* were symptomatic 7.9 years before those who were positive for *GSTM1* (Table 3.11).

Table 3.11. *GSTM1* genotypes for males and females who harbour the C1528T mutation.

GSTM1	Female		Male	
	Positive	Null	Positive	Null
Average age at diagnosis (years)	42.4	-	39.4	31.5
Average age of the unaffected individuals	35.6	33.3	35.5	26.0

P: positive i.e. harbours the *GSTM1* alleleN: null i.e. the *GSTM1* allele is deleted**3.3.4. RFLP analysis**

Seventy-one of the 86 individuals in the C1528T research cohort were suitable for analysis as 15 individuals had ambiguous genotypes for *NAT2* and were excluded from this phase of the study. All of the 71 mutation-positive individuals harboured the *NAT2**4 genotype, which conferred the rapid acetylation status. In total, 15 of the 71 affected individuals were homozygous for the *NAT2**4 allele, whereas the rest were heterozygous for this wild-type allele (Table 3.12).

Table 3.12. *NAT2* genotype frequencies in the C1528T-research cohort.

Disease status	4/4	4/5	4/6	4/7	TOTAL
Unaffected	14	12	15	3	44
Affected	15	8	4	0	27
Total	29	20	19	3	71

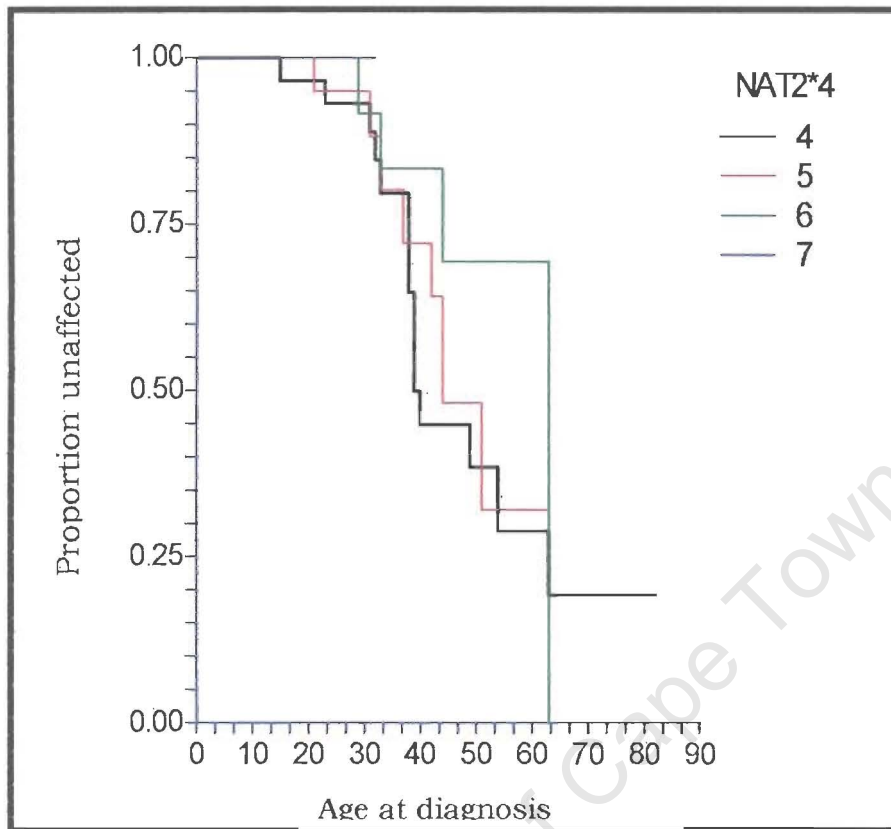


Figure 3.10. Proportion of unaffected individuals in the C1528T-research cohort at each age, stratified by NAT2.

The age at diagnosis of CRC for the various alleles of the *NAT2* gene (4/4, 4/5, 4/6 and 4/7) is graphically represented in Fig. 3.10. No significant difference were observed between the curves for the various *NAT2* genotypes ($p=0.36$). The median age at diagnosis was 39 years for the 4/4 genotype, 44 years for 4/5 genotype and 63 years for the 4/6 genotype. The median age at diagnosis could not be calculated for the 4/7 genotype, because none of the three individuals who carry this genotype were affected at the time of analysis. Neither the RR nor the difference in median age at diagnosis were significant ($p=0.36$).

Combined genotypes:

MTHFR and *GSTM1* were investigated in an attempt to explain why certain individuals who harboured the C1528T mutation in the *hMLH1*

gene were diagnosed at a young age (e.g. 13 years), while others reach the age of 82 years, without manifesting with CRC. The following codes were used to identify the various groups:

FCCN:	Females with the <i>MTHFR</i> CC genotype, <i>GSTM1</i> null.
FCCP:	Females with the <i>MTHFR</i> CC genotype, <i>GSTM1</i> positive.
FCTN:	Females with the <i>MTHFR</i> CT genotype, <i>GSTM1</i> null.
FCTP:	Females with the <i>MTHFR</i> CT genotype, <i>GSTM1</i> positive.
MCCN:	Males with the <i>MTHFR</i> CC genotype, <i>GSTM1</i> null.
MCCP:	Males with the <i>MTHFR</i> CC genotype, <i>GSTM1</i> positive.
MCT-:	Males with the <i>MTHFR</i> CT genotype, <i>GSTM1</i> unknown.
MCTN:	Males with the <i>MTHFR</i> CT genotype, <i>GSTM1</i> null.
MCTP:	Males with the <i>MTHFR</i> CT genotype, <i>GSTM1</i> positive.
MTP:	Males with the <i>MTHFR</i> TT genotype, <i>GSTM1</i> positive.

The number of individuals in each group, and the age at diagnosis for those who are affected, is shown in Table 3.13. The age at diagnosis for those groups who were affected is presented in Fig. 3.11.

Table 3.13. **Number of individuals and average age at diagnosis in the C1528T mutation-positive subjects for the combined gender, GSTM1 and MTHFR genotypes.**

	AFFECTED			AGE AT DIAGNOSIS OF CRC
	NO	YES	TOTAL	
Genotype:				
FCCN	3	-	3	-
FCCP	26	7	33	38.6
FCTN	-	-	-	-
FCTP	7	2	9	56.0
MCCN	1	6	7	31.5
MCCP	13	10	23	37.2
MCT-	1	-	1	-
MCTP	6	3	9	47.3
MTTP	-	1	1	38.0
Total	57	29	86	38.7

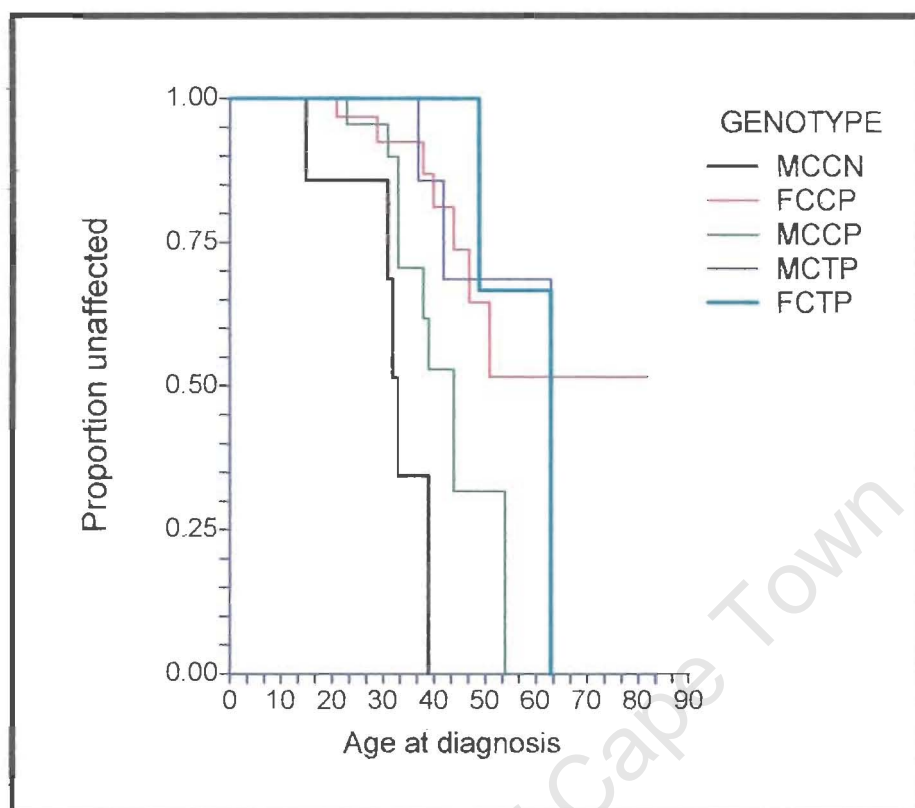


Figure 3.11. Proportion unaffected for each age at diagnosis stratified by combined genotype as defined above.

Significant differences in the age at diagnosis distributions for the various combined genotypes were observed in this cohort ($p < 0.0001$). The genotype associated with the lowest age at diagnosis (39 years) was MCCN (Fig. 3.11). Individuals who carried the FCCP combination of genotypes had the best prognosis i.e. 52% were unaffected by age 51, with the oldest unaffected female at 82 years. The FCCN group had no affected individuals, however the average age was 33.3 years, which was lower than the median age at diagnosis of 49 years for the C1528T-research cohort.

3.4. DISCUSSION AND CONCLUSION

Following the identification of the C1528T mutation in the *hMLH1* gene underlying the predisposition to CRC, it was recognised that initial symptoms presented at a wide range of ages. The present study represents an investigation of several specific polymorphisms within previously reported candidate genes. The modifying effects of the polymorphisms were assessed on a background of subjects who already had a germline disease-causing genetic change in the *hMLH1* gene. This work was an attempt to determine whether these potential modifiers had any effect on the observed phenotype (such as the age at diagnosis).

Since the median age at diagnosis for female and male subjects was 63 and 39 years respectively, it would be sensible to reconsider the screening interval particularly for female subjects, since females became affected at a much older age than males. However, with the age at diagnosis noted to be less than 30 years (youngest 23 years) in some female subjects, this should be done with caution. Further research needs to be conducted on factors that determine the age at diagnosis of disease. Any attempt then at defining subgroups of individuals whose clinical screening may only be warranted at a later age (such as the females who have a median age at diagnosis of 63 years and a mean age at 50.8) is appealing especially in an environment with limited means. Furthermore, males who harbour the C1528T mutation in the *hMLH1* gene are three times more likely to become affected with CRC than female subjects.

At the moment, a considerable effort is made to provide colonoscopic screening by undertaking trips to the hospitals closest to those identified at highest risk in the Northern Cape. This trip is carefully

coordinated between the genetics and surgical team who are responsible for transporting the scopes, light sources, nurses and assistants to help with the endoscopic procedures. Currently, those individuals carrying a disease-causing mutation are advised to undergo endoscopic annual screening after the age of 30 years (once every two years between the ages of 18 and 30 years). Repeated clinical screening is very important in order to identify the disease as early as possible.

3.4.1. *E-cadherin (CDH1)*

Since somatic mutations in the *E-cadherin* gene have been reported in colorectal carcinomas resulting in reduced expression of the CDH1 protein (Kim *et al.*, 1999), the research cohort (belonging to families diagnosed with HNPCC) was screened for the presence of these eight mutations. None of the individuals within the C1528T-research cohort displayed any of the eight germ-line mutations listed in Table 3.1. The affected individuals in this cohort have CRC without any evidence of gastric cancer.

A number of familial inherited cancer syndromes are characterised by gastric cancer. The cumulative risk for developing gastric cancer in HNPCC families who harbour mutations in the MMR genes is 19% (Richards *et al.*, 1999). Germline mutations in *CDH1* were identified in three kindreds from Maori origin who were diagnosed with gastric cancer (Guilford *et al.*, 1998). Subsequently, a study undertaken by Richards *et al.* (1999) identified mutations in eight UK families diagnosed with gastric cancer. Both families presented with gastric cancer, however, one also had an early-onset CRC, suggesting that *CDH1* mutations are associated with extra-gastric tumours. Furthermore, another study reported *CDH1* mutations in subjects diagnosed with diffuse gastric or CRC originating from Korea (Kim *et al.*, 2000).

The current study was by no means an exhaustive search for germline mutations in *CDH1*. As more mutations in *CDH1* are detected, it is becoming increasingly evident that inactivation of the *E-cadherin* complex by various mechanisms plays a significant role in colorectal carcinogenesis.

3.4.2. *Adenomatous polyposis coli* gene (*APC*)

The C1528T-research cohort in the current study was screened for the presence of the I1307K and E1317Q missense variants to determine if they exist and whether they have a modifying effect on the disease phenotype. The absence of the I1307K and E1317Q variants in the current study supports other reports, which have not detected these variants in familial CRC. These variants may not be present in the germline, but may influence the expressivity of disease. Frayling *et al.* (1998) suggest that the existence of these missense mutations could be indicative of additional DNA alterations at sub-polymorphic frequencies that have a significant effect on the incidence of colorectal adenomas and carcinomas. In addition, it is likely that these polymorphisms/variants could only be associated with specific populations and could be absent from the Mixed Ancestry population of South Africa under investigation. It may be possible that an unknown *APC* variant contributes to the manifestation of the disease phenotype in this group of individuals. If so, it would be important to screen the *APC* gene for any possible missense variants that individuals from the C1528T-research cohort might harbour.

3.4.3. *Ataxia telangiectasia* gene (*ATM*)

The C1528T-research cohort was genotyped for the *D1853N* variant in *ATM*. This study has shown that carriers of both the C1528T germline mutation and the *ATM D1853N* variant (i.e. the AG genotype) have a

three times increased risk of developing CRC at a specific age when compared to carriers of the GG genotype. The median age at diagnosis of subjects who harboured the mutated *D1853N* variant was 18 years before those who harboured the wildtype GG allele. Because there were only four individuals with the AG genotype in the cohort, their contribution should be interpreted with caution. It is possible that the sample was too small to show a truly significant difference between the age at diagnosis rates for the genotypes. In addition, it may be possible that the observed difference was an artifact and that no true modifying effect exists.

Mailett *et al.* (2000) found a significant association ($p=0.02$) between the *ATM D1853N* variant and incidence of CRC in their cohort of 167 Swiss subjects from 20 HNPCC families. Germline mutations in either *hMLH1* or *hMSH2* were identified in 67 of these subjects. One hundred and twenty subjects were asymptomatic and 47 developed CRC. None of the subjects were homozygous for the (AA) *ATM D1853N* variant. In addition, the *ATM D1853N* variant was associated with a significantly higher incidence of colonic and extra-colonic malignancies when compared to individuals who were homozygous (GG) for the wildtype ($p=0.02$). Furthermore, patients with a germ-line mutation in either *hMLH1* or *hMSH2*, who also carried the *ATM D1853N* variant, were found to have an eight-times increased risk of developing colorectal or any other HNPCC-associated cancer. The CRC incidence in the Swiss cohort for carriers of both a MMR gene mutation and the *ATM* variant may have reached 100% since the only asymptomatic individual was a 22 year old, who was 20 years younger than the mean age at diagnosis in their cohort. The *ATM D1853N* variant was not associated with a younger age at diagnosis in the swiss study.

Mutations in the *ATM* gene predispose to ataxia-telangiectasia, an autosomal recessive condition. The *ATM* protein has a role in

maintaining the integrity of the DNA. Heterozygous carriers, however, do not exhibit the typical phenotype of the disease and are at a higher risk of developing epithelial malignancies (breast and digestive tract cancers) (Swift *et al.*, 1991). The *D1853N* polymorphism is not a neutral polymorphism and the functional impact of this variant needs further investigation. The inheritance of a disease-predisposing mutation in the *hMLH1* gene together with the *ATM* variant may predispose these individuals to the development of CRC at an earlier age. The epithelial tissue in the colon has a high turnover and is constantly exposed to carcinogens; mutations in genes involved in mismatch repair could understandably predispose to earlier disease as observed in subjects from the C1528T-research cohort.

The data suggested that there was weak association between the *ATM* D1853N polymorphism and the age at diagnosis of cancer in this cohort. This data will eventually be taken into account in devising a more informed management of individuals who carry the mutation in *hMLH1*.

3.4.4. Cyclin D1 (CCND1)

No correlation was observed between the age at diagnosis and the G242A polymorphism in the research cohort under investigation ($p=0.4325$). The median age at diagnosis for individuals who harbour the GG and AG genotypes was similar (47 and 49 years), but much older for the homozygous AA genotype i.e. 63 years. The AA genotype could be protective, as only two of the twelve individuals who harbour it are clinically affected.

The findings of the current study contradict those of previous researchers. This phase of the study differed from those published previously in that all 86 individuals in the C1528T-research cohort

harboured a single disease-causing mutation in *hMLH1*, on which the modifying effect of the variants is being investigated. Kong *et al.* (2000) studied the effect of the *Cyclin D1* polymorphism in 86 subjects from the United States, 57 of who harboured a mutation in the *hMSH2* gene, 28 in the *hMLH1* gene and one in the *hMPS1* gene. Subjects with the variant AA and AG genotype developed CRC 10 years earlier than those with the GG genotype. In addition, the subjects with the AA or AG genotype were 2.46 times more likely to develop CRC at any age than those individuals with the GG genotype.

Bala and Peltomaki (2001) studied 146 individuals from 53 Finnish HNPCC families who harboured mutations in both the *hMLH1* and *hMSH2* genes. The mutations in *hMLH1* included splice-site mutations and large genomic deletions. The age at diagnosis in individuals who harboured the GG and AA genotypes was five to six years lower than in the heterozygote AG carriers. Their suggestions for the difference included:

- i. The predisposing mutations in the MMR genes could “mask the modifying effects of *CCND1* by being strong independent determinants of clinical outcome” and
- ii. Alternately, the MMR genes also participate in the cell cycle control, so they could possibly be interacting with *CCND1*.

The subjects investigated by Kong *et al.* (2000) were mostly carriers of an *hMSH2* mutation, whereas the investigation by Bala and Peltomaki (2001) included predominantly *hMLH1* mutation-carriers. Furthermore, 47% of the subjects studied by Kong *et al.* (2000) harboured missense mutations, whereas the subjects studied by Bala and Peltomaki (2001) harboured mostly truncating mutations. Ethnic or environmental influences could also have a confounding effect on the different results that were observed by these two groups.

3.4.5. Methylenetetrahydrofolate Reductase (MTHFR)

One unaffected individual out of the total cohort of 86 was homozygous for the variant allele (TT) and was ignored for the Kaplan-Meier analyses. No significant difference was observed between the mutated and wildtype genotypes when analysing the cohort as a whole. The median age at which the heterozygote (CT) individuals (male and female) were diagnosed with HNPCC was 24 years later than for the homozygous (CC) males. The median age at diagnosis for females with the CC wildtype alleles could not be calculated, because more than half of them were unaffected at the time of this study. After adjusting for the effect of gender, an individual with the CC genotype was 2.67 times more likely to be affected at a particular age than one who harboured the CT or TT genotype. *MTHFR* was thus an important modifier in the C1528T research cohort.

Data on the dietary folate intake and alcohol consumption amongst subjects was not collected for this study. It would be interesting to investigate alcohol consumption, in both males and females, to determine whether any difference in intake over a specific period does occur. Although field experience over ten years suggests that alcohol consumption was higher in males than females, it is speculative at this stage and has yet to be verified in the cohort under investigation. The next phase of this study will focus on dietary factors and environmental influences for the individuals who harbour the C1528T mutation.

3.4.6. Glutathione-S-transferases (GSTs)

The current study investigated whether the *GSTM1* null genotype played a modifying role in the age at diagnosis of CRC. The C1528T-mutation carriers who were null for *GSTM1* (i.e. the *GSTM1* alleles were deleted) were four times more likely to develop CRC at a specific age than those individuals who harboured the *GSTM1* alleles. On average, the

individuals who were *GSTM1* null had an age at diagnosis of 33 years, whereas the individuals who harboured the *GSTM1* allele were diagnosed at 51 years, an 18-year time difference. *GSTM1* was therefore an important modifier of the age at diagnosis in the C1528T-research cohort under investigation.

An insignificant two-year difference in median age at diagnosis was observed between individuals who were positive and null for *GSTM1* in the research cohort ($p=0.5890$).

Information on diet, environment or lifestyle was not available for the cohort under investigation. Dietary factors play a crucial role in colorectal carcinogenesis as food-derived carcinogens (such as heterocyclic amines and nitrosamines) have been etiologically linked to CRC.

The results of this phase of the study have indicated that the next phase should include an investigation on diet in the C1528T-research cohort. It would then be possible to determine any interaction between diet, lifestyle, the C1528T mutation and *GSTM1* status for this cohort.

3.4.7. *N*-acetyltransferase 2 (*NAT2*)

The current study aimed to determine whether the reported polymorphisms in the *NAT2* gene influenced the age at diagnosis of CRC among the C1528T-mutation carriers.

All the C1528T mutation-positive individuals carried at least one wildtype allele (*NAT2**4), which conferred the rapid acetylation status. Fifteen of the 29 (51%) affected individuals were homozygous for the *NAT2**4 allele, whereas the rest of the subjects were heterozygous for this wild-type allele. Statistical analysis showed no significant

difference between the affected and unaffected groups who harboured the various *NAT2* genotypes ($p = 0.36$). The median age at diagnosis for the 4/4 genotype was 39 years, 44 years for individuals who harboured the 4/5 genotype, 63 years for the 4/6 genotype and unknown for the 4/7 genotype. Thus the observed median age at diagnosis for *NAT2* 4/6 and 4/7 was 24 years later than for the wildtype 4/4. The *NAT2**7 allele was very rare in the population.

In the present study it was not possible to compare fast acetylators to slow acetylators as all the individuals in the current study were classified as fast acetylators, i.e. all carried the *NAT2**4 allele. It was thus decided to test for any significant difference between those individuals who were homozygous for this allele i.e. those who were 4/4 to those who were heterozygous for this allele, i.e. 4/5, 4/6 or 4/7, however a significant difference was not observed.

Studies on the association of acetylation status with risk of CRC have been conflicting. Certain studies have suggested that slow acetylators have a lower risk of developing CRC (Lang *et al.*, 1986; Ilett *et al.*, 1987), whereas others have suggested an increased risk of developing CRC (Heinimann *et al.*, 1999). It is possible that subjects with HNPCC and with defects in the MMR genes may be more sensitive to certain environmental influences, than those without MMR defects (Frazier *et al.*, 2001).

Heinimann *et al.* (1999) investigated 181 individuals from 21 unrelated Swiss HNPCC families diagnosed with cancer of the colon as well as extracolonic cancers. These kindreds all harboured a mutation in either *hMLH1* or *hMSH2*. When the affected and unaffected mutation-carriers were compared, a significant association between the *NAT2**4 allele and individuals with overall cancer was observed ($p < 0.03$). The differences continued to be statistically significant when subjects were

divided into mutation-carriers who have CRC ($p < 0.04$) or extracolonic cancer ($p < 0.05$). Frazier *et al.* (2001) studied 86 individuals with germline mutations in either *hMLH1* or *hMSH2*. Forty-nine subjects were affected with CRC and extracolonic cancers were present. They did not observe any correlation between acetylation status and individuals who harbour mutations in the MMR genes. However the *NAT2*7* allele increased the risk for colorectal and all other cancers in mutation-carriers. They also observed that the *NAT2*5* allele increased the risk for CRC in heterozygous carriers.

The effect of the combined inheritance of the *GSTM1* null genotype and the *MTHFR* variant was further investigated with gender stratification in this study. A significant difference in age at diagnosis was obtained for the grouping of the variants in these two genes ($p < 0.001$). Half of the affected males who carried both the *GSTM1*-null and *MTHFR*-CC (MCCN) genotypes were affected 11 years before the *GSTM1*-positive, *MTHFR*-CC (MCCP) group. The female subjects who were *GSTM1*-positive and *MTHFR*-CC (FCCP) had 52% unaffected by age 51 years.

The combined effect of the inheritance of the C1528T mutation in *hMLH1* with the *GSTM1* and *MTHFR* variants may have on the HNPCC phenotype in this cohort should be taken into account when subjects are undergoing genetic counseling and management for this disorder. Inheritance of a specific combination of variants in these proposed modifier genes would allow for more accurate risk assessment and age at diagnosis calculations in the cohort under investigation.

Information on diet and smoking status was not available from this group of individuals. Previous studies have linked acetylation status and risk of developing CRC with nutrition and lifestyle habits, such as smoking and alcohol consumption. Since the current study was part of a longitudinal investigation and follow-up, it would be valuable to

construct a thorough examination of diet, nutrition and lifestyle to further analyse the influence of environmental and epigenetic factors. Studies such as the present one form the basis on which to devise additional investigations, which will ultimately lead to improved clinical, dietary and ultimately preventative management of these and other subjects predisposed to HNPCC-related cancers.

In conclusion, significant correlations were observed between the C677T variant in *MTHFR* and the null allele of *GSTM1*, and age at diagnosis. In addition, a weak association was observed for the *ATM* D1853N variant and age at diagnosis.

DNA from individuals diagnosed with CRC was collected in the Division of Human Genetics over a period of 10 years with attention to clinical detail and follow-up information, with no additional data regarding lifestyle. However, studies such as the current one investigating the effect of possible modifiers are beneficial to motivate for financial support enabling further investigations amongst this group of individuals with the C1528T mutation. Our observations are based on a population that reside mostly along the West Coast of South Africa to the Namibian border. It is a considerable effort to contact the individuals in this research cohort as many live in fairly remote areas. Heavy-terrain vehicles facilitate travelling to and around some of these regions. Studies such as the present one, conducted with limited field resources, provide a basis for extended studies on diet and lifestyle, especially in individuals who are known to harbour an inherited disease-causing mutation.

The data emerging from the current modifiers in this study must be considered as a pilot, highlighting the value of the C1528T cohort as a “reagent” for testing of a range of further potential genetic and environmental modifiers. This information will contribute towards

individualised clinical surveillance and ultimately personalised therapeutics. Although this is only considered to be preliminary data, the results are exciting and provide strong motivation for expanding this investigation. The expansion may be through (i) increasing the cohort size of C1528T carriers (currently underway as families are extended and new families are identified), and (ii) similar investigations in other cohorts, internationally.

It is clear that multiple mutations are involved in colorectal carcinogenesis. Cancer is a multifactorial disorder, and understanding the factors that determine the transition from adenoma to carcinoma are important in the prevention of CRC in families. If factors such as age at diagnosis, site of disease, the presence of extra-colonic cancers and disease severity can be predicted, it would allow for targeted surveillance of this disorder in subjects which could eventually lead to the reduced prevalence of CRC and morbidity and mortality due to HNPCC.

CHAPTER 4: CONCLUDING REMARKS

University of Cape Town

CHAPTER 4: CONCLUDING REMARKS

4.1. CONCLUDING REMARKS

Amongst other objectives, cancer research is focussed on early detection and effective intervention. The investigation of familial cancers leads to the identification of specific disease-causing mutations in genes. Of immediate benefit is the recognition that blood relatives carry the disease-predisposing factor. Recognition of these high-risk individuals, with adequate clinical surveillance, is a powerful tool in reducing morbidity and mortality due to cancers.

HNPCC is an inherited disorder for which the disease-causing genes have been identified in the majority of families that are clinically diagnosed based on e.g. the Amsterdam or Bethesda criteria. Genetic testing is then available for presymptomatic diagnostic surveillance. With the evolution of molecular genetic technology, it is now possible to screen the major genes, which contribute to the HNPCC phenotype, even in the absence of family history.

The present study was initiated on biological material from subjects who were diagnosed with CRC at a relatively young age, where the diagnosis of FAP was ruled out. The candidate genes chosen for mutation analysis included the MMR genes (*hMLH1*, *hMSH2* and *hMSH6*), *p53* and *APC*. A range of disease-associated mutations, including 13 in the MMR genes, two in *p53* and four in the mutation cluster region of *APC* was identified in the primary research cohort of 100 subjects diagnosed with CRC. The research subjects were further stratified on the basis of the age at diagnosis of disease and the site of the cancer in the colon; a significant proportion (41%) of individuals who harboured mutations in the MMR genes manifested with cancer of the ascending (or proximal) colon. In contrast, those affected

individuals in whom mutations were not identified in the MMR genes presented mostly with rectal cancer (43%). These findings are worth considering in devising a more efficient genetic screening programme i.e. subjects who present with cancer of the proximal colon should initially be screened for mutations in the MMR genes (Vasen *et al.*, 1996). Interestingly, mutations in the *hMLH1* gene were mostly associated with a reduced frequency of extracolonic tumours (Type I HNPCC), whereas mutations in *hMSH2* were associated with Type II. No correlation could be made between specific mutations and the age at diagnosis of disease in this research cohort.

The results obtained in the current study are in alignment with other reports, which demonstrate that mutations in the MMR genes are predominantly associated with cancer of the proximal colon, with an average age at diagnosis of 45 years (Bisgaard *et al.*, 2002). In addition, individuals affected with rectal cancer do not usually have mutations in the MMR genes (Hoogerbrugge *et al.*, 2003).

The frequency of HNPCC is estimated to be between 2-7% of all cases of CRC (Muller and Fishel, 2002). However, in the present study, the frequency of HNPCC, as defined by a germline mutation in a MMR gene was 15% in the subjects screened. It is worth noting, however, that the individuals from the primary research cohort were diagnosed at a relatively young age, mostly with cancer of the proximal colon, which could enrich for HNPCC. On the other hand, the frequency of 15% could be an underestimate of the true frequency in this cohort due to the sensitivity of one of the mutation detection methods employed *viz.* SSCP/HA, which was used to screen the *hMLH1* and *hMSH2* genes. Thus the true frequency of HNPCC is yet to be established with more sensitive techniques including dHPLC analysis and DNA sequencing. A more accurate estimation of the frequency of HNPCC will emerge from a screen of all CRC (not only those affected at a relatively young age) and

the Type II spectrum of tumours in a comprehensive study involving immunohistochemistry of the products of the MMR genes, MSI analysis of tumour tissue and subsequent screening of the appropriate MMR gene. In addition, genes associated with atypical HNPCC such as *EXO1* and *hMSH3*, known to have a function in the MMR cycle, remain to be screened (Peltomaki, 2003). Furthermore, it has been reported that 30% of colorectal cancers investigated by Bardelli *et al.* (2003) harboured mutations in the tyrosine kinase genes. This group of genes could represent a new group that, when mutated, have a role in the pathogenesis of CRC. Germ line mutations have also been reported in the base-excision-repair gene, *MYH*, associated with recessive inheritance of multiple colorectal adenomas (Sieber *et al.*, 2003). Four percent of families investigated by Sampson *et al.* (2003) had bi-allelic mutations of the *MYH* gene.

In the current study, two novel missense mutations were identified in the *p53* gene and four mutations were identified in the *APC* gene—including one that was novel. It is difficult to ascertain if these missense mutations are disease-causing since a family history was not evident, despite interrogation of the index cases. Functional tests need to be conducted to determine the pathogenicity of the missense mutations identified in *p53*, as well as those detected in the *hMLH1* and *hMSH2* genes.

A C1528T mutation in exon 13 of *hMLH1*, initially detected in an individual from the primary research cohort, was identification of 12 additional subjects. Interestingly, all 13 subjects were from families of the Mixed Ancestry population group and originated from the same geographical area in South Africa. The common haplotype observed around the *hMLH1* gene in all 13 families suggested a common founder. Individuals who harbour this mutation are currently enrolled in an

extensive genetic testing and management program in the Northern Cape, Western Cape, Gauteng and Free State Provinces of South Africa.

Two founder mutations have been identified in HNPCC kindreds from Finland, accounting for approximately two thirds of HNPCC cases in that country (Nystrom-Lahti *et al.*, 1995). In addition, a founder mutation identified in the *hMSH2* gene in the Ashkenazi Jewish population accounted for one-third of HNPCC in the study cohort (Foulkes *et al.*, 2002). Other researchers have also reported founder mutations in families that fulfil the Amsterdam criteria (Froggatt *et al.*, 1999). In the South African context, founder mutations have been described in the Afrikaner population for disorders such as pseudoxanthoma elasticum (Le Saux *et al.*, 2002), Fanconi anemia (Tipping *et al.*, 2001), Sclerosteosis (Brunkow *et al.*, 2001) and familial hypercholesterolaemia (Kotze *et al.*, 1991). The Mixed Ancestry population, however, has been neglected as a genetic resource because of the perceived out-breeding of a wide range of population groups, which landed in the Western Cape (Nurse and Jenkins, 1975). It has been recommended more recently in the HapMap project, which is now underway to characterise patterns of linkage disequilibrium (Cardon and Abecasis, 2003), that “mixed” populations will be extremely valuable in defining shared fragments of the genome and mapping of the common disorders.

An investigation involving the extended families of each of the 13 C1528T-positive subjects revealed 86 individuals who harboured this mutation. Analysis of the C1528T research cohort demonstrated that males were approximately three times as likely to develop CRC than women (RR=2.93). It was also noted that the 86 individuals did not present with an identical phenotype with respect to age at diagnosis of disease and the site of cancer. These observations led to the investigation of specific DNA sequence variants that are likely to have

an influence on the initiation and progression of CRC. The variants tested for reside in modifier genes, which include *CDH1*, *APC*, *ATM*, *CCND1*, *MTHFR*, *GSTT1*, *GSTM1* and *NAT2*. Significant correlations were obtained between the age at diagnosis of disease in individuals and various genotypes that were investigated. These included inheritance of the AG genotype in the *ATM* gene ($p=0.0559$), the null genotype for *GSTM1* ($p=0.0018$) and the *MTHFR* C677T variant ($p=0.0014$) (upon gender stratification). Recognition of the influence of modifiers of the HNPCC disease phenotype will improve the capability of predicting an individual's risk of developing the disorder as a function of both the age at diagnosis of the disorder and site of the cancer. This prediction may eventually influence more individualised clinical screening regimes for different family members.

Microsatellite instability analysis of tumour tissue led to the identification of the disease-associated mutations in the MMR genes in a proportion of tumour samples. This study has shown that tumour tissue is an important biological resource when attempting to identify mutations in MMR genes causative of disease. The tumours can be screened for MSI status with the presence of MSI indicating the presence of a mutation in the MMR genes.

Patients who are referred to the colorectal clinic at the Surgical Gastroenterology Unit at Groote Schuur Hospital are usually already clinically affected with CRC and until recently, surgical intervention was the only tool used to aid in the management of the disorder. The identification of a genetic mutation underlying the disease in these individuals would have limited benefit to these already affected individuals e.g. suggesting total colectomy, but could be valuable for any unaffected blood relatives. It is these relatives that would have the advantage of obtaining the result of a predictive genetic test for a

specific mutation and undergoing premonitory screening to limit the morbidity and mortality associated with the disorder.

In order for a genetic screening program to be successful, it would require extensive genealogical interrogation. Detailed information regarding family history was not requested from patients before the implementation of the current project. The nurses and clinicians who are the first to establish contact with the patients have realised the importance of obtaining an accurate detailed family history for every clinically diagnosed CRC case since the identification of disease-causing mutations in families diagnosed with CRC. The clinicians have also engaged in the presymptomatic diagnostic program by setting up endoscopic clinics for the individuals who have been identified with mutations in the MMR genes. Genetic testing for high-risk family members is recommended and extensive families are currently being managed, both genetically and clinically. It is, however, important to monitor the medium and long-term effects of genetic testing. The results of the test may cause the individual harm as well as good. The psychological impact of learning that a 50% risk of carrying the disease gene has now become 100% may cause anger and resentment. This test result might also have to be disclosed in job interviews, medical insurance and/or life insurance. Individuals have to be aware of the implications of testing before they agree to request their results.

The benefit to those who test positive is early detection and ultimate prevention of CRC. Those individuals who are negative for the disease-causing mutation can be excluded from regular screening. In addition, individuals who belong to families who present with Type II HNPCC need a more rigorous follow-up to monitor the tissue in which extracolonic cancers manifest. The delivery of the appropriate service for mutation-positive individuals is crucial to save lives, especially in a developing country where resources are limited and cost-effective

management for illnesses such as cancer is important. The identification of mutations long before disease manifests will benefit the mutation-positive individuals who will not suffer from the debilitating effects of the disease and then ultimately die as a result of this disorder.

The current study represents a realistic scenario where biological material and family history has not been optimally gathered, but disease-causing mutations were identified in a significant proportion of subjects. In addition, when index cases with MMR gene defects were interrogated, it was clear that the seemingly sporadic disease had a strong familial basis. There has been an evolution, however, in how HNPCC has been diagnosed from the Amsterdam and Bethesda criteria to the identification of mutations in the MMR genes. The clinical definition of the syndrome, on the basis of certain criteria (e.g. the Amsterdam criteria), together with an identified germline mutation in one of the MMR genes has been the most effective means of diagnosing the disorder. The current study led to the identification of germline mutations in MMR genes, thus allowing more accurate predictions about the type of CRC that manifests in the patients.

This project has demonstrated the following:

Mutations in the MMR genes are mostly associated with cancer of the proximal colon

Screening of the MMR genes in relatively young individuals diagnosed with a nonpolypotic form of CRC is an acceptable means for the identification of germline mutations in these genes.

There are often large families that are linked to a single proband, in whom the disease manifests at a relatively young age at diagnosis

Modifying genes may play an important role in influencing the variability of the disease phenotype

Genotype/phenotype correlations for HNPCC have not been clearly defined as yet but are likely to evolve with the recognition of epigenetic factors as well as environmental contributions.

Researchers are just beginning to understand hereditary CRC. The recent molecular genetic advances will help to elucidate the underlying factors, leading to an improved understanding of the biology of this group of disorders. Efficient translation of the acquired genetic knowledge into clinical practice will drastically reduce the morbidity and mortality of inherited CRC. The current study has demonstrated how a research project has influenced the implementation of a genetic service for CRC. This service, together with a rigorous clinical endoscopic screening procedure in families known to harbour a disease-causing mutation can improve the chances of survival for the mutation-positive individuals. It is hoped that the progress in molecular genetics to date will impact positively on the frequency, mortality and morbidity of this common and often fatal cancer.

REFERENCES

University of Cape Town

REFERENCES

- Aaltonen, L.A., Peltomaki, P., Leach, F.S., Sistonen, P., Pylkkanen, L., Mecklin, J., Jarvinen, H., Powell, S.M., Jen, J., Hamilton, S.R., Petersen, G.M., Kinzler, K.W., Vogelstein, B. and de la Chapelle, A. (1993). Clues to the pathogenesis of familial colorectal cancer. *Science*, **260**, 812-816.
- Aaltonen, L.A., Salovaara, R., Kristo, P., Canzian, F., Hemminki, A., Peltomaki, P., Chadwick, R.B., Kaariainen, H., Eskelinen, M., Jarvinen, H., Mecklin, J.P. and de la Chapelle, A. (1998). Incidence of hereditary nonpolyposis colorectal cancer and the feasibility of molecular screening for the disease. *N Engl J Med.*, **338**, 1481-1487.
- Aarnio, M., Sankila, R., Pukkala, E., Salovaara, R., Aaltonen, L.A., de la Chapelle, A., Peltomaki, P., Mecklin, J.P. and Jarvinen H.J. (1999). Cancer risk in mutation carriers of DNA-mismatch repair genes. *Int J Cancer.*, **81**, 214-218.
- Ahuja, N., Li, Q., Mohan, A.L., Baylin, S.B. and Issa, J.P. (1998). Aging and DNA methylation in colorectal mucosa and cancer. *Cancer Res.*, **58**, 5489-5494.
- Aktan-Collan, K., Mecklin, J-P., de la Chapelle, A., Peltomaki, P., Uutela, A. and Kaariainen, H. (2000). Evaluation of a counseling protocol for predictive genetic testing for hereditary non-polyposis colorectal cancer. *J Med Genet.*, **37**, 108-113.
- Amundsen, S.A., Myers, T.G. and Fornace, A.J. (1998). Roles for p53 in growth arrest and apoptosis: putting on the brakes after genotoxic stress. *Oncogene*, **17**, 3287-3299.
- Anwar, S., Hall, C., White, J., Deakin, M., Farrell, W., and Elder, J.B. (2000). Hereditary non-polyposis colorectal cancer: an updated review. *Eur J Surg Oncol.*, **26**, 635-645.
- Baba, S. (1997). Hereditary Nonpolyposis Colorectal Cancer-An update. *Dis Colon Rectum.*, **40**, S86-S95.
- Baker, S.J., Markowitz, S., Fearon, E.R., Willson, J.K. and Vogelstein, B. (1990). Suppression of human colorectal carcinoma cell growth by wild-type p53. *Science*, **249**, 912-915.
- Bala, S. and Peltomaki, P. (2001). CYCLIN D1 as a genetic modifier in hereditary nonpolyposis colorectal cancer. *Cancer Res.*, **61**, 6042-6045.

- Ban, C. and Yang, W. (1998). Crystal structure and ATPase activity of MutL: implications for DNA repair and mutagenesis. *Cell*, **95**, 541-552.
- Bardelli, A., Parsons, D., Silliman, N., Ptak, J., Szabo, S., Saha, S., Markowitz, S., Willson, J.K., Parmigiani, G., Kinzler, K.W., Vogelstein, B., and Velculescu, V. (2003). Mutational analysis of the Tyrosine Kinome in Colorectal Cancers. *Science*, **300**, 949.
- Bartlett, S., Straub, J., Tonks, S., Wells, R.S., Bodmer, J.G. and Bodmer, W.F. (2001). Alkaline-mediated differential interaction (AMD1): a simple automatable single-nucleotide polymorphism assay. *Proc Natl Acad Sci*, **98**, 2694-2697.
- Baylin, S.B., Herman, J.G., Graff, J.R., Vertino, P.M. and Issa, J.P. (1998). Alterations in DNA methylation: a fundamental aspect of neoplasia. *Adv Cancer Res.*, **72**, 141-196.
- Beck, N.E., Tomlinson, I.P., Homfray, T., Frayling, I., Hodgson, S.V., Harocopos, C. and Bodmer, W.F. (1997). Use of SSCP analysis to identify germline mutations in HNPCC families fulfilling the Amsterdam criteria. *Hum Genet.*, **99**, 219-224.
- Behrens, J., Jerchow, B.A., Wurtele, M., Grimm, J., Asbrand, C., Wirtz, R., Kuhl, M., Wedlich, D. and Birchmeier, W. (1998). Functional interaction of an axin homolog, conductin, with beta-catenin, APC, and GSK3beta. *Science*, **280**, 596-599.
- Bell, D.A., Taylor, J.A., Paulson, D.F., Robertson, C.N., Mohler, J.L. and Lucier, G.W. (1993). Genetic risk and carcinogen exposure: a common inherited defect of the carcinogen-metabolism gene glutathione S-transferase M1 (GSTM1) that increases susceptibility to bladder cancer. *J Natl Cancer Inst.*, **85**, 1159-64.
- Bellacosa, A., Genuardi, M., Anti, M., Viel, A. and Ponz de Leon, M. (1996). Hereditary nonpolyposis colorectal cancer: Review of clinical, molecular genetics, and counseling aspects. *Am J Med Genet.*, **62**, 353-364.
- Benito, E., Stiggelbout, A., Bosch, F.X., Obrador, A., Kaldor, J., Mulet, M. and Munoz, N. (1991). Nutritional factors in colorectal cancer risk: a case-control study in Majorca. *Int J Cancer.*, **49**, 161-167.
- Beroud, C. and Soussi, T. (1996). APC gene: database of germline and somatic mutations in human tumours and cell lines. *Nucleic Acids Res.*, **24**, 121-124.

- Betticher, D.C., Thatcher, N., Alternatt, H.J., Hoban, P., Ryder, W.D.J. and Heighway, J. (1995). Alternative splicing produces a novel Cyclin D1 transcript. *Oncogene*, **11**, 1005-1011.
- Betticher, D.C. (1996). Cyclin D1, another molecule of the year? *Annals of Oncology*, **7**, 223-225.
- Bisgaard, M.L., Jager, A.C., Myrholm, T., Bernstein, I. and Nielsen, F.C. (2002). Hereditary non-polyposis colorectal cancer (HNPCC): phenotype-genotype correlation between patients with and without identified mutation. *Hum Mutat.*, **20**, 20-27.
- Blair, N.P. and Trempe, C.L. (1980). Hypertrophy of the retinal pigment epithelium associated with Gardner's syndrome. *Am J Ophthalmol.* **90**, 661-667.
- Blount, B.C. and Ames, B.N. (1994). Analysis of uracil in DNA by gas chromatography-mass spectrometry. *Anal Biochem.*, **219**, 195-200.
- Bodmer, W.F., Bailey, C.J., Bodmer, J., Bussey, H.J., Ellis, A., Gorman, P., Lucibello, F.C., Murday, V.A., Rider, S.H., Scrambler, P. *et al.* (1987). Localisation of the gene for familial adenomatous polyposis on chromosome 5. *Nature*, **328**, 614-616.
- Bodmer, W. (1999). Familial adenomatous polyposis (FAP) and its gene, APC. *Cytogenet Cell Genet.*, **86**, 99-104.
- Boland, C.R., Thibodeau, S.N. and Hamilton, S.R. (1998). A national cancer institute workshop on microsatellite instability for cancer detection and familial predisposition: Development of international criteria for the determination of microsatellite instability in colorectal cancer. *Cancer Res.*, **58**, 5248-5257.
- Bonnet, G.C. and Grobler, S.P. (1997). Disease profile of colon and rectum carcinoma at Pelonomi Hospital, Bloemfontein. *SAJS.*, **35**, 63-67.
- Boone, C.W., Kelloff, G.J. and Malone, W.E. (1990). Identification of candidate cancer chemopreventive agents and their evaluation in animal models and human clinical trials: a review. *Cancer Res.*, **50**, 2-9.
- Branch, P., Hampson, R., and Karran, P. (1995). DNA mismatch binding defects, DNA damage tolerance, and mutator phenotypes in human colorectal carcinoma cell lines. *Cancer Res.*, **55**, 2304-2309.

- Briggs, M.D. and Chapman, K.L. (2002). Pseudoachondroplasia and multiple epiphyseal dysplasia: mutation review, molecular interactions, and genotype to phenotype correlations. *Hum Mutat.*, **19**, 465-478.
- 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. *et al.* (1994). Mutation in the DNA mismatch repair gene homologue *hMLH1* is associated with hereditary non-polyposis colon cancer. *Nature*, **368**, 258-261.
- Brunkow, M.E., Gardner, J.C., Van Ness, J., Paepers, B.W., Kovacevich, B.R., Prohl, S., Skonier, J.E., Zhao, L., Sabo, P.J., Fu, Y., Alisch, R.S., Gillett, L., Colbert, T., Tacconi, P., Galas, D., Hamersma, H., Beighton, P. and Mulligan, J (2001). Bone dysplasia sclerosteosis results from loss of the *SOST* gene product, a novel cystine knot-containing protein. *Am J Hum Genet.*, **68**, 577-589.
- Burke, W., Petersen, G., Lynch, P., Botkin, J., Daly, M., Garber, J., Kahn, M.J.E., Offit, K., Thomson, E. and Varicchio, C. (1997). Recommendations for Follow-up Care of individuals with an inherited predisposition to cancer. *JAMA*, **277**, 915-919.
- Burn, J., Chapman, P., Delhanty, J., Wood, C., Laloo, F., Cachon-Gonzalez, M.B., Tsioupra, K., Church, W., Rhodes, M. and Gunn A. (1991). The UK Northern region genetic register for familial adenomatous polyposis coli: use of age of onset, congenital hypertrophy of the retinal pigment epithelium, and DNA markers in risk calculations. *J Med Genet.*, **28**, 289-296.
- Butler, W.J., Ryan, P. and Roberts-Thompson, I.C. (1997). Metabolic genotypes and risk for colorectal cancer increased risk in individuals with glutathione transferase theta 1 (*GSTT1*) gene defect. *Gastroenterology*, **112**, A542.
- Cannon-Albright, L.A., Skolnick, M.H., Bishop, D.T., Lee, R.G. and Burt, R.W. (1988). Common inheritance of susceptibility to colonic adenomatous polyps and associated colorectal cancers. *N Engl J Med.*, **319**, 533-537.
- Cardon, L.R. and Abecasis, G.R. (2003). Using haplotype blocks to map human complex trait loci. *Trends Genet.*, **19**, 135-140..
- Carstens, R.P., Fenton, W.A. and Rosenberg, L.R. (1991). Identification of RNA splicing errors resulting in human ornithine transcarbamylase deficiency. *Am J Hum Genet.*, **48**, 1105-1114.
- Chen, J., Giovannucci, E., Kelsey, K., Rimm, E.B., Stampfer, M.J., Colditz, G.A., Spiegelman, D., Willett, W.C. and Hunter, D.J. (1996). A

methylenetetrahydrofolate reductase polymorphism and the risk of colorectal cancer. *Cancer Res.*, **56**, 4862-4.

Chenevix-Trench, G., Young, J., Coggan, M. and Board, P. (1995). Glutathione S-transferase M1 and T1 polymorphisms: susceptibility to colon cancer and age of onset. *Carcinogenesis*, **7**, 1655-1657.

Cho, Y.; Gorina, S.; Jeffrey, P. D. and Pavletich, N. P. (1994). Crystal structure of a p53 tumor suppressor-DNA complex: understanding tumorigenic mutations. *Science*, **265**, 346-355.

Chung, D.C and Rustgi, A.K. (2003). The Hereditary Nonpolyposis Colorectal Cancer Syndrome: Genetics and Clinical Implications. *Ann Intern Med.*, **138**, 560-570.

Clurman, E.B. and Roberts, J.M. (1995). Cell cycle and cancer. *J Natl Cancer Inst.*, **87**, 1499-1501.

Cohen A.M. and Winawer S.J.,(eds) Friedman M.A., Gunderson, L.L. (1995). Cancer of the colon, rectum and anus. McGraw-Hill Inc.©

Cotton, S.C., Sharp, L., Little, J. and Brockton, N. (2000). Glutathione S-transferase Polymorphisms and Colorectal Cancer: A HuGE Review. *Am J Epidemiology.*, **151**, 7-31.

Deakin, M., Elder, J., Hendrickse, C., Peckham, D., Baldwin, D., Pantin, C., Wild, N., Leopard, P., Bell, D.A., Jones, P., Duncan, H., Brannigan, K., Aldersea, J., Fryer, A.A. and Strange, R.C. (1996). Glutathione S-transferase GSTT1 genotypes and susceptibility to cancer: studies of interactions with GSTM1 in lung, oral, gastric and colorectal cancers. *Carcinogenesis*, **17**, 881-884.

Decker RA, Peacock ML. (1998). Occurrence of MEN 2a in familial Hirschsprung's disease: a new indication for genetic testing of the RET proto-oncogene. *J Pediatr Surg.*, **33**, 207-214.

Deltas, C.C. (2001). Mutations of the human polycystic kidney disease 2 (PKD2) gene. *Hum Mutat.* **18**, 13-24.

De Rosa, M., Fasano, C., Panariello, L., Scarano, M.I., Belli, G., Iannelli, A., Ciciliano, F. and Izzo, P. (2000). Evidence for a recessive inheritance of Turcot's syndrome caused by compound heterozygous mutations within the PMS2 gene. *Oncogene*, **19**, 1719-1723.

deWeese, T.L., Shipman, J.M., Larrier, N.A., Buckley, N.M., Kidd, L.R., Groopman, J.D., Cutler, R.G., te Riele, H. and Nelson, W.G. (1998). Mouse embryonic stem cells carrying one or two defective Msh2 alleles

- respond abnormally to oxidative stress inflicted by low-level radiation. *Proc Natl Acad Sci.*, **95**, 11915-11920.
- deWind, N., Dekker, M., van Rossum, A., van der Valk, M. and te Riele, H. (1998). Mouse models for hereditary nonpolyposis colorectal cancer. *Cancer Res.*, **58**, 248-255.
- Dipple, K.M. and McCabe, E.R.B (2000). Modifier genes convert "simple" mendelian disorders to complex traits. *Molecular genetics and metabolism*, **71**, 43-50.
- Doll, R. and Peto, R. (1981). The causes of cancer: quantitative estimates of avoidable risks of cancer in the United States today. *J Natl Cancer Inst.*, **66**, 1191-1308.
- Donnellan, R. and Chetty, R. (1998). Cyclin D1 and human neoplasia. *Mol Pathol.*, **51**, 1-7.
- Dunlop, M.G., Farrington, S.M., Carothers, A.D., Wyllie, A.H., Sharp, L., Burn, J., Liu, B., Kinzler, K.W. and Vogelstein B. (1997). Cancer risk associated with germline DNA mismatch repair gene mutations. *Hum Mol Genet.*, **6**, 105-110.
- Evertsson, S., Lindblom, A. and Sun, X.F. (2001). APC I1307K and E1317Q variants are rare or do not occur in Swedish colorectal cancer patients. *Eur J Cancer.*, **37**, 499-502.
- Farrington, S.M., Lin-Goerke, J., Ling, J., Wang, Y., Burczak, J.D., Robbins, D.J. and Dunlop, M.G. (1998). Systematic analysis of hMSH2 and hMLH1 in young colon cancer patients and controls. *Am J Hum Genet.*, **63**, 749-759.
- Fearnhead, N.S., Britton, M.P. and Bodmer, W.F. (2001). The ABC of APC. *Hum Mol Genet.*, **10**, 721-733.
- Fields, S. and Jang, S. K. (1990). Presence of a potent transcription activating sequence in the p53 protein. *Science*, **249**, 1046-1049.
- Figer, A., Kaplan, A., Geva, R., Flex, D., Yaron, M., Levy, T., Sapir, E.E., Fidder, H.H., Friedman, E. (2002). Mutational analysis of the hMSH6 gene in familial and early-onset colorectal and endometrial cancer in Israeli patients. *Genet Test.*, **6**, 323-326.
- Finlay, C.A., Hinds, P.W., Tan, T.H., Eliyahu, D., Oren, M. and Levine, A.J. (1988). Activating mutations for transformation by p53 produce a gene product that forms an hsc70-p53 complex with an altered half-life. *Mol Cell Biol.* **8**:531-539.

- Finlay, I.G. and Pickford, I.R. (1994). Colorectal carcinoma. Lawson D.H. ed. Current Medicine. 4th ed. Churchill Livingstone, UK.
- Fink, D., Nebel, S., Norris, P.S., Baergen, R.N., Wilczynski, S.P., CoSta, J.M., Haas, M., Cannistra, S.A., and Howell, S.B. (1998). Enrichment for DNA mismatch repair-deficient cells during treatment with cisplatin. *Int J cancer*, **77**, 741-746.
- Fishel, R., Lescoe, M.K., and Rao, M.R.S. (1993). The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer. *Cell*, **75**, 1027-1038.
- Fishel, R. (2001). The selection for Mismatch repair defects in hereditary nonpolyposis colorectal cancer: revising the mutator hypothesis. *Cancer Res.*, **61**, 7369-7374.
- Fitzgibbons, R.J., Lynch, H.T., Stanislav, G.V., Watson, P.A., Lanspa, S.J., Markus, S.N., Smyrk, T., Krieger, M.D., and Lynch, J.F. (1987). Recognition and treatment of patients with hereditary non-polyposis colorectal cancer. *Ann Surg.*, **206**, 289-295.
- Foord, O.S., Bhattacharya, P., Reich, Z. and Rotter, V. (1991). A DNA binding domain is contained in the C-terminus of wild type p53 protein. *Nucleic Acids Res.*, **19**, 5191-5198.
- Foulkes, W.D., Thiffault, I., Gruber, S.B., Horwitz, M., Hamel, N., Lee, C., Shia, J., Markowitz, A., Figer, A., Friedman, E. *et al.* (2002). The founder mutation MSH2*1906G-->C is an important cause of hereditary nonpolyposis colorectal cancer in the Ashkenazi Jewish population. *Am J Hum Genet.*, **71**, 1395-412.
- Frank, T.S., Svoboda-Newman, S.M., and Hsi, E.D. (1996). Comparison of methods for extracting DNA from formalin-fixed paraffin sections for nonisotopic PCR. *Diagn Mol Pathol.*, **5**, 220-224.
- Frayling, I.M., Beck, N.E., Ilyas, M., Dove-Edwin, I., Goodman, P., Pack, K., Bell, J.A., Williams, C.B., Hodgson, S.V., Thomas, H.J., Talbot, I.C., Bodmer, W.F. and Tomlinson, I.P. (1998). The APC variants I1307K and E1317Q are associated with colorectal tumors, but not always with a family history. *Proc Natl Acad Sci.*, **95**, 10722-10727.
- Frazier, M.L., O'Donnell, F.T., Kong, S., Gu, X., Campos, I., Luthra, R., Lynch, P.M. and Amos, C.I. (2001). Age-associated risk of cancer among individuals with N-acetyltransferase 2 (NAT2) mutations and mutations in DNA mismatch repair genes. *Cancer Res.*, **61**, 1269-1271.

- Friedl, W., Meuschel, S., Caspari, R., Lamberti, C., Krieger, S., Sengteller, M. and Propping, P. (1996). Attenuated familial adenomatous polyposis due to a mutation in the 3' part of the APC gene. A clue for understanding the function of the APC protein. *Hum Genet.*, **97**, 579-584.
- Friedrich, C.A. (2001). Genotype-phenotype correlation in von Hippel-Lindau syndrome. *Hum Mol Genet.*, **10**, 763-767.
- Froggatt, N.J., Green, J., Brassett, C., Evans, D.G., Bishop, D.T., Kolodner, R. and Maher, E.R. (1999). A common MSH2 mutation in English and North American HNPCC families: origin, phenotypic expression, and sex specific differences in colorectal cancer. *J Med Genet.*, **36**, 97-102.
- Frosst, P., Blom, H.J., Milos, R., Goyette, P., Sheppard, C.A., Matthews, R.G., Boers, G.J., den Heijer, M., Kluijtmans, L.A., van den Heuvel, L.P. *et al.* (1995). A candidate genetic risk factor for vascular disease: a common mutation in methylenetetrahydrofolate reductase. *Nat Genet.*, **10**, 111-113.
- Gannon, J.V., Greaves, R., Iggo, R. and Lane, D.P. (1990). Activating mutations in p53 produce a common conformational effect. A monoclonal antibody specific for the mutant form. *EMBO J.*, **9**, 1595-1602.
- Gatti, R.A., Boder, E., Vinters, H.V., Sparkes, R.S., Norman, A. and Lange, K. (1991). Ataxia-telangiectasia: an interdisciplinary approach to pathogenesis. *Medicine*, **70**, 99-117.
- Giardiello, F.M., Offerhaus, G.J., Krush, A.J., Booker, S.V., Tersmette, A.C., Mulder, J.W., Kelley, C.N. and Hamilton, S.R. (1991). Risk of hepatoblastoma in familial adenomatous polyposis. *J Pediatr.*, **119**, 766-768.
- Giardiello, F.M., Offerhaus, G.J., Traboulsi, E.I., Graybeal, J.C., Maumenee, I.H., Krush, A.J., Levin, L.S., Booker, S.V. and Hamilton, S.R. (1991). Value of combined phenotypic markers in identifying inheritance of familial adenomatous polyposis. *Gut*, **32**, 1170-1174. (a).
- Giardiello, F.M., Offerhaus, G.J., Lee, D.H., Krush, A.J., Tersmette, A.C., Booker, S.V., Kelley, N.C. and Hamilton, S.R. (1993). Increased risk of thyroid and pancreatic carcinoma in familial adenomatous polyposis. *Gut*, **34**, 1394-1396.
- Giles, R.H., van Es, J.H. and Clevers, H. (2003). Caught up in a Wnt storm: Wnt signaling in cancer. *Biochim Biophys Acta.*, **1653**, 1-24.

- Gille, J.J., Hogervorst, F.B., Pals, G., Wijnen, J.T., van Schooten, R.J., Dommering, C.J., Meijer, G.A., Craanen, M.E., Nederlof, P.M., de Jong, D., McElgunn, C.J., Schouten, J.P. and Menko, F.H. (2002). Genomic deletions of MSH2 and MLH1 in colorectal cancer families detected by a novel mutation detection approach. *Br J Cancer*, **87**, 892-897.
- Giovannucci, E., Rimm, E.B., Ascherio, A., Stampfer, M.J., Colditz, G.A. and Willett, W.C. (1995). Alcohol, low-methionine--low-folate diets, and risk of colon cancer in men. *J Natl Cancer Inst*, **87**, 265-273.
- Goldberg, P.A., Madden, M.V., Harocopos, C., Felix, R., Westbrook, C. and Ramesar, R.S. (1998). In a resource-poor country, mutation identification has the potential to reduce the cost of family management for hereditary nonpolyposis colorectal cancer. *Dis Colon Rectum*, **41**, 1250-1253.
- Graff, J.R., Herman, J.G., Lapidus, R.G., Chopra, H., Xu, R., Jarrard, D.F., Isaacs, W.B., Pitha, P.M., Davidson, N.E. and Baylin, S.B. (1995). E-cadherin expression is silenced by DNA hypermethylation in human breast and prostate carcinomas. *Cancer Res*, **55**, 5195-5199.
- Green, R.C., Green, A.G., Simms, M., Pater, A., Robb, J.D. and Green, J.S. (2003). Germline hMLH1 promoter mutation in a Newfoundland HNPCC kindred. *Clin Genet*, **64**, 220-227.
- Greenlee, R.T., Murray, T., Bolden, S. and Wingo, P.A. (2000). Cancer statistics 2000. *CA Cancer J Clin*, **50**, 7-33.
- Grilley, M., Welsh, K.M., Su, S.S. and Modrich, P. (1989). Isolation and characterization of *Escherichia coli mutL* gene Product. *J Biol Chem*, **264**, 1000-1004.
- Groden, J., Thliveris, A., Samowitz, W., Carlson, M., Gelbert, L., Albertsen, H., Joslyn, G., Stevens, J., Spirio, L., Robertsen, M., Sargeant, L., Krapcho, K., Wolff, E., Burt, R., Hughes, J.P., Warrington, J., McPherson, J., Wasmuth, J., Le Paslier, D., Abderrahim, H., Cohen, D., Leppert, M. and White, R. (1991). Identification and characterisation of the familial adenomatous polyposis coli gene. *Cell*, **66**, 589-600.
- Grompe, M (1993). The rapid detection of unknown mutations in nucleic acids. *Nature Genet*, **5**, 111-117.
- Gross, E., Arnold, N., Goette, J., Schwarz-Boeger, U. and Kiechle, M. (1999). A comparison of BRCA1 mutation analysis by direct sequencing, SSCP and DHPLC. *Hum Genet*, **105**, 72-78.

- Groves, C., Lamlum, H., Crabtree, M., Williamson, J., Taylor, C., Bass, S., Cuthbert-Heavens, D., Hodgson, S., Phillips, R. and Tomlinson, I. (2002). Mutation cluster region, association between germline and somatic mutations and genotype-phenotype correlation in upper gastrointestinal familial adenomatous polyposis. *Am J Pathol.*, **160**, 2055-2061.
- Gryfe, R. and Gallinger, S. (2001). Microsatellite instability, mismatch repair deficiency, and colorectal cancer. *Surgery*, **130**, 17-20.
- Gsur, A., Haidinger, G., Hinteregger, S., Bernhoffer, G., Schatzl, G., Madersbacher, S., Marberger, M., Vutuc, C. and Micksche M (2001). Polymorphisms of Glutathione S-Transferase genes (GSTP1, GSTM1 and GSTT1) and prostate cancer risk. *Int J Cancer.*, **95**, 152-155.
- Guerrette, S., Acharya, S., and Fishel, R. (1999). The interaction of the human MutL homologues in hereditary nonpolyposis colon cancer. *J Biol Chem.*, **274**, 6336-6341.
- Guilford, P., Hopkins, J., Harraway, J., McLeod, M., McLeod, N., Harawira, P., Taite, H., Scoular, R., Miller, A. and Reeve, A.E. (1998). E-cadherin germline mutations in familial gastric cancer. *Nature*, **392**, 402-405.
- Hackman, P., Tannergard, P., Osei-Mensa, S., Chen, J., Kane, M.F., Kolodner, R., Lambert. B., Hellgren. D and Lindblom. A. (1997). A human compound heterozygote for two MLH1 missense mutations. *Nat Genet.*, **17**, 135-136.
- Haggitt, R.C. and Reid, B.J. (1986). Hereditary gastrointestinal polyposis syndromes. *Am J Surg Pathol.*, **10**, 871-887.
- Hall, M. and Peters, G. (1996). Genetic alterations of cyclins, cyclin-dependent kinases, and Cdk inhibitors in human cancer. *Adv Cancer Res.*, **68**, 67-108.
- Hamilton, S.R., Liu, B., Parsons, R.E., Papadopoulos, N., Jen, J., Powell, S.M., Krush, A.J., Berk, T., Cohen, Z., Tetu, B., Burger, P.C., Wood, P.A., Taqi, F., Booker, S.V., Petersen, G.M., Offerhaus, G.J.A. *et al.* (1995). The molecular basis of Turcot's syndrome. *N Engl J Med.*, **332**, 839-847.
- Hampel, H. and Peltomaki, P. (2000). Hereditary colorectal cancer: risk assessment and management. *Clin Genet.*, **58**, 89-97.
- Han, H.J., Maruyama, M., Baba, S., Park, J.G. and 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**, 237-242.

Harlow, E., Williamson, N.M., Ralston, R., Helfman, D.M. and Adams, T.E. (1985). Molecular cloning and in vitro expression of a cDNA clone for human cellular tumor antigen p53. *Mol Cell Biol.*, **5**, 1601-1610.

Harper, P.S. (1991). Huntington's disease. London Saunders., UK.

Hawn, M.T., Umar, A., Carethers, J.M., Marra, G., Kunkel, T.A., Boland, C.R. and Koi, M. (1995). Evidence for a connection between the mismatch repair system and the G2 cell cycle checkpoint. *Cancer Res.*, **55**: 3721-3725.

Hayes, J.D. and Pulford, D.J. (1995). The glutathione S-transferases supergene family: regulation of GST and the contribution of the isoenzymes to cancer chemoprevention and drug resistance. *Crit Rev Biochem Mol Biol.*, **30**, 445-600.

Heichman, K.A. and Roberts JM. (1994). Rules to replicate by. *Cell*, **79**, 557-562.

Heinimann, K., Scott, R.S., Chappius, P., Weber, W., Muller, H., Dobbie, Z. and Hutter, P. (1999). N-acetyltransferase 2 influences cancer prevalence in hMLH1/hMSH2 mutation carriers. *Cancer Res.*, **59**, 3038-3040.

Heinimann, K., Thompson, A., Locher, A., Furlanetto, T., Bader, E., Wolf, A., Meier, R., Walter, K., Bauerfeind, P., Marra, G., Muller, H., Foernzler, D. and Dobbie, Z. (2001). Nontruncating APC germ-line mutations and mismatch repair deficiency play a minor role in APC mutation-negative polyposis. *Cancer Res.*, **61**, 7616-7622.

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

HMSO. Cancer statistics. Registrations. England and Wales. London, UK.: HMSO 1989.

HMSO. Mortality statistics. Cause. England and Wales. London, UK: HMSO, 1992.

Hoogerbrugge, N., Willems, R., Van Krieken, H.J., Kiemeney, L.A., Weijmans, M., Nagengast, F.M., Arts, N., Brunner, H.G. and Ligtenberg,

- M.J. (2003). Very low incidence of microsatellite instability in rectal cancers from families at risk for HNPCC. *Clin Genet.*, **63**, 64-70.
- Houlston, R.S., Murday, V., Harocopous, C., Williams, C.B., and Slack, J. (1990). Screening and genetic counseling for relatives of patients with colorectal cancer in a cancer family clinic. *BMJ.*, **301**, 366-368.
- Houlston, R.S. and Tomlinson, I.P. (2001). Polymorphisms and colorectal tumor risk. *Gastroenterology*, **121**, 282-301.
- Houlston, R., Crabtree, M., Phillips, R and Tomlinson, I. (2001). Explaining differences in the severity of familial adenomatous polyposis and the search for modifier genes. *Gut*, **48**, 1-5.
- Huggins, M., Bloch, M., Wiggins, S., Adam, S., Suchowersky, O., Trew, M., Klimek, M., Greenberg, C.R., Eleff, M., Thompson, L.P. *et al.* (1992). Predictive testing for Huntington disease in Canada: adverse effects and unexpected results in those receiving a decreased risk. *Am J Med Genet.*, **42**, 508-515.
- Hulsken, J., Birchmeier, W. and Behrens, J. (1994). E-cadherin and APC compete for the interaction with beta-catenin and the cytoskeleton. *J Cell Biol.*, **127**, 2061-2069.
- Hunter, T. and Pines, J. (1994). Cyclins and cancer 2: cyclinD and CDk inhibitors come of age. *Cell*, **79**, 573-582.
- Hussein, M.R. and Wood, G.S. (2002). Building bridges in cancer. Mismatch repair and microsatellite instability. *Am J Dermatopathol.*, **24**, 76-81.
- ICG-HNPCC web site: <http://www.nfdht.nl>
- Ilett, K.F., David, B.M., Detchon, P., Castleden, W.M. and Kwa, R. (1987). Acetylation phenotype in colorectal carcinoma. *Cancer Res.*, **47**, 1466-1469.
- Ilett, K.F., Ingram, D.M., Carpenter, D.S., Teitel, C.H., Lang, N.P., Kadbular, F.F. and Minchin, R.F. (1994). Expression of monomorphic and polymorphic N-acetyltransferase in human colon. *Biochem Pharmacol.*, **47**, 914-917.
- Ilyas, M. and Tomlinson, I.P. (1997). The interactions of APC, E-cadherin and beta-catenin in tumour development and progression. *J Pathol.*, **182**, 128-137.
- Ilyas, M., Straub, J., Tomlinson, I.P., and Bodmer, W.F. (1999). Genetic pathways in colorectal and other cancers. *Eur J Cancer.*, **35**, 335-351..

- Jacob, S. and Praz, F. (2002). DNA mismatch repair defects: role in colorectal carcinogenesis. *Biochimie*, **84**, 27-47.
- Jarvinen, H.J. (1992). Epidemiology of familial adenomatous polyposis in Finland: impact of family screening on the colorectal cancer rate and survival. *Gut*, **33**, 357-360.
- Jass, J.R., Smyrk, T.C., Stewart, S.M., Lane, M.R., Lanspa, S.J., and Lynch, H.T. (1994). Pathology of hereditary non-polyposis colorectal cancer. *Anticancer Res.*, **14**, 1631-1634.
- Jass, J.R. (2000). Pathology of hereditary nonpolyposis colorectal cancer. *Ann N Y Acad Sci.*, **910**, 62-73.
- Jones, P.A. and Laird, P.W. (1999). Cancer epigenetics comes of age. *Nat Genet.*, **21**, 163-167.
- Joslyn, G., Richardson, D.S., White, R. and Alber, T. (1993). Dimer formation by an N-terminal coiled coil in the APC protein. *Proc Natl Acad Sci.*, **90**, 11109-11113.
- Katoh, T., Nagata, N., Kuroda, Y., Itoh, H., Kawahara, A., Kuroki, N., Ookuma, R. and Bell, D.A. (1996). Glutathione S-transferase M1 (GSTM1) and T1 (GSTT1) genetic polymorphism and susceptibility to gastric and colorectal adenocarcinoma. *Carcinogenesis*, **17**, 1855-1859.
- Kariola, R., Otway, R., Lonnqvist, K.E., Raevaara, T.E., Macrae, F., Vos, Y.J., Kohonen-Corish, M., Hofstra, R.M. and Nystrom-Lahti, M. (2003). Two mismatch repair gene mutations found in a colon cancer patient--which one is pathogenic? *Hum Genet.*, **112**, 105-9.
- Karran, P. and Hampson, R. (1996). Genomic instability and tolerance to alkylating agents. *Cancer Surv.*, **28**, 69-85..
- Kern, S.E., Kinzler, K.W., Bruskin, A., Jarosz, D., Friedman, P., Prives, C. and Vogelstein, B. (1991). Identification of p53 as a sequence-specific DNA-binding protein. *Science*, **252**, 1708-1711.
- Kim, H.C., Wheeler, J.M., Kim, J.C., Ilyas, M., Beck, N.E., Kim, B.S., Park, K.C. and Bodmer, W.F. (2000). The E-cadherin gene (CDH1) variants T340A and L599V in gastric and colorectal cancer patients in Korea. *Gut*, **47**, 262-267.
- Kinzler, K.W., Nilbert, M.C., Su, L-K., Vogelstein, B., Bryan, T.M., Levy, D.B., Smith, K.J., Preisinger, A.C., Hedge, P., McKechnie, D., Finnear, R., Markham, A., Groffen, J., Boguski, M.S., Altschul, S.F., Horii, A., Ando, H., Miyoshi, Y., Miki, Y., Nishisho, I. and Nakamura, Y. (1991).

Identification of FAP locus genes from chromosome 5q21. *Science*, **253**, 661-665.

Kirlin, W.G., Ogolla, F., Andrews, A.F., Trinidad, A., Ferguson, R.J., Yerokun, T., Mpezo, M. and Hein, D.W. (1991). Acetylator genotype-dependent expression of arylamine N-acetyltransferase in human colon cytosol from non-cancer and colorectal cancer patients. *Cancer Res.* **51**, 549-555.

Kirsner, J.B., and Shorter, R.G.(eds) (1988). Diseases of the colon, rectum and anal canal. Williams and Wilkins.

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

Kolodner, R.D., Hall, N.R., Lipford, J., Kane, M.F., Roa, M.R.S., Morrison, P., Wirth, L., Finan, P.J., Burn, J., Chapman, P., Earabino, C., Merchant, E., and Bishop, D.T. (1994). Structure of the human MSH2 locus and analysis of two Muir-Torre kindreds for msh2 mutations. *Genomics*, **24**, 516-526.

Kondo, E., Horii, A. and Fukushige, S. (2001). The interacting domains of three MutL heterodimers in man: hMLH1 interacts with 36 homologous amino acid residues within hMLH3, hPMS1 and hPMS2. *Nucleic Acids Res.*, **29**, 1695-1702.

Kong, S., Amos, C.I., Luthra, R., Lynch, P.M., Levin, B. and Frazier, M.L. (2000). Effects of cyclin D1 polymorphism on age of onset of hereditary nonpolyposis colorectal cancer. *Cancer Res.*, **60**, 249-252.

Korinek, V., Barker, N., Morin, P.J., van Wichen, D., de Weger, R., Kinzler, K.W., Vogelstein, B. and Clevers, H. (1997). Constitutive transcriptional activation by a beta-catenin-Tcf complex in APC-/- colon carcinoma. *Science*, **275**, 1784-1787.

Kotze, M.J., Langenhoven, E., Warnich, L., du Plessis, L. and Retief, A.E. (1991). The molecular basis of familial hypercholesterolaemia in South African Afrikaners. *Ann Hum Genet*, **55**, 115-121.

Kowalski, L.D., Mutch, D.G., Herzog, T.J., Rader, J.S. and Goodfellow, P.J. (1997). Mutational analysis of MLH1 and MSH2 in 25 prospectively-acquired RER+ endometrial cancers. *Genes Chromosomes Cancer*. **18**, 219-27.

Kulczycki, L.L., Kostuch, M, and Bellanti, J.A. (2003). A clinical perspective of cystic fibrosis and new genetic findings: relationship of

- CFTR mutations to genotype-phenotype manifestations. *Am J Med Genet.*, **116A**, 262-267.
- Laird, P.W. and Jaenisch, R. (1994). DNA methylation and cancer. *Hum Mol Genet.*, **3**, 487-495.
- Laken, S.J., Petersen, G.M., Gruber, S.B., Oddoux, C., Ostrer, H., Giardiello, F.M., Hamilton, S.R., Hampel, H., Markowitz, A., Klimstra, D., Jhanwar, S., Winawer, S., Offit, K., Luce, M.C., Kinzler, K.W. and Vogelstein B. (1997). Familial colorectal cancer in Ashkenazim due to a hypermutable tract in APC. *Nat Genet*, **17**, 79-83.
- Lane, D.P. (1992). Cancer. p53, guardian of the genome. *Nature*, **358**, 15-16.
- Lang, N.P., Chu, D.Z., Hunter, C.F., Kendall, D.C., Flammang, T.J. and Kadlubar, F.F. (1986). Role of aromatic amine acetyltransferase in human colorectal cancer. *Arch Surg.*, **121**, 1259-1261.
- Laurent-Puig, P., Beroud, C. and Soussi, T. (1998). APC gene: database of germline and somatic mutations in human tumors and cell lines. *Nucleic Acids Res.*, **26**, 269-270.
- 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 non-polyposis colorectal cancer. *Cell*, **75**, 1215-1225.
- Lerman, C., Rimer, B.K. and Engstrom, P.F. (1991). Cancer risk notification: psychosocial and ethical implications. *J Clin Oncol.*, **9**, 1275-1282.
- Le Saux, O., Beck, K., Sachsinger, C., Treiber, C., Goring, H.H., Curry, K., Johnson, E.W., Bercovitch, L., Marais, A.S., Terry, S.F., Viljoen, D.L., Boyd, C.D. (2002). Evidence for a founder effect for pseudoxanthoma elasticum in the Afrikaner population of South Africa. *Hum Genet.*, **111**, 331-338.
- Leung, W.K., Yu, J., Ng, E.K., To, K.F., Ma, P.K., Lee, T.L., Go, M.Y., Chung, S.C. and Sung, J.J. (2001). Concurrent hypermethylation of multiple tumor-related genes in gastric carcinoma and adjacent normal tissues. *Cancer*, **91**, 2294-2301.
- Levine, A.J. (1997). p53, the cellular gatekeeper for growth and division. *Cell*, **88**, 323-331.

- Levison, G., and Gutman, G.A. (1987). Slipped-strand mispairing: a major mechanism for DNA sequence evolution. *Mol Biol Evol.*, **4**, 202-221.
- Lewis, R. (2001). Human genetics, Concepts and Applications. WCB/McGraw-Hill., New York, USA
- Li, G.-M., Wang, H. and Romano, L.J., (1996). Human MutSa specifically binds to DNA containing aminofluorene and acetylaminofluorene adducts. *J Biol Chem.*, **271**, 24084-24088.
- Lin, D., Meyer, D.J., Ketterer, B., Lang, N.P. and Kadlubar, F.F. (1994). Effects of human and rat glutathione S-transferases on the covalent DNA binding of the N-acetoxy derivatives of heterocyclic amine carcinogens in vitro: a possible mechanism of organs specificity in their carcinogenesis. *Cancer Res.*, **54**, 4920-4926.
- Lin, K.M., Shashidharan, M., Ternent, C.A., Thorson, A.G., Blatchford, G.J., Christensen, M.A., Lanspa, S.J., Lemon, S.J., Watson, P. and Lynch, H.T. (1998). Colorectal and extracolonic cancer variations in MLH1/MSH2 hereditary nonpolyposis colorectal cancer kindreds and the general population. *Dis Colon Rectum.*, **41**, 428-433.
- Lindblom, A., Tannergard, P., Werelius, B. and Nordenskjold, M. (1993). Genetic mapping of a second locus predisposing to hereditary nonpolyposis colon cancer. *Nature Genet.*, **5**, 279-282.
- Litt, M., and Luyt, J.A. (1989). A hypervariable microsatellite revealed by *in vitro* amplification of a dinucleotide repeat within the cardiac muscle actin gene. *Hum Mol Genet.*, **44**, 397-401.
- Lohmann, D., Horsthemke, B., Gillessen-Kaesbach, G., Stefani, F.H., and Hofler, H. (1992). Detection of small RB1 gene deletions in retinoblastoma by multiplex PCR and high resolution gel electrophoresis. *Hum Genet.*, **89**, 49-53.
- Lu, S-H., Kawabata, M., Imamura, T., Akiyama, Y., Nomizu, T., Miyazono, K. and Yuasa, Y. (1998). HNPCC associated with germline mutation in the TGF- β type II receptor gene. *Nat Genet.*, **19**, 17-18.
- Lynch, H.T., Shaw, M.W., Magnuson, C.W., Larsen, A.L. and Krush, A.J. (1966). Hereditary factors in cancer. Study of two large midwestern kindreds. *Arch Intern Med.*, **117**, 206-212.
- Lynch, H.T., Cavalieri, R.J. and Lynch, J.F. (1992). Gynaecologic cancer clues to Lynch syndrome II diagnosis: a family report. *Gynecol Oncol.*, **44**, 198-203.

- Lynch, H.T., Smyrk, T.C., Watson, P., Lanspa, S.J., Lynch, J.F., Lynch, P.M., Cavalieri, R.J. and Boland, R. (1993). Genetics, Natural History, Tumor spectrum, and Pathology of Hereditary Nonpolyposis Colorectal Cancer: An updated review. *Gastroenterology*, **104**, 1535-1549.
- Lynch, H.T., Smyrk, T., McGinn, T., Lanspa, S., Cavalieri, J., Lynch, J., Slominski-Caster, S., Cayouette, M.C., Priluck, I. and Luce, M.C. (1995). A phenotypically and genotypically distinctive variant of FAP. *Cancer*, **76**, 2427-2433.
- Lynch, H.T., Smyrk, T., and Lynch, J.F. (1996). Overview of natural history, pathology, molecular genetics and management of HNPCC (Lynch syndrome). *Int J Cancer*, **69**, 38-43.
- Lynch, H.T., Smyrk, T. and Lynch, J. (1997). An update of HNPCC (Lynch syndrome). *Cancer Genet Cytogenet.*, **93**, 84-99.
- Lynch, H.T. and Lynch, J. (1998). Genetics of colonic cancer. *Digestion*, **59**, 481-492.
- Lynch, H.T. and de la Chapelle, A. (1999). Genetic susceptibility to non-polyposis colorectal cancer. *J Med Genet.*, **36**, 801-818.
- Ma, J., Stampfer, M.J., Christensen, B., Giovannucci, E., Hunter, D.J., Chen, J., Willett, W.C., Selhub, J., Hennekens, C.H., Gravel, R. and Rozen, R. (1999). A polymorphism of the methionine synthase gene: association with plasma folate, vitamin B12, homocyst(e)ine, and colorectal cancer risk. *Cancer Epidemiol Biomarkers Prev.*, **8**, 825-829.
- Maillet, P., Chappuis, P.O., Vaudan, G., Dobbie, Z., Muller, H., Hutter, P. and Sappino, A.P. (2000). A polymorphism in the ATM gene modulates the penetrance of hereditary non-polyposis colorectal cancer. *Int J Cancer*, **88**, 928-931.
- Maisto, O.E., and Bremner, C.G. (1981). Cancer of the colon and rectum in the coloured population of Johannesburg. *S Afr Med J.*, **60**, 571-573.
- Maliaka, Y.K., Chudina, A.P., Belev, N.F., Alday, P., Bochkov, N.P. and Buerstedde, J.M (1996). CpG dinucleotides in the hMSH2 and hMLH1 genes are hotspots for HNPCC mutations. *Hum Genet.*, **97**, 251-255.
- Malkin, D.; Li, F. P.; Strong, L. C.; Fraumeni, J. F., Jr.; Nelson, C. E.; Kim, D. H.; Kassel, J.; Gryka, M. A.; Bischoff, F. Z.; Tainsky, M. A.; Friend, S. H. (1990). Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. *Science*, **250**, 1233-1238.

- Matthias, C., Branigan, K., Jahnke V., Leter, K., Haas, J., Heighway, J., Jones, P.W., Strange, R.C., Fryer, A.A. and Hoban, P.R. (1998) Polymorphisms within the cyclin D1 gene is associated with prognosis in patients with squamous cell carcinoma of the head and neck. *Clin Cancer Res.*, **4**, 2411-2418.
- Marx, J. (1994). DNA repair comes into its own. *Science*, **266**, 728-730.
- Mecklin J-P., and Jarvinen, H.J. (1991). Tumour spectrum in cancer family syndrome (hereditary nonpolyposis colorectal cancer). *Cancer*, **68**, 1109-1112.
- Menko, F.H.(ed) 1993. Familial and hereditary colorectal cancer. In: Genetics of colorectal cancer for clinical practice, 1st ed, Kluwer Academic Publishers, Netherlands, **58**-111.
- Milner, J. and Medcalf, E.A. (1991). Cotranslation of activated mutant p53 with wild type drives the wild-type p53 protein into the mutant conformation. *Cell*, **65**, 765-774.
- Miyaki, M., Konishi, M., Kikuchi-Yanoshita, R., Muraoka, M., Yasuno, M., Igari, T., Koike, M., Chibu, M. and Mori, T. (1997). Germline mutation of *MSH6* as the cause of hereditary nonpolyposis colorectal cancer. *Nat Genet.*, **17**, 271-272.
- Miyoshi, Y., Nagase, H., Ando, H., Horii, A., Ichii, S., Nakatsuru, S., Aoki, T., Miki, Y., Mori, T., Nakamura, Y. (1992). Somatic mutations of the APC gene in colorectal tumors: mutation cluster region in the APC gene. *Hum Mol Genet.*, **1**, 229-233.
- Modrich, P. (1997). Strand-specific Mismatch repair in Mammalian Cells (Review). *J Biol Chem.*, **272**, 24727-24730.
- Moisio, A.L., Sistonen, P., Mecklin, J.P., Jarvinen, H. and Peltomaki, P. (1998). Genetic polymorphisms in carcinogen metabolism and their association to hereditary nonpolyposis colon cancer. *Gastroenterology*, **115**, 1387-1394.
- Moon, R.T. and Miller, J.R. (1997). The APC tumor suppressor protein in development and cancer. *Trends Genet.*, **13**, 256-258.
- Morgan, D.O. (1995). Principles of CDK regulation. *Nature*, **374**, 131-134.
- Mori, H., Hattori, N. and Mizuno, Y. (2003). Genotype-phenotype correlation: familial Parkinson disease. *Neuropathology*. **23**, 90-94.

- Morton, R.A., Ewing, C.M., Nagafuchi, A., Tsukita, S. and Isaacs, W.B. (1993). Reduction of E-cadherin levels and deletion of the alpha-catenin gene in human prostate cancer cells. *Cancer Res.*, **53**, 3585-3590.
- Moslein G. (2003). Clinical implications of molecular diagnosis in hereditary nonpolyposis colorectal cancer. *Recent Results Cancer Res.*, **162**, 73-78.
- Muller, H. Hj., Heinimann, K. and Dobie, Z. (2000). Genetics of hereditary colon cancer-a basis for prevention? *Eur J Cancer.*, **36**, 1215-1223.
- Muller, A. and Fishel, R. (2002). Mismatch repair and the Hereditary Non-polyposis Colorectal Cancer Syndrome (HNPCC). *Cancer Invest.*, **20**, 102-109.
- Munemitsu, S., Souza, B., Muller, O., Albert, I., Rubinfeld, B. and Polakis, P. (1994). The APC gene product associates with microtubules in vivo and promotes their assembly in vitro. *Cancer Res.*, **54**, 3676-3681.
- Munemitsu, S., Albert, I., Souza, B., Rubinfeld, B. and Polakis, P. (1995). Regulation of intracellular beta-catenin levels by the adenomatous polyposis coli (APC) tumor-suppressor protein. *Proc Natl Acad Sci.*, **92**, 3046-3050.
- Nadeau, J.H. (2003). Modifier genes and protective alleles in humans and mice. *Curr Opinion Genet & Dev.*, **13**, 290-295.
- Nagar, B., Overduin, M., Ikura, M. and Rini, J.M. (1996). Structural basis of calcium-induced E-cadherin rigidification and dimerization. *Nature*, **380**, 360-364.
- Nakamura, Y. (1993). The role of the adenomatous polyposis coli (APC) gene in human cancers. *Adv Cancer Res.*, **62**, 65-87.
- Nebert, D.W. (2002). Transcription factors and cancer: an overview. *Toxicology*, **181-182**, 131-141.
- Nicolaides, N.C., Papadopoulos, N., Liu, B., Wei, Y.F., Carter, K.C., Ruben, S.M., Rosen, C.A., Haseltine, W.A., Fleishmann, R.D., Fraser, C.M. *et al.* (1994). Mutations of two PMS homologues in hereditary nonpolyposis colon cancer. *Nature*, **371**, 75-80.
- Nicum, S., Midgley, R. and Kerr, D.J. (2003). Colorectal cancer. *Acta Oncol.*, **42**, 263-275.

- Nollet, F., Berx, G., Molemans, F. and van Roy, F. (1996). Genomic organization of the human beta-catenin gene (CTNNB1). *Genomics*, **32**, 413-424.
- Northover M.A. and Kettner, J.D.(1992). Bowel Cancer-the facts. Oxford University Press, UK.
- Nurse, G.T. and Jenkins, T. The Griqua of Campbell, Cape Province, South Africa. (1995). *Am J Phys Anthropol.*, **43**, 71-78.
- Nystrom-Lahti, M., Kristo, P., Nicolaides, N.C., Chang, S.Y., Aaltonen, L.A., Moisio, A.L., Jarvinen, H.J., Mecklin, J.P., Kinzler, K.W., Vogelstein, B., *et al.* (1995). Founding mutations and Alu-mediated recombination in hereditary colon cancer. *Nat Med.*, **1**, 1203-1206.
- Oettle, G.J. and Strijdom, J.G. (1997). Invited comment. *SAJS.*, **35**, 65-67.
- Ohtsubo, M. and Roberts, J.M. (1993). Cyclin-dependent regulation of G1 in mammalian fibroblasts. *Science*, **259**, 1908-1912.
- Orita, M., Iwahana, H., Kanazawa, H., Hayashi, K., and Sekiya, T. (1989). Detection of polymorphisms of human DNA by by gel electrophoresis as single strand conformation polymorphisms. *Proc Natl Acad Sci USA.*, **86**, 2766-2770.
- Overduin, M., Harvey, T.S., Bagby, S., Tong, K.I., Yau, P., Takeichi, M. and Ikura, M. (1995). Solution structure of the epithelial cadherin domain responsible for selective cell adhesion. *Science*, **267**, 386-389.
- Oyama, T., Kanai, Y., Ochiai, A., Akimoto, S., Oda, T., Yanagihara, K., Nagafuchi, A., Tsukita, S., Shibamoto, S., Ito, F., *et al.* (1994). A truncated beta-catenin disrupts the interaction between E-cadherin and alpha-catenin: a cause of loss of intercellular adhesiveness in human cancer cell lines. *Cancer Res.*, **54**, 6282-6287.
- 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. *et al.* (1994). Mutation of a mutL homologue in hereditary colon cancer. *Science*, **263**, 1625-1629.
- Pardee, A.B. (1989). G1 events and regulation of cell proliferation. *Science*, **246**, 603-608.
- Parker, S.L., Tong, T., Bolden, S. and Wingo, P.A. (1999). Cancer statistics, 1996. *CA Cancer J Clin.*, **46**, 5-27.

- Parsons, R., Myerhoff, L., Liu, B., Wilson, J.K.V., Markowitz, S., Kinzler, K.W., and Vogelstein, B. (1995). Microsatellite instability and mutations in the transforming growth factor β type II receptor gene in colorectal cancer. *Cancer Res.*, **55**, 5548-5550.
- Peeters, A.V. and Kotze, M.J. (1994). Improved heteroduplex detection of single base substitutions in PCR-amplified DNA. *PCR Methods Appl.*, **4**, 188-190.
- Peifer, M. (1996). Regulating cell proliferation: as easy as APC. *Science*, **272**, 974-975.
- Peltomaki, P., Lothe, R.A., Aaltonen, L.A., Pylkkanen, L., Nystrom-Lahti, M., Seruca, R., David, L., Holm, R., Ryberg, D., Haugen, A., Brogger, A., Borresen, A-L and de la Chapelle. (1993). Microsatellite instability is associated with tumours that characterize the hereditary nonpolyposis colorectal carcinoma syndrome. *Cancer Res.*, **53**, 5853-5855.
- Peltomaki, P. (2001). DNA mismatch repair and cancer. *Mutat Res.*, **488**, 77-85.(a).
- Peltomaki, P. (2001). Deficient DNA mismatch repair: a common etiologic factor for colon cancer. *Hum Mol Genet.*, **10**, 735-740 (b).
- Peltomaki, P. (2003). Role of DNA mismatch Repair Defects in the Pathogenesis of Human Cancer. *J Clin Oncol.*, **21**, 1174-1179.
- Pemble, S., Schroeder, K.R., Spencer, S.R., Meyer, D.J., Hallier, E., Bolt, H.M., Ketterer, B. and Taylor, J.B. (1994). Human glutathione S-transferase theta (GSTT1): cDNA cloning and the characterization of a genetic polymorphism. *Biochem J.*, **300**, 271-276.
- Percesepe, A., Borghi, F., Menigatti, M., Losi, L., Foroni, M. Di Gregorio, C., Rossi, G., Pedroni, M., Sala, E., Vaccina, F., Roncucci, L., Benatti, P., Viel, A., Genuardi, M., Marra, G., Kristo, P., Peltomaki, P. and Ponz de Leon, M. (2001). Molecular screening for hereditary nonpolyposis colorectal cancer: a prospective, population-based study. *J Clin Oncol.*, **19**, 3944-3950.
- Perucho, M. (1996). Microsatellite instability: the mutator that mutates the other mutator. *Nat Med.*, **2**, 630-631.
- Pietenpol, J.A., Tokino, T., Thiagalingam, S., el-Deiry, W.S., Kinzler, K.W. and Vogelstein, B. (1994). Sequence-specific transcriptional activation is essential for growth suppression by p53. *Proc Natl Acad Sci.*, **9**, 1998-2002.

- Pistorius, S.R., Nagel, M., Kruger, S., Plaschke, J., Kruppa, C., Wehrmann, U., Schackert, H.K. and Saeger, H.D. (2001). Combined molecular and clinical approach for decision making for surgery in HNPCC patients: a report on three cases in two families. *Int J Colorectal Dis.*, **16**, 402-407.
- Potter, J.D. (1999). Colorectal cancer: molecules and populations. *J Natl Cancer Inst.*, **91**, 916-32.
- Quaresima, B., Alifano, P., Tassone, P., Avvedimento, E., Costanzo, F. and Venuta, S. (2003). Human mismatch repair protein MutL homologue 1 (MLH1) interacts with Escherichia coli MutL and MutS in vivo and in vitro: a simple genetic system to assay MLH1 function. *Biochem J.*, **371**, 183-189.
- Quelle, D.E., Ashmun, R.A., Shurtleff, S.A., Kato, J.Y., Bar-Sagi, D., Roussel, M.F. and Sherr, C.J. (1993). Overexpression of mouse D-type cyclins accelerates G1 phase in rodent fibroblasts. *Genes Dev.*, **7**, 1559-1571.
- Raycroft, L., Wu, H.Y. and Lozano, G. (1990). Transcriptional activation by wild-type but not transforming mutants of the p53 anti-oncogene. *Science*, **249**, 1049-1051.
- Rebbeck, T.R. (1997). Molecular epidemiology of the human glutathione S-Transferase genotypes GSTM1 and GSTT1 in cancer susceptibility. *Cancer Epidemiol Biomarkers Prev.*, **6**, 733-743.
- Resnitsky, D., Gossen, M. and Bujard, H (1994). Acceleration of the G1/S phase transition by expression of cyclins D1 and E with an inducible system. *Mol Cell Biol.*, **14**, 1669-1679.
- Ricciardone, M.D., Ozcelik, T., Cevher, B., Ozdag, H., Tuncer, M., Gurgey, A., Uzunalioglu, O., Cetinkaya, H., Tanyeli, A., Erken, E. and Ozturk, M. (1999). Human MLH1 deficiency predisposes to hematological malignancy and neurofibromatosis type 1. *Cancer Res.*, **59**, 290-293.
- Richards, F.M., McKee, S.A., Rajpar, M.H. (1999). Germline E-cadherin gene (CDH1) mutations predispose to familial gastric cancer and colorectal cancer. *Hum Mol Genet.*, **8**, 607-610.
- Rodriguez-Bigas, M.A. (1996). Prophylactic colectomy for gene carriers in hereditary nonpolyposis colorectal cancer-has the time come? *Cancer*, **78**, 199-201.
- Rodriguez-Bigas, M.A., Boland, C.R., Hamilton, S.R., Henson, D.E., Jass, J.R., Khan, P.M., Lynch, H., Perucho, M., Smyrk, T., Sobin, L. and

- Srivastava, S. (1997). A National Cancer Institute Workshop on Hereditary Nonpolyposis Colorectal Cancer Syndrome: meeting highlights and Bethesda guidelines. *J Natl Cancer Inst.*, **89**, 1758-1762. (a).
- Rodriguez-Bigas, M.A., Vasen, H.F.A., Mecklin, J.P., Myrhoj, T., Rozen, P., Bertario, L., Jarvinen, H.J., Jass, J.R., Kunitomo, K., Nomizu, T., and Driscoll, D.L. (1997). Rectal cancer risk in hereditary non polyposis colorectal cancer after an abdominal colectomy. International Group on HNPCC. *Ann Surg.*, **225**, 202-207. (b).
- Roest, P.A., Roberts, R.G., Sugino, S., van Ommen, G.J. and den Dunnen, J.T. (1993). Protein truncation test (PTT) for rapid detection of translation-terminating mutations. *Hum Mol Genet.*, **2**, 1719-1721.
- Rogers, S., Wells, R. and Rechsteiner, M. (1986). Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. *Science*, **234**, 364-368.
- Rubinfeld, B., Souza, B., Albert, I., Muller, O., Chamberlain, S.H., Masiarz, F.R., Munemitsu, S. and Polakis, P. (1993). Association of the APC gene product with beta-catenin. *Science*, **262**, 1731-1734.
- Rubinfeld, B., Albert, I., Porfiri, E., Fiol, C., Munemitsu, S. and Polakis, P. (1996). Binding of GSK3beta to the APC-beta-catenin complex and regulation of complex assembly. *Science*, **272**, 1023-6.
- Rustgi, A.K. (1994). Hereditary gastrointestinal polyposis and nonpolyposis syndromes. *N Engl J Med.*, **331**, 1694-1702.
- Saletti, P., Edwin, I.D., Pack, K., Cavalli, F. and Atkin, W.S. (2001). Microsatellite instability: Application in hereditary non-polyposis colorectal cancer. *Ann Oncol.*, **12**, 151-160.
- Samowitz, W.S., Curtin, K., Lin, H.H., Robertson, M.A., Schaffer, D., Nichols, M., Gruenthal, K., Leppert, M.F. and Slattery, M.L. (2001). The colon cancer burden of genetically defined hereditary nonpolyposis colon cancer. *Gastroenterology*, **121**, 830-838.
- Sampson, J.R., Dolwani, S., Jones, S., Eccles, D., Ellis, A., Evans, D.G., Frayling, I., Jordan, S., Maher, E.R., Mak, T., Maynard, J., Pigatto, F., Shaw, J. and Cheadle, J.P. (2003). Autosomal recessive colorectal adenomatous polyposis due to inherited mutations of MYH. *Lancet*. **362**, 39-41.
- Sandoval, N., Platzer, M., Rosenthal, A., Dork, T., Bendix, R., Skawran, B., Stuhmann, M., Wegner, R.D., Sperling, K., Banin, S., Shiloh, Y., Baumer, A., Bernthaler, U., Sennefelder, H., Brohm, M., Weber, B.H.

- and Schindler, D. (1999). Characterization of ATM gene mutations in 66 ataxia telangiectasia families. *Hum Mol Genet.*, **8**, 69-79.
- Savitsky, K., Bar-Shira, A., Gilad, S., Rotman, G., Ziv, Y., Vanagaite, L., Tagle, D.A., Smith, S., Uziel, T., Sfez, S. *et al.* (1995). A single ataxia telangiectasia gene with a product similar to PI-3 kinase. *Science*. **268**, 1749-1753.
- Sawa, H., Ohshima, T.A., Ukita, H., Murakami, H., Chiba, Y., Kamada, H., Hara, M. and Saito, I. (1998). Alternatively spliced forms of cyclin D1 modulate entry into the cell cycle in an inverse manner. *Oncogene*, **16**, 1701-1712.
- Selhub, J. and Miller, J.W. (1992). The pathogenesis of homocysteinemia: interruption of the coordinate regulation by S-adenosylmethionine of the remethylation and transsulfuration of homocysteine. *Am J Clin Nutr.*, **55**, 131-138.
- Selivanova, G., Iotsova, V., Kiseleva, E., Strom, M., Bakalkin, G., Grafstrom, R.C. and Wiman, K.G. (1996). The single-stranded DNA end binding site of p53 coincides with the C-terminal regulatory region. *Nucleic Acids Res.*, **24**, 3560-3567.
- Service, R. (1994). Stalking the start of colon cancer. *Science*, **263**, 1559-1560.
- Shannon, B., Gnanasampathan, S., Beilby, J. and Iacopetta, B. (2002). A polymorphism in the methylenetetrahydrofolate reductase gene predisposes to colorectal cancers with microsatellite instability. *Gut*, **50**, 520-524.
- Shapiro, L., Fannon, A.M., Kwong, P.D., Thompson, A., Lehmann, M.S., Grubel, G., Legrand, J.F., Als-Nielsen, J., Colman, D.R. and Hendrickson, W.A. (1995). Structural basis of cell-cell adhesion by cadherins. *Nature*, **374**, 327-337.
- Sherr, C.J. (1994). G1 phase progression: cycling on cue. *Cell*, **79**, 551-555.
- Sherr, C.J. (1996). Cancer Cell cycles. *Science*, **274**, 1672-1677.
- Shinmura, K., Kohno, T., Takahashi, M., Sasaki, A., Ochiai, A., Guilford, P., Hunter, A., Reeve, A.E., Sugimura, H., Yamaguchi, N. and Yokota, J. (1999). Familial gastric cancer: clinicopathological characteristics, RER phenotype and germline p53 and E-cadherin mutations. *Carcinogenesis*, **20**, 1127-1131.

- Shirasawa, S., Urabe, K., Yanagawa, Y., Toshitani, K., Iwama, T. and Sasazuki, T. (1991). p53 gene mutations in colorectal tumors from patients with familial polyposis coli. *Cancer Res.*, **51**, 2874-2878.
- Sieber, O.M., Tomlinson, I.P. and Lamlum, H. (2000). The adenomatous polyposis coli (APC) tumour suppressor--genetics, function and disease. *Mol Med Today.*, **6**, 46246-46249.
- Sieber, O.M., Lamlum, H., Crabtree, M.D., Rowan, A.J., Barclay, E., Lipton, L., Hodgson, S., Thomas, H.J., Neale, K., Phillips, R.K., Farrington, S.M., Dunlop, M.G., Mueller, H.J., Bisgaard, M.L., Bulow, S., Fidalgo, P., Albuquerque, C., Scarano, M.I., Bodmer, W., Tomlinson, I.P. and Heinimann, K. (2002). Whole-gene APC deletions cause classical familial adenomatous polyposis, but not attenuated polyposis or "multiple" colorectal adenomas. *Proc Natl Acad Sci U S A.*, **99**, 2954-2958.
- Sieber, O.M., Lipton, L., Crabtree, M., Heinimann, K., Fidalgo, P., Phillips, R.K., Bisgaard, M.L., Orntoft, T.F., Aaltonen, L.A., Hodgson, S.V., Thomas, H.J., Tomlinson, I.P. (2003). Multiple colorectal adenomas, classic adenomatous polyposis, and germ-line mutations in MYH. *N Engl J Med.*, **348**:791-799.
- Siedergard, J., Pero, R., Willer, D.G. and Beattie, E.J. (1986). A glutathione transferase inhuman leukocytes as a marker for the susceptibility to lung cancer. *Carcinogenesis*, **7**, 751-753.
- Sitas, F. (1994). Histologically diagnosed cancers in South Africa, 1988. *S Afr Med J.*, **84**, 344-348.
- Sitas, F., Madhoo, J. and Wessie, J. (1998). Cancer in South Africa, 1993-1995. National Cancer Registry of South Africa, South African Institute for Medical Research, Johannesburg.
- Slattery, M.L., Potter, J.D., Samowitz, W., Schaffer, D. and Leppert, M. (1999). Methylenetetrahydrofolate reductase, diet, and risk of colon cancer. *Cancer Epidemiol Biomarkers Prev.*, **8**, 513-518.
- Smith, G., Stanley, L.A., Sim, E., Strange, R.C. and Wolf, C.R. (1995). Metabolic polymorphisms and cancer susceptibility. *Cancer Surv.*, **25**, 27-65.
- Smith, C.A., Wadelius, M., Gough, A.C., Harrison, D.J., Wolf, C.R. and Rane, A. (1997). A simplified assay for the arylamine N-acetyltransferase 2 polymorphism validated by phenotyping with isoniazid. *J Med Genet.*, **34**, 758-760.

- Spirio, L., Olschwang, S., Groden, J., Robertson, M., Samowitz, W., Joslyn, G., Gelbert, L., Thliveris, A., Carlson, M., Otterud, B., *et al.* (1993). Alleles of the APC gene: an attenuated form of familial polyposis. *Cell*, **75**, 951-957.
- Srivastava, S.; Zou, Z.; Pirollo, K.; Blattner, W.; Chang, E. H. (1990). Germ-line transmission of a mutated p53 gene in a cancer-prone family with Li-Fraumeni syndrome. *Nature*, **348**, 747-749.
- Strachan, P. and Read, A. (1999). Human Molecular Genetics 2. Bios Scientific Publishers Ltd., UK
- Strange, R.C., Matharoo, B., Faulder, G.C., Jones, P., Cotton, W., Elder, J.B. and Deakin, M. (1991). The human glutathione S-transferases: a case-control study of the incidence of the GST1 0 phenotype in patients with adenocarcinoma. *Carcinogenesis*, **12**, 25-28.
- Sturzeneker, R., Haddad, L. A., Bevilacqua, R.A.U., Simpson, A.J.G. and Pena, S.D.J. (1998). Polarity of mutations in tumour associated microsatellite instability. *Hum Genet.*, **102**, 231-235.
- Steele, R.J.C., Thompson, A.M., Hall, P.A. and Lane, D.P. (1998). The p53 tumour suppressor gene. *Br J Surg.*, **85**, 1460-1467.
- Steinmeyer, K. and Deppert, W. (1988). DNA binding properties of murine p53. *Oncogene*, **3**, 501-507.
- Sturzbecher, H.W., Brain, R., Addison, C., Rudge, K., Remm, M., Grimaldi, M., Keenan, E. and Jenkins, J.R. (1992). A C-terminal alpha-helix plus basic region motif is the major structural determinant of p53 tetramerization. *Oncogene*, **7**, 1513-1523.
- Su, S.S., and Modrich, P. (1986). *Escherichia coli mutS*-encoded protein binds to mismatched DNA base pairs. *Proc Nat Acad Sci.*, **83**, 5057-5061.
- Su, L.K., Johnson, K.A., Smith, K.J., Hill, D.E., Vogelstein, B. and Kinzler, K.W. (1993). Association between wild type and mutant APC gene products. *Cancer Res.*, **53**, 2728-2731.
- Sulekova, Z., Reina-sanchez, J., and Ballhausen, W.G. (1995). Multiple APC messenger RNA isoforms encoding exon 15 short open reading frames are expressed in the context of a novel exon 10A-derived sequence. *Int J cancer.*, **63**, 435-441.
- Swift, M., Morrell, D., Massey, R.B. and Chase, C.L. (1991). Incidence of cancer in 161 families affected by ataxia-telangiectasia. *N Engl J Med.*, **325**, 1831-1836.

- Talbot, C.C., Avamopoulos, D., Gerken, S., Chakravati, A., Armour, A.J., Matsunami, N., Withe, R. and Antonarakis, S.E. (1995). The tetranucleotide repeat polymorphism D21S1245 demonstrates hypermutability in germline and somatic cells. *Hum Mol Genet.*, **4**, 1193-1199.
- Thliveris, A., Albertsen, H., Tuohy, T., Carlson, M., Groden, J., Joslyn, G., Gelbert, L., Samowitz, W., Spirio, L. and White, R. (1996). Long-range physical map and deletion characterization of the 1100-kb NotI restriction fragment harbouring the APC gene. *Genomics*, **34**, 268-270.
- Tipping, A.J., Pearson, T., Morgan, N.V., Gibson, R.A., Kuyt, L.P., Havenga, C., Gluckman, E., Joenje, H., de Ravel, T., Jansen, S. and Mathew, C.G. (2001). Molecular and genealogical evidence for a founder effect in Fanconi anemia families of the Afrikaner population of South Africa. *Proc Natl Acad Sci.*, **98**, 5734-5739.
- Trimble, E.L., Karlan, B.Y., Lagasse, L.D. and Hoskins, W.J. (1991). Diagnosing the correct ovarian cancer syndrome. *Obstet Gynecol.*, **78**, 1023-1026.
- van der Luijt, R.B., Meera Khan, P., Vasen, H.F., Breukel, C., Tops, C.M., Scott, R.J. and Fodde, R. (1996). Germline mutations in the 3' part of APC exon 15 do not result in truncated proteins and are associated with attenuated adenomatous polyposis coli. *Hum Genet.*, **98**, 7277-7234.
- Varley, J. M. (2003). Germline TP53 mutations and Li-Fraumeni syndrome. *Hum. Mutat.*, **21**, 313-320.
- Vasen, H.F.A., Offerhaus, G.J.A., den Hartog Jager, F.C.A., Menko, F.H., Nagengast, F.M., Griffioen, G., van Hogezaand, R.B., and Heintz, A.P. (1990). The tumour spectrum in hereditary nonpolyposis colorectal cancer: a study of 24 kindreds in the Netherlands. *Int J Cancer.*, **46**, 31-34.
- Vasen, H.F.A., Mecklin, J-P., Meera-Khan, P., and Lynch, H.T. (1991). The International Collaborative Group on Hereditary Nonpolyposis Colorectal Cancer (ICG-HNPCC). *Dis Colon Rectum.*, **34**, 424-425.
- Vasen, H.F.A., Taal, B.G., Griffioen, G., Nagnenstat, F.M., Cats, A., Menko, F.H., Oskam, W., Kleibeuker, J.H., Offerhaus, G.J. and Khan, P.M. (1994). Clinical heterogeneity of familial colorectal cancer and its influence on screening protocols. *Gut*, **35**, 1262-1266.
- Vasen, H.F.A, Wijnen, J. Th, Menko, F.H., Kleibeuker, J.H., Taal, B.G., Griffioen, G., Nagengast, F.M., Meijers-Heijboer, E.H., Bertario, L.,

- Varesco, L. *et al.* (1996). Cancer risk in families with hereditary nonpolyposis colorectal cancer diagnosed by mutation analysis. *Gastroenterology*, **110**, 1020-1027.
- Vasen, H.F.A., van Ballegooijen, M., Buskens, E., Kleibeuker, J.K., Taal, B.G., Griffioen, G., Nagengast, F.M., Menko, F and Meera-Khan, P. (1998). A cost effectiveness analysis of colorectal screening of hereditary nonpolyposis colorectal carcinoma gene carriers. *Cancer*, **82**, 1632-1637.
- Vasen, H.F., Watson, P., Mecklin, J.P. and Lynch, H.T. (1999). New clinical criteria for hereditary nonpolyposis colorectal cancer (HNPCC, Lynch syndrome) proposed by the International Collaborative Group on HNPCC. *Gastroenterology*, **116**, 1453-1456.
- Vatsis, K.P., Martell, K.J. and Weber, W.W. (1991). Diverse point mutations in the human gene for polymorphic N-acetyltransferase. *Proc Natl Acad Sci*, **88**, 6333-6337.
- Vilkki, S., Tsao, J-L, Loukola, A., Poyhonen, M., Vierimaa, O., Herva, R., Aaltonen, L.A. and Shibata, D. (2001). Extensive somatic microsatellite mutations in normal human tissue. *Cancer Res.*, **61**, 4541-4544.
- Vogelstein, B. and Kinzler, K.W. (1993). The multistep nature of cancer. *Trends Genet.*, **9**, 138-141.
- Vogelstein, B., Lane, D and Levine, A.J. (2000). Surfing the p53 network. *Nature*, **408**, 307-310.
- Wagner, A., Barrows, A., Wijnen, J.T., van der Klift, H., Franken, P.F., Verkuijlen, P., Nakagawa, H., Geugien, M., Jaghmohan-Changur, S., Breukel, C., Meijers-Heijboer, H., Morreau, H., van Puijenbroek, M., Burn, J., Coronel, S., Kinarski, Y., Okimoto, R., Watson, P., Lynch, J.F., de la Chapelle, A., Lynch, H.T. and Fodde, R. (2003). Molecular analysis of hereditary nonpolyposis colorectal cancer in the United States: high mutation detection rate among clinically selected families and characterization of an American founder genomic deletion of the MSH2 gene. *Am J Hum Genet.* **72**, 1088-1100.
- Walén, R. and Boyer, T. (1998). Human Glutathione S-Transferases. *Seminars in Liver Disease*, **18**, 345-357.
- Walker, A.R.P., and Segal, I. (1988). Colorectal cancer-some aspects of epidemiology, risk factors, treatment, screening and survival. *S Afr Med J.*, **73**, 653-657.

- Wang, Q., Lasset, C., Desseigne, F., Frappaz, D., Bergeron, C., Navarro, C., Ruano, E and Puisieux, A. (1999). Neurofibromatosis and early onset of cancers in *hMLH1*-deficient children. *Cancer Res.*, **59**, 294-297.
- Warthin, A.S. (1913). Hereditary with reference to carcinoma: as shown by the study of the cases examined in the pathological laboratory of the University of Michigan, 1895-1913. *Arch Intern Med.*, **12**, 546-555.
- Weber, J.L., and May, P.E. (1989). Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *Am J Hum Genet.*, **44**, 388-396.
- Weber, J.L. and Wong, C. (1993). Mutation of human short tandem repeats. *Hum Mol Genet.*, **2**, 1123-1128.
- Weber, T. (1996). Clinical surveillance recommendations adopted for HNPCC. *Lancet*, **348**, 465.
- Weber, T.K., Chin, H.M., Rodriguez-Bigas, M., Keitz, B., Gilligan, R., O'Malley, L., Urf, E., Diba, N., Pazik, J. and Petrelli, N.J. (1999). Novel *hMLH1* and *hMSH2* germline mutations in African Americans with colorectal cancer. *JAMA.*, **281**, 2316-2320.
- Welfare, M., Adeokun Monesola, A., Bassendine, M.F. and Daly, A.K. (1999). Polymorphisms in *GSTP1*, *GSTM1* and *GSTT1* and susceptibility to Colorectal Cancer. *Cancer Epidemiol Biomarkers Prev.*, **8**, 289-292.
- Welsh, K.M., Lu, A.L., Clark, S. and Modrich, P. (1987). Isolation and characterization of the *Escherichia coli mutH* Gene product. *J Biol Chem.*, **262**, 15624-15629.
- Wheeler, J.M., Loukola, A., Aaltonen, L.A., Mortensen, N.J. and Bodmer, W.F. (2000). The role of hypermethylation of the *hMLH1* promoter region in HNPCC versus MSI+ sporadic colorectal cancers. *J Med Genet.*, **37**, 588-592.
- Wijnen, J., Khan, P.M., Vasen, H., Menko, F., van der Klift, H., van den Broek, M., van Leeuwen-Cornelisse, I., Nagengast, F., Meijers-Heijboer, E.J., Lindhout, D. *et al.* (1996). Majority of *hMLH1* mutations responsible for hereditary nonpolyposis colorectal cancer cluster at the exonic region 15-16. *Am J Hum Genet.*, **58**, 300-307.
- Windmill, K.F., Gaedigk, A., de la M. Hall, P., Samaratunga, H., Grant, D.M. and McManus, M. (2000). Localisation of N-acetyltransferases NAT1 and NAT2 in human tissues. *Toxicol Sci.*, **54**, 19-29.

- Wu, Y., Berends, M.J.W., Mensink, R.G.J., Kempinga, C., Sijmons, R.H., van der Zee, A.G.J., Hollema, H., Kleibeuker, J.H., Buys, C.H. and Hofstra, R.M.W. (1999). Association of hereditary nonpolyposis colorectal cancer-related tumors displaying low microsatellite instability with MSH6 germline mutations. *Am J Hum Genet*, **65**, 1291-1298.
- Wu, Y., Berends, M.J.W., Sijmons, R.H., Mensink, R.G.J., Verlind, E., Kooi, K.A., van der Sluis, T., Kempinga, C., van der zee, A.G.J., hellema, H., Buys, C.H., Kelibeuker, J.H. and Hofstra, R.M. (2001). A role for MLH3 in hereditary nonpolyposis colorectal cancer. *Nat Genet.*, **29**, 137-138.
- Xia, L., St Denis, K.A. and Bapat, B. (1995). Evidence for a novel exon in the coding region of the adenomatous polyposis coli (APC) gene. *Genomics*, **28**, 589-591.
- Yagi, O.K., Akiyama, Y., Nomizu, T., Iwama, T., Endo, M. and Yuasa, Y. (1998). Proapoptotic gene BAX is frequently mutated in hereditary nonpolyposis colorectal cancers but not in adenomas. *Gastroenterology*, **114**, 268-274.
- Yamada, M., O'Regan, E., Brown, R., Karran, P (1997). Selective recognition of cis-platin DNA adduct by human mismatch repair proteins. *Nucleic Acids Res.*, **25**, 491-495.
- Yamamoto, H., Sawai, H., Weber, T.K., Rodriguez-Bigas, M.A. and Perucho, M. (1998). Somatic frameshift mutations in DNA mismatch repair and proapoptosis genes in hereditary nonpolyposis colorectal cancer. *Cancer Res.*, **58**, 997-1003.
- Yoshida, T., Kanegane, H., Osato, M., Yanagida, M., Miyawaki, T., Ito, Y. and Shigesada K. (2003). Functional analysis of RUNX2 mutations in cleidocranial dysplasia: novel insights into genotype-phenotype correlations. *Blood Cells Mol Dis.*, **30**, 184-193.
- Young, G.P., Rozen, P. and Levin, B. (eds). (1996). Prevention and early detection of colorectal cancer. Bath Press, UK.
- Zhang, H., Ahmad, A., Arbman, G., Zdolsek, J., Carstensen, J., Nordenskjold, B., Soderkvist, P. and Sun, X-F. (1999). Glutathione S-transferase T1 and M1 genotypes in normal mucosa, transitional mucosa and colorectal adenocarcinoma. *Int J Cancer.*, **84**, 135-138.
- Zhao, H., Bailey, L.A., Elsas, L.J., Grinzaid, K.A. and Grabowski, G.A (2003). Gaucher disease: in vivo evidence for allele dose leading to neuronopathic and nonneuronopathic phenotypes. *Am J Med Genet.*, **116A**, 52-56.

Zhong, S., Wyllie, A.H., Barnes, D., Wolf, C.R. and Spurr, N.K. (1993). Relationship between the GSTM1 genetic polymorphism and susceptibility to bladder, breast and colon cancer. *Carcinogenesis*, **14**, 1821-1824.

University of Cape Town

APPENDICES

University of Cape Town

APPENDICES

University of Cape Town

APPENDIX A

- A.1. Division of Human Genetics request form for biological specimens.
- A.2. Information brochure for individuals enrolled in the HNPCC counseling programme.
- A.3. Pedigree of NPC 145, displaying haplotypes for the *hMSH2* gene

University of Cape Town

A.1. DNA Consent form

University of Cape Town



REQUEST FOR MOLECULAR STUDIES (DNA)



Molecular Laboratory
Department of Human Genetics
1st Floor, Anatomy Building
UCT Medical School, Observatory 7925

Tel: (021) 406 6425 Fax: (021) 448-0906

Blood should be drawn in 2 plastic EDTA Tubes
(Purple top) +/- 10ml each using a yellow barrel.
Each tube should be inverted to mix and should be
clearly labelled with the patient's name and DOB
Keep blood in fridge at 4°C until able to send to laboratory

Please **DO NOT** send specimens on ice or frozen.

Please fill in all the information requested:

Surname: _____ First Name(s): _____

New Family: Yes ☐ No ☐ Please fill in Family name / Lab Nr: _____

Medical aid: _____ Medical aid No: _____

Sex: M ☐ F ☐ Date of Birth: Year: _____ Month: _____ Day: _____

Ethnic Origin: Black ☐ Indian ☐ Mixed ancestry ☐ Caucasian ☐ Other ☐ _____

Contact Address: _____ Fax: _____

Tel: _____

Cell phone: _____

E-mail: _____

Referring Doctor/Sister: _____

Hospital or Address: _____ Fax: _____
Tel: _____

Bloods taken for: Research HNPCC under 45 ☐ Research HNPCC Family History ☐ Research FAP ☐

Predictive HNPCC ☐ Predictive FAP ☐

Confirmation ☐

Forward my predictive test results to: _____ Address: _____

Disease Status
Affected ☐ At Risk ☐ Spouse ☐ Unknown ☐

For Laboratory use only:

DNA number: _____ Vol. Blood: _____ (ml) Other: _____

Date Received: Year: _____ Month: _____ Day: _____ Computer Index No: _____

CONSENT FOR DNA ANALYSIS AND STORAGE

I, Mr / Ms / Mrs / _____
request that an attempt be made to assess the probability that I, my children or my unborn child might have inherited the gene for:

I understand that the DNA for analysis is to be obtained from:

a) blood cells ☐ b) skin sample ☐ c) other ☐ _____

The methods and risks of obtaining the specimens have been explained to me.

I also understand that:

1. The test procedure is specific to the genetic condition mentioned above and cannot determine the complete genetic makeup of an individual.
2. Lack of co-operation by key relatives in providing blood samples may decrease the accuracy of the test result for linkage analysis and/or ability to perform the test.
3. Sometimes the genetic pattern in a family renders the test results "uninformative" (not clarifying the point in question).
4. Even under the best conditions, current technology of this type is imperfect and may lead to incorrect results.
5. I may withdraw my consent at any time without giving a reason and without this affecting my future medical care.
6. The DNA bank is under an obligation to respect medical confidentiality.
7. The DNA bank may request clinical information from my Doctor(s).

- A portion of the DNA obtained may be stored to enable later testing. The following conditions apply to my DNA that is being stored: (circle any and all that you wish).

A) To be used for linkage analysis for members of my immediate family.

B) To be used for research purposes without my or my family member's knowledge, at the discretion of the Head of Human Genetics, UCT, provided that any information from such research will remain confidential.

C) To be used without any restrictions.

D) Other: _____

I understand that I may change my mind and withdraw my consent for DNA banking at any time.

ALL OF THE ABOVE HAS BEEN EXPLAINED TO ME IN A LANGUAGE THAT I UNDERSTAND AND MY QUESTIONS

ANSWERED BY: _____ DATE: _____

Patient signature _____

Witnessed consent _____

PEDIGREE:

HNPCCFORM

HNPCCFORM

A.2. Information brochures

University of Cape Town

GENERAL INFORMATION ON THE HNPCC PREDICTIVE GENETIC TESTING PROGRAMME

1. Cost for the research leading to the finding of the genetic defect predisposing to colorectal cancer has been borne by the Division of Human Genetics. The (confirmation of diagnosis / predictive) laboratory test is available to members in families where the pre-genetic defect has been detected at an approximate cost of R596.40, which is generally payable through a medical aid scheme. Testing will not be denied to individuals who are unable to afford the cost.
2. We strongly recommend that you inform your family doctor of your decision to undertake the test. If you do not have a family doctor we recommend you find one. As part of the policy of this programme, we believe that the on-going medical care and support your doctor is able to give is very important to you and your family. A letter will be sent to your doctor (with your permission) to inform him/her about the programme after the 2 / 3rd meeting.
3. Support is essential and we therefore strongly advise that you choose a family member or a trusted friend to accompany you to all the meetings. (optional)
4. The final results will be given to you approximately 1 month after your blood samples have been taken and will be strictly confidential. **No results will be given to you by telephone.** With your permission your family doctor will be contacted and written to regarding the results.
5. If at any stage in the programme you decide you do not wish to continue, the decision is entirely yours. Your decision will in no way prejudice our relationship with you or your family. We will be happy to continue to offer you the support and help you need, within our capabilities.
6. Reactions to predictive genetic testing might vary widely. Some people who might have the predisposing genetic defect may suffer a sense of shock and grief however well they may have been prepared beforehand. People whose test is negative may feel relief but at the same time suffer guilt and anxiety. A health care professional will be available to discuss any questions / problems you might encounter after entering into the HNPCC predictive genetic testing programme.
7. No children under the age of 18 will be included in the HNPCC predictive genetic test programme.
8. Those individuals who carry the mutation for HNPCC will be counselled with regard to best practice regarding regular colonoscopic and other relevant screenings for cancer – in private or provincial settings.

CONTACT PERSONS: Sister Elize Pietersen 021-4066373
Sister Ursula Algar 021-4045499



INFORMATION AND PROTOCOL FOR PREDICTIVE TESTING
HEREDITARY NON-POLYPOID COLORECTAL CANCER (HNPCC)

Genes provide the information for growth, development and functions of our bodies. When a gene is altered, called a mutation, an incorrect message is produced. A mutated or changed gene can cause problems with development or functioning of different body systems or organs. Our genes are inherited from our parents who inherit their genes from their parents and so on. When the inheritance pattern from generation to generation is known, it is possible to provide families and individuals with information regarding the chance or risk of a particular condition and how this will affect themselves, family members or future generations.

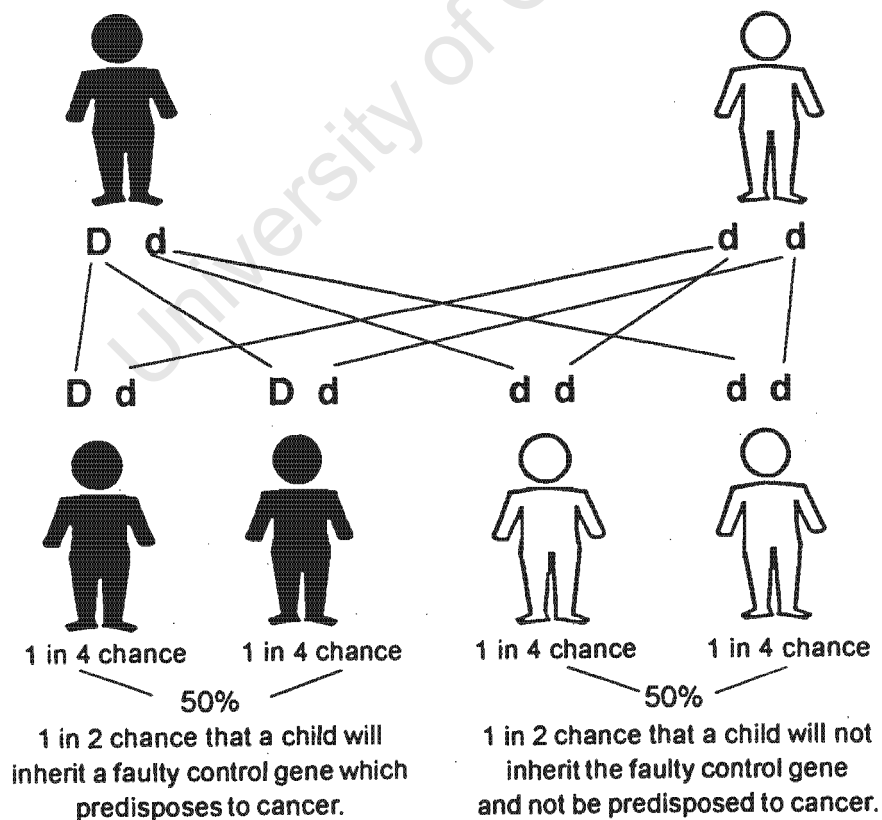
AUTOSOMAL DOMINANT INHERITANCE PATTERN

HNPCC has a dominant inheritance pattern. (Picture) A dominant inheritance pattern means that the disorder is transmitted from a parent to the next generation. The child will either receive the correct (d – in picture) or mutant copy (D – in picture) of the gene from the affected parent. It is thus clear from the picture that there is a 1 in 2 chance (50%) in *every* pregnancy that the DOMINANT mutant gene will be transmitted to the child by the affected parent. The other parent can only pass on the correct copies of the gene. The reverse of this also applies. There is also a 50% chance in *every* pregnancy that the child will inherit the CORRECT copy of the gene. In this case the child will not be affected with the disorder and cannot transmit the disorder to any of his/her children. Men and women are equally affected.

INHERITANCE OF A FAULTY (MUTATED) CELL DIVISION CONTROL GENE WHICH PREDISPOSES TO CANCER.

Parent predisposed to cancer. Has a faulty control gene copy in the germ cell.

Parent not predisposed to cancer. Has correct control gene copies in the germ cell.



GENETIC TESTING

Genetic testing is an examination of the DNA (basic material of hereditary) of an individual. Results of these tests may disprove or confirm a suspected fault or change (mutation) in the DNA. Genetic testing of DNA is performed on a blood sample collected from an individual.

Predictive genetic testing is a means of knowing one's genetic status with regard a particular condition. To undergo a predictive genetic test implies that one is forewarned about one's risk of developing a particular disorder before the signs and symptoms of that condition manifest it in an individual

HNPCC is a dominant inherited disorder with the result that all family members of an affected individual (first degree relative) are at 50% risk of developing cancer at an early age. Predictive genetic test for HNPCC is a relatively recent option available to individuals at risk for having inherited the gene for HNPCC. This test allow an individual a chance of knowing whether he / she has the mutation. Knowledge of one's genetic status, with regard to HNPCC, allows the individual to make informed decisions about commencing preventative screening for cancer.

To take the test is a very serious decision. Therefore, it is important that subjects are well informed and understand the programme and procedures of predictive genetic testing that are necessary before finally getting the result. If after careful consideration, you decide to take the test, you will be requested to come to the Department of Human Genetics on at least 2 occasions to see the geneticist involved in running the programme. Arrangements can be made to link you with the clinical team for screening if needed.

PROGRAMME FOR PREDICTIVE TESTING - HNPCC

1. Telephonic contact with family members re interest in predictive genetic testing. Setting up date(s) for meeting with geneticist.

2. **1st Meeting with geneticists:** (Group session) Information and implication of predictive genetic testing.

The outcome of this meeting can be:

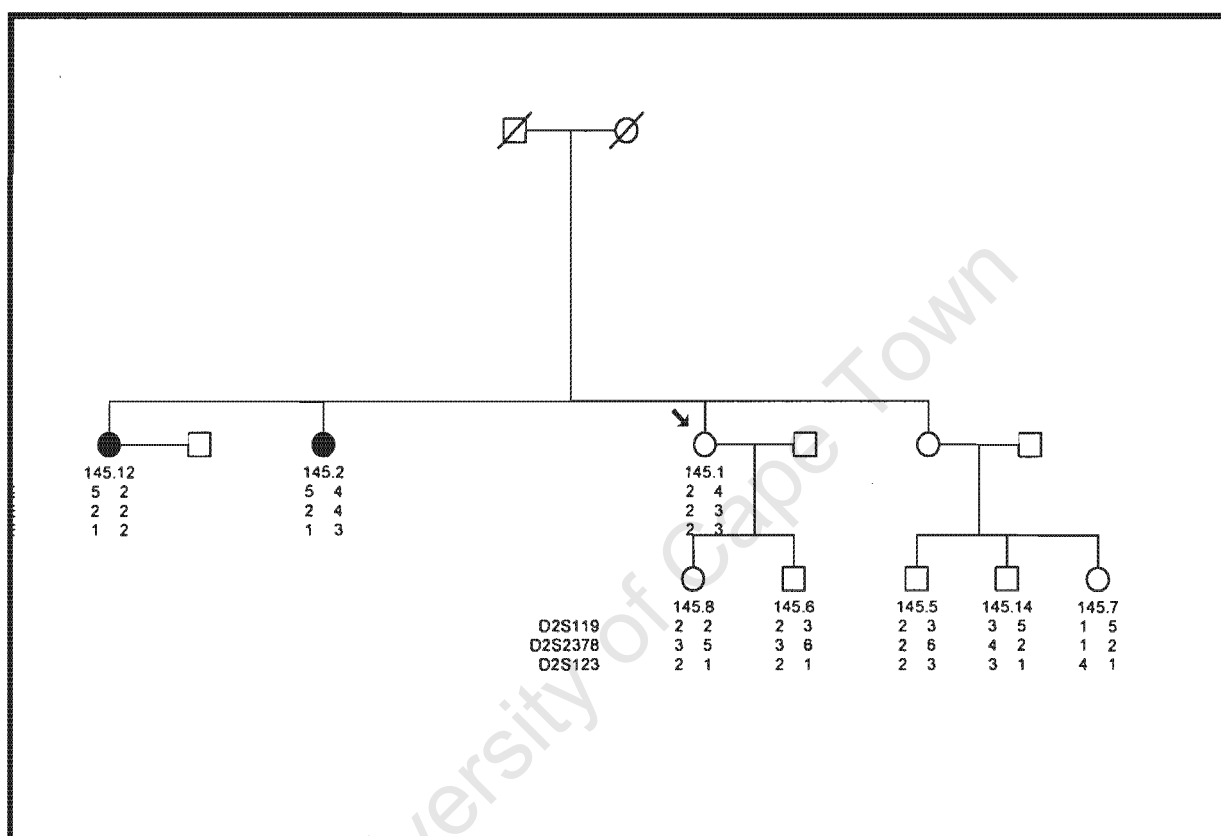
- You need time to assimilate what you have heard. The registered nurses will make contact with you within a week to discuss your decision.
- You have opted to know your genetic status on HNPCC -- 1st Blood samples collected, with informed consent.

3. **2nd Meeting with geneticist:**

- **Those who have opted to know their genetic status during the 1st meeting:** meet individually with the geneticist and receive their genetic information. They have the opportunity of asking further questions pertinent to their specific situation. A second blood sample will be collected for confirmatory analysis. Contact numbers of the support team will be given. Plan of action with regard to preventative screening will be discussed. Appointments (support team + screening) can be arranged.
- **Those who have decided to know their genetic status subsequent to the 1st meeting:** meet individually with geneticist to clear up any uncertainties. Meet with registered nurse to take 1st blood specimen.

4. **3rd Meeting with geneticist:**

- **Those who have opted to know their genetic status during the 2nd meeting:** meet individually with the geneticist and receive their genetic information. They have the opportunity of asking further questions pertinent to their specific situation. A second blood sample will be collected for confirmatory analysis. Contact numbers of the support team will be given. Plan of action with regard to preventative screening will be discussed. Appointments (support team + screening) can be arranged.

A.3. Pedigree of NPC 145

Haplotype analysis of NPC 145 using microsatellite markers in and around the *hMSH2* gene.

APPENDIX B

Techniques:

- B.1. DNA isolation from lymphocytes
- B.2. DNA isolation from paraffin-embedded tumour tissue
- B.3. DNA isolation from fresh tumour tissue
- B.4. Single strand conformation polymorphism (SSCP) analysis
- B.5. DNA sequencing
- B.6. Centri-Sep Spin Column Clean-out procedure
- B.7. DNA oligonucleotides
- B.8. ARMS PCR conditions
- B.9. PCR amplification for GSTT1 and GSTM1
- B.10. PCR amplification for NAT2
- B.11. Restriction digest analysis of NAT2

B.1. DNA Extraction from Lymphocytes

- Up to 10ml of EDTA anti-coagulated blood can be processed when it has been frozen for more than 2hrs at -20°C .
- After complete thawing of the blood at 37°C , 2 volumes of frozen blood washing solution (BWS) (not modified) was added and mixed thoroughly and centrifuged at 10 000 revolutions per minute (rpm) (Beckman J2-21, *Beckman*, USA) for 20 minutes (min) to spin down nuclei and cell debris (Step A).
- The supernatant was discarded without disturbing the pellet.
- Step 'A' was repeated using 7ml BWS. The supernatant was discarded.
- The nuclei were resuspended by vortexing in 1ml of BWS and 1ml of dH_2O . Four ml of lysis solution was added and the mixture was incubated at 68°C for 10min.
- The sample was transferred to a gel barrier tube, 6.5ml of chloroform was added and the 2 phases were mixed by inverting the tube vigorously.
- The sample was centrifuged for 25min at 8000rpm at 10°C and the upper, cleared, aqueous phase was poured into a DEPC treated 30ml Corex tube.
- 8ml of precipitating solution was added to the aqueous phase and the tube was mixed until a filamentous precipitate of DNA formed.
- A cut-off Pasteur pipette was used to remove the DNA clump and this was transferred to a clean 30ml Corex tube.
- The sediment was resuspended in 4ml of Ionic Exchange Solution and mixed. The tubes were shaken gently to completely resuspend any sediment and left overnight at room temperature.

- 8ml of ethanol was added and mixed until a precipitate was visible. A cut-off Pasteur pipette was used to remove the DNA and a 2 X 70% ethanol wash was performed. The DNA was transferred to a 1.5ml microfuge tube with a safelock lid (*Eppendorf*, Germany) and centrifuged for 2min at 13 000rpm. The washings were removed and the tube was inverted on filter paper to drain all liquid.
- The pellet was resuspended in 200 μ l of sterile, distilled water (dH₂O).
- The concentration of the DNA was calculated on a spectrophotometer at A₂₆₀ and A₂₈₀. The reading was taken at an absorbance of 260nm, which allowed for the calculation of nucleic acid in the sample. The ratio of the readings at 260nm and 280nm $\left(\frac{OD_{260}}{OD_{280}}\right)$ provided an estimate of the purity of the DNA. Pure preparations of DNA had an OD₂₆₀/OD₂₈₀ of approximately 1.8. If there was contamination of the DNA sample with protein or phenol, the OD₂₆₀/OD₂₈₀ values would be less and an accurate quantitation of the amount of nucleic acid was not possible (Sambrook *et al.*, 1989).

All gloves, pipette tips, blood tubes, paper towel and microfuge tubes that were contaminated with blood were placed in a biohazard bag that was autoclaved when filled. This bag was placed into a special waste disposal box that was collected by a waste disposal company (*Wastetech*, South Africa). Care was taken not to contaminate equipment, clothing and ones laboratory coat with blood or its components.

B.2. DNA Isolation from Paraffin-embedded tumour tissue:

- DNA was extracted from unstained 10µm paraffin embedded tissue sections
- Using a microtome, 10µm serial sections were cut off the paraffin block. This was placed onto a glass slide and the tumour tissue was demarcated on the slide
- To dissolve the paraffin, the unstained sections were transferred from the glass slides to 1.5ml Safelock tubes containing 100µl Xylene (*Merck*, South Africa)
- 100µl of 100% ethanol (*Merck*, South Africa) was added to facilitate pelleting
- The samples were gently mixed and spun on a countertop Sigma-201M microcentrifuge (*Labaratory and Scientific Equipment Company*, South Africa) at 12 000rpm for 10min.
- The xylene-ethanol mixture was decanted and the tissue pellets were dried at 55°C in a hybridisation oven (*Bachofer*, Germany).
- The pellet was resuspended in 250µl of 50mM Tris, pH8.3 (*ICN Biochemicals*, USA) containing 200ng/µl proteinase K (*Boehringer Mannheim*, Germany).
- Following overnight incubation at 37°C, the tubes were immersed in boiling water for 8min. This facilitated proteinase K inactivation and denaturation of DNA.
- The extracted DNA was placed on ice for 5mins and used for PCR, without any further purification.

B.3. DNA isolation from Fresh Tumour Tissue

- 200 μ l of tissue lysis buffer [4M urea, 200mM Tris, 100mM NaCl, 200mM EDTA, pH 7.4 (25°C)] and 40 μ l *Proteinase K* (20mg/ml) was added to 25-50mg of tumour tissue sample and incubated for 1hour at 55°C. The tumour sample was cut into tiny pieces with a disposable scalpel to increase the yield of nucleic acids.
- 200 μ l binding buffer [6M guanidine-HCl, 10mM urea, 10MmTris-HCl, 20% Triton X-100 (v/v), pH8.4 (25°C)] was added, mixed and incubated for 10min at 72°C in a water bath.
- 100 μ l isopropanol (*AnalaR BDH*, UK) was added. The *High Pure* filter tube was combined with a collection tube and the sample was pipetted into the upper reservoir.
- The sample was centrifuged for 1min at 8000rpm in a standard table top centrifuge.
- The flow-through was discarded and combined to the filter tube and the used collection tube.
- 500 μ l wash buffer [20mM NaCl, 2mM Tris-HCl, pH 7.5, (25°C)] was added to the upper reservoir and centrifuged for 1min at 8000rpm.
- The flow-through was discarded and the filter tube was again combined to the used collection tube. Five hundred μ l washing buffer [20mM NaCl and 2mM Tris-HCl, pH7.5 (25°C)] was added to the upper reservoir and centrifuged for 1 min at 8000 rpm. Finally, the sample was centrifuged for 10s at 13000rpm to remove residual wash buffer.

- The collection tube was discarded and the filter tube was inserted into a clean 1.5ml microfuge reaction tube with a safelock lid.
- 200µl of prewarmed (70°C) elution buffer (10mM Tris, pH 8.5) was added for the elution of the nucleic acids. Elution buffer was added to the filter tube and centrifuged for 1min at 8000rpm.

The nucleic acids are stable and could be used directly or stored at 4°C for later analysis.

B.4. Single Stranded Conformational Polymorphism Analysis

The exonic sequences of genes were PCR amplified in 200-300 bp fragments. An equal amount of SSCP loading dye was added and the samples were loaded onto acrylamide gels.

B.4.1.Preparation of Electrophoresis Plates

The glass plates were cleaned with ethanol. The back plate was wiped with acetone to avoid the gel from adhering to it. The front plate was coated with 250 µl plate glue (500 µl tMTMS(t-methacryloxypropyltrimethoxysilane), 50ml absolute ethanol and 22.5 µl 10% acetic acid and left to dry for 3min. The plate was then bathed twice in 100% ethanol to remove all excess plate glue and assembled using 0.375mm spacers and a 0.375 mm well-forming comb. The glass plates used were 380 cm long and 300cm wide. Electrophoresis took place on a standard sequencing system purchased from *Omeg Scientific* (South Africa).

B.4.2.Electrophoresis Conditions

PCR products were denatures at 95°C for 5 min, chilled on ice for 5 min and loaded onto polyacrylamide gels. The gels were electrophoresed

overnight at 2 to 10 W (depending on fragment size). Two experimental conditions were used for mutation analysis (PAGE+10% Urea, RT and PAGE+5%glycerol, RT).

B.4.3. Visualisation

Gels were stained over two stages:

- i. 0.1% AgNO₃ solution for 20 min
- ii. 1.5% NaOH/ 1.5% formaldehyde for 10 min

The gels were washed in distilled water between each solution. The banding pattern was noted and the gels were transferred onto chromatography paper.

B.4.4. SSCP Gel Matrix

40% Stock solution of polyacrylamide (1% crosslinking):

99% Acrylamide

1% Bisacrylmide

10% Urea gel:

7.5% Urea

10% polyacrylamide (from the 40% stock)

1.5X TBE

dH₂O to 60 ml

600 µl 10% ammonium persulphate (APS) [BDH, UK]

60 µl N'N'N'N'-tetramethylethylenediamine (TEMED) [BDH, UK]

A 1.5XTBE was used as a running buffer

5% Glycerol gel

5% Glycerol

0.5X TBE

10% polyacrylamide solution (from the 40% stock)

dH₂O to 60 ml

600 µl 10% ammonium persulphate (APS) [BDH, UK]

60 µl N'N'N'N'-tetramethylethylenediamine (TEMED) [BDH, UK]

A 0.5XTBE was used as a running buffer

10X TBE: 0.89 M Tris/ 0.89 M Boric Acid/ 0.02 M EDTA pH8.0

B.5. DNA SEQUENCING

Direct DNA sequencing was performed by one of three different methods:

- The Sequenase II sequencing Kit (Amersham, USA)
- The SequiTherm EXCEL II DNA Sequencing Kit (Epicentre Technologies, USA)
- ABI PRISM Dye Primer Cycle Sequencing Ready Reaction Kit (Perkin-Elmer, USA).

In each case, the protocols were followed according to the manufacturer's instructions. In the first two methods, [$\alpha^{32}\text{P}$] dCTP was incorporated into the sequencing reactions, which were subsequently electrophoresed on 6% denaturing polyacrylamide gels at 60 W in 1X TBE. After the run, the gels were dried at 80°C and autoradiographed at -70°C. For the ABI Kit, the sequencing reactions were separated by automated capillary electrophoresis and analysed by an ABI PRISM Genetic Analyser (Perkin Elmer, USA).

B.6. Centri-Sep Spin Column Clean-out.

B.6.1. Column hydration:

- The column was tapped insure that the dry gel was settled in the bottom of the spin column
- The column cap was removed and reconstituted by adding 0.80 ml of reagent grade water or buffer. The column end stopper was left in place so the column could stand up by itself. The column cap was replaced and the gel was hydrated by sheking and inverting the column or vortexing briefly

- At least 30 min of room temperature hydration was allowed before using the columns. Reconstituted columns may be stored refrigerated at 4°C for several days. Longer storage may be accomplished in 10mM sodium azide

B.6.2. Removal of Interstitial Fluid

- Air bubbles were removed from the column gel by inverting the column and the gel to slurry was allowed to settle to the opposite end of the column. The column up was placed upright and the gel was allowed to settle in a microtube rack
- After the gel has settled and is free of bubbles, first remove the top column cap, and then remove the column end stopper from the bottom
- The excess column fluid drained away (gravity) into a WASH TUBE (2ml). Approximately 200-250ul was drained from the column. This fluid was discarded
- The column was centrifuged and the tube was washed in a variable speed centrifuge at 750 X g for 2 min to remove interstitial fluid
- Approximately 300ul of fluid was removed. Any drops at the end of the column was blotted dry.

B.6.3. Sample processing

- 20 µl of DyeDeoxy™ terminator reaction mixture was added to the top of the gel. The sample must be dispensed directly onto the centre of the gel bed at the top of the column, without disturbing the gel surface

- The column was placed into the sample collection tube (1.5ml) and placed into the rotor. The column and collection tube were centrifuged at 750 X g for 2 min. The purified sample was collected at the bottom of the sample collection tube. The spin column was discarded. Proceed with the ABI sample preparation procedure
- The sample was dried in a vacuum centrifuge, no heat was applied.

B.7. DNA OLIGONUCLEOTIDES

All DNA oligonucleotides used in this study as primers for PCR were obtained from the following sources:

- Research Genetics (USA)
- Department of Biochemistry at University of Cape Town, South Africa
- Imperial Cancer Research Fund, London, United Kingdom.

B.8. ARMS PCR conditions

All PCR were performed with 25 ng of genomic DNA template in a final reaction volume of 15 μ l. The reaction mixture consisted of

- 2 X buffer (67 mM Tris-HCl (pH8.4), 16 mM Ammonium sulphate, 2 mM magnesium Chloride, 0.01%(vol/vol) Tween 20)
- 200 μ M of each dNTP
- 1 μ M of each primer
- 0.5 units Amplitag Gold (*Perkin-Elmer*, UK)

Thermal cycling was performed in a MJ Research machine PTC-225 (Cambridge, MA) by using 96-well thermo-fast plates (*Advanced Biotechnologies*, UK). Each well was sealed with a layer of mineral oil and the plate was sealed with an adhesive sealing sheet to prevent evaporation and cross-contamination.

The cycling conditions were as follows: One cycle of 95°C for 15 minutes, followed by 5 cycles of 95°C for 25sec, 70°C for 45 sec and 72°C for 30 sec; followed by 27 cycles of 95°C for 25 sec, 50-65°C for 45sec and 72°C for 30sec; followed by a final extension of 72°C for 10 min.

B.9. PCR amplification for GSTT1 and GSTM1

All PCR were performed with 200 ng of genomic DNA template in a final reaction volume of 25 µl. The reaction mixture for *GSTT1* and *GSTM1* consisted of:

- 10 µM GSTT1F or GSTM1R (Table 3.8)
- 10 µM GSTT1R or GSTM1R (Table 3.8)
- 10 µM β-globinF (Table 3.8)
- 10 µM β-globinR (Table 3.8)
- 200 µM dNTPs (*Gibco BRL*, UK)
- 10 x Promega B+ buffer (*Promega*, UK)
- 15 mM MgCl₂
- 0.1Units picTaq
- dH₂O to 25 µl

The reactions were subjected to a denaturation step at 95°C and processed through 30 cycles of 95°C for 30 seconds, 60°C for 30 seconds (*GSTT1*)/52°C for 30 seconds (*GSTM1*) and 72°C for 40 seconds, followed by a final cycle of 72°C for 5mins.

B.10. PCR amplification for NAT2

All PCR were performed with 25 ng of genomic DNA template in a final reaction volume of 50 μ l. The reaction mixture consisted of:

- 10 μ M NAT2F (Table 3.9)
- 10 μ M NAT2R (Table 3.9)
- 8 mM dNTPs (*Gibco*, UK)
- 10 x buffer (100 mM Tris/HCl pH 8.4, 500 mM KCl)
- 15 mM $MgCl_2$
- 0.2 Units Amplitaq Gold (*Perkin Elmer*, United Kingdom)
- dH_2O to 50 μ l

The reactions were subjected to a denaturation step at 95°C for 4 mins and processed through 30 cycles of 95°C for 30 seconds, 60°C for 30 seconds and 72°C for 40 seconds, followed by a final cycle of 72°C for 5mins.

B.11. Restriction digest analysis of NAT2

10 units Restriction endonuclease

10 x PCR buffer

8ul PCR product

Incubate at the recommended temperature for 2 hours.

APPENDIX C

Media, Buffers and Solutions

- C.1. 10X TBE
- C.2. Formamide loading dye
- C.3. 6% denaturing polyacrylamide gels
- C.4. SSCP loading dye

University of Cape Town

C.1. TBE (10X)

0.89 M Tris

0.89 M Boric Acid

0.02 M EDTA pH 8.0

C.2. Formamide Loading Dye

80% formamide

10 mM EDTA (pH 8.0)

0.1% xylene cyanol

0.1% bromophenol blue

C.3. 6% denaturing polyacrylamide gels

5.7% acrylamide

0.3% bis-acrylamide

7.65 M urea

1X TBE

C.4. SSCP loading dye

95% formamide

20 mM EDTA (pH 8.0)

10 mM NaOH

0.02 % xylene cyanol

0.02 % bromophenol blue

APPENDIX D

Tables of primer sequences used for PCR amplification

University of Cape Town

Table D.1. Oligonucleotide sequences for hMSH2 for the purpose of PCR based mutation detection.

PRIMER	SEQUENCE (5' -> 3')	SIZE OF PCR FRAGMENT (bp)
1F	CTTCAACCAGGAGGTGAGGAGGT	301
1R	GAAAGGAGCCGCGCCACAAG	
2F	ATGTAATATCTCAAATCTGTAATGT	227
2R	ATAAGTAAATTAAAAAGGAAGATAA	
3F	GCTTATAAAATTTTAAAGTATGTTT	359
3R	GCCTTTCCTAGGCCTGGAATCTCC	
4F	TTCATTTTTTGCTTTTCTTATTCC	224
4R	ATATGACAGAAATATCCTTC	
5F	CCAGTGGTATAGAAATCTTCG	252
5R	CCAATCAACATTTTAAACCC	
6F	ATTGAGCTTGCCATTCTTTCTA	220
6R	TGCAGGTTACATAAACTAACGAA	
7F	AATATTTTACATTAAATCAAGTTAA	284
7R	ATATATTGTATGAGTTGAAGGAA	
8F	AATGAGATCTTTTATTTGTTTGTT	182
8R	ACTGCTTAAATTAAAAAGTATATTG	
9F	GGATTTTGTGAGTTTGTCTGTTTG	179
9R	TTCCAACCTCCAATGACCCATTC	
10F	ATACTTTTCTTTTCTTCTTGATTA	210
10R	GGTTAAAAATATAATAACGACTTG	
11F	TTAATAAACTGTTATTTTCGATTTG	169
11R	AGCCAGGTGACATTCAGAACATTAT	
12F	TTTCTGTTTTTATTTTTATACACG	309
12R	AAACGTTACCCCCACAAAG	
13F	AAACTTGCTTTCTGATATAATTTGT	273
13R	CATTTCTATCTTCAAGGGACTAGGA	
14F	TTCAGGTAATTATGTGCTTCA	302
14R	TACTGAATTTAGAGTACTCCAATAGT	
15F	TCTCATGCTGCTCCCTCACG	247
15R	TAAGTTAACTATGAAAACAACTG	
16F	ATTTTAATTACTAATGGGACATT	246
16R	GCTTATCAATATTACCTTCATTC	

(Primer sequences taken from Kolodner et al., 1994; Han et al., 1995; Beck et al., 1997)

Table D.2. Oligonucleotide sequences for hMLH1 for the purpose of PCR-based mutation detection.

PRIMER	SEQUENCE (5' -> 3')	SIZE OF PCR FRAGMENT(bp)
1F 1R	GACGTTTCCTTGGCTCTTCTG CCGTTAAGTCGTAGCCCTTAAGT	193
2F 2R	TGATTTGCCAGTTTAGATGCA TGACTCTTCCATGAAGCGC	154
3F 3R	GAGATTTGGAAAAATGAGTAACA CAACAGGAGGATATTTTACACA	200
4F* 4R*	TGACCCAGCAGTGAGTTTTTC* ACTCTGAGACCTAGGCAAAAA*	191
5F 5R	TTCTCTTTTCCCCTTGGGAT ATTCTGAACGATCTGGGTCC	155
6F 6R	GACATCTTGGGTTTTATTTTCA GACAAATCTCAGAGACCCAC	200
7F 7R	TAGTGTGTGTTTTTGGCAACTC CATCCCCCATAAACCAAGAAC	120
8F 8R	CCTTGTGTCTTCTGCTGTTTG GATTTTTTTATATAGGTTATCGACA	146
9F 9R	GCTTCAGAATCTCTTTTCTA GTGGATTTCCCATGTGGTTC	169
10F 10R	CCCTCAGACAGTTTTGAACT GGGATTCACCTCTGATTGTACT	349
11F 11R	TCTAAGGTAATTGTTCTCTCTTA AAGTAGCTGGATGAGAAGCG	216
12AF* 12AR*	TTTGGGACCTGTATATCTATACTCT* CTGCAGAAATGCATCAAGC*	249
12BF* 12BR*	AACAGAAGCTTGATGCATTTTC* TTTCCCAAAGCCATACCT*	257
13F 13R	GGTTCATTCACAGCTCTGTAG CACAGCGTTTAGTACCCTCA	292
14F 14R	AAGTGGGGTTGGTAGGATTCT CTCCCTGGACCATGTTGTAG	225
15F 15R	ATTTGTCCCAACTGGTTGTATCTC ACTATACAATACAGCAACTATCCT	269
16F 16R	CTTCATGTTCTTGCTTCTTCC CAGAAGTATAAGAATGGCTGTC	214

PRIMER	SEQUENCE (5' -> 3')	SIZE OF PCR FRAGMENT(bp)
17F 17R	TGTCCTTTTTCCTGCAAGC TTTCCCTCCAGCACACATG	172
18F 18R	AGGTATTGAATTTCTTTGGACC GTGTGCATCACCACTGTACC	157
19F 19R	CAGGGAGGCTTATGACATC CAGAGAAAGAAGAACACATCC	233

*(Primer sequences taken from Beck et al., 1997; * = own design)*

University of Cape Town

All the oligonucleotide sequences presented in Table D.3 to D.6 were designed for this phase of the current study.

Table D.3. Oligonucleotide sequences for hMSH6 for the purpose of PCR-based mutation detection.

PRIMER	SEQUENCE (5' -> 3')	SIZE OF PCR FRAGMENT(bp)
Exon 1F	CCCCAGATTTCCCGCCAGCAG	440
Exon 1R	CCCAACCCCCTGTGCGAGCCT	
Exon 2F	GGAATTGTTTATTTGTAGGTAAGTGC	412
Exon 2R	CCACCCCTTGCAATTTACTG	
Exon 3F	GAGATGGGGTTTGCTATGTTG	299
Exon 3R	CTAAGTAACCTAACTTATCCTACAGTCC	
Exon 4AF	GGCAAGATCATCTAACACAAAG	462
Exon 4AR	AACCTTGACAGGGCTGTTT	
Exon 4BF	GAAGCAGTGATGAAATAAGCAGTGG	392
Exon 4BR	GAGGAAATCCTCAGGCACATAGAG	
Exon 4CF	GGCTTAAGGAGGAAAAGAGAAGAG	425
Exon 4CR	GATCTCCCTCCTCACCCTCTATC	
Exon 4DF	AGTAGCACGAGTGGAACAGACTG	476
Exon 4DR	GGAGAGTTCTCAAAGTTTGGATG	
Exon 4EF	GAAAAAGGAAATCTCTCAAAGGAAAC	421
Exon 4ER	CATCTAGCACCATTTCGTTGATAGG	
Exon 4FF	CTATCTTCACCAAAGCCTATCAACG	373
Exon 4FR	GTTTCTTCATACATTATAGCCCTGCTG	
Exon 4GF	GGCTACTCAGTAAAATTCATAATGTTG	391
Exon 4GR	CTCTGTTTCAATTTCTCTTATGTCAGC	

Exon 4HF	GACTTATTACTCCCAAAGCAGGC	538
Exon 4HR	CCTAGTCAGGATGCAAATATGTATATGC	
Exon 5F	GATCGTTGGACTGTAATTGAAAG	436
Exon 5R	CACAAATGACTTTCTGATAACAAAAC	
Exon 6F	GGATTTTCAGAACAGAACCAACG	299
Exon 6R	CTGAATGAGAACTTAAGTGGGAAAC	
Exon 7F	GCCAATAATTGCATAGTCTCTTAATG	349
Exon 7R	GCTCTAAAAACATTCATATTGTTTAAATG	
Exon 8F	GCTGCTAAGCAGACTCGTGTA	300
Exon 8R	GCAACATTTTGAGAATAATCTTCTA	
Exon 9F	CAACTCACTACCATTTCATTAGTAGAAG	371
Exon 9R	GTTAGTTAGTTACCGAAATAATCGTAGT	
Exon 10F	GGAAGGGATGATGCACTATGAA	374
Exon 10R	GTATAAAACAGCCTGAAAGTGTATGC	

Table D.4. Oligonucleotide sequences for PCR amplification of the MCR of APC for dHPLC Wave analysis.

PRIMER	SEQUENCE (5' -> 3')	SIZE OF PCR FRAGMENT(bp)
APC15GFOR APC15GREV	AAGAAACAATACAGACTTATTGTG ATGAGTGGGGTCTCCTGAAC	385
APC15HFOR APC15HREV	ATCTCCCTCCAAAAGTGGTG TCCATCTGGAGTACTTTCTGTG	421
APC15IFOR APC15IREV	AGTAAATGCTGCAGTTCAGAGG CCGTGGCATATCATCCCCC	508
APC15JFOR APC15JREV	CCCAGACTGCTTCAAATTACC GAGCCTCATCTGTACTTCTGC	317

APC15KFOR	CCCTCCAAATGAGTTAGCTGC	352
APC15KREV	TTGTGGTATAGGTTTACTGGTG	
APC15LFOR	ACCCAACAAAAATCAGTTAGATG	415
APC15KREV	GTGGCTGGTAACTTTAGCCTC	

Table D.5. Oligonucleotide sequences for PCR amplification of β -catenin for dHPLC Wave analysis.

PRIMER	SEQUENCE (5' -> 3')	SIZE OF PCR FRAGMENT(BP)
Beta-cat exon 3F	ATAGCTGATTTGATGGAGTT	238
Beta-cat exon 3R	TCTTACCAGCTACTTGTCTT	
Beta-cat exon 5F	TACAGGTGGTGGTTAATAAG	251
Beta-cat exon 5R	TTCTTACCCAAGCATTTTCA	
Beta-cat exon 6F	CAGTTCACCAGTGGATTCT	209
Beta-cat exon 6R	TTACCTTGCTTTCTTGGTTG	

Table D.6. Oligonucleotide sequences for PCR amplification of exons 4 to 9 of p53 for dHPLC Wave analysis.

PRIMER	SEQUENCE (5' -> 3')	SIZE OF PCR FRAGMENT(BP)
P53 exon 4F	GCTCTTTTCACCCATCT	335
P53 exon 4R	AGTCTCATGGAAGCCAG	
P53 exon 5&6F	TCCTTCCTCTTCCTACAGTACTCCCCT	411
P53 exon 5&6R	AGTTGCAAACCAGACCTCAGG	
P53 exon 7F	ATCTTGGGCCTGTGTTATCT	319
P53 exon 7R	GGTCCCAAAGCCAGAGA	
P53 exon 8&9F	GGTGGTTGGGAGTAGAT	460
P53 exon 8&9R	CGGCATTTTGAGTGTTAGACTG	

Table D.7. ARMS primers for the amplification of the P60R and S315C mutations in p53.

PRIMER	SEQUENCE (5' -> 3')	SIZE OF PCR FRAGMENT (bp)
P60RFW1	GCCAGAGGCTGCTCCCCG	335
P60RFM1	GCCAGAGGCTGCTCCCC	
P60RREV	GAGTCAGAGATCACACATTAAGTGGGTAAAC	
S315CFW1	CTAGCACTGCCCAACAACACCAGCTCCTC	139
S315CFM1	CTAGCACTGCCCAACAACACCAGCTCCTG	
S315CRev	CGGCATTTTGAGTGTTAGACTGGAACTTTC	

Table D.8. Microsatellite markers used for MSI analysis.

MARKER	CHROMOSOMAL LOCATION	SEQUENCE (5' -> 3')	SIZE OF PCR FRAGMENT (bp)
BAT25	4q12	TCGCCTCCAAGAATGTAAGT TCTGCATTTTAACTATGGCTC	90
BAT26	2p	TGACTACTTTTGACTTCAGCC AACCATTCAACATTTTAAACCC	80-100
D2S123	2p16	AAACAGGATGCCTGCCTTTA GGACTTTCCACCTATGGGAC	197-227
D5S346	5q21	ACTCACTCTAGTGATAAATCG AGCAGATAAGACAGTATTACTAGTT	96-122
D17S250	17q11.2-q12	GGAAGAATCAAATAGACAAT GCTGGCCATATATATATTAAACC	151-169

Table D.9. ARMS primers for PCR amplification of eight mutations in the CDH1 gene.

PRIMER	SEQUENCE	SIZE OF PCR FRAGMENT (BP)
CDH(49-2)FW1 CDH(49-2)FM1 CDH(49-2)REV	CCATCTACCTTTCCCCACCCCA CCATCTACCTTTCCCCACCCCG CCCAGGTTTTTGAGTCTGAAATGAGAGA	455
CDHG59AFW1 CDHG59AFM1 CDHG59AREV	CCACCCCAGGTCTCCTCTTG CCACCCCAGGTCTCCTCTTA GTGTCGTTTGAGCCAAGGAGG	245
G62VFW1 G62VFM1 G62VREV	GAATTTTGAAGATTGCACCGG GAATTTTGAAGATTGCACCGT TCAGAAAAATGCCAACATACCTG	243
CDH7FW1 CDH7FM1 CDH7REV	CACTGGGCTGGACCGAGAG CACTGGGCTGGACCGAGAT CAGAGAGATCGACACAACCAAGTC	291
T340AFW1 T340AFM1 T340AREV	CGATCTCTCTGCAGAGTTTCCCTA CGATCTCTCTGCAGAGTTTCCCTG TACTAAAAGACAACCTGCCATCTCAAGAT	311
L649VFW1 L649VFM1 L649VREV	CATCTCCCTTCACAGCAGAAC CATCTCCCTTCACAGCAGAAAG GGAGTGGAGGGTGGCTAGA	330
CDH13FW1 CDH13FM1 CDH13REV	CGGCGTCTGTAGGAAGGCAC CGGCGTCTGTAGGAAGGCAT CAAACCTTCTATTTTATGGAGGCTGGC	248
CDH15FW1 CDH15FM1 CDH15REV	AACCCTCATGAGTGTCCCCCG AACCCTCATGAGTGTCCCCCG GCAGAGGTTGTGGTGAGCAAAGATT	321

Table D.10. ARMS primers for PCR amplification of the I1307K and E1317Q missense variants in the APC gene.

PRIMER	SEQUENCE	SIZE OF PCR FRAGMENT (BP)
APC1307FW1 APC1307FM1 APC1307REV	CTAATACCCTGCAAATAGCAGAAAT CTAATACCCTGCAAATAGCAGAAAA TGAGTGGGGTCTCCTGAACATA	252
APC1317FW1 APC1317FM1 APC1317REV	AGAAAAGATTGGAAGTAGGTCAGCTG AGAAAAGATTGGAAGTAGGTCAGCTC TGAGTGGGGTCTCCTGAACATA	224

Table D.11. ARMS primers for PCR amplification of the D1853N variant in the ATM gene.

PRIMER	SEQUENCE	SIZE OF PCR FRAGMENT (BP)
D1853NF D1853NRW1 D1853NRM1	GCAGTATGTTGAGTTTATGGCAG GATTTCTCCATGATTCATTTGTATC GATTTCTCCATGATTCATTTGTATTT	228

Two reverse primers (D1853NRW1 and D1853NRM1) were used for each PCR, one specific for the wild type allele and the second specific for the mutated allele.

Table D.12. ARMS primers for PCR amplification of the G242A variant in the CCND1 gene.

PRIMER	SEQUENCE	SIZE OF PCR FRAGMENT (BP)
CyclinD1WTF CyclinD1MUTF CyclinD1REV	CAGAGTGATCAAGTGTGACCCA CAGAGTGATCAAGTGTGACCCG GAAACTCTACGGCTCCTGGAAG	276

Table D.13 ARMS primers for PCR amplification of the C677T variant in the MTHFR gene.

PRIMER	SEQUENCE	SIZE OF PCR FRAGMENT (BP)
MTHFRFW1 MTHFRFM1 MTHFRREV	GAAGGAGAAGGTGTCTGCGGGAGC GAAGGAGAAGGTGTCTGCGGGAGT GAGAGTGGGGTGGAGGGAGCTTATG	187

Table D.14 Primer sequences for the PCR amplification of the GSTM1 and GSTT1 polymorphisms.

PRIMER	SEQUENCE	SIZE OF PCR FRAGMENT (BP)
GSTT1F GSTT1R	TTCCTTACTGGTCCTCACATCTC TCACCGGATCATGGCCAGCA	461
GSTMG5 GSTMG6	GAA CTC CCT GAA AAG CTA AAG C GTT GGG CTC AAA TAT ACG GTG G	215
β-GlobinF β-GlobinR	CAA CTT CAT CCA CGT TCA CC GAA GAG CCA AGG ACA GGT AC	268

Table D.15. Primer sequences for generating PCR fragment for NAT2 polymorphisms.

PRIMER	SEQUENCE	SIZE OF PCR FRAGMENT (BP)
NAT2F NAT2R	GCTGGGTCTGGAAGCTCCTC TTGGGTGATACATACACAAGGG	547

University of Cape Town