THE HISTOPATHOLOGY AND IMMUNOHISTOCHEMICAL EXPRESSION OF CELL CYCLE REGULATORS AND MISMATCH REPAIR GENE PROTEINS IN COLORECTAL CARCINOMA: A COMPARATIVE STUDY

RAVEENDRA SOOKHAYI
Student number: SKHRAV001

University of Cape Town

A dissertation submitted in partial fulfilment of
the requirements for the degree of
Master of Medicine (Anatomical Pathology)

Division of Anatomical Pathology
Department of Clinical Laboratory Sciences
Faculty of Health Sciences
UNIVERSITY OF CAPE TOWN

Date of submission: 16 February 2015

Supervisor: Professor D Govender
Division of Anatomical Pathology,
University of Cape Town

University of Cape Town

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.

Declaration

I, Dr Raveendra Sookhayi, hereby declare that the work on which this dissertation is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university.

I empower the university to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever.

Signature: Signed by candidate ……………………………..

Date: 16 February 2015

Acknowledgements

My supervisor, Professor Govender for all his guidance and perseverance with this project.

Alvera A Vorster and Professor Ramesar from the Division of Genetics for providing the cases, in both the mutation positive and negative cohorts.

Shamal Sookhayi, my loving wife for all her support and hours helping with the layout and taking care of our beautiful baby boy, Vheer during my long nights.

Subash Govender for her excellent and timeous work with the immunohistochemical staining.

Henri Carrara, Anelisa Jaca and Pumza Philips for their help, support and input with the statistical analysis.

Riyaadh Roberts and Sharon Fenwick for their help with the formatting and layout of the document.

The NHLS Research Trust for providing funding for this study.

Abstract

Introduction: It has been reported that HNPCC colorectal carcinomas demonstrate a better prognosis compared to sporadic carcinoma, however the exact mechanism for this is still uncertain. It is possible that tumour morphology, location and cell cycle markers may be indicators of the underlying molecular mechanism. In a resource limited setting these factors may help to stratify which cases need further molecular testing and genetic counselling.

Aims and objectives: To characterise the macroscopic and microscopic pathology in three cohorts of patients. The cohorts include (1) patients with CRCs that are < 50 years and mutation negative, (2) < 50 years and DNA mismatch repair gene mutation positive and (3) more than 50 years (sporadic).

To investigate the immunoexpression of the cell cycle regulators (p21, p27, p53, c-myc, cyclin D1 and cyclin E) and MMP-7 in each cohort.

To compare the immunoexpression of each marker between cohorts.

To correlate the immunoexpression of each marker with tumour type, stage and grade.

Materials and methods: In total, 17 mutation negative, 15 mutation positive and 28 sporadic adenocarcinoma resection cases were available for study. The histopathological features of all cases were reviewed. The cases were stained with antibodies against p21, p27, cyclin D1, cyclin E, p53, c-myc MMP-7, MLH1, MSH2 and MSH6. Results were considered statistically significant if $P < 0.05$, and $P<0.017$ if 3 pairs of medians were compared.

Results: The mutation positive tumours were more frequently right sided tumours and showed mucinous differentiation, tumour infiltrating lymphocytes and an expanding border. The sporadic and mutation negative cohorts showed similar morphology. In the sporadic cohort, the five tumours that were MLH1 negative demonstrated morphological features of MSI-H tumours.

MLH1 mutations were the commonest. MLH1 immunoexpression was lost in the mutation positive tumours and was statistically significant when compared to the other two cohorts. There was no statistical significance among the three cohorts for MSH2 and MSH6 immunoexpression.

There was no statistically significant difference in immunoexpression for p21, p27, p53 and MMP-7 among the three cohorts. Furthermore, there was no

association with tumour type and stage. Cyclin D1 expression was increased in the mutation positive cohort and was statistically significant when compared to the mutation negative cohort only. Cyclin E expression was also increased in the mutation positive cohort and was statistically significant when compared to the sporadic cohort only.

Conclusion: The morphological features of colorectal carcinomas can be helpful in identifying MSI-H tumours and cases requiring further molecular studies. The cell cycle marker expressions did not explain the expected differences in patient outcome and prognosis. The mutation negative cohort in our population continues to remain enigmatic and further testing at the molecular level is required, that may reveal another novel pathway of colorectal carcinogenesis or other novel mutations in mismatch repair genes.

TABLE OF CONTENTS

		Page
	TITLE PAGE	i
	DECLARATION	ii
	ACKNOWLEDGEMENTS	iii
	ABSTRACT	iv
	TABLE OF CONTENTS	vi
	TABLES	x
	FIGURES	xi
	APPENDICES	xii
	LIST OF ABBREVIATIONS	xiii
	CHAPTERS	
1	INTRODUCTION AND LITERATURE REVIEW	1
1.1	Introduction	1
1.2	Epidemiology	2
1.2.1	Incidence	2
1.2.2	Mortality rates and trends	2
1.2.3	Non-modifiable risk factors	2
1.2.3.1	Age	2
1.2.3.2	History of adenomatous polyps	3
1.2.3.3	History of Inflammatory Bowel Disease	3
1.2.3.4	Family history of CRC or adenomatous polyps	3
1.2.3.5	Inherited genetic risk	3
1.2.4	Environmental risk factors	3
1.2.4.1	Nutritional Factors	3
1.2.4.2	Lifestyle factors and Obesity	4
1.3	Molecular Pathways of colorectal carcinoma	4
1.3.1	Chromosomal Instability Pathway (CIN)	4
1.3.1.1	*KRAS* oncogene	5
1.3.1.2	Loss of 5q allele	6
1.3.1.3	APC and the β-catenin-dependent Wnt Signalling Pathway	6
1.3.1.4	*MCC* gene	7
1.3.1.5	Loss of 8p allele	8

1.3.1.6	Loss of 17p allele	8
1.3.1.7	Loss of 18q allele	8
1.3.2	Microsatellite Instability (MSI)	8
1.3.2.1	MSI in sporadic vs hereditary CRC	10
1.3.2.2	Lynch syndrome/HNPCC	11
1.3.3	CpG Island Methylator Phenotype (CIMP)	12
1.3.4	Alternative serrated neoplastic pathway	14
1.3.5	MicroRNAs	15
1.3.6	Epigenetic changes	16
1.3.6.1	Histone modification	17
1.3.6.2	DNA methylation	17
1.4	Hereditary colorectal cancer syndromes	19
1.4.1	Familal Adenomatous Polyposis	20
1.4.2	Lynch syndrome	20
1.4.3	MYH-associated polyposis	20
1.4.4	Harmatomatous polyposis syndromes	21
1.4.4.1	Peutz–Jeghers syndrome	21
1.4.4.2	Juvenile polyposis syndrome	21
1.4.4.3	Cowden syndrome	21
1.5	Regulation of the cell cycle	21
1.5.1	Cyclins D1 and E	22
1.5.2	Cell cycle Inhibitors – p21,p27 and p53	23
1.6	C-myc	24
1.7	Matrix metalloproteinase 7 (MMP-7)	24
1.8	Clinical features and Special Investigation	25
1.9	Pathology	25
1.9.1	Macroscopic pathology	25
1.9.2	Microscopic pathology	26
1.10	Aims and objectives	28
2	MATERIALS AND METHODS	29
2.1	Ethics approval	29
2.2	Sample selection	29
2.3	Antibodies	30

2.4	Immunohistochemistry Method	31
2.5	Assessment of immunohistochemical stains	32
2.6	Clinical data and pathological features	33
2.6.1	Macroscopic Features	33
2.6.2	Microscopic Features	33
2.6.3	Grading of tumours	34
2.6.4	Staging of tumours	35
2.7	Statistical analysis	36
3	RESULTS	38
3.1	Clinical and Pathological characteristics	38
3.1.1	Cohort 1 – Mutation negative (n = 17)	39
3.1.2	Cohort 2 – Mutation positive (n = 15)	39
3.1.3	Cohort 3 – Sporadic (n = 28)	39
3.1.4	Immunohistochemical Analysis	41
3.1.4.1	p21	41
3.1.4.2	p27	42
3.1.4.3	p53	43
3.1.4.4	Cyclin D1	44
3.1.4.5	Cyclin E	45
3.1.4.6	C-myc	47
3.1.4.7	MMP-7	47
3.1.4.8	MLH1	50
3.1.4.9	MSH2	51
3.1.4.10	MSH6	52
4	DISCUSSION	54
4.1	Clinical and Morphological Features	54
4.2	Cell cycle Inhibitors – p21, p27 and p53	55
4.3	C-myc	56
4.4	Cyclins D1 and E	57
4.5	MMP-7	58
4.6	Mismatch repair proteins – MLH1, MSH2 and MSH6	60
4.7	Limitations	61
5	CONCLUSION	63

6	REFERENCES	65
	APPENDICES	72

TABLES **Page**

Table 1	Genetics of inherited colorectal tumour syndromes	19
Table 2	Primary antibodies	30
Table 3	Kits used for the IHC protocol	31
Table 4	Characteristics of Colorectal carcinomas in the three cohorts.	38
Table 5	Staining of MMP-7 at invasive tumour margin	49

FIGURES Page

Figure 1	Molecular events in CRC	5
Figure 2	Model of APC and β-catenin function	7
Figure 3	Histopathological features of colorectal adenocarcinomas	40
Figure 4	p21 immunohistochemistry	41
Figure 5	p27 immunohistochemistry	42
Figure 6	p53 immunohistochemistry	43
Figure 7	Cyclin D1 immunohistochemistry	45
Figure 8	Cyclin E immunohistochemistry	46
Figure 9	c-myc immunohistochemistry	47
Figure 10	MMP-7 immunohistocemistry	48
Figure 11	Invasive tumour margin (TM) in a low grade adenocarcinoma	48
Figure 12	MLH1 immunohistochemistry	50
Figure 13	MSH2 immunohistochemistry	51
Figure 14	MSH6 immunohistochemistry	52

APPENDICES Page

Appendix 1	Cohort 1 (Mutation negative) Morphological features	73
Appendix 2	Cohort 1 (Mutation negative) Immunohistochemical results	74
Appendix 3	Cohort 2 (Mutation positive) Morphological features	75
Appendix 4	Cohort 2 (Mutation positive) Immunohistochemical and gene mutation results	76
Appendix 5	Cohort 3 (Sporadic) Morphological features	77
Appendix 6	Cohort 3 (Sporadic) Immunohistochemical results	78

List of abbreviations

APC	Adenomatous polyposis coli
CRC	Colorectal carcinoma
CpG	Cytosine phosphate guanine
CIMP	CpG island methylator phenotype
CIMP-H	CIMP high
CIMP-L	CIMP low
CIN	Chromosomal Instability
DNA	Deoxyribonucleic acid
DVL	Dishevelled
EBV	Epstein Barr virus
FAP	Familial adenomatous polyposis
FZ	Frizzled
GIT	Gastrointestinal tract
GCHP	Goblet cell rich hyperplastic polyp
GSK3β	Glycogen synthase kinase 3 beta
H&E	Haematoxylin and eosin
HNPCC	Hereditary non-polyposis colorectal carcinoma
HP	Hyperplastic polyp
IHC	Immunohistochemical
LOH	Loss of heterozygosity
MAP	MYH-associated polyposis
MPHP	Mucin poor hyperplastic polyp
MVHP	Microvesicular hyperplastic polyp
MLH1	MutL homolog 1
MMed	Master of Medicine
MMR	Mismatch repair
MSH2	MutS homolog 2
MSH6	MutS homolog 6
MSI	Microsatellite instability
MSI-H	Microsatellite instability – high
MSI-L	Microsatellite instability – low
MSS	Microsatellite stable
NHLS	National Health Laboratory Service
NOS	Not otherwise specified
Obj mag	Objective magnification
PBS	Phosphate buffered saline
PMS2	post meiotic segregation 2
TCF/LEF	T cell factor/lymphoid enhancer
TIMP	Tissue inhibitors of metalloproteinases
TNF-α	Tumour necrosis factor alpha
TSG	Tumour suppressor gene
SA	South Africa
SSA	Sessile serrated adenoma
UCT	University of Cape Town
USA	United States of America
WHO	World Health Organization

CHAPTER 1

1. INTRODUCTION AND LITERATURE REVIEW

1.1 Introduction

Colorectal carcinoma (CRC) in South Africa has been increasing over the past decade, and was recently ranked as one of the top five cancers, fifth amongst males and third amongst females (1). Worldwide, CRC is a major cause of morbidity and mortality.

The main pathways that are implicated in carcinogenesis include the chromosomal instability pathway (CIN), microsatellite instability pathway (MSI), and the CpG island methylator phenotype (CIMP) pathway. In recent times, the serrated neoplastic pathway showing some overlapping features with the above pathways has been described. Finally, microRNAs are emerging as important adjuncts in the pathway of carcinogenesis. It is known that HNPCC positive patients (part of the MSI pathway) have a more favourable prognosis when compared with sporadic CRC.

Pathological features, tumour site and cell biomarkers are helpful in guiding the pathologist to the underlying molecular aberrations. This may aid the pathologist in choosing the appropriate molecular diagnostic tests, especially in a resource limited setting. In addition, the appropriate patients may be referred for further genetic counselling and testing.

This study investigated cell cycle regulators, mismatch repair proteins and matrix metalloproteinase 7 (MMP-7) together with morphology in three cohorts of patients. The cohorts included patients older than 50 years with sporadic CRC, patients younger than 50 years with CRCs that are mutation positive for one of the mismatch repair (MMR) genes and patients younger than 50 years with CRCs that are mutation negative for MMR genes. The study was conducted to determine possible differences in expression of the markers mentioned above in correlation with the three patient cohorts and pathological features, to help further

understand the progression of colorectal carcinoma and to elucidate possible mechanisms for the differences in behaviour of CRC in HNPCC.

1.2 Epidemiology

1.2.1 Incidence

CRC is the commonest type of gastrointestinal tract (GIT) cancer and the lifetime risk of developing CRC is 1 in 17 for men and 1 in 26 for women (2). It accounts for 9% of all cancers, with the highest incidence seen in Australia, New Zealand, Canada, parts of Europe and the USA. The countries that are low risk include India, China, parts of Africa and South America. Therefore, it is mainly a disease of developed countries which contribute 63% of all cases. However, the incidence rates may be biased as there may be a high degree of underreporting in developing countries (2, 3). The crude incidence in Sub-Saharan Africa was found to be 4.04 per 100 000 population for both sexes, with South Africa having the highest rates (4).

Interestingly, in some regions like the USA, Northern and Western Europe, the incidence rates have been stabilising and declining. This is due to screening programmes that detect precancerous polyps (3).

1.2.2 Mortality rates and trends

The worldwide mortality due to CRC is approximately half that of the incidence, with just over 1 million new cases recorded in 2002, and 530 000 deaths due to CRC. This accounts for ~8% of all cancer related deaths. It is the fourth most common cause of cancer related deaths worldwide. Mortality rates have declined in North America, New Zealand, Australia and Western Europe (2, 3).

1.2.3 Non-modifiable risk factors

1.2.3.1 Age

The vast majority, 90% of CRCs occur after 50 years of age. The incidence rate increases after 40 years, rising sharply after 50. The rates are also increasing among younger persons (3).

1.2.3.2 History of adenomatous polyps

Adenomas are neoplastic proliferations of glandular epithelium that display dysplastic morphology. In the USA, the lifetime risk of developing an adenoma is 19% in the general population (3). The prevalence of these lesions in the USA is approximately 25% by age 50 years and 60% by age 70 years (5). An individual with a history of adenomas is at a higher risk of developing CRC than an individual without a history of adenomas.

1.2.3.3 History of Inflammatory Bowel Disease

Crohn's disease and ulcerative colitis increase an individual's overall risk of developing CRC. The relative risk of CRC in both these diseases has been estimated to be between 4-to 20-fold (2, 3).

1.2.3.4 Family history of CRC or adenomatous polyps

Individuals with a history of CRC or adenomatous polyps in one or more first-degree relatives are at increased risk. The vast majority of CRC cases occur without a family history, with only 20% (excludes inherited genetic risk) of people with CRC having other family members affected by the disease. The increased risk may be due to inherited genes or shared environmental factors (2).

1.2.3.5 Inherited genetic risk

Approximately 5% of CRCs are due to inherited mutations, which are highly penetrant single gene mutations (3, 6). The most common are familial adenomatous polyposis (FAP) and hereditary non-polyposis colorectal cancer (HNPCC), also known as Lynch syndrome.

1.2.4 Environmental risk factors

1.2.4.1 Nutritional Factors

Diets high in animal fat and meat, and low in dietary fibre have been implicated in CRC. The high dietary fat has been linked to the growth of intestinal bacterial flora which degrade bile salts to carcinogenic compounds (3). The mechanisms for the association with meats include haem iron in red meat and cooking at high temperatures resulting in heterocyclic amines and polycyclic aromatic hydrocarbons. In addition, differences in dietary fibre account for geographic

differences in CRC. Increased fibre increases faecal bulk and decreases the transit time (3).

1.2.4.2 Lifestyle factors and Obesity

Lack of exercise and obesity are contributing factors in about 25 to 30% of CRC. The carcinogens found in cigarettes increase cancer growth in the colon and rectum. Regular alcohol consumption is also a factor in the development of CRC, especially in younger patients (3).

1.3 Molecular pathways in CRC

Fearon and Vogelstein put forward the first model for colorectal carcinoma more than 20 years ago (7). The sequential pathway outlined four steps:

1. *APC* gene inactivation → adenoma
2. *KRAS* mutations → promotes adenoma growth
3. Alterations of chromosome 18q → progression
4. p53 inactivation → triggers final transition to carcinoma

Colorectal cancer encompasses a complex disease process, and new findings require further refinement on this initial model. Four distinct molecular pathways have been identified, reflecting different phenotypes. There is dysregulation of cell metabolism, proliferation, differentiation, survival and apoptosis secondary to genetic and epigenetic alterations (8, 9).

The four pathways are the chromosomal instability (CIN) pathway, microsatellite instability (MSI) pathway, the CpG island methylator phenotype (CIMP) pathway and the alternative serrated neoplastic pathway (**Figure 1**). Besides genomic instability, other molecular changes involve microRNAs, epigenetic factors, growth factors and epithelial-mesenchymal transitions (9). The latter two are mainly responsible for metastases. The pathways are not exclusive, and some tumours can exhibit more than one pathway.

1.3.1 CIN Pathway

This pathway, also known as the suppressor pathway, is the most common, ranging from 50-85% of sporadic cancers according to recent literature (9). It is characterised by gains and/or losses of whole or portions of chromosomes

resulting in karyotypic variability. It results from defects in chromosome segregation and microtubule dysfunction, with subsequent aneuploidy (imbalance in chromosome number), telomere dysfunction, telomerase overexpression, or defects in the DNA damage response mechanisms (8, 9). The effect is aneuploidy, genomic amplifications and high frequency loss of heterozygosity (LOH). Broad amplifications have been identified on chromosomes 7, 8q, 13q, 20 and X. Broad deletions have been identified on 1, 4, 5, 8p, 14q, 15q, 17p, 18, 20p and 22q (8). Chromosomes 1, 5, 8, 17 and 18 account for between 46-78% of allelic loss (8). Focal gains and losses have been identified in regions harbouring important genes e.g. *MYC*, *VEGF*, *PTEN* and others. The net effect is mutations in proto-oncogenes and tumour suppressor genes (TSG). *KRAS* and *APC* genes are the most common genetic alterations in CRC (8).

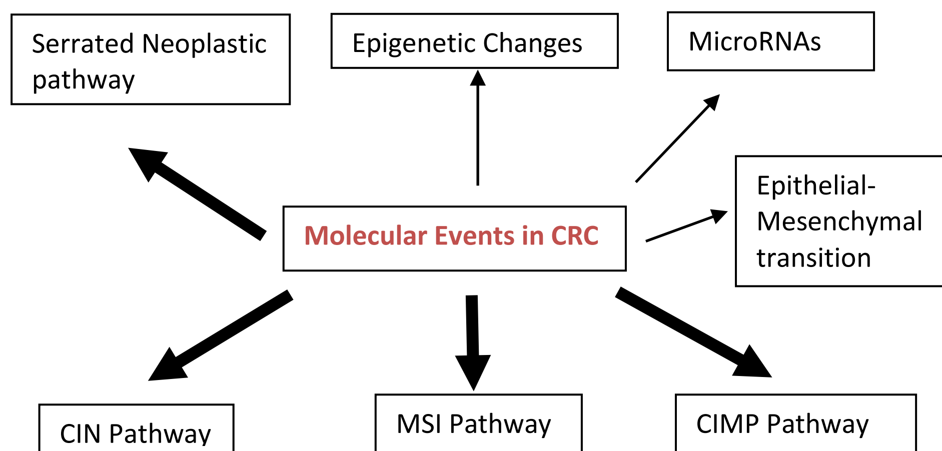

Figure 1: Molecular events in CRC

1.3.1.1 *KRAS* oncogene

The RAS-RAF-MAPK pathway activates transcription factors that express proteins with a role in cellular proliferation, differentiation and survival (9). The RAS family is composed of three members KRAS, HRAS and NRAS (5). *KRAS* alterations, which are implicated in many human cancers, are seen in 30-60% of CRCs and large adenomas (8). The gene product is a 21 kDa membrane bound protein that is activated in response to extracellular signals. Single nucleotide point mutations are usually found which lock the protein in its active form due to impaired GTPase activity. The downstream targets include Bcl2, H2AFZ,

RAP1B, TBX19, E2F4 and MMP1. Mutations are usually seen at codons 12 (82-87%), 13 (13-18%) and 61 (% unknown) in human cancers (9).

1.3.1.2 Loss of 5q allele

Allelic loss of chromosome 5q is seen in 20-50% of sporadic CRCs. The two important genes located on 5q are *APC* and the mutated in colorectal cancer (*MCC*) (8, 9).

Mutations of *APC*, the "gatekeeper" of cellular proliferation, have the same frequency in very small adenomas as in advanced carcinomas and are even found in the earliest lesions composed of a few dysplastic crypts (9). Hence, it suggests that *APC* mutations are a rate-limiting event in carcinogenesis. In addition, biallelic *APC* gene defects are needed to promote tumour development (Knudson's two hit hypothesis) (5).

1.3.1.3 APC and the β-catenin-dependent Wnt signalling pathway

The *APC* gene encodes a protein that regulates cell-cell adhesion, cell migration, chromosomal segregation and apoptosis in colonic crypts. It is a negative regulator of β-catenin, by inducing its degradation.

In the Wnt pathway, in the absence of Wnt ligand signalling, APC binds to the scaffold protein Axin, followed by phosphorylation by casein kinase 1 alpha (CK1a) and glycogen synthase kinase 3β (GSK3β) (**Figure 2**). The net effect is ubiquitination and degradation of β-catenin (8, 10).

In the presence of Wnt signalling, the Wnt ligand binds to a complex of Frizzled (Fz) proteins and a member of the low-density lipoprotein receptor-related protein family (e.g. LRP5 or LRP6), resulting in inhibitory effects on GSK3β and Axin (**Figure 2**). This is also aided by the action of dishevelled (Dvl). The net result is stabilisation of the β-catenin pool, accumulating in the cytoplasm and eventually binding with DNA-binding proteins TCF/LEF (T cell factor/lymphoid enhancer family) (5, 10). Wnt target genes include proto-oncogenes such as *c-myc* and cyclin D1 (*CCND1*), membrane factors such as matrix metalloproteinase 7 (MMP-7)/Matrilysin, membrane type 1 MMP-1, laminin-5 gamma 2 chain,

CD44, angiogenic factors, apoptotic factors, growth factors such as FGF20 and FGF9 (5). In addition, Wnt feedback regulators (e.g. Axin2, Wnt inhibitory factor) are also affected. Therefore, when both alleles of the *APC* gene are inactivated, the process mimics activation of the Wnt pathway.

Mutations in other canonical Wnt pathway components, include β-catenin mutations (*CTNNB1* gene), which render it resistant to degradation, mutations in Axin1/Axin2 and mutations involving *TCF4* (5).

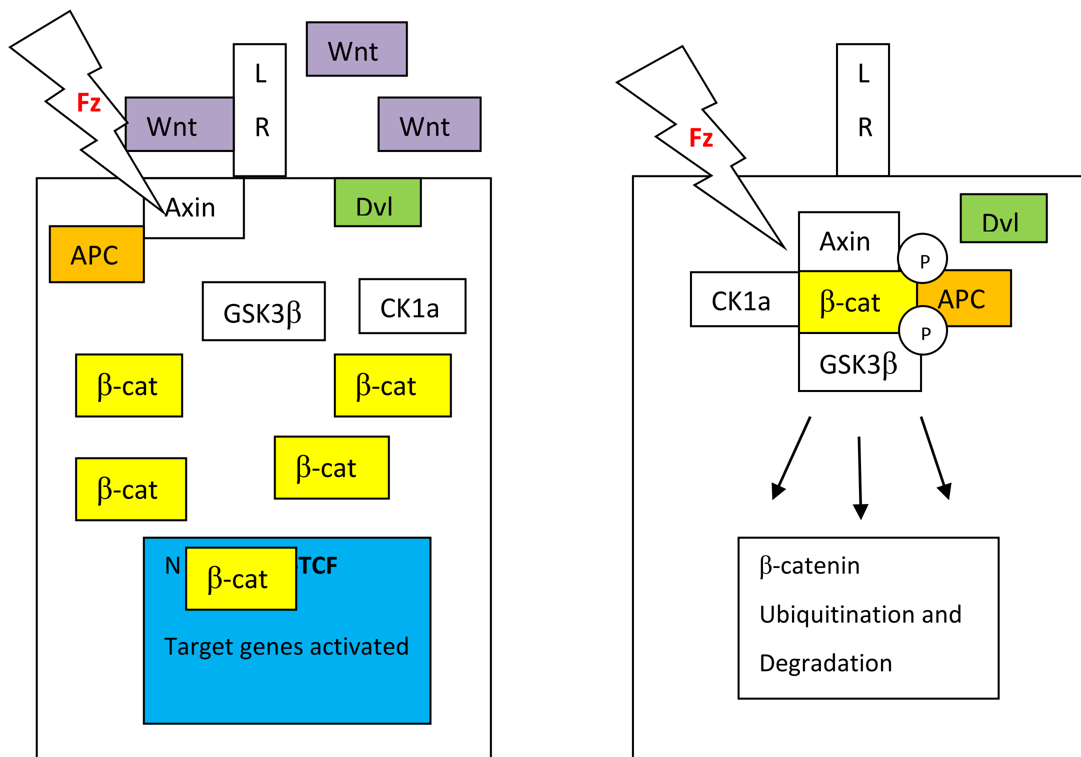

Figure 2: Model of APC and β-catenin function: Binding of Wnt to Frizzled (Fz) receptor activates Dishevelled (Dvl) and blocks the destruction complex formed by Axin. In the absence of Wnt, the axin/conductin complex phosphorylates β-catenin by GSK3β and CK1a, resulting in β-catenin being degraded in proteasomes.

1.3.1.4 *MCC* gene

The *MCC* gene, located on 5q21, is commonly inactivated through promoter hypermethylation. The protein product is a cell cycle regulator that induces cell

cycle arrest following DNA damage. A recent study also suggested an inhibitory function on Wnt/β-catenin signal transduction (8).

1.3.1.5 Loss of 8p allele

In over 50% of CRCs, allelic loss of 8p has been identified. It is thought that TSGs are found in these regions of deletions. Loss of locus 8p21-22 also increases susceptibility for metastases (8).

1.3.1.6 Loss of 17p allele

This segmental loss is seen in 75% of CRCs but not in adenomas, suggesting a late event in CRC carcinogenesis (5, 8). It houses the *p53* gene that is responsible for cell cycle arrest, apoptosis, senescence, autophagy and cellular metabolism. This TSG will be discussed further below.

1.3.1.7 Loss of 18q allele

Loss of heterozygosity of 18q is seen in approximately 70% of CRCs, approximately 50% of large adenomas and fewer than 10% of small early adenomas. A number of TSGs are found in the long arm of chromosome 18. These include *Cables*, Deleted in Colorectal Cancer (*DCC*), *Smad2* and *Smad4*. The Cables protein is involved with regulation of the cell cycle and interacts with CDK2, CDK3 and CDK5. *DCC* codes for an immunoglobulin superfamily protein (190kda), that plays a role in cell adhesion, migration and induces apoptosis in the absence of its ligand netrin-I. The Smad proteins are transcription factors involved in the transforming growth factor β pathway, and target genes include *c-myc*, *CBFA1*, *FLRF* and *furin*. In addition, these proteins downregulate claudin1, in a TGF-β independent manner, which is upregulated in metastatic lesions (9).

1.3.2 Microsatellite instability (MSI) pathway

This pathway first described in the background of HNPCC is seen in approximately 15-20% of sporadic CRCs (11, 12). Microsatellites are short repeat nucleotide sequences (1-4 bp in length) that are spread out over the entire genome. These areas in the genome are prone to reading errors and the cell has in place, the DNA mismatch repair (MMR) system that detects and repairs base pair

mismatches and insertion/deletion loops (IDLs) during replication. IDLs are functionally equivalent as an insertion on one strand which results in the same effect as a deletion of the same number of nucleotides on the other strand (13). Inactivation of the DNA MMR system results in microsatellites becoming unstable when they become abnormally long or short due to gains or losses of repeat units. This inactivation may be genetic (e.g. Lynch syndrome) or acquired. Proto-oncogenes and tumour suppressor genes (TSGs) that contain repetitive DNA sequences are the major targets as loss of MMR proteins mediates an increase in mutation rate that produces a 'mutator phenotype'. Genes affected by MSI include *TGFβRII*, *EGFR*, *BAX*, regulators of proliferation (e.g. *GFB1*, *Axin-2*, *CDX*, *GRB1*), the cell cycle or apoptosis (e.g. *BCL10*, *PTEN*, *caspase-5*, *FAS*) and DNA repair. The genes that do not appear to be affected include *p53* and *KRAS* (9).

The MMR system includes the following genes, *MLH1*, *MLH2*, *MLH3*, *MSH2*, *MSH3*, *MSH4*, *MSH5*, *MSH6*, *PMS1*, *PMS2*, and *Exo1*. Two heterodimeric complexes have been described, the MutS- and MutL-related proteins. The MutL complex contains MLH1 protein plus one of the other binding partners (PMS2, MLH3, PMS1 or MLH2). Similarly, MutS contains the MSH2 protein plus one of the other binding partners usually MSH6 or MSH3. These complexes together with replication factors and other proteins allow excision and resynthesis of the affected DNA to proceed. MSH2 and MLH1 are common components of these complexes, therefore loss of either one will result in no MMR activity, whereas loss of one of the partner components (e.g. MSH6) will only diminish activity of the MMR pathway. This might explain why most of the Lynch syndrome patients have germline mutations in either *MSH2* or *MLH1* (~90%) (13, 14).

The International Workshop, held in Bethesda on Microsatellite Instability sponsored by the National Cancer Institute in 1997, recommended a panel of five microsatellite loci for the identification of MSI (15). This involves extraction of tumour DNA and amplification of the microsatellite loci using polymerase chain reaction (PCR). The PCR products are then compared to normal tissue or blood and using a fluorescent DNA analyser MSI can be identified.
Based on the above test MSI can be divided into 3 cohorts:

1. Microsatellite stable (MSS) – no evidence of MSI
2. MSI found at high frequency (MSI-H) - Defined as instability in 2 or more markers.
3. MSI found at low frequency (MSI-L) – Instability in one marker.

Immunohistochemical staining can be used to detect the MMR proteins. Abnormal MMR protein expression can be detected in two patterns, either complete loss of expression or patchy weak expression (13). Complete loss of expression may be due to no expression of the protein or expression of a truncated protein to which the antibody does not bind. Patchy weak staining may be seen when a prematurely truncated but variably stable protein is generated, or a protein with alterations in the epitope recognized by the antibody. The changes in expression of MMR proteins follow the pattern of their heterodimeric complex formation, e.g. abnormal PMS2 expression is associated with abnormal MLH1 expression (13, 16).

MSI-H tumours tend to have a distinct phenotype characterised by location and histopathological features. These tumours are usually located proximal to the splenic flexure, mainly in the right colon. Histologically, these tumours have increased lymphocytic infiltration, mucinous histology and poor differentiation with medullary growth pattern (17).

1.3.2.1 MSI in sporadic vs hereditary CRC

Mutations or deletions or epigenetic silencing of both copies of one of the MMR genes results in loss of function. MSI is the hallmark of HNPCC (Lynch syndrome), where one germline mutation is inherited and subsequent inactivation of the wild-type allele by mutation, or loss or epigenetic inactivation (second hit), with no expression of the functional protein. Approximately 15-20% of sporadic CRCs are due to somatically acquired changes in the MMR proteins, and is almost entirely caused by hypermethylation of the *MLH1* promoter region that results in gene silencing. A high frequency of MSI is found in adenomas of HNPCC and is only rarely found in sporadic adenomas (13, 17).

BRAF^{V600E} mutations are seen frequently in sporadic MSI-H CRC but not in HNPCC. The RAS-RAF-MAPK pathway activates transcription factors that

express proteins with roles in cell differentiation and proliferation. The RAF family of molecules includes ARAF, RAF1 and BRAF. Similar to *KRAS* mutations, *BRAF* mutations render it continually active and usually involves the point mutation that results in substitution of thymidine to adenine at nucleotide 1799, resulting in substitution of valine to glutamic acid (V600E mutation) (9, 17).

1.3.2.2 Lynch syndrome/HNPCC

Dr Alfred Warthin described the first case of hereditary CRC, over a hundred years ago, which he designated family G. In 1971, Lynch and Krush subsequently updated studies on family G with what came to be known as Lynch syndrome (HNPCC). It is an autosomal dominant condition caused by germline mutations in MMR genes. These include *MLH1* (3p21), MSH2 (2p21-22), *MSH6* (2p16) and *PMS2* (7p22). A recent germline mutation in the *TACSTD1* gene which encodes the epithelial cell adhesion molecule (Ep-CAM) has also been identified (6, 18). It is located upstream of *MSH2* on chromosome 2 and is a rare mutation accounting for 1-3% of cases which results in epigenetic silencing of *MSH2*.

Patients with HNPCC usually develop multiple tumours between 20-30 years of age in the colon and are at higher risk of developing extra colonic tumours including endometrial, ovarian, pancreatic, small bowel, hepatobiliary, skin, brain and urethral tumours (18). The lifetime risk of developing CRC in HNPCC is approximately 85% (6).

Guidelines have been developed to identify patients at risk of HNPCC. It is recommended that MSI testing be performed on all patients fulfilling the criteria with referral to a geneticist. Examples of the guidelines include the Revised Bethesda criteria summarized below (12, 18).

Revised Bethesda Guidelines

1. CRC diagnosed in patients <50 years
2. Presence of synchronous, metachronous colorectal, or other HNPCC associated tumour regardless of age
3. MSI-H histology in a patient <60 years of age

4. CRC diagnosed in ≥ 1 first degree relative with an HNPCC related tumour, with one of the cancers diagnosed <50 years
5. CRC diagnosed in ≥ 2 first or second degree relatives with HNPCC related tumour regardless of age

A large family of approximately 500 individuals, who are at increased risk for HNPCC syndrome on account of a transversion mutation in *MLH1* has been identified in the remote rural village of Kleinsee in the Northern Cape province of South Africa, 500km from Cape Town. This family demonstrates anatomic site specific HNPCC, in that increased risk for malignancy in family members is largely restricted to the colon (19).

1.3.3 CpG island methylator phenotype (CIMP) pathway

This is one of the newer pathways that was originally grouped with MSI and involves epigenetic alterations (methylation or histone modification) that affect gene expression. DNA methylation commonly occurs at the 5'-CG-3' (CpG), usually in promoter regions resulting in gene silencing providing another mechanism of inactivation of TSGs. Increased methylation is seen with increasing age, smoking, injury and chronic inflammation (8, 20).

The CpG islands are regions in the genome rich in cytosine-phosphate-guanine (CpG) dinucleotides where cytosine DNA methylation does not covalently modify. CIMP positivity is found in 35-40% of CRCs and has been identified in adenomas. Genes silenced by hypermethylation include *APC*, *MLH1*, *MCC*, *MGMT*, *p16*, *p14* and several others, resulting in repression of these important tumour suppressor genes (8, 9, 20).

CIMP characterisation is based on analysis of various genes but the consensus for the exact panel of CpG sites for CIMP analysis has not been reached. A review by Curtin *et al* demonstrated 8 different studies using different CIMP panel markers and in all 8 *BRAF* ^{V600} mutation had been observed (21).

The "traditional"/classic panel to assess CIMP included: *MLH1*, *CDKN2a* (p16), methylated in tumours (*MINTS*) 1, 2 and 31. The cases were divided as follows on assessment of CIMP status in serrated polyps in the study (22):

No loci positive:	CIMP negative
One locus positive:	CIMP low
Two or more loci positive:	CIMP positive

CIMP status has been implicated in two pathways in sporadic tumours: CIMP high (CIMP-H) related to *BRAF* mutations and CIMP low (CIMP-L) related to *KRAS* mutations. Tumours characterized as CIMP positive based on the classic panel include both CIMP-H and CIMP-L.

In the study by Weisenberger *et al* the "new" CIMP panel included: *CACNA1G*, *IGF2*, *NEUROG1*, *RUNX3* and *SOCS1*, based on the stepwise screen of 195 markers. CIMP positive was defined by methylation of at least three markers (8).

Finally in two comprehensive DNA methylation studies, several CpG sites (202 CpG sites and 20 panel markers) were analysed (21). The emergence of two methylation panels as classifiers were proposed: the first to identify high methylated tumours (strongly associated with *BRAF*) and to distinguish between intermediate (associated with *KRAS*) and low methylation groups.

CIMP-high CRC accounts for 15-20% of sporadic CRCs and has overlapping phenotypic features with MSI-H tumours. It is often poorly differentiated, of mucinous or signet ring histology and seen in the proximal colon (8). In addition, these tumours harbour *KRAS* and/or *BRAF* mutations as mentioned above (9).

The prognostic significance of CIMP remains ill defined. CIMP+ tumours confer a worse prognosis than MSI-H tumours. However, if both MSI-H and CIMP positivity are present, the prognosis improves compared with MSI-/CIMP+ tumours (9). Jass suggested a classification based on MSI and CIMP status and included the following 5 molecular subtypes (23):

1. CIMP high/MSI high – originates in sessile serrated adenomas (see below) together with *BRAF* mutations and *MLH1* methylation.
2. CIMP high/MSI low – originates in sessile serrated adenomas together with *BRAF* mutations and methylation of multiple genes.

3. CIMP low/MSI low or MSS – Occurs in tubulovillous adenomas or traditional serrated adenomas and is characterised by CIN, *KRAS* mutation and *MGMT* methylation.
4. CIMP-/MSS – originates in the traditional adenoma and is characterized by CIN.
5. CIMP negative/MSI-H – seen in HNPCC; negative for *BRAF* mutations.

In summary, serrated polyps are the precursors of types 1 and 2 CRCs while types 4 and 5 evolve through the adenoma-carcinoma sequence, while type 3 CRCs may arise within either type of polyp. However, it is thought that the traditional serrated adenoma does not belong to the serrated neoplastic pathway (see below) (24).

1.3.4 Alternative serrated neoplastic pathway

It is thought that between 20 and 35% of sporadic colorectal carcinomas arise from this alternative pathway, developing from a "serrated polyp" and accounts for at least 15% of all CRCs (24). These polyps are composed of epithelial infoldings that create a saw-tooth appearance. The nomenclature continues to be controversial in this group. This pathway is seen in elderly women (average age of 61 years), particularly in the right side of the colon. It is characterised by excessive methylation, propensity for *BRAF* mutations and MSI. It is associated with sessile serrated adenomas/polyps, which are different from hyperplastic polyps. The WHO classifies serrated polyps into three categories: hyperplastic polyps, sessile serrated adenomas and traditional serrated adenomas (24, 25).

Hyperplastic polyps, which include goblet cell hyperplastic polyps (GCHPs), microvesicular hyperplastic polyps (MVHPs) and mucin poor hyperplastic polyps (MPHP) are recognized by histology and classified according to the morphology of the lining cells, GCHPs are associated with *KRAS* mutations while MVHPs are linked to *BRAF* mutations. The polyps are small, distally located and have little or no risk of progression to malignancy (2, 24).

Sessile serrated adenomas (SSAs) are located predominantly in the right colon. These polyps differ from hyperplastic polyps, in that the crypts are distorted,

branched and dilated resembling an "L" or inverted "T". SSAs may have cytologic atypia and only rarely display dysplasia. However, again the nomenclature is still controversial being disputed among experts (24).

Traditional serrated adenomas (TSAs) are located in the distal colorectum and are rare accounting for 1% of polyps. Morphologically, there is greater villous or tubulovillous pattern with ectopic crypt foci and epithelial cells displaying eosinophilic cytoplasm and "pencillate" nuclei. The ectopic crypt foci, help in distinguishing TSAs from SSAs. Foci resembling conventional tubulovillous adenomas have been described, the exact significance is uncertain. Although this polyp shares its name with SSA, it is not felt to belong to part of this pathway. These lesions are different morphologically and pathogenetically (24).

Two molecular pathways have been proposed (9, 24):
1. *BRAF*-mut with CIMP-H and MSI or MSS cancers. There is usually promoter methylation of *MLH1* in the MSI CRCs. This is seen in the SSA pathway. MVHPs are thought to be precursors of SSAs (24).
2. *KRAS*-mut with CIMP-L and MSS are seen in TSAs which are thought not to belong to the serrated neoplastic pathway by some authors. *BRAF* may also be mutated. GCHPs are thought to be precursors of TSAs (24).

Overall, the detection and pathological identification are not yet optimised. Pathologists are confused over the exact nomenclature due to the different fields of thought, and need to be aware of this minor yet significant contribution the serrated neoplastic pathway has in CRCs. This will allow for greater recognition of these polyps and help standardise reporting.

1.3.5 MicroRNAs

MicroRNAs (miRNA) are short non-coding RNA sequences that bind to mRNA and control translation of complementary genes. MiRNA dysregulation is suggested to play a role in carcinogenesis when miRNA targets TSGs or oncogenes through silencing or overexpression respectively. MiRNAs that allow tumour growth may be upregulated through transcriptional activation and/or amplification of the miRNA encoding gene. Those that are decreased and

potentially unfavourable to growth are silenced by deletion or epigenetic modifications. There appears to be a role for miRNAs in CRC initiation, progression and development. In addition, miRNAs are much more stable molecules than DNA or RNA, due to their small size allowing detection in blood, and formalin fixed tissue (26, 27).

Some of the miRNA findings in CRC carcinogenesis include [23,24]:
- miRNA-143 and miRNA-145 are reduced indicating tumour suppressive activity.
- miRNA-21 is elevated and has important roles in initiation, progression and metastasis. Elevated levels of miRNA-21 are found in adenomas.
- miRNA-135a and miRNA-135b downregulate the *APC* gene in the Wnt pathway.
- miRNA-92 and miRNA-17 were elevated in preoperative samples of plasma from CRC patients and decreased following removal of the cancers.

In the future, miRNAs may be potential molecular classifiers, early detection biomarkers and therapeutic targets for CRC.

1.3.6 Epigenetic changes

These changes, seen in approximately 40% of CRCs, alter genetic expression, without changes to the DNA nucleotide sequences and are conveyed during cellular division. The DNA helix is wrapped around histone proteins into chromatin that is regulated by epigenetic factors, involving methylation, histone modification and non-coding RNAs. It influences expression of genes by inducing modifications in the chromatin at gene promoter areas, affecting the accessibility of transcription factors (28).

The genetic code provides the blueprint for cellular protein, while the epigenetic code controls elaboration of the blueprint (9, 20). The two examples that will be discussed are histone modification and DNA methylation.

1.3.6.1 Histone modification

Histones are highly alkaline proteins that are responsible for packaging DNA into nucleosomes, which are the basic units of chromatin. It consists of a 146-bp loop of DNA wrapped over an octamer of core histones (H2A/H2B dimers and H3/H4 tetramer). Modifications of histones include acetylation, methylation, phosphorylation and ubiquitinylation, which are reversible and can change inactive heterochromatin to active euchromatin, and vice versa. The modifications occur at the N-terminal or C-terminal domains of the histones. The enzymes involved in the modifications include histone acetyltransferases (HATs), histone deacetylases (HDACs) and histone methyltransferases. It is the balance of these enzymes and related proteins that allow for normal functioning of cells (20).

The modifications that have been identified in aberrant TSGs in CRC include deacetylation and methylation of lysine 9 residue of core histone H3. Acetylation results in gene expression whereas methylation is associated with gene silencing. These modifications are mutually exclusive. A pattern of change in core histone H4 (monoacetylation from lysine 16 and trimethylation from lysine 20) has also been proposed as a universal marker for malignant transformation in CRCs. Histones may also be replaced with variants e.g. H3.3 for the canonical H3 and H2A.Z variant for H2A (20).

1.3.6.2 DNA methylation

The enzymatic modification of cytosine bases in DNA to form 5-methylcytosine is the most widely investigated and understood mechanism of epigenetic modification.

In the human genome, methylation of cytosine residues commonly occurs at the CpG islands but can occur at 5’-CA-3’ or 5’-CT-3’ residues. Approximately 70% of all CpG dinucleotides in the human genome are heavily methylated, and the remaining CpG rich regions are present in promoters and first exons of genes. Hypermethylation of CpG islands have been discussed in the CIMP pathway.

In addition to hypermethylation, global DNA hypomethylation can also occur, first noted in the 1970s, and found in lesions across the spectrum from hyperplastic polyps, adenomatous polyps to carcinomas. Highly repeated DNA sequences include short interspersed transposable elements (SINEs) and long interspersed transposable elements (LINEs) which are seen in approximately 40% of human DNA and are normally methylated but become hypomethylated in CRC. The epigenetic change increases chromosome breakage, disrupting gene function or activating oncogenes (20).

1.4 Hereditary colorectal cancer syndromes

Table 1: Genetics of inherited colorectal tumour syndromes (adapted from reference 5)

Syndrome	Features	Gene(s)
FAP	Multiple adenomatous polyps (>100) and carcinomas of the colon and rectum	*APC*
Gardner	Same as FAP with desmoids tumours and mandibular osteomas	*APC*
Turcot	Adenomatous polyposis, colorectal cancer and brain tumours, usually glioblastomas	*APC*, *MLH1*, *PMS2*
Attenuated FAP	Less than 100 adenomatous polyps	*APC* (5'mutations)
HNPCC	High risk of colorectal cancer without extensive polyposis. Increased risk of extraintestinal malignancies.	*MSH2*, *MLH1*, *PMS2*, *MSH6*, *GTBP*
Peutz-Jeghers	Hamartomatous polyps throughout the GIT with mucocutaneous pigmentation. Risk of GI and non-GI cancers are increased.	*LKB1* (*STK11*)
Cowden	Multiple hamartomatous polyps involving the GIT. Increased risk of breast, thyroid and uterine cancers. Risk of GI cancer unclear	*PTEN*
Juvenile polyposis syndrome	Multiple hamartomatous / juvenile polyps. Variable risk in colon and stomach cancers	*DPC*, *BMPR1a* *PTEN*
MAP	Multiple adenomatous GI polyps, often with *KRAS* mutations	*MYH*

The hereditary syndromes arising from germline mutations account for less than 5% of colorectal carcinomas (8) (**Table 1**). Familial adenomatous polyposis and Lynch syndrome account for the majority of syndrome associated colorectal cancer.

1.4.1 Familial Adenomatous Polyposis (FAP)

Familial adenomatous polyposis is an autosomal dominant disorder which affects 1 in 12 000 individuals and accounts for 0.5-1% of all CRCs (5, 8). It involves a germline mutation in the *APC* gene, a component of the Wnt pathway. The lifetime incidence of CRC is 100% and the mean age of diagnosis is 36 years. Hence, prophylactic removal of the colon early in the patient's life is the treatment of choice. An attenuated form of FAP exists, that is less aggressive and characterised by fewer polyps and later onset of disease (5).

Most of the *APC* mutations associated with FAP are frameshift or nonsense mutations that lead to a premature truncated protein. Frameshift mutations are more common than nonsense mutations (3:1 ratio). In addition, the distribution of the mutations is different in the attenuated variant of FAP. Mutations in codons 1250 and 1464 are associated with profuse polyposis in contrast to mutations that involve the N-terminal of codon 157 or near the C-terminus which lead to attenuated polyposis. Two hotspots, codons 1061 and 1309, account for 35% of mutations identified in FAP (5).

Extracolonic manifestations include, osteomas, dental anomalies, cutaneous manifestations, hypertrophy of the retinal pigment epithelium, desmoids tumours and extracolonic carcinomas. Colonic polyposis together with brain tumours (glioblastomas) occur in Turcot syndrome. Polyposis together with epidermoid cysts, desmoid tumours and osteomas are seen in Gardner syndrome (8).

1.4.2 Lynch syndrome (Hereditary non-polyposis colorectal cancer)

This syndrome Is discussed under the MSI pathway.

1.4.3 MYH-associated polyposis (MAP)

This is an autosomal recessive disorder resulting from biallelic mutations in the *MYH* gene, located on chromosome 1p35. It is a base excision repair gene following oxidative damage to DNA, thereby leading to increased GC-to-AT base pair transversion mutations when defective.

The patients with MAP generally have between 10 to 100 colonic polyps, with CRC developing in approximately 65% of patients, but at a later age (mean age 45 years) than FAP. Gastric and small intestinal polyps can occur but other extraintestinal manifestations are less common than FAP (8).

1.4.4 Hamartomatous polyposis syndromes

1.4.4.1 Peutz–Jeghers syndrome

This is an autosomal dominant syndrome characterized by hamartomatous polyps throughout the gastrointestinal tract as well as mucocutaneous melanin deposition. There is also an increased risk of adenomatous polyps with a predisposition to gastric, intestinal and extra-intestinal cancers. There is a germline mutation in the *STK11/LKB1* tumour suppressor gene, which leads to increased activation of the mammalian target of rapamycin (mTOR) pathway, a major regulator of cell growth (8).

1.4.4.2 Juvenile polyposis syndrome

This autosomal dominant syndrome is associated with numerous juvenile polyps throughout the GIT and is ten times less common than FAP. These patients are at increased risk of developing GIT and pancreatic cancers. In 50-60% of patients *Smad4* or *BMPR1A* mutations are identified (5, 8).

1.4.4.3 Cowden syndrome

This syndrome sometimes referred to as PTEN hamartoma syndrome involves hamartomas consisting of derivatives of all three germ layers. Hamartomatous polyps can occur throughout the GIT. In addition, these patients are at increased risk of extraintestinal malignancies, including breast, thyroid and endometrial carcinoma (8).

1.5 Regulation of the cell cycle

Dividing eukaryote cells pass through a series of stages known collectively as the cell cycle: two gap phases (G1 and G2); an S (synthesis) phase, where the genetic material is duplicated; and the M phase in which mitosis occurs. In order to undergo replication cells have to move from the G0 or “resting phase” of the cell cycle to the G1 phase. The progression through the various phases of the cell

cycle are driven by cyclins, cyclin dependent kinases (CDK) and their inhibitors (e.g. p21 and p27) (29, 30).

1.5.1 Cyclins D1 and E

G1 phase cyclins include cyclin D and E, activation of which results in the proliferation of cells. Cyclin D1 appears in the mid G1 phase and is no longer detectable in the S phase. It acts together with CDK4 and 6, that form active complexes that promote G1 to S phase progression through the inactivation of retinoblastoma protein (pRb) via phosphorylation. In the G1 phase of the cell cycle, cyclin D1-CDK4 complex releases active E2F by phosphorylating pRb, which in its hypophosphorylated form is bound to E2F. Free active E2F in turn activates the transcription of genes important in S phase progression (30). Usually following the mutation of the *APC* gene, cyclin D1 is activated via the Wnt-β-catenin pathway discussed earlier. In addition, recent studies have shown that cyclin D1 regulates other transcription factors (31). It is an unstable molecule and is degraded through the ubiquitin-proteasome pathway.

Cyclin E associates with CDK2, resulting in diminished requirement for growth factors, accelerating the G1 phase, which suggests that cyclin E is responsible for G1 phase progression. Cyclin E/CDK2 complexes phosphorylate pRb and thus release more active E2F, which in turn stimulates, amongst others, its own transcription (5).

Usually pRb requires phosphorylation of cyclin D1/CDK4 prior to the action of cyclin E/CDK2. However, some studies demonstrated that malignant cells use one cyclin molecule to phosphorylate pRb preferentially over the other, making the other cyclin redundant (29).

High copy amplification of the cyclin E gene is found in only 5% of CRCs. This is due to inactivation of the *FBXW7* gene, which encodes an F-box protein. This protein forms part of the SCF ubiquitin ligase complex and may also regulate other oncogenic factors like c-myc, c-jun and Notch proteins. It is inactivated in approximately 20% of CRCs (5).

1.5.2 Cell cycle inhibitors – p21, p27 and p53

The two main classes of CDK inhibitors are the Cip/Kip and the INK4/ARF families. There are three main components of the Cip/Kip family: p21, p27 and p57 that inactivate the complexes between cyclins and CDKs (30).

p27 responds to antimitogenic signals like the growth suppressor, transforming growth factor β (TGFβ). It regulates G0 to S phase in the cell cycle and inhibits cyclin E/CDK2 and regulates function of cyclin D1/CDK4 and cyclin D1/CDK6. Regulation of p27 is achieved by the ubiquitin mediated protein degradation. Interestingly, the F-box protein (mentioned above), SKP2 has been identified as the substrate recognition component that targets p27 for degradation (32). Studies in CRCs show mixed results regarding prognosis, with 9 out of 13 retrospective multivariate analyses demonstrating a 1.43-11 fold increase in disease recurrence or death with reduced expression of p27 (32).

p21 is under the control of p53, a tumour suppressor that is mutated in a large proportion of human cancers. p21 has a pro-apoptotic as well as an anti-apoptotic role and may also facilitate invasion and metastasis, through p21-activated kinase 1 (PAK-1). Cellular senescence and aging of stem cells have also been linked to p21. Most studies show no prognostic value of p21 in CRC (33).

p53 functions as a transcriptional regulator of genes that encode proteins which function as cell cycle check points. It promotes apoptosis, cell cycle arrest and antiangiogenic pathways in cells with wild type p53 function, secondary to DNA stresses like strand breakage, telomere erosion and hypoxia. It is negatively regulated by MDM2, E3 ubiquitin ligase and MDM4 which promote ubiquitination (5).

p53 is the target of 17p LOH (as discussed above) and is seen in 70% of CRCs with the remaining *p53* allele demonstrating a somatic mutation. The mutations include missense defects, often occurring at codons 175, 245, 248, 273 and 282. Only a few sporadic CRCs and most adenomas lack 17p LOH, implying that LOH is closely associated with adenoma-carcinoma transition (5).

1.6 c-myc

The *c-myc* proto-oncogene encodes a transcription factor that regulates genes involved with cell cycle progression, cell adhesion, survival and cellular metabolism. It is expressed in numerous human malignancies. High copy amplification of the *c-myc* gene is seen in 5-10% of CRCs while moderate increases are seen in approximately 30% of CRCs.

Activation of the *c-myc* gene expression is central to signal transduction via the APC tumour suppressor protein which negatively regulates β-catenin. β-catenin is able to directly activate c-myc expression, so that when APC is inactivated, activation of β-catenin results (5).

1.7 Matrix Metalloproteinase 7 (MMP-7)

The MMP family of zinc-dependent proteinases are composed of 23 members, and based on their structure are divided into secreted-type MMP and membrane-anchored MMP. MMPs are thought to be responsible for pathological tissue destruction due to the ability to catabolize ECM components. However, they are also able to break down growth factors, cytokines and their receptors. Tissue inhibitors of metalloproteinases (TIMP) of which there are various subtypes, are the major physiological inhibitors of MMPs (34). The MMPs are overexpressed in human cancers and are associated with an aggressive malignant phenotype and adverse prognosis (35).

MMP-7 (Matrilysin, pump-1) is the smallest known member of the MMP family and its protein product was first identified in the culture medium of a human rectal carcinoma cell line. MMP-7 degrades various ECM proteins including proteoglycans, fibronectin, entactin, laminin, gelatin, type IV collagen, type X collagen and elastin (34, 36). It therefore supports tumourigenesis and progression. It also has the capability of degrading MMP-2 and MMP-9. It is elevated in numerous primary cancers including breast, ovary, endometrial, lung, stomach carcinomas and oesophageal squamous cell carcinomas. In addition, MMP-7 is a target gene of the Wnt signalling pathway described above.

MMP-7 is overexpressed both in dysplasia and CRCs and MMP-7 mRNA expression correlated with the stage of CRC progression in a study by Polistena *et al* (36). In CRC hepatic metastases, a study has shown variable degrees of MMP-7 overexpression in the metastatic carcinoma when compared to normal liver tissue (37). In addition, studies have shown an increased expression of MMP-7, following neoadjuvant radiotherapy for rectal cancer, suggesting a role for tissue remodelling after injury (36).

1.8 Clinical features and special investigations

Usually patients experience a change in bowel habits, constipation being common, because of the obstructive effect of the mass, especially in the left colon. Tumours are also prone to ulceration, and haematochezia and anaemia can be presenting features. Abdominal distension, bowel obstruction and perforation can be late presentations of the tumour, sometimes encountered as a surgical emergency. Tenesmus and rectal bleeding are symptoms seen in lesions in the rectosigmoid area. Non-specific symptoms include abdominal pain, fever and weight loss (2).

Imaging studies allow for detection and staging of the tumour. Computer assisted tomography (CT) has largely replaced barium enemas in first world settings. Magnetic resonance imaging (MRI), cross sectional imaging by CT and transrectal ultrasound can also estimate depth of invasion and metastases. Colonoscopy allows observation of the entire mucosal surface, identifying the lesion and other benign polyps. Biopsy or therapeutic removal of superficial adenomas and carcinomas can also achieved by snare polypectomy, mucosal resection or submucosal dissection (2).

1.9 Pathology

1.9.1 Macroscopic pathology

The majority (65%) of the carcinomas occur in the rectosigmoid region of the large bowel followed by the ascending colon (2). The molecular pathology described above also has site differences. Tumours with high levels of microsatellite instability (MSI-H) and CIMP-H are frequently located in the caecum, ascending colon and transverse colon (2, 9).

The majority of tumours are ulcerating producing a degree of stenosis and obstruction. The carcinomas can be exophytic/fungating with predominantly intraluminal growth or annular with circumferential involvement of the wall resulting in stenosis and dilatation. The diffusely infiltrative/linitis plastica pattern is extremely rare. The carcinomas in the ascending and transverse colons tend to be exophytic and those in the descending colon, annular and constricting (2).

1.9.2 Microscopic pathology

The typing of CRCs is based on the predominant histomorphology of the tumour. The subtypes according to the WHO classification include (2):

- Adenocarcinoma, NOS
- Adenocarcinoma special types:
  Cribriform comedo type adenocarcinoma, medullary carcinoma, micropapillary carcinoma, mucinous adenocarcinoma, serrated adenocarcinoma and signet ring cell adenocarcinoma.
- Adenosquamous carcinoma
- Spindle cell carcinoma
- Squamous carcinoma
- Undifferentiated carcinoma

The vast majority of the carcinomas are adenocarcinomas and its variants (>90%). These carcinomas are positive with cytokeratin 20, negative with cytokeratin 7 and express the transcription factor CDX2.

The tumour is designated mucinous when more than 50% of the lesion is composed of pools of extracellular mucin (2, 38). If the mucinous component forms less than 50% of the tumour it is categorized as having a mucinous component. The epithelial component can include acinar structures, layers of tumour cells, single tumour cells or signet ring cells. Most mucinous adenocarcinomas are MSI-H and low grade. The mucinous carcinomas that are MSS or MSI-L behave as high grade lesions (2).

The signet ring variant of CRC is defined by the presence of >50% signet ring cells. Similar to the mucinous tumours, those that are MSI-H are low grade while those that are MSS and MSI-L behave aggressively (2).

The medullary variant is rare and displays a sheet-like growth of malignant cells, with a prominent intraepithelial lymphocytic infiltrate (2, 38). The cells have vesicular nuclei with prominent nucleoli. This variant is almost exclusively seen in MSI-H tumours with a favourable prognosis.

The serrated variant has architecture similar to a serrated polyp and may have mucinous, cribriform and trabecular areas. The molecular alterations associated with this histological type include MSI-L, MSI-H, *BRAF* mutations and CpG island methylation. The cells have clear to eosinophilic cytoplasm, with vesicular nuclei with limited or absent tumour necrosis (<10% necrosis) (39).

The cribriform comedo-type adenocarcinomas have large cribriform glands with central necrosis similar to that seen in the breast. These tumours are usually MSS with CpG island hypermethylation (2).

The micropapillary adenocarcinomas have clusters of tumour cells with stromal spaces that mimic vessels (pseudovascular). These malignant cells are usually positive for MUC1 on immunohistochemistry (2).

The adenosquamous carcinoma has both adenocarcinoma and squamous cell carcinoma components. A pure squamous cell carcinoma is extremely rare. Spindle cell carcinoma is usually a biphasic tumour with a spindle cell/sarcomatoid component that shows patchy immunoreactivity with cytokeratins. Undifferentiated carcinomas lacking morphological, immunohistochemical and molecular evidence of differentiation beyond that of an epithelial tumour are also rare.

1.10 Aims and Objectives

Aim: To determine the comparative histopathology and immunohistochemical expression of cell cycle regulators and MMR proteins in different cohorts of CRC patients.

The specific objectives of the study were:

- To characterise the macroscopic and microscopic pathology in three cohorts of patients'. The cohorts include patients' with CRCs that are < 50 years and mutation negative, < 50 years and mutation positive and more than 50 years (sporadic).
- To investigate the immunoexpression of the cell cycle regulators (p21, p27, p53, c-myc, cyclin D1 and cyclin E) and MMP-7 in each cohort
- To compare the immunoexpression of each marker between cohorts.
- To correlate the immunoexpression of each marker with tumour type, stage and grade.

CHAPTER 2

2. MATERIALS AND METHODS

2.1 Ethics approval

Ethics approval was obtained from the University of Cape Town (UCT) Faculty of Health Sciences Human Research Ethics Committee (Reference number 050/2007). The research proposal was approved by the Department of Clinical Laboratory Sciences Research Committee and the Faculty of Health Sciences Postgraduate Committee at UCT. Funding for this study was obtained from the National Health Laboratory Service Research Trust.

2.2 Sample selection

An electronic search of the laboratory information system database of the Division of Anatomical Pathology, National Health Laboratory Service–Groote Schuur hospital was performed. The search included resection specimens of colorectal carcinoma occurring in patients older than 50 years between January 1996 and December 2013. In addition, mutation positive and mutation negative resection cases, confirmed by the Division of Human Genetics were also provided. The following mutations were tested:

MSH2 EXON8 c.1340_1341insGG,

MSH2 EXON7 c.1221_1222delCT,

MSH2 EXON3 c.387_388delTC,

MSH2 EXON15 c.2502_2508delTAATTTC and

MLH1 Exon13 c.1528C>T.

Archived stained slides of the cases were reviewed. The diagnosis in each case was confirmed and the morphological data recorded (including lymph node status) and further information from the reports, also recorded. During review of the haematoxylin and eosin (H&E) sections a suitable block of tumour tissue was selected for immunohistochemical staining. The block containing predominant tumour, and minimal haemorrhage and necrosis was selected. In most cases adjacent morphologically normal mucosa was present. Tissue blocks were retrieved from the archives of Division of Anatomical Pathology and 4µm

sections were cut for immunohistochemistry. The cases were allocated random case numbers after retrieval of reports and patient names were not used.

A retrospective, comparative study between three different patient cohorts was undertaken. The three patient cohorts included:
Cohort 1: Patients younger than 50 years of age without known mismatch repair gene mutations (n=17).
Cohort 2: Patients younger than 50 years of age with confirmed mismatch repair gene mutations (n=15).
Cohort 3: Patients older than 50 years of age with sporadic colorectal carcinoma (n=28).

2.3 Antibodies

Table 2: Primary antibodies

Primary antibody (Mouse/Rabbit) (mono/poly)	Clone	Supplier	Antigen Retrieval	Dilution	Incubation time	Positive control
p21 (M)(m)	Waf1	Novocastra	EDTA pH8	1:30	1 hour	tonsil
p27 (M)(m)	1B4	Novocastra	EDTA pH8	1:20	1 hour	tonsil
p53 (M)(m)	DO-7	Dako	EDTA pH8	1:100	1 hour	tonsil
Cyclin D1 (R)(m)	SP4	Abcam	EDTA pH8	1:400	1 hour	Mantle cell lymphoma
Cyclin E (M)(m)	13A3	Dako	EDTA pH8	1:60	1 hour	Placenta
c-myc (M)(m)	9E11	Novocastra	EDTA pH8	1:200	1 hour	Tonsil
MMP7 (M)(m)	ID2	Abcam	Tris EDTA	1:50	1 hour	Colon
MLH1 (M)(m)	M1	Ventana	EDTA pH8	prediluted	20 min	Colon
MSH2 (M)(m)	FE11	Calbiochem	EDTA pH8	1:250	1 hour	Colon
MSH6 (M)(m)	44/MSH6	BD Bioscience	EDTA pH8	1:500	1 hour	Colon

(Key: M – mouse, R-rabbit, m – monoclonal and p - polyclonal)

A negative reagent control, in which the primary antibody was replaced by PBS (buffer), was run simultaneously for each antibody (**Table 2**).

PMS2 was not included in the immunohistochemical analysis of MMR proteins since MLH1 forms heterodimers with PMS2 and MLH1 mutations are more common. Another important factor was the limited funding and the need to work within a tight budget.

2.4 Immunohistochemistry Method

Table 3: Kits used for the IHC protocol

KITS	SUPPLIER
Envision HRP System Labelled Polymer Anti-mouse	Dako - CA, USA
Envision HRP System Labelled Polymer Anti-rabbit	Dako - CA, USA
Liquid DAB + Substrate chromogen system	Dako - CA, USA

- Four micron tissue sections were cut from paraffin wax embedded tissue blocks, picked up onto Histobond slides (Marienfeld-Germany) and heat fixed on a hotplate for 10-15 minutes.
- Sections were dewaxed through xylene, cleared in ethanol and rehydrated in water.
- Endogenous peroxidase activity was blocked by treating the slides with a 3% hydrogen peroxide (H_2O_2) solution for 10 minutes.
- Slides were washed well in water.
- Antigen retrieval was performed by pressure-cooking slides in either Tris-EDTA (pH9) (MMP7 only) or EDTA (pH8) (p21, p27, p53, cyclin D and E, c-myc, MLH1, MSH2 and MSH6) for 1 minute 30 seconds at full pressure.
- The sections were allowed to cool for 10 minutes.
- This was followed by washing in tap water.
- Thereafter, slides were rinsed with phosphate buffered saline solution (PBS pH 7.6) (Oxoid-Hampshire, England).
- Non-specific binding was blocked by treating slides with a 5% goat serum solution (DAKO- Denmark).

- Serum was then drained off and sections were incubated with primary antibody at room temperature at specified times and dilutions (**Table 2**).
- The slides were then washed well with PBS buffer.
- This was followed by incubation with DAKO Envision labelled Polymer, HRP (DAKO- USA) for 30 minutes at room temperature (**Table 3**).
- Sections were washed well with PBS buffer.
- Positivity was developed by applying the chromogenic substrate 3.3 diaminobenzidine (DAB) (DAKO- USA) for 5-10 minutes.
- Slides were washed in running tap water and counterstained with Mayers haematoxylin for approximately 3 minutes.
- After washing in running tap water, sections were blued in ammoniated water. Finally, the slides were then dehydrated through alcohols, cleared with xylene and mounted with Entellan (MERCK- Germany) and cover-slipped.

2.5 Microscopic assessment of immunohistochemical stains

The proportion of positive cells as well as the intensity of staining was recorded for each stain and for each case. The location of staining, nuclear and/or cytoplasmic was recorded.

Proportion score	**Proportion of cells staining**
0	Staining in <5% of tumour cells
1+	Staining in 6-25% of tumour cells
2+	Staining in 26-50% of tumour cells
3+	Staining in 51-75% of tumour cells
4+	Staining in >75% of tumour cells

Intensity score	**Intensity of staining**
0	No staining
1+	Weak staining
2+	Intermediate (moderate) staining
3+	Strong staining

After determining the intensity and percentage of cells staining, the H-score (HS) was calculated, using the multiplication of the two scores, similar to that used in breast carcinomas (HS = Intensity of stain x Proportion of cells staining) (40). However, for the MSI markers, i.e. MLH1, MSH2 and MSH6, the tumours were either designated positive or negative. According to the College of American Pathologists guidelines, weak and patchy nuclear staining in the tumour was also designated positive. Expression of loss, was only made if the internal control nuclei (stromal cells, inflammatory cells and epithelial cells) were positive, else the case was designated indeterminate/non assessable (38).

2.6 Clinical data and pathological features

The clinical data recorded were age and gender.

2.6.1 Macroscopic Features

- Anatomic site of the tumour: rectum, left colon (distal to the splenic flexure and proximal to the rectum) or right colon.
- The greatest dimension of the tumour, shape/configuration (annular/constricting, exophytic, mucinous) were recorded from the pathology reports.
- The tumours were staged as per the TNM staging system.
- Presence of colonic polyps was documented, including type, anatomical location and number.
- Synchronous tumours are defined as two or more primary tumours in the excision specimen.

2.6.2 Microscopic Features and Study Definitions

The following variables were assessed:

- Tumour grade (low or high)
- Special variant of adenocarcinoma if present as per the 2007 WHO classification
- Mucinous differentiation was also noted and defined as any tumour that had a mucinous component between 5% and 100%.

- Tumour infiltrating lymphocytes (TILs) were defined as a lymphoid infiltrate where more than 30 lymphocytes were present in 10 HPFs. The exact nature of the lymphocytes (B-cell vs T-cell) was not determined, although it is noted in studies that TILs are usually cytotoxic T lymphocytes (41, 42).
- Crohn's reaction was defined as a lymphoid reaction with lymphoid follicles, with and without germinal centres.
- Tumour margins as defined by Jass (43). An infiltrative margin does not allow a clear demarcation of the tumour invasive front. On low power examination, there is dissection of malignant glands through structures of the bowel wall. In contrast, an expanding/pushing border is seen when there is clear demarcation of the tumour invasive front. The malignant glands form a "round" circumscribed front.

The tumours were classified according to the 2007 WHO classification (2).
Adenocarcinoma, NOS
Special variants:

Cribriform-comedo type,
Medullary carcinoma,
Micropapillary carcinoma,
Mucinous adenocarcinoma,
Serrated adenocarcinoma and
Signet ring cell adenocarcinoma.

Adenosquamous carcinoma
Spindle cell carcinoma
Squamous carcinoma
Undifferentiated carcinoma

2.6.3 Grading of tumours

In the past, CRCs were graded as well-, moderately, poorly and undifferentiated based on the percentage of gland formation and cytomorphology. However, adenocarcinomas are now better classified as "low grade" and "high grade", due to the similar behaviour of well- and moderately differentiated tumours (2, 38). In addition, a two-tiered grading system allows for greater reproducibility.

Low grade - Greater than or equal to 50% gland formation.
High grade - Less than 50% gland formation.

The grading system only applies to "Adenocarcinoma, NOS" (2). The other variants carry their own prognostic significance and grading does not apply (2). In addition, when the tumour is heterogeneous, the grading is based on the least differentiated component.

2.6.4 Staging of tumours

CRCs in this study were staged according to the 2009 7th edition TNM tumour staging system, which characterises the extent of the lesion according to the parameters of depth of invasion of the primary tumour (T), regional lymph node involvement (N) and distant metastatic spread (M) (2).

T - Primary tumour

Tis	Carcinoma in situ: intraepithelial or invasion of lamina propria
pT1	Tumour invades submucosa
pT2	Tumour invades muscularis propria
pT3	Tumour invades subserosa or into non-peritonealised pericolic or perirectal tissues.
pT4	Tumour perforates visceral peritoneum and/or directly invades other organs or structures
T4a	Tumour perforates visceral peritoneum
T4b	Tumour directly invades other organs or structures

N - Regional lymph nodes

pN0	No regional lymph node metastasis
pN1	Metastasis in 1-3 regional lymph nodes
N1a	Metastasis in 1 regional lymph node
N1b	Metastasis in 2 to 3 regional lymph nodes
N1c	Tumour deposit(s), i.e. satellites, in the subserosa, or in non-peritonealised pericolic or perirectal soft tissue without regional lymph node metastasis

pN2 Metastasis in 4 or more regional lymph nodes

M - Distant metastasis

M0 No distant metastasis

M1 Distant metastasis present

M1a Metastasis confined to one organ

M1b Metastasis in more than one organ or peritoneum

Stage Grouping

Stage	**T**	**N**	**M**
Stage 0	Tis	N0	M0
Stage I	T1, T2	N0	M0
Stage II	T3, T4	N0	M0
Stage IIA	T3	N0	M0
Stage IIB	T4a	N0	M0
Stage IIC	T4b	N0	M0
Stage III	Any T	N1, N2	M0
Stage IIIA	T1, T2	N1	M0
	T1	N2a	M0
Stage IIIB	T3, T4a	N1	M0
	T2, T3	N2a	M0
	T1, T2	N2b	M0
Stage IIIC	T4a	N2a	M0
	T3, T4a	N2b	M0
	T4b	N1, N2	M0
Stage IVA	Any T	Any N	M1a
Stage IVB	Any T	Any N	M1b

2.7 Statistical analysis

Statistical analysis was performed using Stata version 13.1 (StataCorpLP, 4905 Lakeway Drive, College Station, TX 77845, USA). The distribution of continuous data was assessed graphically and using the Shapiro Wilk test. Normally distributed data were summarised using means and standard deviations

and analysed using parametric methods (independent or paired t-tests as appropriate, or analysis of variance for the comparison of 3 or more means). Skewed data were summarised using medians and interquartile ranges and analysed using non-parametric methods (Mann-Whitney-U-tests or Wilcoxon signed rank tests for paired data as appropriate, or the Kruskal Wallis test for the comparison of 3 or more medians). A P-value of < 0.05 was considered statistically significant for analyses. An adjustment for multiple comparisons i.e. 3 pairs of medians compared, then the P-value for statistical significance was $0.05 \div 3 = 0.017$ i.e. $P<0.017$.

CHAPTER 3

3. RESULTS

3.1 Clinical and Pathological characteristics

Table 4: Characteristics of colorectal carcinomas in the three cohorts

	Cohort 1 (n=17)	**Cohort 2 (n=15)**	**Cohort 3 (n=28)**
Mean age, years	38.2	42.8	67
Male , n (%)	9 (53)	8 (53)	15 (54)
Female, n (%)*	8 (47)	7 (47)	13 (46)
Synchronous tumours n (%)	1 (6)	3 (20)	0
Site (includes synchronous tumours), n (%):			
Left	4 (22)	3(17)	15(54)
Right	8 (45)	14(78)	4(14)
Rectum	6 (33)	1(5)	9(32)
Histologic type, n (%)			
Adenocarcinoma, NOS	15 (88)	8 (54)	22 (78)
Differentiation, n (%):			
Low grade	14 (93)	9 (100)	21 (95)
High grade	1 (7)	0	1 (5)
Adenocarcinoma – special types			
Signet ring cell	0	1 (7)	1 (4)
Medullary	0	1 (7)	1 (4)
Serrated	0	0	1 (4)
Mucinous	2 (12)	5 (33)	3 (10)
Mucinous differentiation, n (%) **	2 (12)	9 (60)	9 (32)
Margins, n(%):			
Infiltrating	9 (53)	1 (7)	16 (57)
Expanding	8 (47)	14 (93)	12 (43)
Tumour infiltrating lymphocytes, %:	1 (6)	5 (33)	2 (7)
Crohns-like reaction, %:	5 (29)	8 (53)	12 (43)
Background Adenomatous Polyps, %:	2 (12)	5 (33)	1 (14)
pT stage, n(%):			
T1	2 (12)	0	0
T2	2 (12)	2 (13)	4 (14)
T3	8 (47)	7 (47)	17 (61)
T4	5 (29)	6 (40)	7 (25)
pN stage,n(%):			
N0	7 (41)	7 (47)	14 (50)
N1	8 (47)	3 (20)	10 (36)
N2	2 (12)	5 (33)	4 (14)
Stage, n(%):			
I	2 (12)	1 (7)	4 (14)
II	5 (29)	6 (40)	10 (36)
III	9 (53)	8 (53)	13 (46)
IV	1 (6)	0	1 (4)

* n = number of cases, ** Mucinous differentiation is defined as any tumour with 5-100% mucinous component

3.1.1 Cohort 1 – Mutation negative (n = 17)

The patients' ages in this cohort ranged from 25 to 49 years, with a mean age of 38.2 years. The gender distribution was near equal, with 8 females and 9 male patients. The tumours were frequently seen in the right colon (45%), followed by the rectum (33%) and then the left colon (22%). The majority of the adenocarcinoma, NOS tumours was low grade (93%) and mucinous differentiation was seen in 2 (12%) cases only. A Crohns like reaction was seen in 5 (29%) cases with tumour infiltrating lymphocytes seen in 1 case (6%). Both expanding and infiltrating margins were seen with no marked difference (47% infiltrating vs 53% circumscribed). The majority of the cases were stage III (53%).

3.1.2 Cohort 2 – Mutation positive (n = 15)

Thirteen patients had mutations of the *MLH1* gene and two had mutations of the *MSH2* gene (**Appendix 4**). The patients' ages ranged from 29 to 50 years, with a mean age of 42.8 years. The gender distribution was near equal and similar to cohort 1. The tumours were frequently seen in the right colon (75%). All of the adenocarcinomas, NOS were low grade and 9 (60%) cases demonstrated mucinous differentiation (**Figure 3B**). In addition, tumour infiltrating lymphocytes were most frequent in this cohort (33%) (**Figure 3F**). All except one tumour had an expanding tumour edge (**Figure 3C**) and a Crohns like reaction was present in 53% of cases (**Figure 3E**). Only 3 cases demonstrated synchronous tumours. The majority of the cases were stage III (53%).

3.1.3 Cohort 3 – Sporadic (n = 28)

The patients' ages ranged from 53 to 83 years, with a mean age of 67 years. There were 15 males and 13 females. The tumours were frequently seen in the left colon (54%), followed by the rectum (32%) and then the right colon (14%). Again, as in cohort 1, the adenocarcinomas, NOS were predominantly low grade (95%) (**Figure 3A**). However, mucinous differentiation was seen in 32% of cases. Again both infiltrating (57%) (**Figure 3D**) and expanding margins (43%) were noted. A Crohns like reaction was seen in 43% of cases and tumour infiltrating lymphocytes were infrequent (7%). The majority of the cases were stage III (46%).

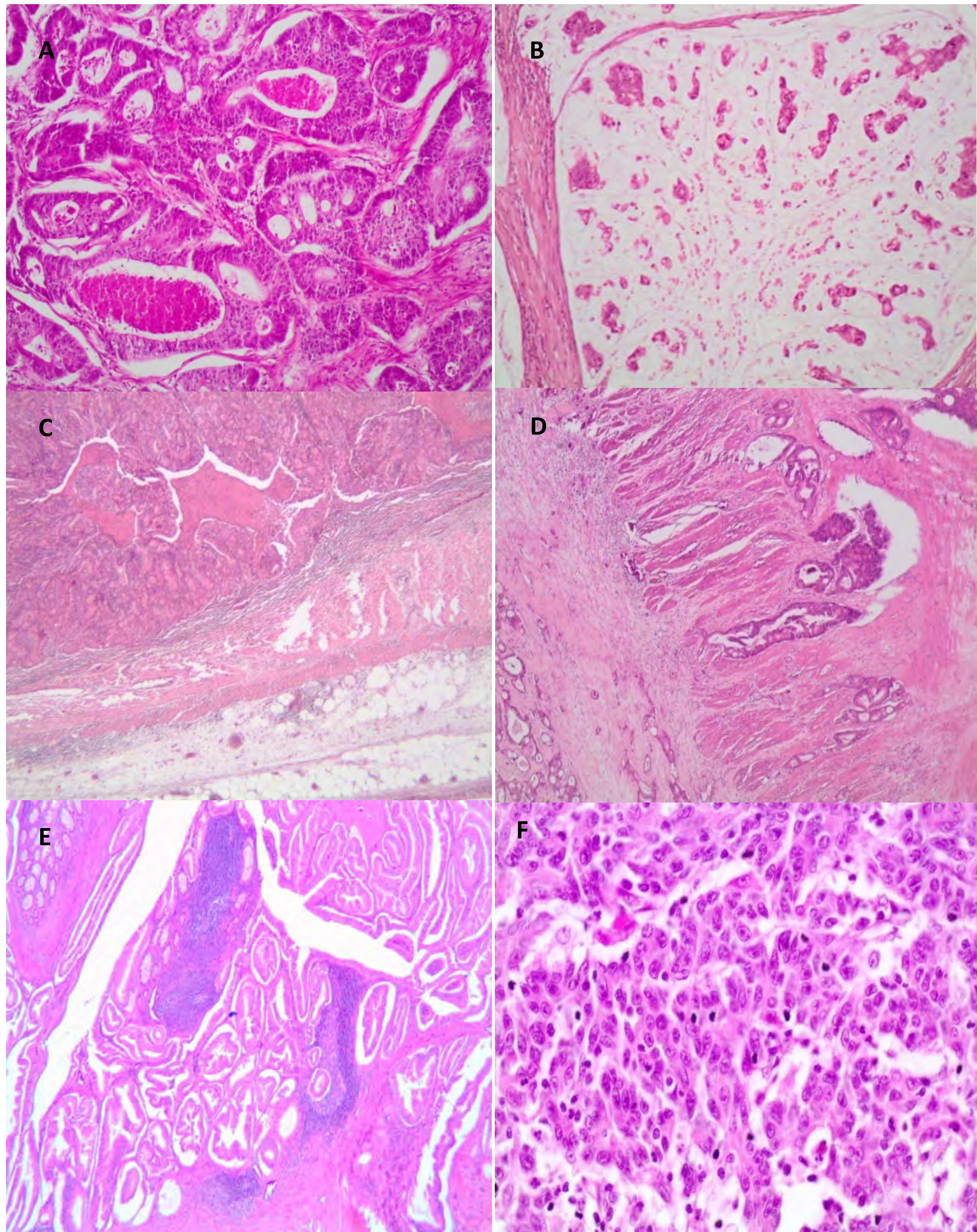

Figure 3: Histopathological features of colorectal adenocarcinomas, (A) Low grade colorectal adenocarcinoma, NOS 40x obj mag, (B) Mucinous adenocarcinoma 40x obj mag, (C) Circumscribed pushing border 2.5x obj mag, (D) Infiltrating border 4x obj mag, (E) Crohns like reaction 2.5 obj mag and (F) Tumour infiltrating lymphocytes 40x obj mag in medullary carcinoma.

3.1.4 Immunohistochemical analysis

3.1.4.1 p21

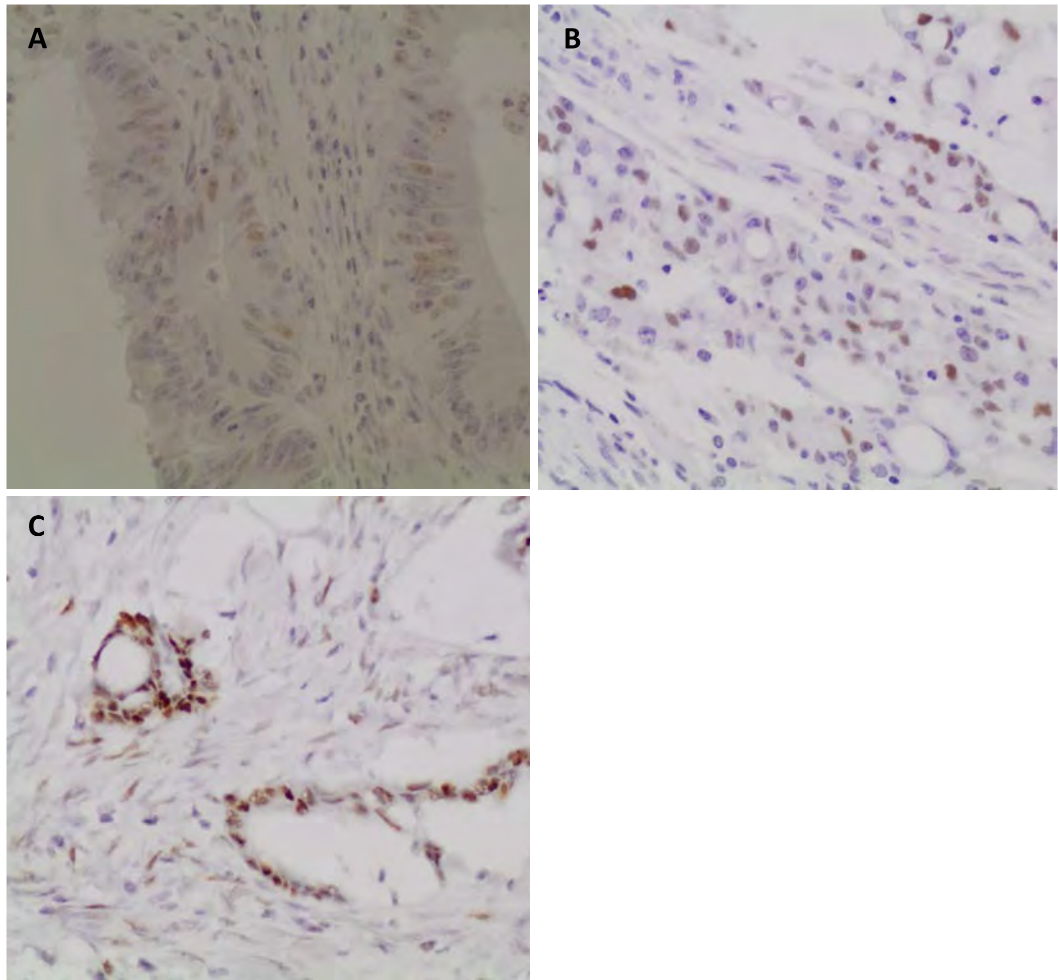

Figure 4: p21 immunohistochemistry: (A) 1+ intensity 40x obj mag, in a low grade adenocarcinoma, (B) 2+ intensity 40x obj mag in the solid component of a mucinous adenocarcinoma and (C) 3+ intensity 40x obj mag in a low grade conventional adenocarcinoma

The expression of p21 was localised to the nuclei and the intensity of staining was variable (**Figure 4A to C**). The stromal cells and the epithelial cells were negative in the majority of cases. The tumours showed marked intratumoural heterogeneity, with very limited expression in the 3 cohorts. Fifty one of the 60 cases were completely negative with only 9 positive cases, of which three were mucinous adenocarcinomas and the rest low grade conventional adenocarcinomas, NOS. Two of the cases were stage I tumours, 2 cases were stage II tumours and 5 were stage III adenocarcinomas.

There was no statistical significance for p21 expression among the three cohorts (Kruskal –Wallis equality-of-populations rank test, *P*=0.5209).
There was no statistical significance in p21 expression and stage in cohort 1 (*P*=0.8811), cohort 2 (*P*=0.3499) and cohort 3 (*P*=0.8235).
There was no statistical significance in p21 expression and tumour type in cohort 1 (*P*=0.4230), cohort 2 (*P*=0.8204) and cohort 3 (*P*=0.5261) .

3.1.4.2 p27

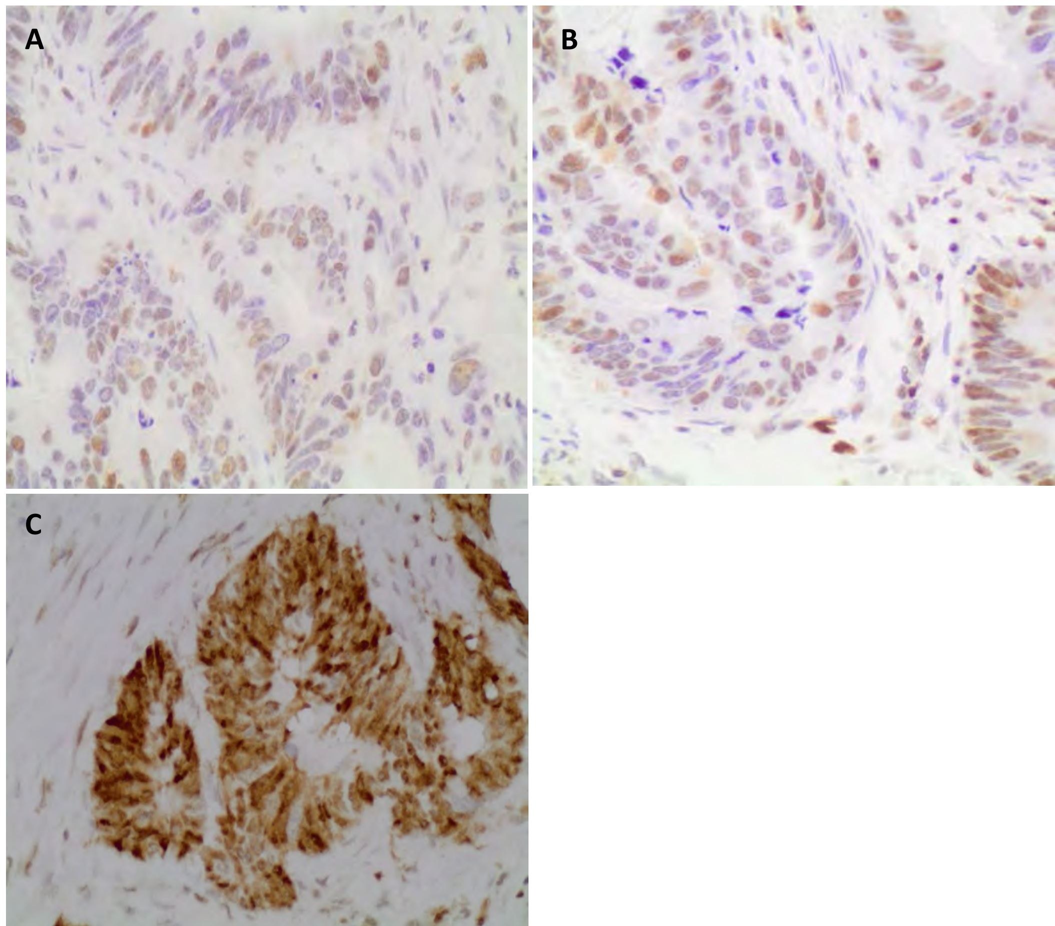

Figure 5: p27 immunohistochemistry: (A) 1+ intensity of nuclear expression 40x obj mag, (B) 2+ intensity of nuclear expression 40x obj mag and (C) 3+ intensity of nuclear and cytoplasmic expression 40x obj mag in 3 low grade conventional adenocarcinomas, NOS

Expression of p27 was localised in the nucleus with occasional cytoplasmic expression (**Figure 5A to 5C**). All of the tumours stained positive for p27. The background lymphocytes were also diffusely positive and the normal mucosal

epithelium showed greater heterogeneity of staining than the invasive tumour. These served as inbuilt controls.

There was no statistical significance for p27 expression among the three cohorts (Kruskal –Wallis equality-of-populations rank test, *P*=0.3048).
There was no statistical significance in p27 expression and stage in cohort 1 (*P*=0.5036), cohort 2 (*P*=0.5415) and cohort 3 (*P*=0.5641) .
There was no statistical significance in p27 expression and tumour type in cohort 1 (*P*=0.7812), cohort 2 (*P*=0.3186) and cohort 3 (*P*=0.4086) .

3.1.4.3 p53

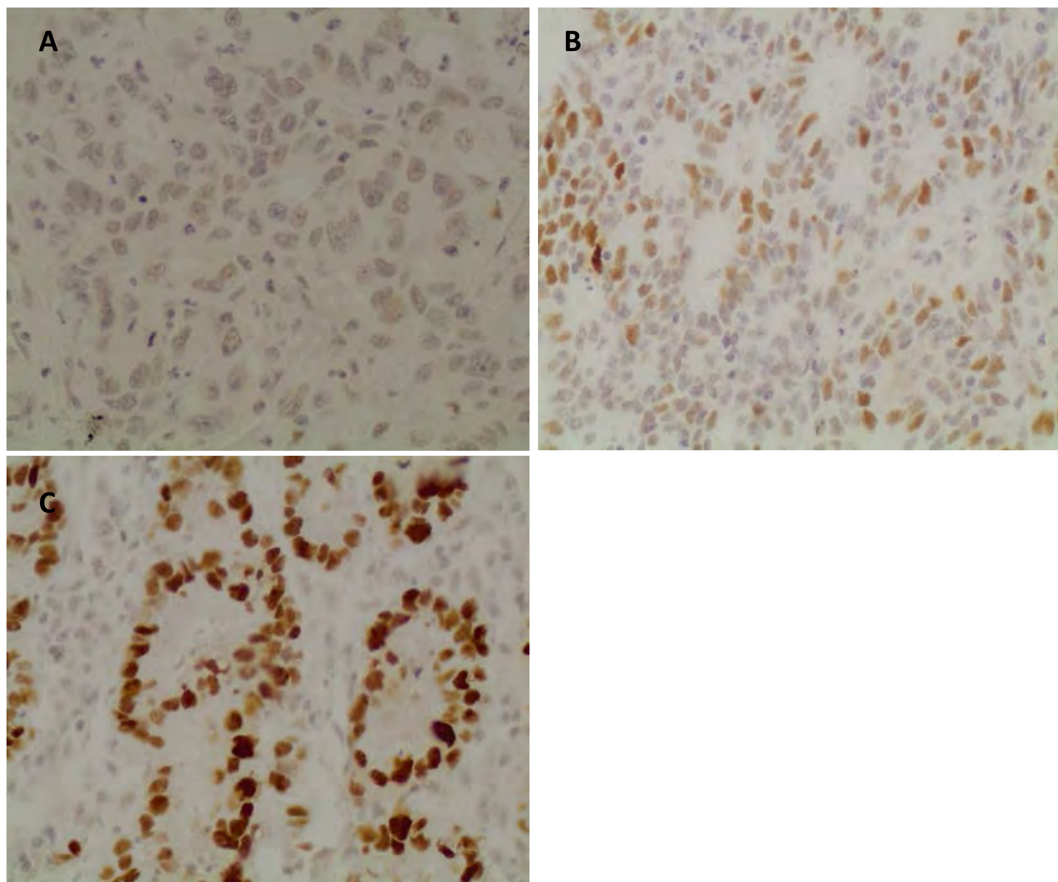

Figure 6: p53 immunohistochemistry: (A) 1+ intensity 40x obj mag, (B) 2+intensity 40x obj mag and (C) 3+ intensity 40x obj mag in 3 low grade conventional adenocarcinomas, NOS

p53 was expressed in the nuclei of the malignant cells (**Figure 6A to 6C**). There was no statistical significant difference in p53 expression among the 3 cohorts (Kruskal –Wallis equality-of-populations rank test, *P*=0.3652). One case was

negative in each of the mutation positive and mutation negative cohorts, while ten cases were negative in the sporadic cohort. Interestingly, the highest scores were recorded in the mutation negative and sporadic cohorts with 9/17 and 13/28 scoring an H-score of 12 respectively, in contrast only 3/15 cases scored a total of 12 in the mutation positive cohort.

There was no statistical significance in p53 expression and stage in cohort 1 (*P*=0.5693), cohort 2 (*P*=0.4046) and cohort 3 (*P*=0.6583).
There was no statistical significance in p53 expression and tumour type in cohort 1 (*P*=0.1934), cohort 2 (*P*=0.7692) and cohort 3 (*P*=0.3580) .

If one considered wild type staining as the expression in less than 50% of cells (i.e proportion 1 and 2) and mutant p53 as more than 50% expression in the cells (i.e. proportions 3 and 4) in the 3 cohorts, the results are as follows:

	Negative	Wild-type p53	Mutant p53
Cohort 1 (n=17)	1	3	13
Cohort 2 (n=15)	1	6	8
Cohort 3 (n=28)	10	4	14

It appears the mutation negative cohort has the most cases with mutant p53 expression, with 76% of the cases positive, followed by the mutation positive cohort (53%) and then the sporadic cohort of tumours (50%).

3.1.4.4 Cyclin D1

Cyclin D1 demonstrated heterogeneous nuclear staining in most cases, with varying intensities (**Figure 7A to 7C**). On further analysis, the difference in cyclin D1 expression was statistically significant between the mutation positive and mutation negative cohorts, with a *P*=0.0024 (Two-sample Wilcoxon rank-sum test). There were no statistical differences between the mutation positive and the sporadic cohorts (Two-sample Wilcoxon rank-sum test – p=0.0343) and the mutation negative and sporadic cohorts (Two-sample Wilcoxon rank-sum test – *P*=0.1347). All the mutation positive cases expressed some degree of cyclin D1 compared with 8/28 sporadic cases and 10/17 mutation negative cases being

completely negative. In addition, of all the cyclin D1 negative cases i.e. 18/60 cases, 11 were also negative for cyclin E.

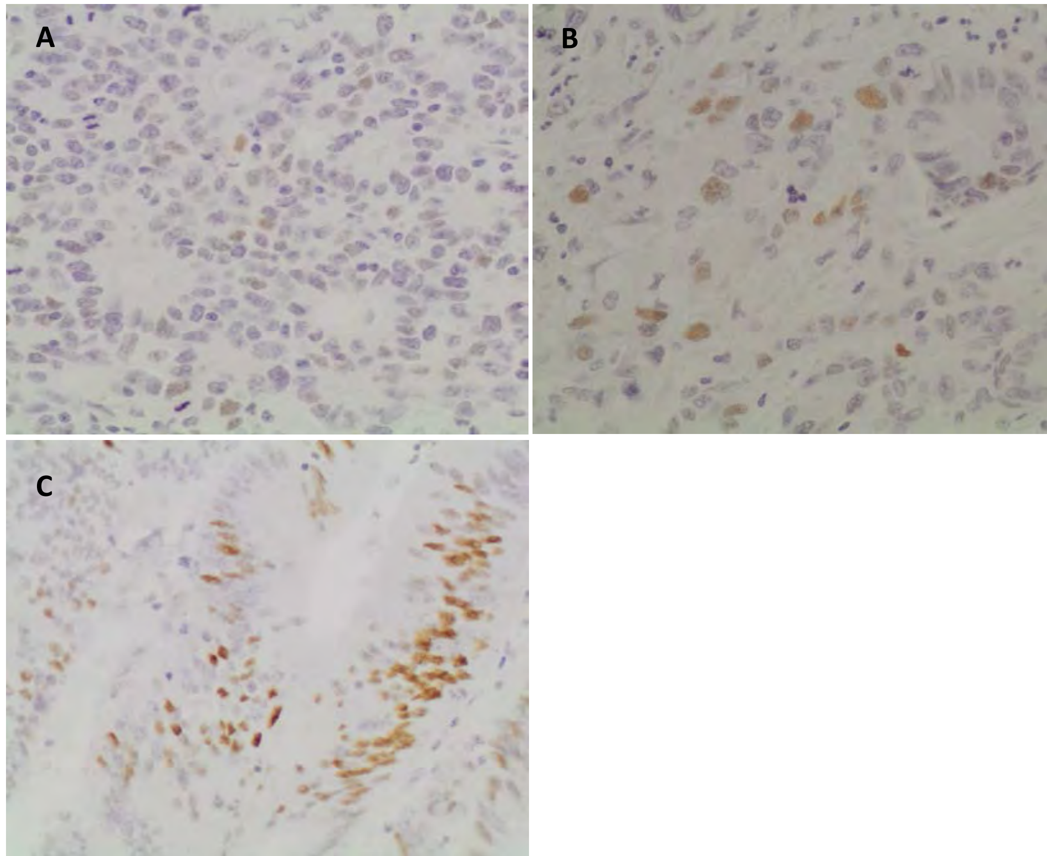

Figure 7: Cyclin D1 immunohistochemistry: (A) 1+ intensity 40x obj mag (B) 2+ intensity 40x obj mag and (C) 3+ intensity 40x obj mag in 3 low grade conventional adenocarcinomas, NOS

There was no statistical significance in cyclin D1 expression and stage in cohort 1 (*P*=0.1202), cohort 2 (*P*=0.7204) and cohort 3 (*P*=0.9986).
There was no statistical significance in cyclin D1 expression and tumour type in cohort 1 (*P*=0.2409), cohort 2 (*P*=0.3482) and cohort 3 (*P*=0.1594) .

3.1.4.5 Cyclin E

Cyclin E demonstrated heterogeneous nuclear staining in most cases, with intensities as illustrated in **Figure 8A to 8C**. Cyclin E expression was statistically significant only between the mutation positive and the sporadic cohort (*P*=0.0012, Two-sample Wilcoxon rank-sum test).

There were no statistically significant differences in expression between mutation positive and mutation negative cohorts (Two-sample Wilcoxon rank-sum test – *P*=0.232) and mutation negative and sporadic cohorts (Two-sample Wilcoxon rank-sum test – *P*=0.2612).

Again, the majority of the mutation positive cases 13/15 were immunoreactive for cyclin E. There were 17/28 cases in the sporadic cohort and 7/17 cases in the mutation negative cohort that were completely negative.

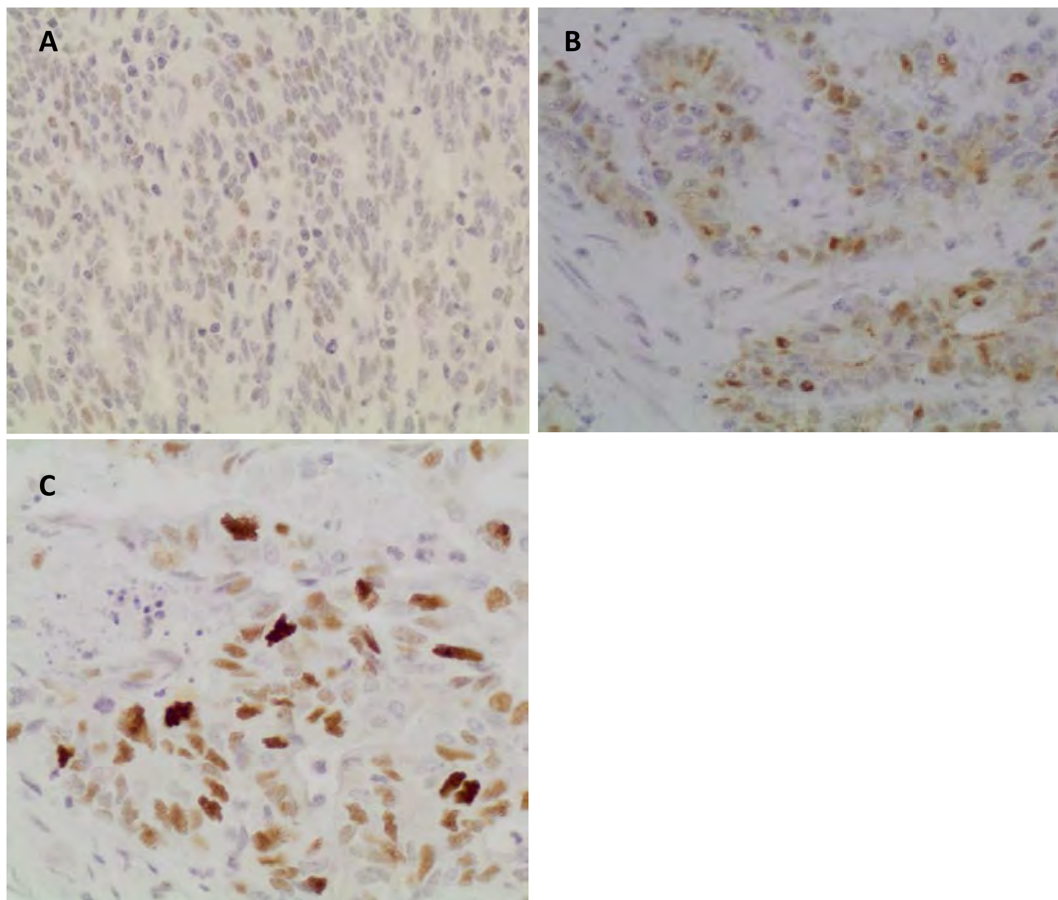

Figure 8: Cyclin E immunohistochemistry: (A) 1+ intensity 40x obj mag, (B) 2+ intensity 40x obj mag in an adenocarcinoma, NOS and (C) 3+ intensity 40x obj mag in 3 conventional adenocarcinomas, NOS

There was no statistical significance in cyclin E expression and stage in cohort 1 (*P*=0.2738), cohort 2 (*P*=0.3936) and cohort 3 (*P*=0.4986).

There was no statistical significance in cyclin E expression and tumour type in cohort 1 (*P*=0.4829), cohort 2 (*P*=0.4880) and cohort 3 (*P*=0.1978) .

3.1.4.6 C-myc

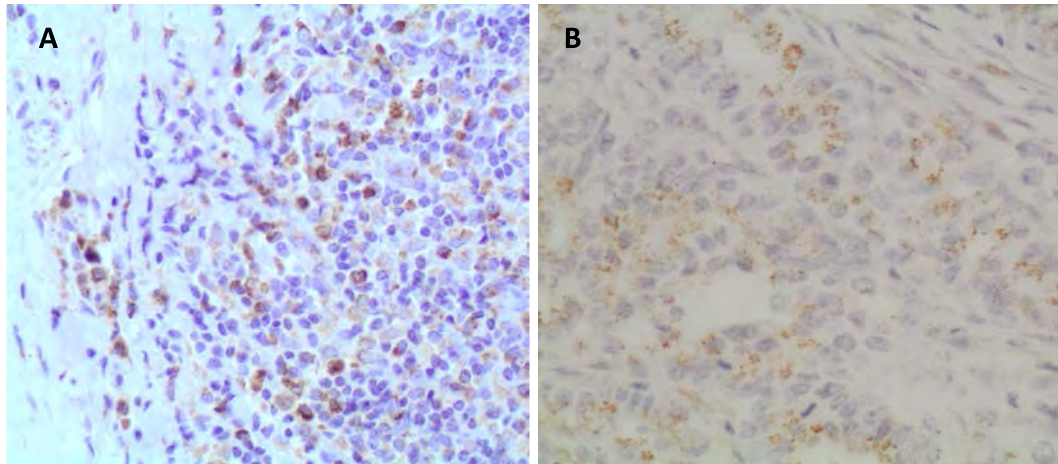

Figure 9: C-myc immunohistochemistry: (A) Normal tonsil control demonstrating nuclear and occasional cytoplasmic staining in lymphocytes and (B) granular cytoplasmic staining of the malignant cells in a low grade conventional adenocarcinoma, NOS. Nuclei are negative.

No nuclear expression was seen in any of the nuclei of the 60 cases examined. However, cytoplasmic granular staining (**Figure 9**) of varying intensities was seen. This included 8/17 cases (47%) in cohort 1, 9/15 cases (60%) in cohort 2 and 17/28 cases in cohort 3 (61%). Since only nuclear staining was considered positive, no statistical analysis was done as the exact significance was uncertain. Although nuclear staining was observed in lymphocytes in the tonsil control tissue, cytoplasmic staining with peri-nuclear accentuation was also noted in lymphocytes.

.

3.1.4.7 MMP-7

A fine to coarse granular staining was noted in the cytoplasm of the malignant cells (**Figure 10**). There was no staining in the stromal fibroblasts. There was no statistical significance for MMP-7 among the three cohorts (Kruskal –Wallis equality-of-populations rank test, p=0.6587).

There was no statistical significance in MMP-7 expression and stage in cohort 1 (P=0.5305), cohort 2 (P=0.2703) and cohort 3 (P=0.8950).

There was no statistical significance in MMP-7 expression and tumour type in cohort 1 (P=0.9372), cohort 2 (P=0.8204) and cohort 3 (P=0.6289) .

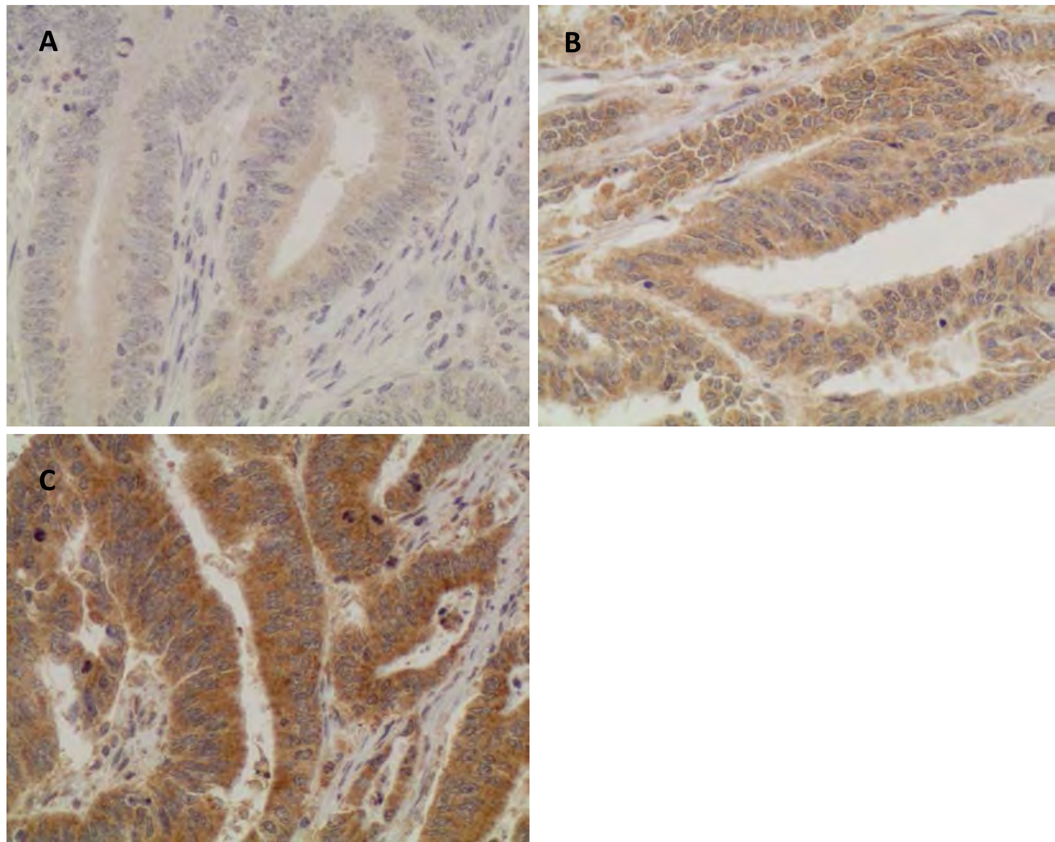

Figure 10: MMP-7 immunohistocemistry (A) 1+ intensity 40x obj mag, (B) 2+intensity 40x obj mag and (C) 3+ intensity 40x obj mag in three low grade conventional adenocarcinomas, NOS

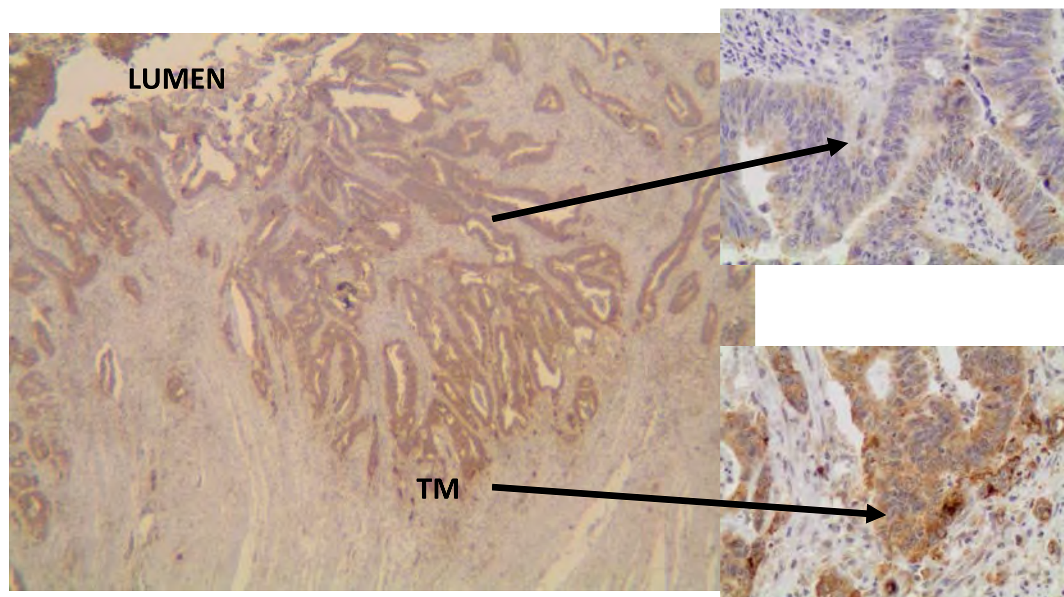

Figure 11: Invasive tumour margin in a low grade adenocarcinoma, NOS demonstrating increased MMP-7 immunoexpression with accentuation at the infiltrating edge.

There was a variation in staining at the invasive tumour margin (**Figure 11**). The staining was either uniform, increased or in a few cases decreased at the invasive margin (**Table 5**). In addition, the cytoplasmic staining appeared much coarser at the invasive front when increased. This was assessed away from the edge of the tissue sections as far as possible to avoid artefactual increase in staining (edge artefact).

Table 5: Staining of MMP-7 at invasive tumour margin

	Increased staining at tumour margin	Decreased staining at tumour margin	Uniform staining throughout the tumour
Cohort 1 (n=17)	8	0	9
Cohort 2 (n=15)	5	1	9
Cohort 3 (n=28)	15	2	11

3.1.4.8 MLH1

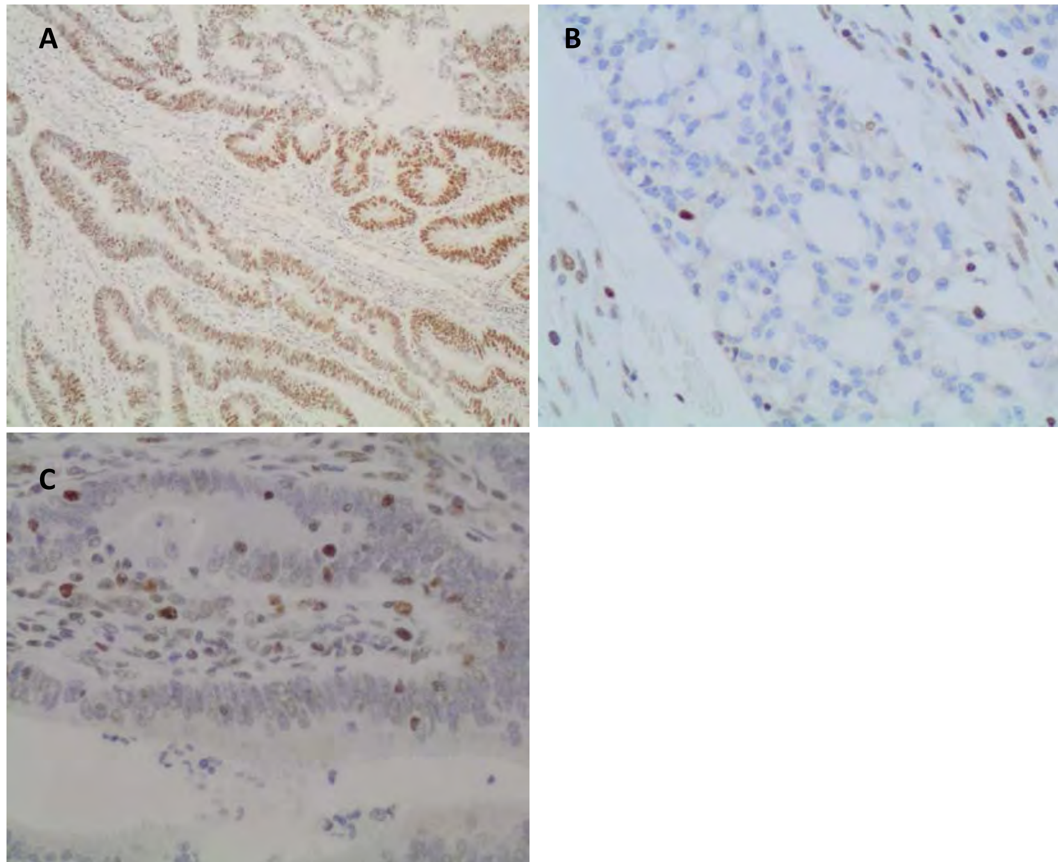

Figure 12: MLH1 immunohistochemistry: (A) Nuclear expression in the malignant glands 10x obj mag of a low grade adenocarcinoma, NOS, (B) loss of expression in the solid component of a mucinous carcinoma, note the positive internal control in stromal cells and lymphocytes 40x obj mag and (C) loss of expression in a low grade adenocarcinoma, NOS with positive TILs in a mutation positive case 40x obj mag.

In the negative cases, the background stromal cells, lymphocytes and normal glandular epithelial cells were positive (**Figure 12**). In cases, where there was an increase in tumour infiltrating lymphocytes, the positive MLH1 staining served as a good marker for the lymphocytes, making it easier to count.

All 13 cases that showed *MLH1* mutations were negative for MLH1 immunohistochemistry, a 100% concordance. In contrast, a smaller percentage of cases in the mutation negative and sporadic cohorts showed loss of MLH1. This was statistically significant between the mutation positive cohort and the sporadic cohort ($P<0.0001$) as well as between the mutation positive cohort and the

mutation negative cohort (*P*<0.0001). There was no statistical significance between the mutation negative cohort and the sporadic cohort (*P*=0.0676).

Only five cases were negative for MLH1 immunoexpression in our cohort of sporadic CRCs. The interesting findings were that all five of the MLH1 negative tumours were on the right side of the colon and all were special subtypes, i.e. 2 mucinous, 1 signet ring cell, 1 medullary and 1 serrated adenocarcinoma. All the mutation negative cases expressed MLH1 by immunohistochemistry (17/17).

3.1.4.9 MSH2

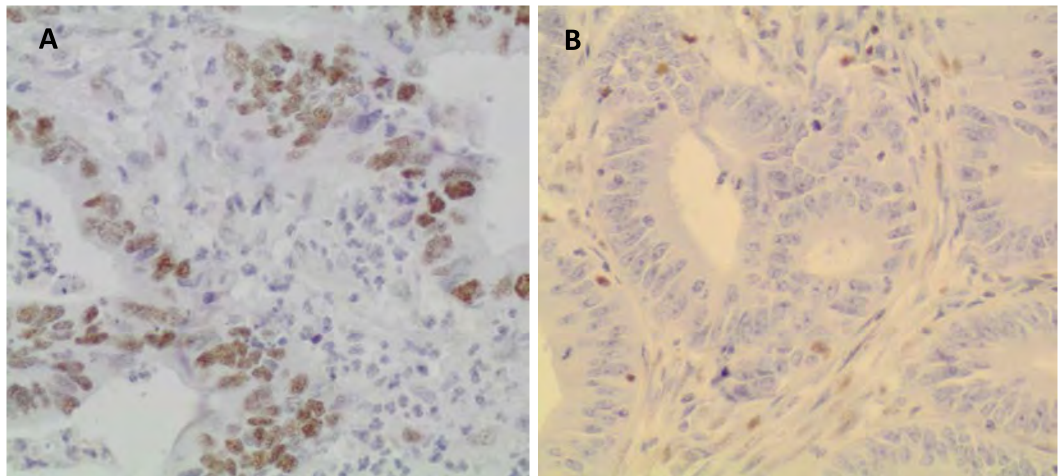

Figure 13: MSH2 immunohistochemistry: (A) Nuclear expression in the malignant glands 10x obj mag of a low grade adenocarcinoma, (B) a low grade adenocarcinoma showing loss of expression, note positive internal control in stromal cells and lymphocytes 40x obj mag

Tumours that were negative for MSH2 immunohistochemistry showed positive inbuilt controls similar to that of MLH1 (**Figure 13B**). There was no statistical difference in MSH2 immunoexpression among the three cohorts.

Two cases in the sporadic cohort were indeterminate as both tumour and inbuilt controls were negative, and were not included in the statistical analysis. The indeterminate cases were re-stained and remained indeterminate. The negative case was a serrated adenocarcinoma, where all three MMR markers were negative. In the latter case the following cell cycle regulators were positive: p27, cyclin D1 and cyclin E.

There were 3 cases from the mutation positive cohort that were negative. Of the 3 cases, 2 were positive for MLH1 and both were negative for MSH6. These two cases had confirmed *MSH2* mutations. The third case was a stage 3 mucinous carcinoma and negative for all three markers.

The mutation negative cohort had three cases which were negative for MSH2 only (i.e. these cases expressed both MLH1 and MSH6), 2 were from the right colon with conventional adenocarcinoma histology and the third was from the left colon, with high grade histology and tumour infiltrating lymphocytes.

3.1.4.10 MSH6

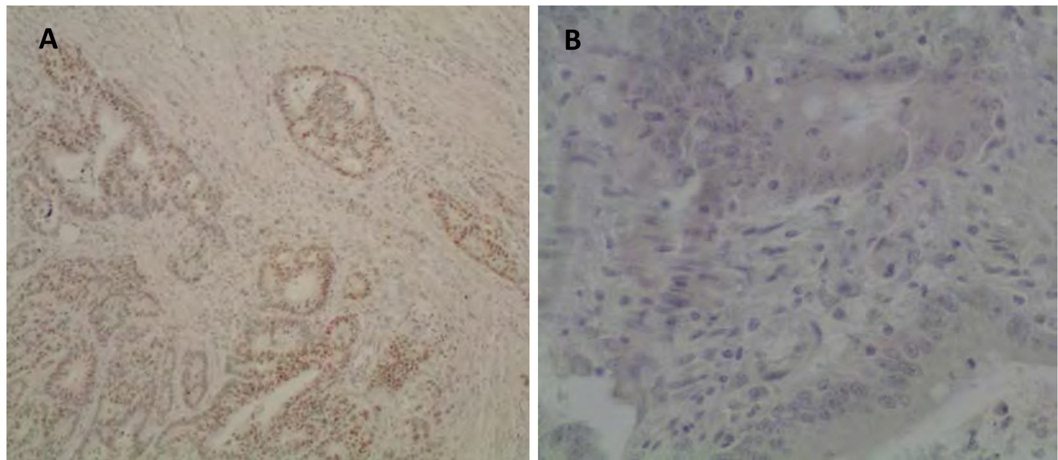

Figure 14: MSH6 immunohistochemistry: (A) A positive low grade adenocarcinoma, NOS demonstrating nuclear staining 10x obj mag and (B) an MSH6 negative low grade adenocarcinoma, NOS 40x obj mag.

Of the 60 cases, 3 were indeterminate and were excluded from the statistical analysis. There was no statistically significant difference in expression among the 3 cohorts.

The mutation negative cohort showed 16/17 positive cases, with 1 case being indeterminate. There were 6 cases in the mutation positive cohort which did not express MSH6 on immunohistochemistry: 3 negative MLH1 cases, 2 negative MSH2 cases and 1 case negative for both MSH2 and MLH1.

In the sporadic cohort, there were 3 cases which were negative and a further 2 cases were indeterminate. Of the 3 negative cases, one case was negative for MLH1 only and the second case negative for both MLH1 and MSH2 while the third case was positive for MLH1 but the MSH2 was indeterminate. Of the 2 indeterminate cases, 1 was positive for MLH1 and indeterminate for MSH2 while the second case was positive for MLH1 and MSH2. All indeterminate cases were re-stained but there was no change in the expression, normal and tumour cells were still completely negative.

CHAPTER 4

4. DISCUSSION

4.1 Clinical and Morphological Features

The mutation positive cases (cohort 2) showed a predominance of right sided tumours, more frequent mucinous differentiation, Crohns like reaction and tumour infiltrating lymphocytes, in keeping with previously published findings (17, 44). Of all the parameters described, tumour infiltrating lymphocytes and an expanding/"circumscribed" margin was more predominant in cohort 2 when compared to the other cohorts, which may contribute to the improved prognosis in HNPCC patients. The lymphocytes are usually cytotoxic T lymphocytes, which are associated with tumour cell apoptosis but these were not tested for in our study (13). A study by Greenson *et al*, which investigated 52 microsatellite unstable tumours, found that TILs were the best predictor for microsatellite unstable tumours. In addition, the absence of "dirty necrosis" and/or focal mucinous differentiation together with TILs identified all 52 tumours in that study, which showed morphology to be 100% sensitive (45). Although mucinous tumours that are MSI-H have a better prognosis than MSI-L and MSS mucinous CRCs (2), there was no correlation with the cell cycle regulators and MMP-7 when compared with the tumour type in each patient cohort. This may be the result of the low numbers of mucinous adenocarcinomas in the cohorts.

The patients in cohort 1 (younger than 50 years and mutation negative) also demonstrated a predominance of tumours on the right side of the colon but less than that of the mutation positive cases. The morphology in this cohort was similar to that of the sporadic cohort, with no significant differences. However, this cohort had the youngest mean age of 38 years, which suggests a possible underlying germline mutation.

Overall, the mutation negative cases appeared to have similar morphological profiles to the sporadic cohort, with the only major discrepancies being tumour site and Crohns like reaction. The latter can be quite subjective as the literature suggests, because criteria are not well defined in studies.

In the sporadic cohort, the interesting finding was that all 5 adenocarcinomas that tested negative for MLH1 immunohistochemistry were special subtypes, all associated with MSI-H tumours. Therefore, histology of the colorectal carcinomas may be important in screening for possible MSI-H tumours.

The majority of tumours in all three cohorts showed low grade histology and stage III tumours, with nodal metastasis seen in 59% (mutation negative), 53% (mutation positive) and 50% (sporadic) of the cases.

4.2 Cell cycle inhibitors – p21, p27 and p53

The immunoexpression of p21, p27 and p53 did not show statistical significance among the three cohorts. p21 expression was downregulated in all three cohorts, with only 10 of the 60 cases (18%) demonstrating some degree of positivity among the cohorts. The downregulation of p21 is usually associated with mutations in *p53*. A study by Sinicrope *et al* demonstrated a significant reduction of p21 in sporadic colorectal carcinoma when compared to HNPCC cases, with 80% of cases (12/15 cases) retaining p21 expression (46). Their findings were attributed to a possible inverse relationship to p53 i.e. mutant p53 was seen in 27% of the HNPCC cases and 69% of the sporadic cases compared to p21 seen in 31% of sporadic CRCs and 80% of HNPCC cases. Similar findings were also seen in MSI-H tumours compared to MSS tumours in a study by Edmonston *et al* (47). In another study of a Danish population of HNPCC patients, upregulation of p21 was present with down regulation of p53 (48).

The findings of our study do not support the inverse relationship of p21 and mutant p53 in the mutation positive cohort (cohort 2) i.e. there was no significant p21 expression and 53% of all cases expressed mutant p53. There was a reduction in p21 expression in all three cohorts compared to p53. In our study cohort, of the 8 cases that expressed p21, 7 had also expressed p53, suggesting an alternative pathway of activation of p21, this is similar to that described in the literature (49, 50). Two other factors WISp39 and Hsp90 have been shown to regulate p21 (51). An alteration of any one of these factors causes downregulation of p21, irrespective of the presence of wild-type p53.

Antibodies can detect both mutant and wild type p53 and since wild-type p53 has low detection levels and a short half-life, most p53 antibodies are able to detect abnormal p53 (52). If one had to consider staining of p53 in less than 50% of cells as wild-type staining (52), then the sporadic tumours have the least expression of p53 (43%) from the three cohorts with a 10% difference from the mutation positive cohort (53%). This finding is different from the studies mentioned above (47, 48) and a further study which demonstrated decreased *p53* gene mutations in HNPCC patients which was much lower than the sporadic tumours (53). However, there is also a study in the literature demonstrating no statistical significance in p53 staining between sporadic (70 cases) and HNPCC (46 cases) CRCs (54).

The p27 immunohistochemical marker was positive in all 60 cases, suggesting an upregulation in CRCs, independent of p53 and p21. Other studies have revealed high levels of p27 expression in lower-stage CRCs compared to low levels of p27 in advanced stage CRCs. Our study showed no significance in p27 expression between the three cohorts and findings are similar to that seen in a study by Edmonston *et al* (47).

Overall, there was no significant correlation of the cell cycle inhibitors (p21, p27 and p53) with tumour type and stage in all three cohorts.

4.3 C-myc

The proto-oncogene *c-myc* is thought to act through the nucleus. C-myc did not show any nuclear staining in any of the three cohorts, which was considered significant. However, some cases did demonstrate granular cytoplasmic staining (different intensities) of uncertain significance.

Cytoplasmic expression has been described in other tumours like aggressive lymphomas (55). Royds *et al* demonstrated pancytoplasmic staining of c-myc in colorectal carcinomas (compared to adenomas), with decreased nuclear staining. In that study, confirmation of c-myc in the cytoplasm was accomplished using electron microscopy which showed binding to polyribosomes (56). The features suggested that alterations in the C-terminus allowed for accumulation of c-myc in

the cytoplasm because of ineffective binding to the nucleus. Another study demonstrated that subcellular localization was highly dependent on the type and timing of tissue fixation. They found that in frozen tissue c-myc expression localised to the nucleus while the same tissue following formalin fixation showed loss of nuclear staining (57). However, the antibodies were not validated. Some of the tissues in our study did date back to the 1980s and showed comparable staining to more recent cases, confirming that protocols for fixing excision specimens have remained constant in our laboratory.

4.4 Cyclins D1 and E

Cyclins D1 and E were both increased in the mutation positive cohort. When expression in the mutation positive cohort was compared to the other cohorts, only cyclin E demonstrated a statistical difference with the sporadic cohort and cyclin D1 with the mutation negative cohort. Overall, since the cell cycle has to be activated in all three cohorts for cells to proliferate, the expression of both cyclins should be increased, in theory. It appears the mutation negative cohort preferentially expresses cyclin E and the sporadic cohort preferentially expresses cyclin D1 at the expense of the other cyclin in our study. Overall, the combination of cyclins were activated in 49 cases with 11 cases negative for both cyclin D1 and cyclin E.

In a Danish study population of 34 HNPCC CRCs, both cyclins D1 and E expression were significantly increased in 68% and 32% of cases, respectively (48). However, overexpression of p53 and p21 was found in 21% and 71% of the study cases respectively, the reverse of what was seen in our study population. Since p21 regulates the cyclins, one would expect expression of cyclin D1 and E to be increased. However, p21 expression was decreased in our study and p27 was upregulated which inhibits cyclin E. Therefore other factors involved in the regulation of the cell cycle and Wnt signalling pathway may be implicated in the expression of the cyclins.

The mutation in the *APC* gene allows for accumulation of β-catenin within the cell, which in turn activates cyclin D1 and cyclin E (58). Therefore increased expression of cyclin D1 and cyclin E is expected with *APC* mutations as seen in

sporadic CRCs. However, increased frequency of *CTNNB1* mutations (encodes β-catenin) have been described in patients with MSI-H and HNPCC (59, 60). In addition, *APC* mutations have also been reported in HNPCC CRCs but with much less frequency when compared to sporadic tumours (60). These HNPCC tumours with *CTNNB1* mutations demonstrated β-catenin accumulation in the nuclei of tumour cells with immunohistochemistry. Overall this suggests that the Wnt-pathway still plays a major role in carcinogenesis irrespective of the MSI status, as seen in this cohort of patients and described in the literature (59, 60).

The elevated expression of cyclin E in the mutation positive cohort may be linked to downregulation of the *FBXW7* gene in this cohort. However, a study by Miyaki *et al* did not demonstrate appreciable difference in *FBXW7* gene mutations between HNPCC and sporadic cases of CRCs (61). In addition, another study showed that low expression of mRNA *FBXW7* was associated with a poor prognosis compared to the high expression group. This is contradictory to HNPCC patients in general, where the patients have a good prognosis (62). Therefore increases in cyclin E expression together with cyclin D1 may be linked to aberrations in the Wnt pathway. There are limited studies with cyclin E and HNPCC in the literature, and the activation of this cell cycle regulator needs to be explored further.

Finally, there was no statistically significant correlation of cyclin D1 and cyclin E with tumour type and stage in all 3 cohorts.

4.5 MMP-7

MMP-7 is linked to increased frequency of metastasis in CRCs in some studies (36, 63). In our study there was no appreciable difference in MMP-7 expression among the three cohorts and there was no statistically significant difference between MMP-7 expression and stage of tumour.

MMP-7 is a target gene of the Wnt pathway and the expression of β-catenin and MMP-7 is correlated in CRCs (64). MMP-7 allows for the breakdown of the E-cadherin-β-catenin complex, allowing β-catenin to accumulate in the cytoplasm. A recent study of epithelial mesenchymal transition demonstrated that the

expression of β-catenin, mismatch repair proteins, E-cadherin, MMP-7, TGFβRII and tissue inhibitors of metalloproteinases-2 (TIMP-2) was significantly associated with depth of invasion and metastasis (64). Unfortunately, we were unable to investigate β-catenin immunoexpression because of limited funding. Gu *et al* also demonstrated lower expression levels of hMSH2, hMLH1, cytoplasmic β-catenin, TGFβRII and MMP-7 with higher levels of E-cadherin, TIMP-2 and membrane β-catenin in HNPCC tumours when compared to sporadic CRCs. In another gene microarray study, 11 microsatellite unstable tumours, demonstrated dowregulation of the MMP-7 gene when compared to MSI-L and MSS tumours (65).

The β-catenin that accumulates following mutations in the *APC* gene in sporadic CRCs, results in accumulation of cytoplasmic β-catenin and increasing expression of MMP-7, resulting in invasion and metastasis. However in HNPCC, mutations in *TGFβRII* seen especially with *MLH1* and *MSH2* mutations, reduces the ability of the E-cadherin and plasma membrane β-catenin to disassemble, so that plasma membrane β-catenin increases in HNPCC CRCs (64). This may explain the better prognosis and fewer metastases in Lynch syndrome patients.

Our study did not reveal a significant decrease in expression of MMP-7 in the mutation positive cohort and no statistical differences were seen among the three cohorts or when compared to the stage in the individual cohorts. Although our study number was small, most of the cases demonstrated uniform staining in the mutation positive cohort i.e. 9 cases (60%). Uniform staining was also seen in 53% of the mutation negative cases and 39% of the sporadic cases. All the cases (n=60) expressed MMP-7, which allows for the tumour cells to invade and breakdown the stroma. Also, nodal metastases were seen in 5 cases (33%) in the mutation positive cohort. The increase in expression at the tumour margins is interesting, as there appears to be some “ordered communication” of the malignant cells at the tumour edge, allowing for tumour progression. The “microenvironment” at the edge of the tumour may possibly be disrupted by TILs (secreting interleukins), decreasing tumour progression.

4.6 Mismatch repair proteins – MLH1, MSH2 and MSH6

MLH1 demonstrated significant differences in expression among the three cohorts. There was a strong concordance of MLH1 loss on immunohistochemistry and *MLH1* mutation positive tumours. Groote Schuur hospital offers a screening service to the Kleinsee population with our laboratory receiving the specimens. MLH1 is by far the commonest mutation seen in numerous studies (44).

The sporadic tumours in our cohort that were MLH1 negative were all special subtypes of adenocarcinoma that are also seen in HNPCC and MSI-H cases. These possible sporadic MSI-H tumours may be the result of hypermethylation of DNA and *BRAF* mutations, which are commonly seen in sporadic MSI tumours, and also being more common in female patients. Of these 5 sporadic cases in our cohort, 4 were female patients. However, one should not completely rule out a case of HNPCC in patients over 50 years. Some patients can have late-onset CRCs in Lynch families (>50 years) (66).

The MLH1 immunoexpression in the mutation positive cohort was consistent with that of the germline analysis in that all cases that showed loss of immunoexpression tested positive for *MLH1* mutations (**Appendix 3**). The two cases that tested negative for MSH2 also demonstrated mutations in the *MSH2* gene.

The 2 cases that were negative for MSH2 were also negative for MSH6 and positive for MLH1 in the mutation negative cohort. This could be consistent with *MSH2* mutations not tested in our centre, rare *MSH6* mutations or mutations involving factors or proteins controlling expression of the *MSH2* gene. There was a single case that was negative for all three MMR proteins. Our findings are consistent with Rigau *et al* who conducted a meta-analysis of MMR proteins to assess MSI in 11 series and found that using only 2 antibodies, MLH1 and MSH2 allowed for a diagnosis of MSI with high sensitivity and specificity (44).

Mutations in MSH6 and PMS2 are seen in rare cases of HNPCC. Our cohort showed 6 cases negative for the MSH6 protein in the mutation positive cohort.

Of the 6 cases, three were associated with loss of MLH1 and intact MSH2, two were associated with loss of MSH2 and one was associated with loss of all three proteins. The 2 cases associated with loss of MSH2 can be explained by the *MSH2* mutations, since there is heterodimeric pairing of the MMR proteins i.e. MSH2 forms a complex with MSH6, so if one is mutated, the other is degraded. The loss of both these proteins has also been well documented in other studies (67, 68). However, the loss of both MLH1 and MSH6 would be unusual, since these two proteins do not form a complex. The full mutational analysis of all four MMR genes for these three tumours and the tumour with all three markers negative would be useful to determine which results are truly negative. In the sporadic cohort, two cases were negative for MSH6, the first a serrated adenocarcinoma demonstrating loss of all three markers. This suggests hypermethylation of the DNA resulting in downgrading of all three proteins. The other case was a conventional adenocarcinoma, with only MSH6 loss and an indeterminate MSH2 expression.

In this study, we were limited to three MMR markers rather than four markers which includes PMS2. This was based on the recommendations of Rigau *et al* (44) and the known heterodimeric complexes formed by MLH1 and PMS2, so that if there is inactivation of the one protein, there is usually decreased expression of the other.

4.7 Limitations

Finally, some of the drawbacks of this study included low case numbers in cohorts 1 and 2. This was due to the limited excision specimens that were less than 50 years with and without proven MMR gene mutations in our database. Another factor was the non-availability of tissue blocks for many of the mutation positive cases where surgery was done at other hospitals.

Furthermore, there was limited or no survival data available for the majority of patients in all the cohorts. This limited the outcome analysis to tumour stage. Grade of tumour, could not be used as a reliable prognostic variable as there were only two cases with high grade histology in all 60 cases.

Pre-analytic variables regarding optimal fixation of CRC excision specimens in formalin, may have had an influence on the expression of immunohistochemical markers and also needs to be taken into consideration. In this laboratory colorectal resection specimens are usually opened, fixed in formalin overnight and then processed the next day. All of the cases demonstrated staining in one or more of the immunohistochemical markers in this study, implying intact tissue integrity. The immunohistochemical markers themselves have their own sensitivities and specificities which was accounted and optimised for in the protocol.

CHAPTER 5

5. Conclusion

Although major advances have been made in the various pathways of colorectal carcinogenesis, it still remains complex, with various interactions at the molecular level.

Overall, there were statistically significant differences between the mutation positive and both the sporadic and mutation negative cohorts with respect to morphology and MLH1 immunoexpression. The morphological features included TILs, expanding margin, Crohns like reaction and mucinous differentiation. These features in combination can be helpful in identifying MSI-H tumours and cases requiring further molecular studies. In our population, *MLH1* mutations are the commonest.

This study found that cell cycle markers were unlikely to be solely or independently responsible for the differences in behaviour and prognosis of HNPCC CRCs versus sporadic CRCs. Overall, the cyclins are increased in all three cohorts, the upregulation dependent on cell cycle regulators and factors in the Wnt signalling pathway. Cyclin D1 showed some difference in immunoexpression between the mutation positive and negative cohorts and cyclin E between the mutation positive and sporadic cohorts. The mutation negative cohort preferentially expressed cyclin E and the sporadic cohort preferentially expressed cyclin D1, with both these cell cycle markers upregulated in the mutation positive cohort. Overall, the cell cycle markers did not explain the improved prognosis of the mutation positive cohort. The cell cycle regulators and MMP-7 did not show any significant correlation with tumour stage or tumour type in the three cohorts.

It is worth further exploring the expression of MMP-7 in CRCs to determine whether distribution of expression has any significance on invasiveness and metastatic potential of the primary neoplasm.

Overall, the mutation negative cohort in our population continues to remain enigmatic and further testing at the molecular level is required, that may reveal another novel pathway of colorectal carcinogenesis or other novel mutations in mismatch repair genes. Since only tests for mutations in *MLH1* and *MSH2* are only available at our centre, some of these cases could still harbour a mutation in another MMR gene.

This study has raised the possibility that the Wnt signalling pathway may have a more integral role in CRC behaviour. Further investigation of this pathway in relation to the MMR profile is warranted.

The complex intracellular signalling pathways most likely hold the key to understanding and predicting CRC behaviour and prognosis. Furthermore, identifying some of these key biomarkers may lead to new therapeutic interventions, such as, targeted therapy.

REFERENCES

1. Cronje L, Paterson AC, Becker PJ. Colorectal cancer in South Africa: a heritable cause suspected in many young black patients. *S Afr Med J*. 2009;99:103-6.
2. WHO Classification of Tumours of the Digestive System. 4 ed. Bosman FT, Carneiro F, Hruban RH, Theise ND, editors: International Agency for Research on Cancer; 2010.
3. Haggar FA, Boushey RP. Colorectal cancer epidemiology: incidence, mortality, survival, and risk factors. *Clin Colon Rectal Surg*. 2009;22:191-7.
4. Graham A, Adeloye D, Grant L, Theodoratou E, Campbell H. Estimating the incidence of colorectal cancer in Sub-Saharan Africa: A systematic analysis. *J Glob Health*. 2012;2:1-14
5. Fearon ER. Molecular genetics of colorectal cancer. *Annu Rev Pathol*. 2011;6:479-507.
6. Schlussel AT, Gagliano RA Jr., Seto-Donlon S, Eggerding F, Donlon T, Berenberg J*, et al.* The evolution of colorectal cancer genetics-Part 1: from discovery to practice. *J Gastrointest Oncol*. 2014;5:326-35.
7. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. *Cell*. 1990;61:759-67.
8. Al-Sohaily S, Biankin A, Leong R, Kohonen-Corish M, Warusavitarne J. Molecular pathways in colorectal cancer. *J Gastroenterol Hepatol*. 2012;27:1423-31.
9. Kanthan R, Senger JL, Kanthan SC. Molecular events in primary and metastatic colorectal carcinoma: a review. *Pathol Res Int*. 2012;2012:1-14
10. Schneikert J, Behrens J. The canonical Wnt signalling pathway and its APC partner in colon cancer development. *Gut*. 2007;56:417-25.
11. Bedeir A, Krasinskas AM. Molecular diagnostics of colorectal cancer. *Arch Pathol Lab Med*. 2011;135:578-87.
12. Geiersbach KB, Samowitz WS. Microsatellite instability and colorectal cancer. *Arch Pathol Lab Med*. 2011;135:1269-77.
13. Poulogiannis G, Frayling IM, Arends MJ. DNA mismatch repair deficiency in sporadic colorectal cancer and Lynch syndrome. *Histopathology*. 2010;56:167-79.

14. Arends MJ. Pathways of colorectal carcinogenesis. *Appl Immunohistochem Mol Morphol*. 2013;21:97-102.
15. Boland CR, Thibodeau SN, Hamilton SR, Sidransky D, Eshleman JR, Burt RW*, et al.* 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*. 1998;58:5248-57.
16. Overbeek LI, Ligtenberg MJ, Willems RW, Hermens RP, Blokx WA, Dubois SV*, et al.* Interpretation of immunohistochemistry for mismatch repair proteins is only reliable in a specialized setting. *Am J Surg Pathol*. 2008;32:1246-51.
17. Geiersbach KB, Samowitz WS. Microsatellite Instability and Colorectal Cancer. *Arch Pathol Lab Med*. 2011;135:1269-77.
18. Schlussel AT, Gagliano RA, Jr., Seto-Donlon S, Eggerding F, Donlon T, Berenberg J*, et al.* The evolution of colorectal cancer genetics-Part 2: clinical implications and applications. *J Gastrointest Oncol*. 2014;5:336-44.
19. Ramesar RS, Madden MV, Felix R, Harocopos CJ, Westbrook CA, Jones G*, et al.* Molecular genetics improves the management of hereditary non-polyposis colorectal cancer. *S Afr Med J*. 2000;90:709-14.
20. Wong JJ, Hawkins NJ, Ward RL. Colorectal cancer: a model for epigenetic tumorigenesis. *Gut*. 2007;56:140-8.
21. Curtin K, Slattery ML, Samowitz WS. CpG island methylation in colorectal cancer: past, present and future. *Pathol Res Int*. 2011;2011:1-8.
22. Park SJ, Rashid A, Lee JH, Kim SG, Hamilton SR, Wu TT. Frequent CpG island methylation in serrated adenomas of the colorectum. *Am J Pathol*. 2003;162:815-22.
23. Jass JR. Classification of colorectal cancer based on correlation of clinical, morphological and molecular features. *Histopathology*. 2007;50:113-30.
24. Haque T, Greene KG, Crockett SD. Serrated neoplasia of the colon: what do we really know? *Curr Gastroenterol Rep*. 2014;16:380.
25. Batts KP. The pathology of serrated colorectal neoplasia: practical answers for common questions. *Mod Pathol*. 2015;28 Suppl 1:S80-7.
26. Tsang AH, Cheng KH, Wong AS, Ng SS, Ma BB, Chan CM*, et al.* Current and future molecular diagnostics in colorectal cancer and colorectal adenoma. *World J Gastroenterol*. 2014;20:3847-57.

27. Schetter AJ, Okayama H, Harris CC. The role of microRNAs in colorectal cancer. *Cancer J*. 2012;18:244-52.
28. Stefansson OA, Esteller M. Epigenetic modifications in breast cancer and their role in personalized medicine. *Am J Pathol*. 2013;183:1052-63.
29. Donnellan R, Chetty R. Cyclin E in human cancers. *Faseb J*. 1999;13:773-80.
30. McKay JA, Douglas JJ, Ross VG, Curran S, Ahmed FY, Loane JF*, et al.* Expression of cell cycle control proteins in primary colorectal tumors does not always predict expression in lymph node metastases. *Clin Cancer Res*. 2000;6:1113-8.
31. Wangefjord S, Manjer J, Gaber A, Nodin B, Eberhard J, Jirstrom K. Cyclin D1 expression in colorectal cancer is a favorable prognostic factor in men but not in women in a prospective, population-based cohort study. *Biol Sex Differ*. 2011;2:10.
32. Chu IM, Hengst L, Slingerland JM. The Cdk inhibitor p27 in human cancer: prognostic potential and relevance to anticancer therapy. *Nat Rev Cancer*. 2008;8:253-67.
33. Pasz-Walczak G, Kordek R, Faflik M. P21 (WAF1) expression in colorectal cancer: correlation with P53 and cyclin D1 expression, clinicopathological parameters and prognosis. *Pathol Res Pract*. 2001;197:683-9.
34. Shiomi T, Lemaitre V, D'Armiento J, Okada Y. Matrix metalloproteinases, a disintegrin and metalloproteinases, and a disintegrin and metalloproteinases with thrombospondin motifs in non-neoplastic diseases. *Pathol Int*. 2010;60:477-96.
35. Malemud CJ. Matrix metalloproteinases (MMPs) in health and disease: an overview. *Front Biosci*. 2006;11:1696-701.
36. Polistena A, Cucina A, Dinicola S, Stene C, Cavallaro G, Ciardi A*, et al.* MMP7 expression in colorectal tumours of different stages. *In Vivo*. 2014;28:105-10.
37. Zeng ZS, Shu WP, Cohen AM, Guillem JG. Matrix metalloproteinase-7 expression in colorectal cancer liver metastases: evidence for involvement of MMP-7 activation in human cancer metastases. *Clin Cancer Res*. 2002;8:144-8.
38. Washington MK, Berlin J, Branton P, Burgart LJ, Carter DK, Fitzgibbons PL*, et al.* Protocol for the examination of specimens from patients with primary carcinoma of the colon and rectum. *Arch Pathol Lab Med*. 2009;133:1539-51.

39. Garcia-Solano J, Perez-Guillermo M, Conesa-Zamora P, Acosta-Ortega J, Trujillo-Santos J, Cerezuela-Fuentes P*, et al.* Clinicopathologic study of 85 colorectal serrated adenocarcinomas: further insights into the full recognition of a new subset of colorectal carcinoma. *Hum Pathol*. 2010;41:1359-68.
40. McCarty KS Jr., Szabo E, Flowers JL, Cox EB, Leight GS, Miller L*, et al.* Use of a monoclonal anti-estrogen receptor antibody in the immunohistochemical evaluation of human tumors. *Cancer Res*. 1986;46:4244s-8s.
41. Dolcetti R, Viel A, Doglioni C, Russo A, Guidoboni M, Capozzi E*, et al.* High prevalence of activated intraepithelial cytotoxic T lymphocytes and increased neoplastic cell apoptosis in colorectal carcinomas with microsatellite instability. *Am J Pathol*. 1999;154:1805-13.
42. Michael-Robinson JM, Biemer-Huttmann A, Purdie DM, Walsh MD, Simms LA, Biden KG*, et al.* Tumour infiltrating lymphocytes and apoptosis are independent features in colorectal cancer stratified according to microsatellite instability status. *Gut*. 2001;48:360-6.
43. Koelzer VH, Lugli A. The tumor border configuration of colorectal cancer as a histomorphological prognostic indicator. *Front Oncol*. 2014;4:29.
44. Rigau V, Sebbagh N, Olschwang S, Paraf F, Mourra N, Parc Y*, et al.* Microsatellite instability in colorectal carcinoma. The comparison of immunohistochemistry and molecular biology suggests a role for hMSH6 immunostaining. *Arch Pathol Lab Med*. 2003;127:694-700.
45. Greenson JK, Bonner JD, Ben-Yzhak O, Cohen HI, Miselevich I, Resnick MB*, et al.* Phenotype of microsatellite unstable colorectal carcinomas: Well-differentiated and focally mucinous tumors and the absence of dirty necrosis correlate with microsatellite instability. *Am J Surg Pathol*. 2003;27:563-70.
46. Sinicrope FA, Roddey G, Lemoine M, Ruan S, Stephens LC, Frazier ML*, et al.* Loss of p21WAF1/Cip1 protein expression accompanies progression of sporadic colorectal neoplasms but not hereditary nonpolyposis colorectal cancers. *Clin Cancer Res*. 1998;4:1251-61.
47. Edmonston TB, Cuesta KH, Burkholder S, Barusevicius A, Rose D, Kovatich AJ*, et al.* Colorectal carcinomas with high microsatellite instability: defining a distinct immunologic and molecular entity with respect to prognostic markers. *Hum Pathol*. 2000;31:1506-14.

48. Therkildsen C, Isinger-Ekstrand A, Ladelund S, Nissen A, Rambech E, Bernstein I*, et al.* Cancer risks and immunohistochemical profiles linked to the Danish MLH1 Lynch syndrome founder mutation. *Fam Cancer*. 2012;11:579-85.
49. Slebos RJ, Baas IO, Clement M, Polak M, Mulder JW, van den Berg FM*, et al.* Clinical and pathological associations with p53 tumour-suppressor gene mutations and expression of p21WAF1/Cip1 in colorectal carcinoma. *Br J Cancer*. 1996;74:165-71.
50. Ogino S, Kawasaki T, Kirkner GJ, Ogawa A, Dorfman I, Loda M*, et al.* Down-regulation of p21 (CDKN1A/CIP1) is inversely associated with microsatellite instability and CpG island methylator phenotype (CIMP) in colorectal cancer. *J Pathol*. 2006;210:147-54.
51. Ogino S, Kawasaki T, Kirkner GJ, Yamaji T, Loda M, Fuchs CS. Loss of nuclear p27 (CDKN1B/KIP1) in colorectal cancer is correlated with microsatellite instability and CIMP. *Mod Pathol*. 2007;20:15-22.
52. Kruschewski M, Mueller K, Lipka S, Budczies J, Noske A, Buhr HJ*, et al.* The prognostic impact of p53 expression on sporadic colorectal cancer is dependent on p21 status. *Cancers (Basel)*. 2011;3:1274-84.
53. Losi L, Ponz de Leon M, Jiricny J, Di Gregorio C, Benatti P, Percesepe A*, et al.* K-ras and p53 mutations in hereditary non-polyposis colorectal cancers. *Int J Cancer*. 1997;74:94-6.
54. Losi L, Fante R, Di Gregorio C, Aisoni ML, Lanza G, Maestri I*, et al.* Biologic characterization of hereditary non-polyposis colorectal cancer. Nuclear ploidy, AgNOR count, microvessel distribution, oncogene expression, and grade-related parameters. *Am J Clin Pathol*. 1995;103:265-70.
55. Ruzinova MB, Caron T, Rodig SJ. Altered subcellular localization of c-Myc protein identifies aggressive B-cell lymphomas harboring a c-MYC translocation. *Am J Surg Pathol*. 2010;34:882-91.
56. Royds JA, Sharrard RM, Wagner B, Polacarz SV. Cellular localisation of c-myc product in human colorectal epithelial neoplasia. *J Pathol*. 1992;166:225-33.
57. Loke SL, Neckers LM, Schwab G, Jaffe ES. c-myc protein in normal tissue. Effects of fixation on its apparent subcellular distribution. *Am J Pathol*. 1988;131:29-37.
58. Tetsu O, McCormick F. Beta-catenin regulates expression of cyclin D1 in colon carcinoma cells. *Nature*. 1999;398:422-6.

59. Shitoh K, Furukawa T, Kojima M, Konishi F, Miyaki M, Tsukamoto T, *et al.* Frequent activation of the beta-catenin-Tcf signaling pathway in nonfamilial colorectal carcinomas with microsatellite instability. *Genes Chromosomes Cancer*. 2001;30:32-7.
60. Miyaki M, Iijima T, Kimura J, Yasuno M, Mori T, Hayashi Y, *et al.* Frequent mutation of beta-catenin and APC genes in primary colorectal tumors from patients with hereditary nonpolyposis colorectal cancer. *Cancer Res*. 1999;59:4506-9.
61. Miyaki M, Yamaguchi T, Iijima T, Takahashi K, Matsumoto H, Mori T. Somatic mutations of the CDC4 (FBXW7) gene in hereditary colorectal tumors. *Oncology*. 2009;76:430-4.
62. Iwatsuki M, Mimori K, Ishii H, Yokobori T, Takatsuno Y, Sato T, *et al.* Loss of FBXW7, a cell cycle regulating gene, in colorectal cancer: clinical significance. *Int J Cancer*. 2010;126:1828-37.
63. Yeh CY, Chiang JM, Hsieh PS, Tsai WS, Tang R, Changchien CR, *et al.* Immunohistochemical study of vascular endothelial growth factor (VEGF) and matrilysin (MMP-7) in T1 adenocarcinoma of the colon and rectum to predict lymph node metastases or distant metastases. *Hepatogastroenterology*. 2011;58:1943-6.
64. Gu GL, Zhu XQ, Wei XM, Ren L, Li DC, Wang SL. Epithelial-mesenchymal transition in colorectal cancer tissue of patients with Lynch syndrome. *World J Gastroenterol*. 2014;20:250-7.
65. Ortega P, Moran A, Fernandez-Marcelo T, De Juan C, Frias C, Lopez-Asenjo JA, *et al.* MMP-7 and SGCE as distinctive molecular factors in sporadic colorectal cancers from the mutator phenotype pathway. *Int J Oncol*. 2010;36:1209-15.
66. Jass JR. What's new in hereditary colorectal cancer? *Arch Pathol Lab Med*. 2005;129:1380-4.
67. Schweizer P, Moisio AL, Kuismanen SA, Truninger K, Vierumaki R, Salovaara R, *et al.* Lack of MSH2 and MSH6 characterizes endometrial but not colon carcinomas in hereditary nonpolyposis colorectal cancer. *Cancer Res*. 2001;61:2813-5.
68. Plaschke J, Kruger S, Pistorius S, Theissig F, Saeger HD, Schackert HK. Involvement of hMSH6 in the development of hereditary and sporadic colorectal

cancer revealed by immunostaining is based on germline mutations, but rarely on somatic inactivation. *Int J Cancer*. 2002;97:643-8.

APPENDICES

KEY:

ADENO	Adenocarcinoma, NOS
ASSOC POLYP	Associated Polyps
CONFIG	Configuration
CROHNS	Crohns like reaction
E	Expanding
G	Gender
GD	Greatest dimension
HS	H-score
I	Intensity
If	Infiltrating
IN	Indeterminate
M	Distant metastases
MC	Mucinous component
MN	Mutation negative
MP	Mutation positive
N	Nodes involved
Neg	Negative
NR	Not recorded
P	Proportion
PERF	Perforation
Pos	Positive
S	Sporadic
T	Tumour size
TIL	Tumour infiltrating lymphocytes
TM	Tumour Margin
X	Unknown

Appendix 1: Cohort 1 (Mutation negative) Morphological features

CASE	G	AGE	SITE	GD	CONFIG	PERF	ASSOC POLYP	TYPE	GRADE	TIL	CROHNS	TM	T	N	M	STAGE
MN1	M	25	RIGHT	50	ANN,CONST	YES	YES	ADENO	LOW	NO	YES	E	4	1	X	3
MN2	F	41	LEFT	35	EXOPHYTIC	NO	NO	ADENO	LOW	NO	YES	E	3	1	X	3
MN3	M	37	RIGHT	55	EXOPHYTIC	NO	NO	MUCINOUS		NO	NO	If	2	1	X	3
MN4	F	34	RECTUM	60	ANN,CONST	NO	NO	ADENO	HIGH	YES	NO	E	4	0	X	2
MN5	M	35	RIGHT	40	ANN,CONST	NO	YES	ADENO	LOW	NO	NO	In	4	1	X	3
MN6	M	40	RECTUM	14	ULCERATED	NO	NO	ADENO	LOW	NO	NO	E	3	1	X	3
MN7	M	42	RECTUM	55	NR	NO	NO	MUCINOUS		NO	NO	E	3	2	X	3
MN8	F	44	LEFT	28	ULCERATED	YES	NO	ADENO	LOW	NO	NO	If	4	1	1	4
MN9	F	43	RECTUM	60	ANN,CONST	NO	NO	ADENO	LOW	NO	NO	If	3	0	X	2
MN10	M	36	RIGHT	80	ANN,CONST	YES	NO	ADENO	LOW	NO	NO	In	2	1	X	3
MN11	M	37	RIGHT	50	ANN,CONST	NO	NO	ADENO	LOW	NO	YES	E	3	0	X	2
MN12	F	33	RIGHT and LEFT	100,43	ANN,CONST	NO	NO	ADENO	LOW	NO	NO	E	4	1	X	3
MN13	M	36	RIGHT	80	EXOPHYTIC	NO	NO	ADENO	LOW	NO	YES	If	3	0	X	2
MN14	M	49	RECTUM	50	ULCERATED	NO	NO	ADENO	LOW	NO	NO	If	1	0	X	1
MN15	F	36	RIGHT	60	ULCERATED	NO	NO	ADENO	LOW	NO	NO	If	1	0	X	1
MN16	F	46	RECTUM	80	EXOPHYTIC	NO	NO	ADENO	LOW	NO	NO	In	3	2	X	3
MN17	F	36	LEFT	40	EXOPHYTIC	NO	NO	ADENO	LOW	NO	YES	E	3	0	X	2

Appendix 2: Cohort 1 (Mutation negative) Immunohistochemical results

CASE	P21			P27			CYCLIN D1			CYCLIN E			p53			c-myc			MMP7			MLH1	MSH6	MSH2
	P	I	HS	P	I	HS	P	I	HS	P	I	HS	P	I	HS	P	I	CYT	P	I	HS			
MN1	0	1	0	2	2	4	0	0	0	0	0	0	4	3	12	0	0	Pos	4	2	8	Pos	Pos	Pos
MN2	0	0	0	3	2	6	0	0	0	0	1	0	4	3	12	0	0		4	2	8	Pos	IN	Pos
MN3	0	0	0	3	3	9	0	1	0	0	1	0	4	3	12	0	0	Pos	2	1	2	Pos	Pos	Pos
MN4	0	0	0	4	2	8	0	2	0	1	1	1	2	3	6	0	0	Pos	4	3	12	Pos	Pos	Neg
MN5	0	2	0	4	2	8	0	1	0	0	0	0	0	0	0	0	0		4	3	12	Pos	Pos	Pos
MN6	0	1	0	4	3	12	0	1	0	0	0	0	4	3	12	0	0		4	2	8	Pos	Pos	Pos
MN7	0	0	0	4	3	12	0	2	0	1	1	1	4	3	12	0	0	Pos	4	3	12	Pos	Pos	Pos
MN8	0	0	0	4	3	12	0	2	0	1	1	1	4	3	12	0	0	Pos	4	2	8	Pos	Pos	Pos
MN9	0	0	0	4	3	12	0	1	0	1	1	1	4	3	12	0	0		2	1	2	Pos	Pos	Pos
MN10	1	3	3	4	3	12	0	1	0	3	1	3	3	2	6	0	0		4	2	8	Pos	Pos	Pos
MN11	0	0	0	4	3	12	2	1	2	1	2	2	1	1	1	0	0		4	2	8	Pos	Pos	Pos
MN12	0	0	0	4	3	12	2	2	4	0	0	0	3	1	3	0	0	Pos	4	1	4	Pos	Pos	Pos
MN13	0	0	0	4	3	12	2	2	4	0	1	0	2	2	4	0	0		4	3	12	Pos	Pos	Neg
MN14	0	1	0	4	3	12	2	3	6	3	2	6	3	2	6	0	0	Pos	2	1	2	Pos	Pos	Pos
MN15	1	1	1	4	3	12	3	2	6	2	1	2	3	2	6	0	0		4	2	8	Pos	Pos	Neg
MN16	2	2	4	4	3	12	2	3	6	4	3	12	4	3	12	0	0	Pos	2	1	2	Pos	Pos	Pos
MN17	1	2	2	4	3	12	3	3	9	1	2	2	4	3	12	0	0		4	3	12	Pos	Pos	Pos

Appendix 3: Cohort 2 (Mutation positive) Morphological features

CASE	G	AGE	SITE	GD	CONFIG	PERF	ASSOC POLYP	TYPE	GRADE	TIL	CROHNS	TM	T	N	M	STAGE
MP18	M	36	RIGHTX2	25,30	ULCERATED	NO	YES	MUCINOUS		NO	YES	E	4	0	X	2
MP19	M	44	RIGHT	70	EXOPHYTIC	NO	YES	MUCINOUS		NO	YES	E	4	2	X	3
MP20	M	47	RECTUM	40	ULCER	NO	NO	MUCINOUS		NO	YES	If	2	2	X	3
MP21	M	50	RIGHT	60	ULCER	NO	NO	ADENO	LOW	NO	NO	E	3	1	X	3
MP22	F	40	RIGHTx2	40,30	ANN CONSTR & POLYPOID	YES	YES	ADENO + MC	LOW	NO	NO	E	4	2	X	3
MP23	F	73	RIGHT	16	EXOPHYTIC	NO	YES	ADENO	LOW	YES	YES	E	2	0	X	1
MP24	F	46	RIGHT and LEFT	65,110	ANN CONSTR	NO	YES	ADENO + MC	LOW	YES	NO	E	3	0	X	2
MP25	M	41	RIGHT	16	EXOPHYTIC	YES	NO	SIGNET RING		NO	YES	E	4	2	X	3
MP26	M	35	RIGHT	60	ANN CONSTR	NO	NO	ADENO	LOW	NO	YES	E	3	0	X	2
MP27	M	29	LEFT	85	ULCERATED	YES	NO	MEDULLARY		YES	NO	E	4	2	X	3
MP28	M	45	RIGHT	60	ULCERATED	NO	NO	ADENO	LOW	YES	YES	E	3	0	X	2
MP29	F	30	RIGHT	78	EXOPHYTIC	YES	NO	ADENO + MC	LOW	NO	NO	E	4	1	X	3
MP30	F	46	RIGHT	65	EXOPHYTIC	NO	NO	ADENO	LOW	NO	NO	E	3	0	X	2
MP31	F	30	RIGHT	100	EXOPHYTIC	NO	NO	ADENO + MC	LOW	NO	NO	E	3	0	X	2
MP32	F	50	LEFT	57	ANN CONSTR	NO	NO	MUCINOUS		YES	YES	E	3	1	X	3

Appendix 4: Cohort 2 (Mutation positive) Immunohistochemical and gene mutation results

CASE	p21			P27			CYCLIN D1			CYCLIN E			p53			c-myc			MMP7			MLH1	MSH6	MSH2	MUTATION
	P	I	HS	P	I	HS	P	I	HS	P	I	HS	P	I	HS	P	I	CYT	P	I	HS				
MP18	0	0	0	4	2	8	1	2	2	0	0	0	1	1	1	0	0		4	3	12	Neg	Neg	Pos	MLH1 1528C>T
MP19	0	2	0	3	2	6	3	2	6	2	2	4	2	3	6	0	0	Pos	4	2	8	Neg	Pos	Pos	MLH1 1528C>T
MP20	4	3	12	4	3	12	3	3	9	2	2	4	4	3	12	0	0		3	2	6	Neg	Neg	Neg	MLH1 1528C>T
MP21	0	0	0	4	3	12	3	3	9	1	2	2	3	3	9	0	0		4	2	8	Neg	Pos	Pos	MLH1 1528C>T
MP22	0	0	0	4	2	8	2	1	2	0	0	0	4	3	12	0	0	Pos	4	2	8	Pos	Neg	Neg	MSH2 EXON7 del CT AT 1219
MP23	2	2	4	4	3	12	3	2	6	3	2	6	3	2	6	0	0	Pos	4	3	12	Neg	Pos	Pos	MLH1 1528C>T
MP24	0	0	0	4	3	12	4	2	8	2	2	4	0	0	0	0	0		3	1	3	Neg	Pos	Pos	MLH1 1528C>T
MP25	0	0	0	4	2	8	1	1	1	1	1	1	2	2	4	0	0		3	1	3	Pos	Neg	Neg	MSH2 EXON3 del TC AT 387
MP26	0	0	0	4	3	12	2	1	2	3	1	3	3	3	9	0	0		2	1	2	Neg	Pos	Pos	MLH1 1528C>T
MP27	0	0	0	4	2	8	2	2	4	2	2	4	3	3	9	0	0	Pos	4	3	12	Neg	Neg	Pos	MLH1 1528C>T
MP28	0	0	0	4	3	12	1	2	2	2	1	2	4	3	12	0	0	Pos	4	2	8	Neg	Neg	Pos	MLH1 1528C>T
MP29	0	0	0	4	3	12	3	3	9	3	2	6	3	2	6	0	0	Pos	4	3	12	Neg	Pos	Pos	MLH1 1528C>T
MP30	0	0	0	2	2	4	3	2	6	2	2	4	1	2	2	0	0	Pos	4	2	8	Neg	Pos	Pos	MLH1 1528C>T
MP31	0	0	0	4	3	12	3	3	9	3	3	9	2	2	4	0	0	Pos	4	1	4	Neg	Pos	Pos	MLH1 1528C>T
MP32	0	0	0	3	3	9	3	3	9	1	1	1	1	2	2	0	0	Pos	4	3	12	Neg	Pos	Pos	MLH1 1528C>T

Appendix 5: Cohort 3 (Sporadic) Morphological features

CASE	G	AGE	SITE	GD	CONFIG	PERF	ASSOC POLYP	TYPE	GRADE	TIL	CROHNS	TM	T	N	M	STAGE
S33	F	53	RECTUM	30	ULCERATED	NO	NO	ADENO	LOW	NO	NO	E	3	1	X	3
S34	M	83	RECTUM	48	UNKNOWN	NO	NO	ADENO	LOW	NO	NO	If	3	2	X	3
S35	F	78	RECTUM	48	EXOPHYTIC	NO	NO	ADENO	LOW	NO	YES	If	2	0	X	1
S36	F	80	RIGHT	45	EXOPHYTIC	NO	YES	MEDULLARY		YES	NO	E	3	1	X	3
S37	F	73	RIGHT	50	EXOPHYTIC	YES	NO	ADENO+MC	LOW	NO	YES	E	4	0	X	2
S38	M	59	LEFT	90	ANN,CONSTR	NO	NO	ADENO	LOW	NO	YES	If	3	0	X	2
S39	M	80	RIGHT	60	ANN,CONSTR	NO	YES	SIGNET RING		NO	NO	If	3	2	X	3
S40	M	72	RECTUM	65	ANN,CONSTR	NO	NO	MUCINOUS		NO	NO	If	3	0	X	2
S41	M	65	RECTUM	52	EXOPHYTIC	NO	NO	ADENO	LOW	NO	NO	If	3	0	X	2
S42	F	65	LEFT	55	EXOPHYTIC	YES	NO	ADENO	LOW	NO	NO	E	4	0	X	2
S43	F	63	RIGHT	70	EXOPHYTIC	NO	NO	MUCINOUS		NO	NO	If	3	1	X	3
S44	F	60	RECTUM	25	ANN,CONSTR	YES	YES	ADENO	LOW	NO	NO	If	3	0	X	2
S45	M	70	RIGHT	45	UNKNOWN	NO	NO	ADENO	LOW	NO	NO	If	3	2	X	3
S46	M	64	RECTUM	85	ANN,CONSTR	NO	NO	ADENO	LOW	NO	NO	E	3	1	X	3
S47	M	65	LEFT	55	EXOPHYTIC	NO	NO	ADENO	LOW	NO	YES	If	3	1	X	3
S48	M	80	RECTUM	30	ULCERATED	NO	NO	ADENO	LOW	NO	YES	E	2	0	X	1
S49	M	59	RIGHT	59	UNKNOWN	YES	NO	ADENO	LOW	NO	NO	If	4	1	X	3
S50	M	73	RIGHT	50	EXOPHYTIC	YES	NO	ADENO+MC	LOW	NO	NO	If	4	0	X	2
S51	M	64	RIGHT	65	EXOPHYTIC	NO	NO	ADENO+MC	LOW	NO	YES	E	2	0	X	1
S52	F	81	RIGHT	40	EXOPHYTIC	NO	NO	ADENO	LOW	NO	YES	E	3	0	X	2
S53	F	60	LEFT	90	ULCERATED	NO	NO	ADENO	HIGH	NO	NO	If	3	2	X	3
S54	M	73	RIGHT	73	UNKNOWN	YES	NO	ADENO	LOW	NO	YES	If	4	0	X	2
S55	F	64	RECTUM	39	ULCERATED	NO	NO	ADENO+MC	LOW	NO	NO	If	3	1	X	3
S56	F	62	RIGHT	105	UNKNOWN	YES	NO	MUCINOUS		NO	YES	E	4	1	1	4
S57	M	53	RIGHT	25	ULCERATED	YES	NO	ADENO	LOW	NO	YES	E	4	1	X	3
S58	M	57	RIGHT	110	EXOPHYTIC	NO	YES	ADENO	LOW	NO	NO	If	3	1	X	3
S59	F	61	RIGHT	60	UNKNOWN	NO	NO	SERRATED+MC		NO	YES	E	2	0	X	1
S60	F	59	RIGHT	90	EXOPHYTIC	NO	NO	ADENO	LOW	YES	YES	E	3	0	X	2

Appendix 6: Cohort 3 (Sporadic) Immunohistochemical results

CASE	P21			P27			CYCLIN D1			CYCLIN E			P53			c-myc			MMP7			MLH1	MSH6	MSH2
	P	I	HS	P	I	HS	P	I	HS	P	I	HS	P	I	HS	P	I	CYT	P	I	HS			
S33	0	0	0	4	2	8	0	0	0	1	1	1	4	3	12	0	0		4	2	8	Pos	Pos	Pos
S34	0	0	0	4	3	12	0	0	0	3	2	6	4	3	12	0	0	Pos	4	1	4	Pos	Pos	Pos
S35	0	0	0	3	3	9	0	0	0	0	0	0	4	3	12	0	0		4	2	8	Pos	IN	IN
S36	0	0	0	3	2	6	3	2	6	2	2	4	1	1	1	0	0		4	2	8	Neg	Pos	Pos
S37	0	0	0	4	3	12	0	0	0	0	1	0	0	0	0	0	0		4	2	8	Pos	Pos	Pos
S38	0	0	0	4	3	12	0	1	0	0	0	0	4	3	12	0	0	Pos	4	1	4	Pos	Pos	Pos
S39	0	0	0	4	3	12	0	2	0	0	0	0	0	2	0	0	0		4	2	8	Neg	Neg	Pos
S40	1	1	1	4	3	12	2	2	4	0	1	0	3	2	6	0	0		4	3	12	Pos	Pos	Pos
S41	0	0	0	4	3	12	2	1	2	0	0	0	4	3	12	0	0	Pos	2	1	2	Pos	Pos	Pos
S42	0	0	0	3	3	9	3	3	9	0	0	0	2	3	6	0	0	Pos	4	1	4	Pos	Pos	Pos
S43	2	3	6	4	3	12	4	3	12	0	0	0	2	3	6	0	0		4	2	8	Neg	Pos	Pos
S44	0	0	0	3	3	9	2	3	6	2	2	4	4	3	12	0	0	Pos	4	2	8	Pos	Pos	Pos
S45	0	0	0	3	3	9	0	0	0	0	0	0	0	0	0	0	0	Pos	2	1	2	Pos	Pos	Pos
S46	0	1	0	4	3	12	3	2	6	1	1	1	4	3	12	0	0		4	1	4	Pos	Pos	Pos
S47	0	0	0	3	2	6	2	2	4	0	1	0	4	3	12	0	0		4	2	8	Pos	Pos	Pos
S48	0	0	0	3	3	9	1	2	2	0	0	0	0	0	0	0	0	Pos	4	2	8	Pos	Neg	IN
S49	0	0	0	2	2	4	0	0	0	0	0	0	4	3	12	0	0	Pos	4	2	8	Pos	Pos	Pos
S50	0	0	0	3	3	9	2	2	4	0	0	0	4	3	12	0	0	Pos	4	1	4	Pos	Pos	Pos
S51	0	0	0	3	2	6	1	2	2	0	0	0	4	3	12	0	0	Pos	4	2	8	Pos	Pos	Pos
S52	0	0	0	3	2	6	3	2	6	2	1	2	4	3	12	0	0	Pos	4	2	8	Pos	Pos	Pos
S53	0	0	0	4	3	12	2	2	4	1	1	1	0	0	0	0	0	Pos	4	2	8	Pos	Pos	Pos
S54	0	0	0	4	3	12	1	1	1	0	0	0	0	0	0	0	0	Pos	4	3	12	Pos	Pos	Pos
S55	0	0	0	4	3	12	1	1	1	1	2	2	4	3	12	0	0		4	2	8	Pos	Pos	Pos
S56	0	0	0	3	2	6	1	2	2	0	0	0	0	0	0	0	0		3	2	6	Neg	Pos	Pos
S57	1	1	1	3	2	6	3	3	9	3	2	6	0	0	0	0	0	Pos	4	1	4	Pos	Pos	Pos
S58	0	0	0	3	3	9	3	2	6	0	1	0	2	2	4	0	0	Pos	4	2	8	Pos	IN	Pos
S59	0	0	0	4	3	12	4	3	12	1	2	2	0	2	0	0	0		4	1	4	Neg	Neg	Neg
S60	0	0	0	3	2	6	1	1	1	1	1	1	0	0	0	0	0	Pos	4	3	12	Pos	Pos	Pos