CHARACTERIZATION OF COLORECTAL CANCER AND IDENTIFICATION OF PROGNOSTIC DETERMINANTS

ANGELA J. HYDE

Characterization of Colorectal Cancer and Identification of Prognostic Determinants

by
© Angela J. Hyde

A thesis submitted to the School of Graduate Studies
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

Discipline of Genetics / Faculty of Medicine
Memorial University of Newfoundland

May 2012

St. John's Newfoundland

Abstract

Newfoundland and Labrador has the highest rate of colorectal cancer (CRC) in Canada, and its rate of familial CRC is among the highest worldwide. The goal of this PhD programme was to further characterize CRC in this province, to gain a better understanding of the heterogeneity of this disease, and to identify clinically relevant differences in the tumours that could be exploited to personalize and improve clinical care.

The molecular genetic, protein expression, and histological heterogeneities of colorectal cancer were evaluated in three large projects. First, the molecular features related to mismatch repair proficiency were assessed and correlated with family history. In a population-based cohort of 296 patients younger than 70 years at diagnosis, 60% met high or intermediate risk of hereditary CRC according to family history criteria, while only 13% of the associated tumours had defects in the mismatch repair system compatible with Lynch Syndrome. Although Newfoundland has a very high proportion of familial colorectal cancer, Lynch Syndrome does not account for the majority of these cancers; there are likely other, novel hereditary factors contributing to the burden of disease.

Next, the heterogeneity in protein expression was assessed, using tissue microarray technology and immunohistochemistry, in two studies. To determine if the expression of selected proteins can predict patient outcome, expression scores for 12 proteins were assessed together in unsupervised hierarchical clustering analysis using data from 280 tumours from population-based patients along with another 111 from unselected

CRC patients diagnosed younger than age 75 years. This analysis identified three tumour subgroups, one of which had statistically significant reduced disease-specific survival after adjustment for known prognostic factors (adjusted hazard ratio 1.82). The conclusions were that expression of specific proteins studied may have prognostic utility, and that the statistical methods used were valuable for identifying prognostically relevant subgroups of CRC.

In the same cohort, lamin A/C protein expression in tumour nuclei was assessed to determine whether it predicted prognosis. Stage III patients whose tumours expressed lamin A/C in >25% of cells were four times more likely to die of their disease when compared to stage III patients with lower levels of expression (adjusted hazard ratio 4.14).

Finally, to determine whether features of tumour histology were good predictors of microsatellite instability, hemotoxylin and eosin-stained tumour sections were studied from a population-based group of 710 incident CRCs. A number of histological features were associated with microsatellite instability, and an algorithm (PREDICT) was developed to predict this status using multivariate modelling. In a validation cohort of 280 population-based incident CRCs, PREDICT had a sensitivity of 96.9% and a specificity of 76.6% for identifying microsatellite instable tumours. Microsatellite instability can be used to assist in the identification of Lynch Syndrome, and in both the prediction of patient outcome and response to standard adjuvant chemotherapy. This algorithm has been adopted into clinical practice in parts of Ontario, and as part of a Community Genetics initiative in Newfoundland and Labrador.

Acknowledgements

This PhD programme, from planning the initial proposal, to collecting and analyzing the data, and writing up the results, was a long process during which many obstacles were encountered and ultimately overcome. Many people were involved in helping me navigate these obstacles, and in supporting me through this degree.

My supervisory committee has changed many times over the years, through changing requirements for expertise as my programme progressed and adapted to the ever-changing research field, and also through a retirement, multiple resignations and a death. Throughout all of these changes, the importance to me of the constant support, the enthusiasm for teaching, and the gentle guidance from my primary supervisor, Dr Ban Younghusband, cannot be understated. I have no doubt that completing this thesis would not have been possible without his dedication to me and to my programme. I will forever be grateful for the many ways that he supported me. From the hours he spent patiently explaining different molecular genetic concepts and teaching me the history and people behind them, to our shared enthusiasm for music, and everything in between. He is not only my supervisor, but also a true mentor and friend. Thank you Ban.

The late Dr. Desmond Robb was instrumental in the development of the initial proposal for my programme. He introduced me to histology and pathology, guided me through the first project in this thesis, and laid the groundwork that made the rest of my programme possible. I will never forget the hours we spent digging through overheated storage rooms full of old histology slides and surgical pathology specimens, or the hours

shared looking through his microscope. His enthusiasm was contagious, and he is missed by many. Thank you Des.

I would like to thank Dr. Dan Fontaine for stepping in as the pathologist on my supervisory committee after Des' death. He always found the time to sit with me and help me score the histology and immunohistochemistry, even going out of his way to come in for many hours late at night when we had deadlines. Throughout his resignations and eventual moving away, he didn't walk away from my project, but remained dedicated to seeing it through for me. Thank you Dan.

I would also like to thank Dr. Pat Parfrey for is constant support and vision throughout my programme, and for stepping up as my co-supervisor upon Ban`s retirement. His ability to see the big picture within a mess of data, as well as his writing and editing talents, and his determination to make things work are all very much appreciated. Thank you Pat.

Since my projects fell under the larger umbrella of the interdisciplinary colorectal cancer research team, my work was never in isolation, but always involved the input and assistance of others. I would like to take the opportunity to thank a number of the people who helped make my projects possible. For help with statistical analyses I would like to thank Susan Stuckless, Peter Campbell, and Tyler Wish. I would like to thank Cathy Searle for her assistance in teaching me how to work with surgical pathology specimens and perform immunohistochemical staining of tissue sections. I would like to thank Dr. Jane Green for her constant support, and for providing me with a quiet desk for writing the bulk of my thesis. I would also like to thanks Dr. Betty Dicks for providing me with another quiet desk for concluding my thesis writing. I would like to thank Fiona Curtis,

Aimee Adams-Francis, Amanda Dohey, Laura Edwards, and Dr. Chris Butt for being my colleagues and friends as we went along this journey together.

For their financial support I would like to thank Memorial University of Newfoundland, the CIHR-IHRT Team Grant and associated educational bursary, as well as the CIHR and Canadian Gene Cure Foundation for my MD/PhD Scholarship. I would particularly like to thank the CIHR and Canadian Gene Cure Foundation, not just for the financial assistance, but also for the support and mentorship that came with it. Through numerous meetings with the other scholarship recipients from across the country I gained a wonderful network of friends and colleagues, through which we developed the Clinical Investigator Trainee Association of Canada. Through these meetings we were also introduced to Dr. Charles Scriver, the namesake of the scholarship, who continues to have a personal interest in supporting us through our training. Dr. Scriver has become another mentor of mine, and I would like to thank him for his inspiration and guidance.

Most importantly, I would like to thank my family. While I have greatly enjoyed my work on this PhD, it has taken a long time, and there have been many challenges along the way. The love and support from my family, and their unwavering belief in my ability to do this, and do it well, mean more to me than I can ever express. Thank you. Thank you to my parents, Bob and Elaine Hyde. Thank you to my brother and sister-in-law, Danny and Cathy Hyde. Thank you to my grandfather, who proudly tells family and friends of my PhD "in DNA". Thank you to my children, Grace and Thomas, and of course, to my husband, Jim Parsons, whose love and support means the world.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

5-FU	5-fluorouracil
6-TG	6-thioguanine
A	adenine
AAA^+	ATPase associated with various cellular activities
ACF	aberrant crypt foci
aCGH	array comparative genomic hybridization
ACI	Amsterdam Criteria I
ACII	Amsterdam Criteria II
ACMAC	Age and Cancer Modified Amsterdam Criteria
aFAP	attenuated Familial Adenomatous Polyposis
APC	adenomatous polyposis coli
ARF	alternate reading frame
ASCO	American Society of Clinical Oncology
AXIN1	axis inhibitor 1
BRRS	Bannayan-Riley-Ruvalcaba Syndrome
BER	base excision repair
BMI	body mass index
C	cytosine
CAP	College of American Pathologists
CCOPEBC	Cancer Care Ontario Program in Evidence-Based Care

CDK	cyclin-dependent kinase
CDKI	cyclin-dependent kinase inhibitor
cDNA	complementary deoxyribonucleic acid
CEA	carcinoembryonic antigen
CEMA	cutting edge matrix assembly
CFR	Cancer Family Registry
CFS	Cancer Family Syndrome
CHRPE	congenital hypertrophy of retinal pigment epithelium
CI	confidence interval
CIHR	Canadian Institutes of Health Research
CIMP	CpG island methylator phenotype
CIN	chromosomal instability
CIP	cyclin dependent kinase interacting protein
CK1	casein kinase 1
CLR	Crohn's-like lymphocytic reaction
CpG	5'-cytosine - guanine-3' dinucleotide
CRC	colorectal cancer
CT	computed tomography
CUP	cancer of unknown primary
DCBE	double contrast barium enema
DCC	deleted in colorectal carcinomas
DNA	deoxyribonucleic acid
EGFR	epidermal growth factor receptor

EMT	epithelial-mesenchymal transition
ER	estrogen receptor
FAP	Familial Adenomatous Polyposis
FCCTX	familial colorectal cancer type X
FdUMP	5-fluoro-2`-deoxyuridylate
FISH	fluorescence *in situ* hybridization
FOBT	faecal occult blood test
G	guanine
GI	gastrointestinal
GIST	gastrointestinal stromal tomour
GSK3	glycogen synthase kinase 3
H & E	haematoxylin and eosin
HER2	human epidermal growth factor receptor 2
HNPCC	Hereditary Non-Polyposis Colorectal Cancer
HPS	Hyperplastic Polyposis Syndrome
IBD	inflammatory bowel disease
IDL	insertion/deletion loop
IHC	immunohistochemistry
IHRT	Interdisciplinary Health Research Team
INK4	inhibitor of cyclin dependent kinase 4
IRR	incidence rate ratio
KIP	cyclin dependent kinase inhibitory protein
LAP	lamin-associated protein

LEF	Lymphoid enhancer factor
LI	labelling index
LOH	loss of heterozygosity
mAb	monoclonal antibody
MALT	mucosa-associated lymphoid tissue
MAP	MUTYH-Associated Polyposis
MAPK	mitogen-activated protein kinase
Mb	megabase
MCM	minichromosome maintenance
MEN	multiple endocrine neoplasia
MET	mesenchymal-epithelial transition
MGMT	O^6-methylguanine DNA methyltransferase
MIN	microsatellite instability
MLPA	multiplex ligation-dependent probe amplification
MMR	mismatch repair
MNNG	N-methyl-N'-nitro-nitrosoguanidine
MSH	mutS homolog
MSI	microsatellite instability
MSS	microsatellite stable
NCI	National Cancer Institute
NFCCR	Newfoundland Familial Colorectal Cancer Registry
NIH	National Institutes of Health
NSAID	non-steroidal anti-inflammatory drug

OCT	optimal cutting temperature
PBS	phosphate buffered saline (not used yet – was, then removed)
PCNA	proliferating cell nuclear antigen
PCR	polymerase chain reaction
PI3K	phosphatidylinositol 3-kinase
PMS	post-meiotic segregation protein
PR	progesterone receptor
pRb	retinoblastoma protein
PSA	prostate specific antigen
RER	replication errors
RBG	Revised Bethesda Guidelines
RNA	ribonucleic acid
RT-PCR	reverse transcriptase polymerase chain reaction
SIFT	sorting intolerant from tolerant
SIR	standardized incidence ratio
T	thymine
TCF	T-cell factor
THBS1	thrombospodin 1
TILs	tumour-infiltrating lymphocytes
TKI	tyrosine kinase inhibitor
TMA	tissue microarray
TNM	tumour nodes metastases
TS	thymidylate synthetase

UC	ulcerative colitis
WHO	World Health Organization
Wnt	Wingless Int

List of Appendices

Chapter 1 – Introduction and Review of Literature

1.1 Introduction and History

Colorectal cancer (CRC) is one of the most common forms of malignancy, representing almost 10% of all new cancer diagnoses and 8% of deaths from cancer worldwide [1]. It is third in overall incidence, after lung and breast, and second in prevalence, after only breast [2]. CRC is a heterogenous disease, with different tumours arising from different causes, developing through different pathways of carcinogenesis, exhibiting different pathological characteristics, and progressing with different levels of aggression. While the majority of CRCs are sporadic, 5% to 10% of CRCs are hereditary, passed from one generation to the next in a Mendelian pattern of inheritance [3]. The two major forms of Mendelian CRC are Familial Adenomatous Polyposis (FAP) and Lynch Syndrome, also known as Hereditary Non-Polyposis Colorectal Cancer (HNPCC). Both are inherited in an autosomal dominant pattern. Lynch Syndrome is the most common form, accounting for up to 3% of all CRC [4-6]. Another 20% of CRC occurs in patients with a familial risk; defined as individuals having two or more first- or second-degree relatives with a history of colorectal cancer [3]. A genetic cause is unknown for the majority of familial clustering, which may also be due in part to shared environmental exposures.

The first published evidence that some cancers could be hereditary came from Aldred Scott Warthin, an American pathologist, in his seminal paper of 1913 [7]. Warthin

conducted the first population-based study of cancer, based on the 3600 neoplasms that were diagnosed in the Pathology Department of the University of Michigan, where he worked, from 1895 to 1913. He studied the available family history data collected from all cases of carcinoma, and made many insightful conclusions that hold true today. His foremost conclusion was that "a marked susceptibility to carcinoma exists in the case of certain family generations and family groups" [7]. He also noted that when a carcinoma was seen in multiple members of one generation of a family there was almost always carcinoma in the preceding generation, and that the cancer burden within a family was greater when there was a history of carcinoma in the families of both parents as compared to just one. Warthin recognized that careful studies of large families were needed to enable accurate descriptions of patterns of inheritance. Finally, he had the foresight to make a subtle, but important distinction in the approach to the emerging concept of the heritability of cancer. While others, such as Isaac Levin, proposed the common inheritance of a resistance to cancer that was absent in families with cancer [8], Warthin proposed the opposite: that resistance to cancer was the normal state, and families manifesting an increased frequency of cancer possessed a trait conferring the increased susceptibility. For this work he has been termed by some as "the father of cancer genetics" [9].

Of the pedigrees presented by Warthin in his 1913 paper, one would eventually prove key in the story of hereditary colorectal cancer. Warthin first became aware of this family in 1895. His seamstress was depressed due to what she considered to be her inevitable fate, that she would die young of cancer, as was common in her family.

Warthin took an interest in her story and researched her family, which he called Family G, as they were of German ancestry [10]. He found a predominance of cancers of the bowel and endometrium, as well as cancers of the stomach and ovary [7,10]. This was the first published record of a family suffering from what is now referred to as Lynch Syndrome, the most common cause of hereditary colorectal cancer.

In the mid-1960s an internal medicine resident named Henry Lynch became fascinated by the story of one of his patients who spoke of the overwhelming history of colorectal and other cancers in his family [11]. Lynch and colleagues carried out a thorough review of the medical and pathology records of the family, and recognized that they were dealing with a type of cancer syndrome that had yet to be defined in literature. The family exhibited an autosomal dominant transmission of a trait which predisposed to early onset carcinoma. The predominant form of cancer was of the colon, however cancers of the endometrium and ovary were also common among the female family members. This family became known as Family N, for Nebraska, the state in which they resided. In 1966 Lynch and colleagues published a report on Family N and another family with a similar cancer history, Family M, from Michigan [12]. This was the first publication specifically describing this new cancer syndrome. By the late 1960s many researchers were reporting families which had cancer manifestations and frequencies that were similar to Families G, M and N (reviewed in [13]). Based on the research of these and other families, in 1971 Lynch and his colleague Anne Krush published the first criteria for "Cancer Family Syndrome" (CFS) [14], which eventually became known as Lynch Syndrome.

In the late 1960s, upon hearing of their work, A. James French of the University of Michigan offered to share with Lynch and Krush the notes of his late predecessor, Aldred Warthin. These notes contained the pedigree, pathological and clinical information pertaining to Warthin's original Family G. Lynch and Krush proceeded to revisit the family, and updated their status in 1971 [15]. They and others continued to follow this family, and in 2000, more than a century after the family was first identified, the mutation predisposing this family to cancer was discovered in the mismatch repair gene *MSH2* [16,17]. The *MSH2* gene had been mapped in 1993 [18] through the identification of close linkage to a locus 4Mb downstream of this gene among affected individuals from two large families with early-onset cancer. One of these families was from Newfoundland. The mapping of this gene was the first proof of a genetic susceptibility to colorectal cancer.

1.2 Colorectal Anatomy and Pathology

The colon and rectum make up the distal end of the human gastrointestinal tract. The colon is approximately 1.5 m long, and the rectum is 12 to 15 cm long [19]. The colon is divided into five anatomic segments, as illustrated in Figure 1.1. The terminal ileum, the most distal portion of the small intestine, is connected via the ileocecal valve to the cecum, the most proximal segment of the colon. The ileocecal valve allows bowel contents to pass from the small to the large intestine, but prevents reflux back into the small intestine. Distal to the cecum is the ascending colon, which turns at the hepatic

flexure to become the transverse colon. The transverse colon then turns at the splenic flexure to become the descending, and then sigmoid colon. The sigmoid colon extends into the s-shaped rectosigmoid junction, to become the rectum. The cecum has the largest diameter, and the colon progressively narrows distally.

The primary functions of the colon and rectum are to reabsorb water and electrolytes from the luminal contents, and store and eliminate waste. The surface epithelium of the colonic mucosa is composed of two main types of cells: columnar cells and goblet cells. The most numerous are the columnar cells, which are involved in absorption of water and salt from the luminal contents. They predominate over the goblet cells, by a ratio of four to one throughout most of the colon. However, approaching the rectum this ratio decreases, with both cell types present in approximately equal numbers [20]. Goblet cells produce and secrete mucus, which acts as a lubricant to aid in the passage of waste. More mucus is required as the waste becomes more solid, at the distal end of the colon, where the numbers of goblet cells is highest. The epithelial cells are organized into numerous straight, tubular glands (crypts of Lieberkühn), which extend from the smooth luminal wall of the colon to the first layer of muscle within the wall of the colon, known as the muscularis mucosa. This is illustrated in Figure 1.2. The layers of the colon wall are illustrated in Figure 1.3.

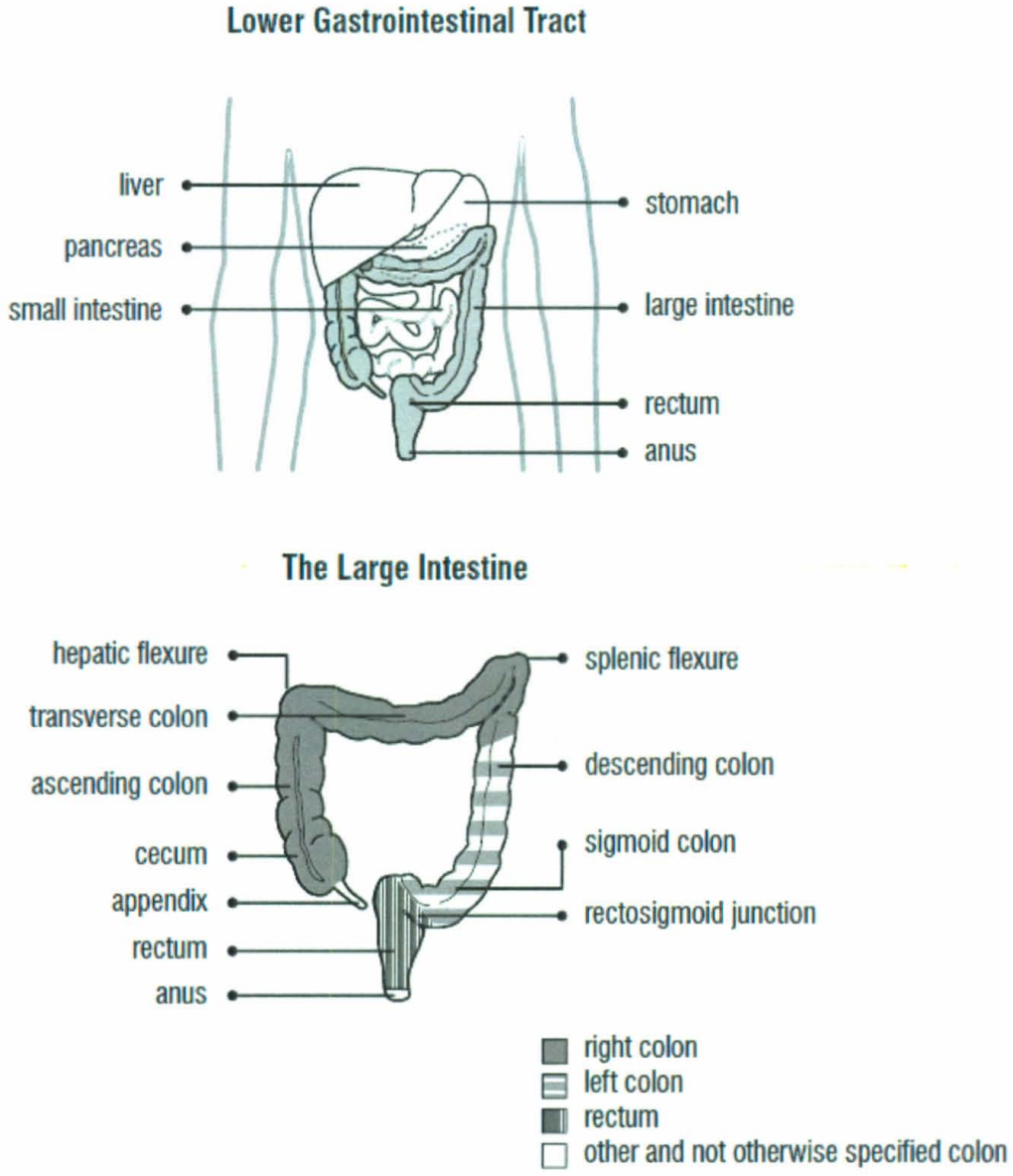


Figure 1.1. Anatomy of the colon and rectum
The upper image illustrates the location of the colon within the abdomen in relation to other organs of the lower gastrointestinal system. The lower image illustrates the different sections and landmarks of the large intestine. Note that while the anus is included in the figure, it is not included in the definition of colorectal cancer used in this thesis. Figure from *Canadian Cancer Statistics 2011* [21].

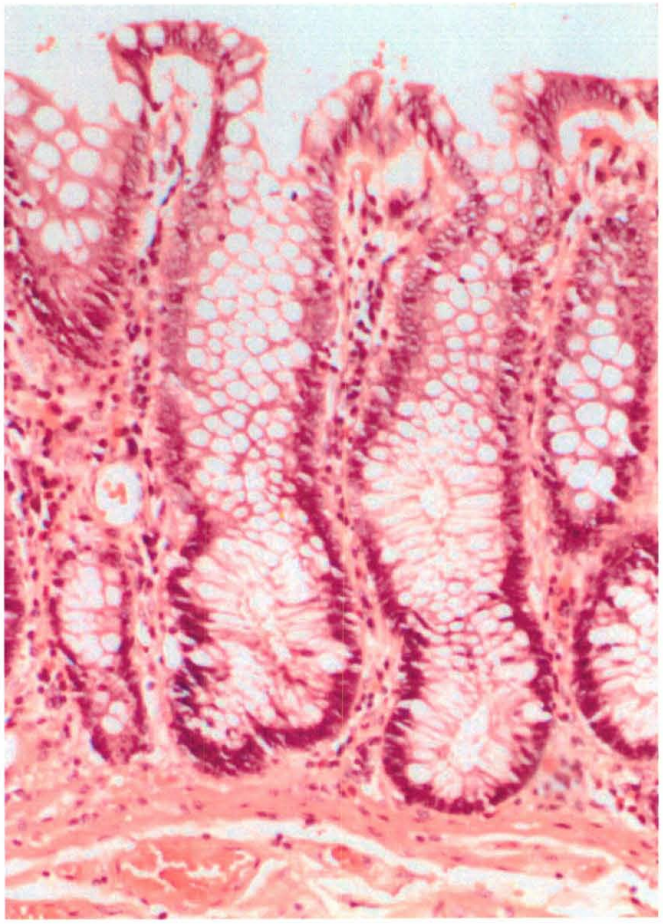


Figure 1.2. Colonic crypts of Lieberkühn

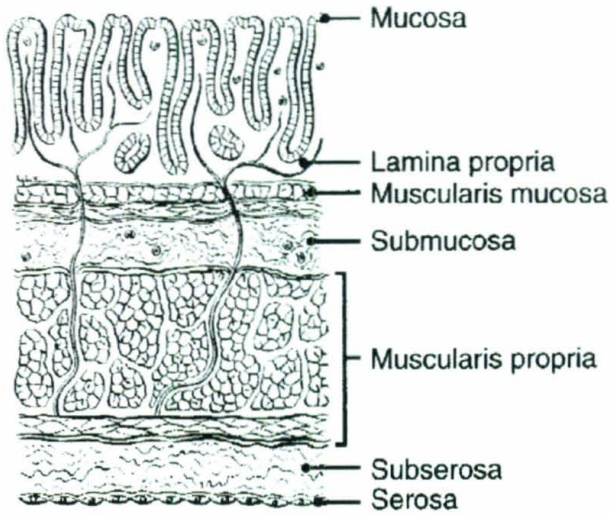


Figure 1.3. The layers of the wall of the colon and rectum (from [22], reprinted with permission from Springer)

Another two types of cells, endocrine and stem cells, are located in the base of the crypts. The endocrine cells are few in number, and they are scattered among the other epithelial cells. They produce and secrete various hormones, which are released along the basal or basolateral surface of the cell, but not on the apical (luminal) side [19]. These hormones modulate normal digestive functions.

The colonic stem cells can differentiate to produce any of the specialized epithelial cells of the colon: the columnar cells, goblet cells, and endocrine cells. The lower one third of the crypt is known as the replicative zone, and all colonic epithelial cells arise there from stem cells. With every cycle of mitosis, a stem cell divides into two cells. They can divide asymmetrically, resulting in: a new stem cell, which remains in the replicative zone of the crypt, and a daughter cell which differentiates and migrates up the

crypt to the luminal surface; or they can divide symmetrically, resulting in either two daughter cells or two stem cells [23]. Cells at the surface are frequently shed due to surface abrasion and senescence. Turnover of the columnar and goblet cells occurs every four to six days, and of the endocrine cells every six weeks [20]. This rapid renewal of the colonic epithelium allows for quick repair of damage, but also leaves the colon particularly vulnerable to replicative error and to agents that interfere with cellular replication [19].

The anatomy and physiology of the colon can affect how a colorectal cancer presents clinically. Typically, colorectal cancers are asymptomatic for a long time, but can become symptomatic as they grow larger. Some patients experience fever, a general feeling of being unwell, weight loss, and abdominal pain. Symptoms depend on the size and location of the tumour. Because the colon progressively narrows distally, and because water is absorbed from bowel contents as it moves distally, the proximal colon has a larger circumference and less solid contents than the distal colon. Therefore, cancers of the proximal colon tend to present in a more indolent manner, with less specific findings, including microcytic anemia, fatigue, and dark, tarry stool (known as melena). Cancers of the distal colon tend to present with more acute findings, including a change in bowel habits, such as passing stools with a smaller than normal caliber. They can also present with constipation, abdominal distension, bowel obstruction, tenesmus (the urgent and painful need to pass stool, often without being able to do so), or fresh blood in the stool, known as hematochezia.[24,25]

1.3 Histological Classification of Colorectal Tumours

The World Health Organization (WHO) classifies colorectal cancers into three main categories: (1) epithelial tumours, (2) non-epithelial tumours, and (3) secondary tumours [25]. Epithelial tumours are those derived from the cells of the epithelium, and can be further divided into carcinomas and carcinoids. Carcinoids are well-differentiated tumours arising from endocrine cells. Carcinomas arise in the other epithelial cells, both columnar and goblet, and are the focus of this thesis. Adenocarcinomas originate from the glandular cells of the colonic crypts, and are the most common subtype of carcinoma. Some of these are mucin-producing, but the majority are not. Adenocarcinomas with mucin composing greater than 50% of the lesion are termed mucinous adenocarcinomas. Tumours with at least 50% of cells containing prominent intracytoplasmic mucin are termed signet-ring carcinomas. Other types of colorectal carcinomas include medullary and undifferentiated carcinomas.

Non-epithelial tumours include lipomas, which originate in adipose cells; leiomyomas and leiomyosarcomas, which originate in smooth muscle cells; gastrointestinal stromal tumours, which arise in stromal cells; malignant melanomas, which arise in melanocytes; malignant lymphomas, which can arise from a variety of lymphatic cells of the immune system; and others.

Malignant melanoma can present in the colon and rectum as either a primary tumour or a metastasis from a distant site. Primary malignant melanoma of the large intestine is rare, accounting for less than one percent of malignancies at this site [26,27].

Secondary tumours of the colon and rectum are tumours which arise in other sites within the body and metastasize there, as well as local recurrences of previous primary cancers of the colon or rectum. Distant, metastatic spread to the colon and rectum can occur from primary lung cancer and malignant melanoma, and rare cases from breast have been reported [25,28]. Cancers of the bladder, ovary, cervix, and prostate can also spread to the colon and rectum, but typically by local extension rather than true metastases.

1.4 Tumour Staging Methods

Staging is an important step after the diagnosis of cancer, to determine the extent of disease for the purpose of guiding treatment and estimating patient prognosis. In 1926, John Percy Lockhart-Mummery published the first attempt to stage rectal cancer [29]. Lockhart-Mummery was a surgeon, and his 1926 paper described the development and use of his new perineal surgical approach to rectal cancer in 200 patients. He argued that it was an improvement over the other surgical approaches of the time, specifically the abdomino-perineal operation that was then method of choice. He classified the patients in his study into three categories, depending on the extent of the primary tumour, and then used these classifications in the analysis of outcome, for the purpose of comparing the effect of the surgical techniques on patient survival. The three categories, in Lockhart-Mummery's words, are presented in Table 1.1.

Table 1.1. Lockhart-Mummery staging method for rectal cancer (from [29])

A	very favourable cases where the growth was small and had not apparently invaded the muscular coat, and no glands were involved.
B	medium cases, where there was involvement of the muscular coat, but where the growth was not unduly fixed and there was no extensive involvement of glands.
C	very bad cases, where the growth was large and fixed, or where there was evidence of extensive involvement of glands. These were borderline cases with bad prognosis.

Of the 200 patients included in the study, 73 were categorized as group A, 96 as group B, and 31 as group C. Considering only those patients for whom there was at least five years of follow-up time since surgery, and excluding patients who had died of other causes or were lost to follow-up, Lockhart-Mummery found a significant difference in outcome in the group A patients when compared to groups B and C. Of the patients in group A, 74% were still alive and cancer free after five years, compared to 44% of patients in each of groups B and C. He stated that he expected that the values for group C were misleading due to the small number of patients (nine) included in analysis. [29]

Four years after the publication of Lockhart-Mummery's paper, modifications to his staging categories were suggested by Cuthbert Dukes in a paper co-authored with Sir Charles Gordon-Watson [30]. Dukes wrote the introduction, titled "The Spread of Cancer of the Rectum", to Gordon-Watson's article "The Treatment of Carcinoma of the Rectum with Radium". Dukes was a pathologist at St Mark's Hospital in London, the same hospital at which Lockhart-Mummery was a surgeon. He proposed subdividing both the B and C categories. He suggested that Group B tumours should be classified as B1 if only the more superficial of the two muscular layers, the circular muscle, was involved,

or as B2 if the deeper, longitudinal muscle layer was involved. Group C tumours should be classified as C1 if lymph nodes were not involved, or C2 if there was metastatic spread to lymph nodes.

In 1932, two years after his paper with Gordon-Watson, Dukes published another article, called "The Classification of Cancer of the Rectum" [31]. The model proposed in this paper has become known as the Dukes staging method, and with modifications, remained the staging model of choice for decades. Although his research was on rectal cancer, he stated that his method could be applied to all intestinal cancers. The model used today, the Tumour Nodes Metastasis (TNM) Classification System, is in fact very similar to the Dukes staging method. Dukes approached the staging of rectal tumours as a two-step process. The first step involved the proper handling and processing of the tumour tissue, and careful collection of pathological data. Dukes stressed the importance of this work, and described the process in detail. The pathological data was then used for the second step, which was to classify the tumours into one of three categories. The three categories are presented in his words in Table 1.2.

Table 1.2. Dukes' 1932 staging method for rectal cancer
(from [31])

A	cases are those in which the carcinoma is limited to the wall of the rectum, there being no extension into the extrarectal tissues and no metastases in lymph nodes,
B	cases those in which the carcinoma has spread by direct continuity to the extrarectal tissues but has not yet invaded the regional lymph nodes, and
C	cases those in which metastases are present in the regional lymph nodes.

In this 1932 paper [31], Dukes made a number of important observations and comments that are still relevant today. His emphasis on the importance of careful attention to the collection of pathological data was key. While his staging model assessed only the extent of tumour spread, Dukes did mention that prognosis could not depend solely on this feature, but also on others such as the patient's physical condition, psychology, and social environment. He also mentioned the prognostic value of grading a tumour through histology to assess the probable rate of growth. Dukes had the foresight to recognize the potential importance of tumour screening and early tumour identification. He observed that patients coming from private practice had a slightly higher proportion of lower stage, category A and B tumours than those coming from hospital practice. He stated that there would be an increase in the proportion of A and B tumours if diagnoses could be made earlier.

While Dukes' model, as proposed in 1932, was valuable for categorizing tumours for the purpose of patient prognosis, it did have some limitations. In his model he referred to the spread of growth to or through the rectal wall; however, he did not define what he considered to be the rectal wall. While most pathologists have interpreted this to mean the muscular component of the wall, there has been some controversy. He also believed at that time that spread to lymph nodes could not occur if tumour growth was limited to the rectal wall. As well, he did not address the importance of distant metastases. Today, most patients with distant metastases are not considered to be curable, and are instead treated palliatively. Dukes did not have a separate classification for palliative cases. Many of these issues were addressed in subsequent modifications to Dukes' staging method.

In 1949, Kirklin and colleagues modified Dukes' staging method by dividing Dukes' A into A and B1 [32]. Tumours confined to the bowel wall were classified as A if they had not spread beyond the mucosa, or B1 if they had spread to the submucosa or to, but not through, the muscularis propria. Dukes' B became B2. An issue with this modification is that the Kirklin-modified group A does not actually represent true invasive carcinoma. Today the neoplasm would be considered a carcinoma *in situ*, since the tumour is confined to the mucosa and has not penetrated the lamina propria.

In 1954, Astler and Coller [33] further modified the Kirklin-modified Dukes' staging method. They recognized that spread to lymph nodes could occur even when the tumour growth was limited to the bowel wall. For this reason they subdivided Dukes' C into C1 and C2. Tumours with metastases to lymph nodes were classified as C1 if the continuous direct spread of the tumour was limited to the bowel wall, and as C2 if it spread beyond the bowel wall.

In 1958 Dukes published an article with H. Bussey [34], in which they summarized a 25 year research project designed to examine the effect of tumour spread on prognosis. During the 25 year period from 1928 to 1952, 3596 patients with rectal cancer were treated at St. Mark's Hospital in London. Of these, 2447 were treated by surgical excision of the tumour. Through rigorous patient follow-up, clinical outcome data was available for all but 28 (1.1%) of these patients. After adjusting for age and sex, they determined that the average five year survival rate was 57.4%. When patients were classified by Dukes' original method, the adjusted five year survival rate for patients with group A tumours was 97.7%, group B 77.6%, and group C 32.0%.

In this paper [34], Dukes and Bussey subdivided group C into C1 and C2. This subclassification differed, however, from that of Astler and Coller [33]. Rather than base the classification of group C tumours on the extent of direct tumour spread, Dukes and Bussey based it on the location of affected lymph nodes [34]. They assigned tumours to group C1 if only regional lymph nodes were affected, while those with more extensive spread, including lymph nodes found along the points of merging of blood vessels, were assigned to group C2. Dukes and Bussey also used their extensive research data to assess the prognostic significance of other pathological features, including histological grade and spread within veins; both of which they found to be valuable. Again, however, palliative cases were not considered in a separate group from potentially curable cases.

A further modification to the Dukes' staging method came in a publication by Turnbull and colleagues in 1967 [35]. Turnbull had previously recognized the potential importance of finding cancer cells in venous blood. He suspected that such cells could cause distant metastases, and believed that they entered the bloodstream due to mechanical manipulation during surgery. He developed a "no-touch isolation" technique, whereby the cancer-containing segment of bowel would not be handled during surgery until the lymphovascular pedicles were divided and ligated, and the colon divided for resection. His 1967 paper [35] compared patient survival rates using this "no-touch isolation" technique and conventional surgical methods, and found the "no-touch" approach to be superior. In comparing the patients in the two treatment arms of the study, they were classified according to a modified version of Dukes' staging, whereby there was a fourth group, D, added to the classification, representing patients with distant

metastasis. This latest version of the Dukes' staging system is the one most commonly associated with the Dukes' name today. It was presented as follows in Table 1.3.

Table 1.3. Dukes' staging method for colorectal cancer as modified by Turnbull (from [35])

A	Tumour confined to the colon and its coats
B	Tumour extension into pericolic fat
C	Tumour metastasis to regional mesenteric lymph nodes, but no evidence of distant spread
D	Tumour metastasis to liver, lung, bone, seeding of tumour, irremovable because of parietal invasion; adjacent organ invasion

While the addition of stage D was important for prognostic value, it did complicate the staging system, in that it included clinical data into an otherwise pathological system. While some distant metastases can be identified pathologically, many cannot. Liver metastases and peritoneal seeding can be identified if observed and biopsied at the time of colorectal surgery. Invasion into adjacent organs can also be identified if portions of these organs were also removed surgically. In other situations, however, distant metastases are identified clinically, and surgery would only be performed palliatively, if at all. There is still some confusion regarding the mixing of clinical and pathological data, and the mixing of palliative and potentially curable cases in this stage category.

Through all of these modifications, the Dukes' method remained the model of choice in clinical practice for staging of colon and rectal tumours, and retained the Dukes' name. This led to tremendous confusion in the comparison of research results and

clinical outcomes between groups, as different versions of the Dukes' method were often used [36]. However, the value of the Dukes' method for colorectal cancer is still recognized as a powerful tool for prognostication.

The Dukes' method has, however, been replaced in clinical practice by the Tumour Nodes Metastases (TNM) classification system [37]. The TNM system was developed by Pierre Denoix, a French surgeon who is recognized as one of the founders of modern oncology [38]. Denoix developed the TNM system from 1943 to 1952 in a series of short publications [39,40]. This system was not based on scientific research, but instead was a method of encoding the pathological and clinical data that was relevant to prognosis. It was simple in design, and could be applied to tumours at all sites within the body, with only small variations. Scores were given for extent of tumour growth and the presence or absence of metastases to lymph nodes and distant sites. In 1953, the system was adopted by the International Union Against Cancer (UICC).

Although simple, the TNM system was cumbersome to use clinically. A score from 0 to 4 was assigned for T, the size or extent of tumour growth; from 0 to 4 for N, the increasing number or extent of lymph node involvement; and of 0 or 1 for M, the absence or presence of distant metastases. Scores could also be given for other prognostic features including histologic grade (G1 through G4), and residual tumour following surgery (R0 through R2) [41].

The American Joint Committee on Cancer (AJCC) was founded in 1959, and became involved with the TNM system. In 1977 they published the first "Manual for Staging Cancer" [41], in which they translated the cumbersome scores of the TNM system

into four stages (I, II, III, and IV), which were identical to the four stages of the original 1932 Dukes method [31] as modified by Turnbull *et al* [35]. Since then the UICC and AJCC have jointly maintained and revised the TNM system as new data has become available. The most recent publications from each group were released in late 2009 [37,42,43], and came into effect on January 1, 2010. For colorectal cancer, the T, N, and M categories were defined as follows in Table 1.4. Illustrations of the T categories are given in Figure 1.4.

Table 1.4. T, N, and M categories for colorectal cancer, as defined by the American Joint Committee on Cancer
(from [43])

a.) Primary Tumour (T):

Tx	Primary tumour cannot be assessed
T0	No evidence of primary tumour
Tis	Carcinoma in situ: intraepithelial or invasion of lamina propria (ie intramucosal)
T1	Tumour invades submucosa
T2	Tumour invades muscularis propria
T3	Tumour invades through the muscularis propria into pericolorectal tissues
T4a	Tumour penetrates to the surface of the visceral peritoneum
T4b	Tumour directly invades or is adherent to other organs or structures

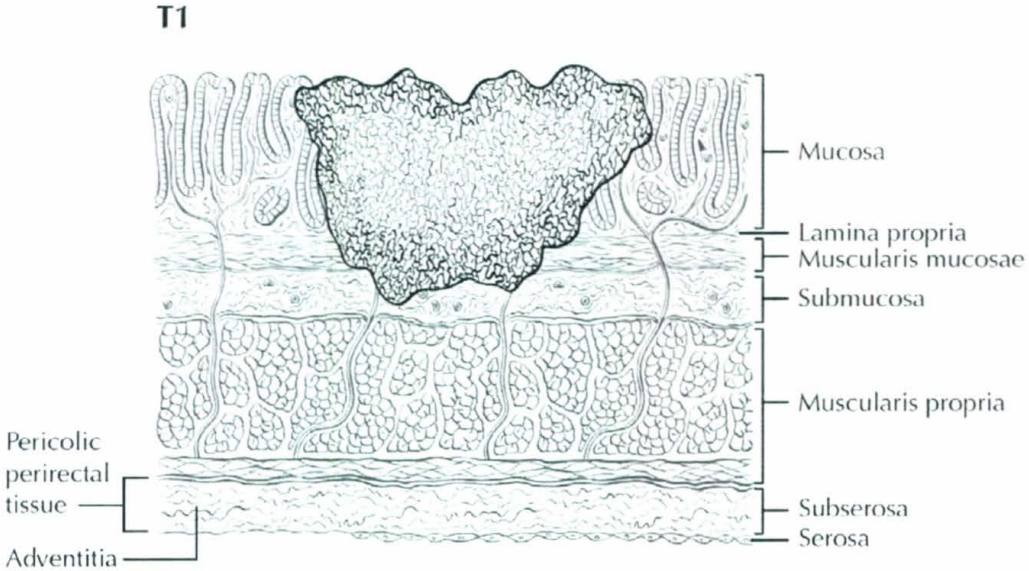
b.) Regional Lymph Nodes (N):

NX	Regional lymph nodes cannot be assessed
N0	No regional lymph node metastasis
N1	Metastasis in 1-3 regional lymph nodes
N1a	Metastasis in one regional lymph node
N1b	Metastasis in 2-3 regional lymph nodes
N1c	Tumour deposit(s) in the subserosa, mesentery, or nonperitonealized pericolic or perirectal tissues without regional nodal metastasis
N2	Metastasis in four or more regional lymph nodes
N2a	Metastasis in 4-6 regional lymph nodes
N2b	Metastasis in seven or more regional lymph nodes

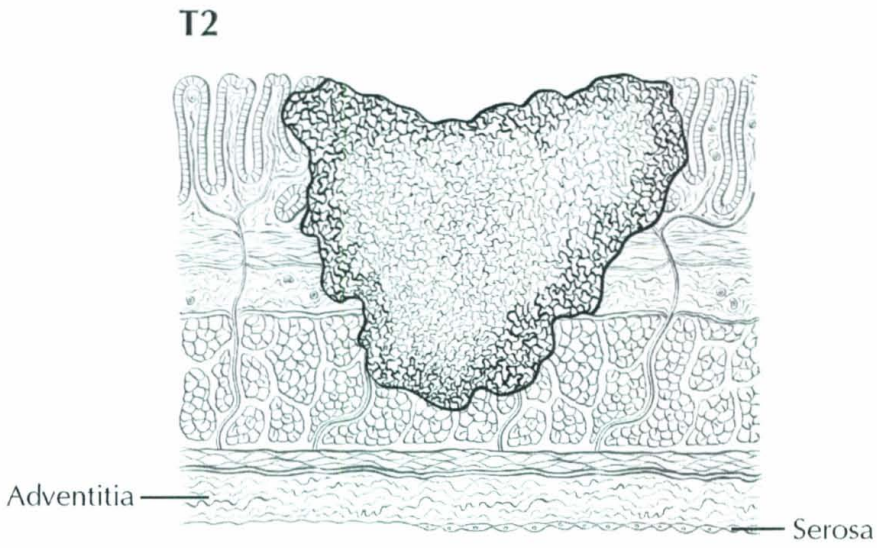
c.) Distant metastasis (M):

M0	No distant metastasis
M1	Distant metastasis
M1a	Metastasis confined to one organ or site (e.g., liver, lung, ovary, nonregional node)
M1b	Metastasis in more than one organ/site or the peritoneum

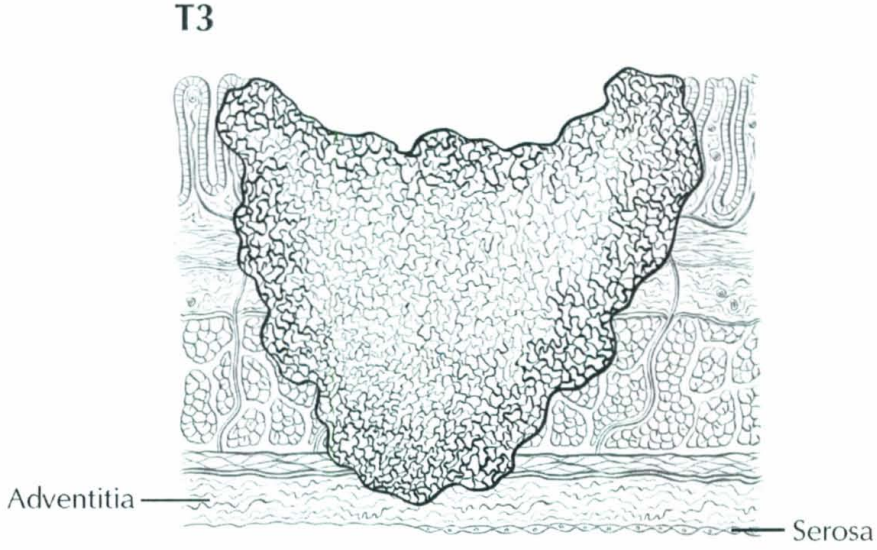
The regional lymph nodes included in the N category include those located adjacent to the colon and rectum (the pericolic and perirectal nodes), and those that follow the course of the major arteries supplying the colon and rectum. They are illustrated in Figure 1.5.



a.) T1 – tumour invades submucosa

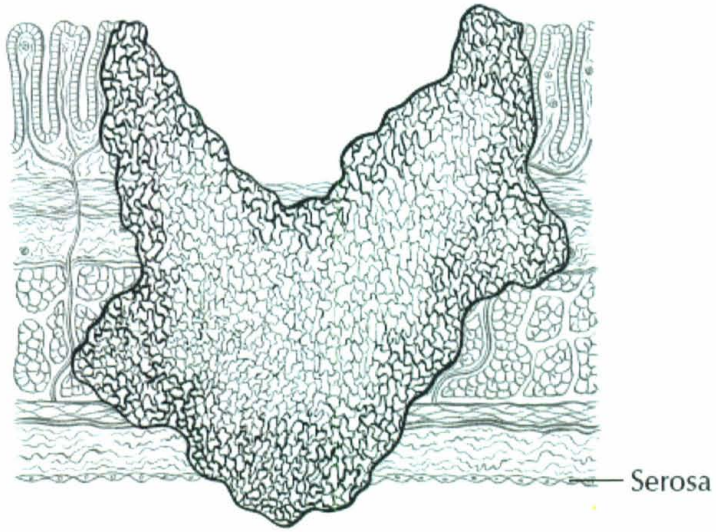


b.) T2 – tumour invades muscularis propria



c.) T3 – tumour invades through the muscularis propria into pericolorectal tissues

T4



d.) T4 – tumour penetrates to the surface of the visceral peritoneum, or directly invades or is adherent to other organs or structures.

Figure 1.4. Illustrations of T, the extent of primary tumour growth
(from [44,45], reprinted with permission from Springer)

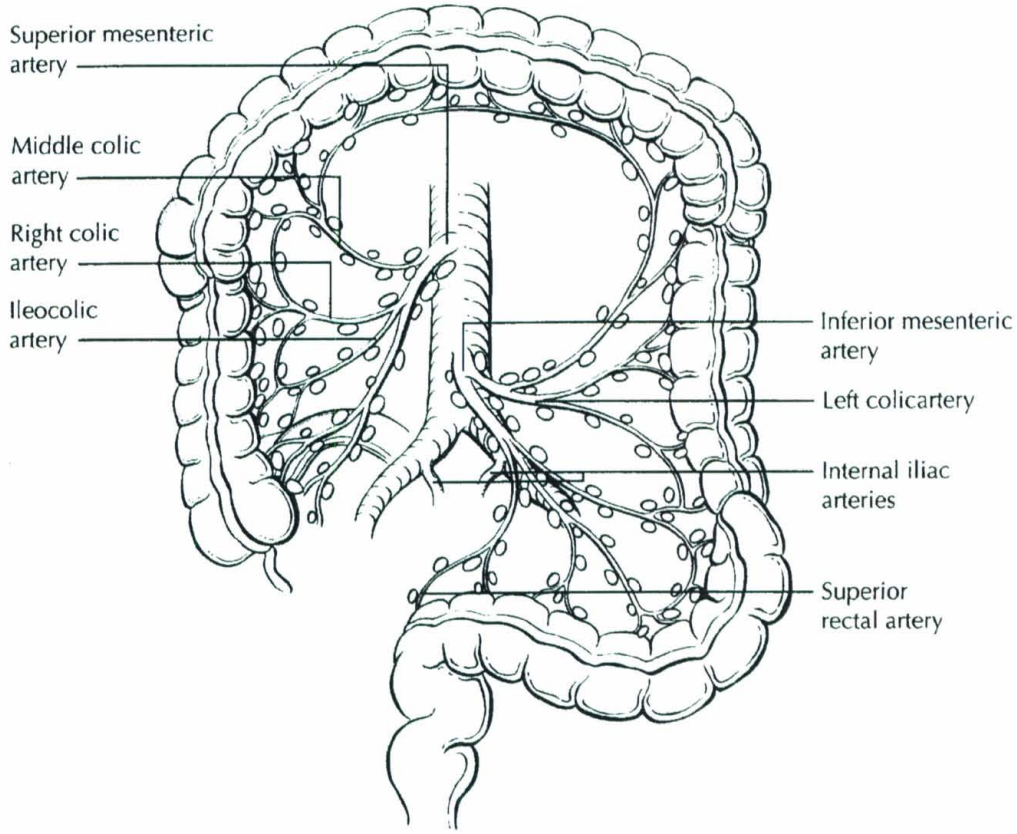


Figure 1.5. The regional lymph nodes of the colon and rectum (from [44,45], reprinted with permission from Springer)

The number of lymph nodes with metastatic disease is of prognostic value, even within the N1 and N2 scores, so these have each been subdivided into a and b categories[43]. The greater the number of involved lymph nodes, the worse the prognosis. It is evident then, that the more lymph nodes that are harvested and examined the more accurate the staging can be. Studies have found that the greater the number of lymph nodes examined, the better the patient outcome will be [46-49]. However, it is important to note that some researchers have found evidence to support the association between

lymphocytic infiltration within a tumour, and the number of lymph nodes harvested [49]. Lymphocytic infiltration is associated with a more positive clinical outcome, and the reactive enlargement of lymph nodes which results from the lymphocytic infiltration may make them easier to find, and therefore harvest. This association may in part explain the better patient outcome observed with increased lymph node harvest.

Some groups have suggested a minimum of 10 [47] or 13 [50] lymph nodes be examined, and that special techniques such as fat clearance be used if necessary to assist the pathologist in identifying an adequate number of nodes [50]. The College of American Pathologists (CAP) currently advises a minimum of 12 lymph nodes be examined, and that the total number be recorded in the pathology report, along with the number containing metastatic tumour [44]. If lymph node harvesting is inadequate, positive nodes could be missed; resulting in assignment of an inappropriately low tumour stage, and subsequent lack of appropriate treatment.

The T, N, and M scores assigned to a tumour are used to determine the tumour stage. For example, a tumour that extends through the muscularis propria, and has metastasised to one regional lymph node, but does not have any distant metastasis, would be classified as T3N1aM0. This would be a stage IIIB tumour. The most current TNM classification guidelines for colorectal cancer are as follows in Table 1.5. [43]. Dukes' staging method, as modified by Turnbull, is included in the Table for comparison.

Table 1.5. Tumour Nodes Metastasis staging guidelines for colorectal cancer (from [43])

Stage	T	N	M	Dukes*
0	Tis	N0	M0	-
I	T1	N0	M0	A
	T2	N0	M0	A
IIA	T3	N0	M0	B
IIB	T4a	N0	M0	B
IIC	T4b	N0	M0	B
IIIA	T1-T2	N1/N1c	M0	C
	T1	N2a	M0	C
IIIB	T3-T4a	N1/N1c	M0	C
	T2-T3	N2a	M0	C
	T1-T2	N2b	M0	C
IIIC	T4a	N2a	M0	C
	T3-T4a	N2b	M0	C
	T4b	N1-N2	M0	C
IVa	Any T	Any N	M1a	D
IVb	Any T	Any N	M1b	D

*Dukes staging method [31], as modified by Turnbull *et al* [35].

The TNM score can be further complicated by the use of prefixes. The use of pathological data is indicated with "p", and clinical data with "c". For example, a tumour could obtain a TNM score of pT3pN2cM1 if the T and N scores were derived by pathology, but a distant metastasis, for example of lung, was identified by staging CT scan. Scores determined from autopsy specimens are indicated with the prefix "a". Tumours, especially rectal, often receive neoadjuvant treatment prior to surgery; TNM scores for such tumours are classified with the prefix "y". Multiple, synchronous primaries in one organ are labelled with "m"; in such cases, the tumour with the highest T score determines the value of T. Finally, the scores for tumours that have recurred after a disease-free interval are labelled with the prefix "r".

Although not part of the TNM classification for colorectal cancer, the AJCC recommends collecting data on histologic grade (G), and the presence of residual tumour (R), since these factors are associated with patient outcome independently of the TNM stage [43]. The scoring system for these features is presented in Tables 1.6 (histologic grade) and 1.7 (residual tumour).

Table 1.6. American Joint Committee on Cancer reporting of tumour grade (G) (from [43])

GX	Grade cannot be assessed
G1	Well differentiated
G2	Moderately differentiated
G3	Poorly differentiated
G4	Undifferentiated

Table 1.7. American Joint Committee on Cancer reporting of residual tumour (R) (from [43])

R0	Complete resection, margins histologically negative, no residual tumour left after resection.
R1	Incomplete resection, margins histologically involved, microscopic tumour remains after resection of gross disease.
R2	Incomplete resection, margins macroscopically involved or gross disease remains after resesction.

1.5 Prognostic and Predictive Features

The terms "prognostic" and "predictive" are often used synonymously, however, they have different meanings. Prognosis refers to the estimate of survival. A prognostic

feature is associated with clinical outcome. A predictive feature, however, is associated with response to therapy or treatment.

1.5.1 Clinically Prognostic Features Recognized by the American Joint Committee on Cancer

The TNM classification groups, as defined above, classify tumours into different prognostic groups. While the T, N, and M features are the most significant for prognosis, the AJCC recommends collecting data on other clinically significant features to further refine prognosis. These features include: serum carcinoembryonic antigen (CEA) levels; the presence and number of satellite tumour deposits, tumour regression grade, circumferential resection margin, microsatellite instability, perineural invasion, and *KRAS* gene mutation status [43].

1.5.1.1 Carcinoembryonic Antigen

CEA was the first tumour marker to be used regularly for any common type of cancer. It was first described in a 1965 paper by Gold and Freedman [51]. They used immuno-diffusion techniques in agar gel to test for the presence of colon tumour-specific antigens in different tissues. They found that the antigens were present only in malignancies from anatomical sites derived from the embryonic entodermal tissue, and not from those of other sites in the gastrointestinal tract that were derived from

ectodermal tissue. The antigens were not present in normal, non-malignant adult tissue. They then tested normal fetal gastrointestinal tissues, and found the antigens to be present, again only in entodermally derived tissues. They named them "carcinoembryonic antigens", and suggested that they were normally present in the developing fetus, but repressed before birth, then reactivated as part of the malignant process [51].

It has since been determined that there is a family of CEA-related genes, which belong to the immunoglobulin superfamily [52]. The clinically-used CEA protein is expressed by the *CEACAM5* gene. It is a 180 kDa cell surface glycoprotein containing seven immunoglobulin domains, and a glycosylphosphatidylinositol anchor [53]. It is a cell adhesion molecule, and may play a role in the metastatic potential of CRC [54]. More sensitive detection methods than those used by Gold and Freedman have shown that CEA is not colon tumour specific. CEA is present in small quantities in normal colonic epithelium, and other epithelial tissues including the tongue, stomach, esophagus, lung, cervix, prostate, and sweat glands [53,54]. Elevations in serum levels can be associated with inflammatory bowel disease, liver cirrhosis, hepatitis, pancreatitis, primary biliary cirrhosis, cholangitis, cholecystitis, renal failure, and other non-malignant processes [54-56]. Serum levels can also be elevated with heavy cigarette smoking [57]. Serum elevation of CEA can also be observed in non-colonic cancers, including pancreatic, bladder, stomach, breast, and lung cancer [52,54,55,58].

While serum CEA level is not as specific for colorectal cancer as once thought, it remains a strong prognositic marker. Serum levels will rise with increasing tumour

burden, and in the absence of CRC, serum CEA levels should be very low. Baseline levels for each patient should be determined pre-operatively, and followed over time. Increasing levels are a strong indicator of recurrent disease.

1.5.1.2 Satellite Tumour Deposits

The AJCC has added tumour deposits as a new prognostic feature in the latest edition of the Cancer Staging Handbook [43]. Tumour deposits are defined as being non-continuous with the leading edge of the primary cancer, and without evidence of a residual lymph node. They are sometimes found in the pericolonic and mesenteric adipose tissue, and are generally palpable, often mistaken for small lymph nodes on gross inspection. They may represent discontinuous spread of disease, or lymphovascular or venous invasion with extravascular spread [43]. A 2000 paper by Goldstein and Turner [59] determined that tumour deposits were an independent indicator of poor prognosis. They identified a trend of decreasing disease-free survival with increasing numbers of tumour deposits, and found that tumour deposits were associated with a significantly worse prognosis than lymph node metastasis [59]. While they also found an association with increasing size of tumour deposits and poor prognosis, it is the number of deposits, and not their size, that the AJCC recommends reporting. A finding of tumour deposits in the pericolonic or perirectal tissue of a pT1 or pT2 tumour is recorded in a separate N category, as N1c [43].

1.5.1.3 Tumour Regression Grade

Neoadjuvant chemoradiation of rectal cancer can lead to regression of the tumour, and subsequent downstaging. The degree of tumour response to neoadjuvant therapy is of prognostic significance. Complete eradication of the tumour, as detected by thorough pathologic examination of the subsequent surgical specimen, is associated with a much better prognosis than a tumour that shows little or no regression as a result of neoadjuvant therapy. The AJCC recommends scoring tumour regression as suggested by the College of American Pathologists (CAP). In the 2009 CAP "Protocol for the Examination of Specimens from Patients with Primary Carcinoma of the Colon and Rectum" [44], they recommend a scoring system of 0 to 3 to report the tumour regression grade. This system is included in Table 1.8.

Table 1.8. College of American Pathologists/American Joint Committee on Cancer scoring system to report tumour regression grade

Description	Tumour Regression Grade
No viable cancer cells	0 (Complete response)
Single cells or small groups of cancer cells	1 (Moderate response)
Residual cancer outgrown by fibrosis	2 (Minimal response)
Minimal or no tumour kill; extensive residual cancer	3 (Poor response)

1.5.1.4 Circumferential Resection Margin

The different anatomical segments of colon and rectum are either totally, partially, or not at all covered by peritoneum (serosa), a serous membrane that lines the abdominal

cavity. For areas not covered by peritoneum, the surgically dissected nonperitonealized surface of the specimen is known as the circumferential margin. The other surgical margins are the distal and proximal margins, and refer to the resected margins of bowel respectively farthest from and closest to the small intestine. For areas that are covered by peritoneum, this layer is considered an anatomical boundary and not a surgical resection, so there is no circumferential margin. For tumours with a circumferential resection margin, it is important for the pathologist to evaluate the deepest point of tumour invasion, and the distance between this leading edge of the tumour and the resection margin. A clearance of greater than 1 mm is considered a negative margin, however, a clearance of less than 1 mm is considered a positive margin. In rectal cancer a positive circumferential margin is considered the strongest predictor of local recurrence, with a 3.5-fold increased risk [43,44]. A positive circumferential resection margin also doubles the risk of death from disease [44].

1.5.1.5 Microsatellite Instability

A high level of microsatellite instability (MSI-H) is a molecular feature present in 10% to 15% of colorectal tumours, and results from defects in the DNA mismatch repair (MMR) system [60-62]. Regions of DNA with tandem repeat sequences, called microsatellites, are particularly prone to errors during replication. Due to DNA repair deficiencies, MSI-H tumours accumulate frameshifting insertion and deletion mutations. MSI is both a prognostic and predictive marker for colorectal cancer. Patients with

tumours exhibiting high levels of microsatellite instability have a better overall prognosis than those without this feature [61,63-69], although the tumours respond less well to standard, 5-fluorouracil (5-FU)-based adjuvant chemotherapy [70,71]. MSI status can be determined through molecular comparison of tumour and normal DNA from the same patient; it can also be estimated through an examination of tumour histology. Certain histological features have been associated with high levels of MSI, including a mucinous, signet ring, or medullary histological subtype, and the presence of tumour infiltrating lymphocytes and Crohn-like lymphocytic reaction [44,72]. Microsatellite instability is discussed further in section 1.7.2.

1.5.1.6 Perineural and Lymphovascular Invasion

Perineural invasion is the invasion by tumour cells into regional nerves. Lymphovascular invasion is the invasion by tumour cells into regional lymphatic channels and vasculature. Both features are independently associated with a poor prognosis, and with an increased risk of metastasis [73-75].

1.5.1.7 *KRAS* Mutation Status

Although listed by the AJCC as a prognostic feature [43], *KRAS* mutation status is a predictive feature for CRC. Metastatic CRCs that are wild-type for the *KRAS* gene

respond to monoclonal antibody therapy targeted against the epidermal growth factor receptor (EGFR), whereas *KRAS* mutant tumours tend not to respond to these therapies [76,77].

1.5.2 Other Prognostic and Predictive Features

In 1931 MacCarty published the 15 most important features to consider for prognosis in any cancer type [78]. The first feature on his list was the presence or absence of glandular (lymph node) involvement, and distant metastasis. He of course associated these with a poor prognosis, as he did the second feature on his list: fixation of the growth of a cancer to surrounding structures. These first two features equate to the T, N, and M categories of the TNM system.

MacCarty's list of prognostic features in cancer also included: location of a tumour (how therapeutically accessible it was, and if early stages caused symptoms which allowed for early identification and treatment), overall health (including renal and cardiac efficiency), the presence of anemia (a feature of poor prognosis, indicating bleeding or reduced diet), and the size of the tumour (increasing size at diagnosis leading to a worse patient outcome). He also included age at diagnosis as a prognostic factor, observing that cancers occurring in young patients were often more aggressive, more quickly leading to death than in older patients. The direction of tumour growth, whether it be in a direction inwards towards the lumen of an organ, or outwards towards adjacent structures was another feature listed by MacCarty. He observed that tumours of the latter category had a

worse prognosis, and more commonly had spread to lymph nodes. Loss of weight was another negative prognostic feature described. He listed duration of disease, but concluded that it was often reported inaccurately, and wasn't of much value without consideration of other features.

Finally, MacCarty listed some histological features of tumours, including cellular differentiation (the TNM "G" score), lymphocytic infiltration, fibrosis, and hyalinization. He concluded that, although there was a correlation between cellular differentiation and the rate of tumour growth, this correlation was of little prognostic significance without considering the other features as well. He presented some data to support the prognostic value of lymphocytic infiltration, fibrosis, and hyalinization. He reported that patients with rectal cancer who had lymphocytic infiltration had a 19% increase in post-operative survival compared to those that did not have this feature. Rectal cancer patients with fibrosis had an 18% increase, and those with hyalinization had a 61% increase in post-operative survival.[78]

In 2000, the CAP published a comprehensive consensus statement on prognostic factors in colorectal cancer [79]. They divided factors into four different categories based on the supportive evidence on their prognostic values. Features included in category I were definitively proven through multiple, large, validated studies, to be important for prognosis. Category II was subdivided into A and B groups. Category IIA included features shown biologically and clinically to be important for prognosis, but that had yet to be validated in statistically robust studies. Category IIB included features that appeared promising in multiple studies, but did not have sufficient data to be included in

categories I or IIA. Category III included features that had not been studied adequately to determine their prognostic value, while category IV included features which had been shown through multiple studies to have no prognostic significance.[79] The features included in each category are summarized in Table 1.9.

Table 1.9. College of American Pathologists consensus statement on prognostic factors in colorectal cancer
(from [79])

Category	Prognostic Factors
I	• the local extent of tumour assessed pathologically (TNM "pT" category) • regional lymph node metastasis (TNM "pN" category) • blood or lymphatic vessel invasion • residual tumour following surgical excision (TNM "R" category) • pre-operative elevation of CEA levels
IIA	• tumour grade (TNM "G" category) • circumferential margin status • residual tumour in the surgical specimen following neoadjuvant therapy (yTNM category)
IIB	• histological type • MSI-associated histological features (including: host lymphocytic response, and medullary or mucinous histological type) • high levels of MSI (MSI-H) • loss of heterozygosity (LOH) at 18q (allelic loss of deleted in colorectal cancer [*DCC*] gene) • tumour border configuration (infiltrating versus pushing)
III	• DNA content • all molecular markers (except MSI and LOH of *DCC*) • perineural invasion • microvessel density • tumour cell-associated proteins or carbohydrates • peritumoural fibrosis • peritumoural inflammatory response • focal neuroendocrine differentiation • nuclear organizing regions • proliferation indices
IV	• tumour size • gross tumour configuration

Since the publication of the CAP consensus statement in 2000 [79], there has been further research and progress in determining the prognostic and predictive values of many factors, and the relationship of these factors to each other. An example is tumour budding: the presence of single tumour cells, or clusters of up to four tumour cells, along the invasive margin [80]. Tumour budding is an indicator of tumour aggression, and is associated with a poor prognosis [80,81]. Tumour budding has been associated with the presence of a *KRAS* mutation [82], expression of p16 at the infiltrative front of invasion [81], and with the absence of E-cadherin expression [83]. Absence of tumour budding has been associated with the presence of both tumour infiltrating lymphocytes and peritumoural lymphocytes, suggesting that the host immune reaction removes any developing tumour buds, leading to a better prognosis [83,84].

Some prognostic and predictive factors that have received a lot of recent research attention include MSI, host immune responses, genetic markers, and protein markers. MSI has already been discussed briefly (section 1.5.1.5), and will be discussed further in a later section (1.7.2).

1.5.2.1 Host Immune Response

Tumours of different types often incite an immune response; immune cells respond to various antigens expressed on the surface of tumour cells. Some mechanisms that lead to expression of these tumour antigens include: mutation, gene activation, and

clonal amplification [85]. Somatic mutations abundant in tumour cells may create novel antigens by producing mutant proteins (for example p53 mutation leads to over-expression of the mutant protein); they can also unmask structures found on the surface of normal cells, thereby exposing these structures as antigens. Gene activation can create tumour antigens by inappropriately expressing normal genes that are usually only active during embryogenesis (oncofetal antigens). An example is CEA, which was previously discussed in section 1.5.1.1. Finally, clonal amplification can lead to over-expression of normal structures, which can elicit an immune response. An example of this is the over-expression of the *HER-2/neu-1* oncogene in some breast and ovarian cancers, which leads to expression of the HER-2/neu tumour antigen on the cells of these tumours.

The frameshifts observed in MSI-H tumours can be antigenic and tumour-specific [86-94]. When these frameshifts occur in the coding sequences of genes it can result in the production of truncated peptides. These peptides can be recognized by the immune system as being foreign, and can incite an immune response, particularly by T-cell lymphocytes. MSI-H tumours generally have a high level of immune cell presence, and the density of tumour-infiltrating lymphocytes is strongly associated with the total number of frameshift mutations [93]. Many groups have suggested that the increased immunogenicity of MSI-H tumours is reflective of an anti-tumour immune response, and is responsible for the improved prognosis seen in these patients [89-91,93,94].

Some groups have been able to identify frameshift peptides in the peripheral blood of individuals with mismatch repair (MMR) gene mutations, in some cases in patients who had not been diagnosed with a tumour, suggesting that these antigens may be

expressed by the precursor lesions or very early tumours [88,92]. This suggests a possible role for these antigens as targets for immunotherapy. There are many predictable antigens created from commonly occurring frameshift mutations in MSI-H tumours. There has recently been some research into using these antigens as targets for a multivalent cancer vaccine for individuals with germline MMR mutations (Lynch Syndrome), and for immunotherapy for patients with MSI-H CRCs [86-89,91-95].

Host immune response, also known as anti-tumour immunity, includes a number of different types of immune infiltration. Immune cells can be found infiltrating tumours in different locations, and with different patterns. Various immune and inflammatory cells can be involved, including neutrophils, macrophages, eosinophils, mast cells, and plasma cells. These cells can infiltrate the stroma surrounding the tumour cells, or cluster along the invasive margin. T cell lymphocytes are the most preponderant of the infiltrating immune cells, and can infiltrate among the tumour epithelial cells. Most colorectal tumours do not have an abundance of infiltrating immune cells, but instead just a low, background level. Approximately 20% of rectal cancers have infiltrating lymphocytes [36]. Patients with CRC whose tumours have an abundant infiltration have an improved prognosis [36,78,96-101]. This improved prognosis is independent of TNM stage [99], yet there is also a linear, inverse correlation between stage and immune cell density [98].

Galon and colleagues [99] studied the type, density, and location of immune cells within colorectal tumours, and the association of these features with clinical outcome. They studied the expression levels of genes related to inflammation, T cell mediated adaptive immunity, and immunosuppression. In a cluster analysis, they found an inverse

relationship between the expression of these genes and tumour recurrence. They also used tissue microarrays (TMAs) (discussed further in chapter 3), to examine the adaptive immune response at the centre of the tumour, and at the tumour margin. They used image analysis computer software to quantify the immunohistochemical staining of the TMAs for: total T lymphocytes (CD3), CD8 T cell effectors and their associated cytotoxic molecule (granzyme B), and memory T cells (CD45RO). They found that for all cell types tested, a strong presence in both tumour regions was associated with good prognosis, regardless of TNM stage. Conversely, a weak immune reaction at both locations was associated with a poor outcome, again regardless of TNM stage. In their study, the immune response was actually a stronger prognostic indicator than the TNM classification. They suggested that the immune cells could modify tumour cells and supporting stroma, decreasing the ability of a tumour to metastasize. They also concluded that the long-lasting effects of the memory T cells could be important in preventing tumour recurrence after surgery.[99]

There are two recognized patterns of immune cell infiltration of the invasive margin of a tumour: Crohn-like lymphocytic reaction (CLR), and peritumoural lymphocytic reaction. CLR was first described in 1990 [102]. It is defined as discrete lymphoid aggregates, some with germinal centres, along the advancing tumour margin, usually in the muscularis propria or subserosal fat [102,103](Figure 1.6). Similar lymphoid aggregates are seen at the same interface of inflamed, non-malignant colon affected by Crohn disease [19], however, when these lymphoid aggregates are seen in colorectal cancer they are unrelated to Crohn disease [102]. They are a positive prognostic feature of

colorectal cancer. Graham and Appelman's 1990 paper [102] that first described CLR found it to be associated with a lower incidence of lymph node metastases and a significant increase in ten year survival. They also found CLR to be associated with an increased overall lymphocytic infiltrate at the advancing tumour edge, also known as peritumoural lymphocytic reaction. Harrison and colleagues published an article in 1995 [103], in which they tested the prognostic utility of CLR in multivariate statistical analysis of 344 right-sided colon cancers. They included eleven pathologic features in their analysis, many of which were well recognized prognostic markers including: tumour grade, growth pattern, depth of invasion, lymph node metastasis, and peritumoural lymphocytic reaction. In their analysis, CLR proved to be an independent indicator of prognosis for right-sided colon cancer.

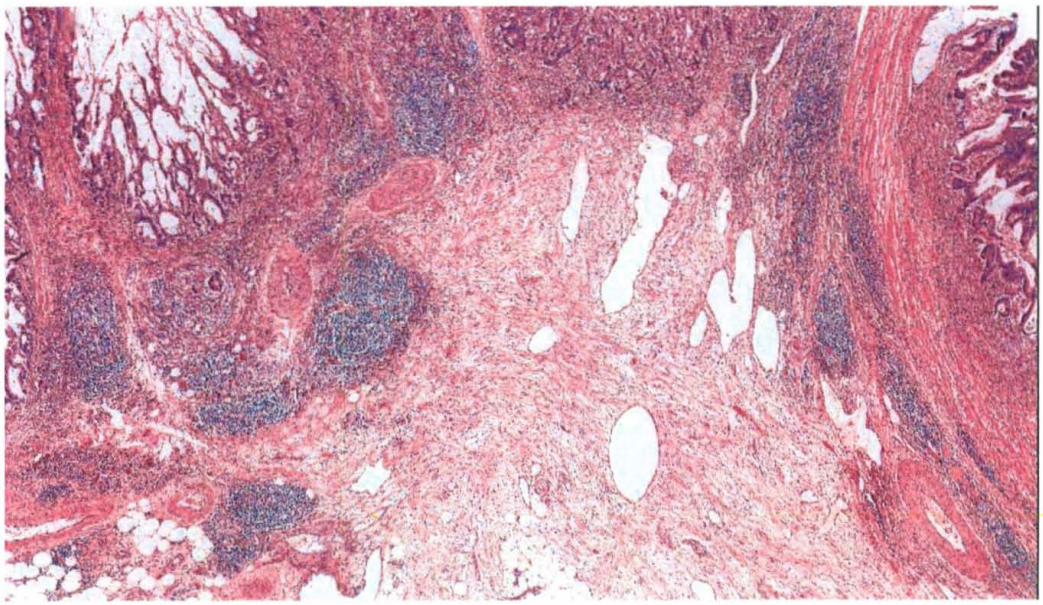


Figure 1.6. Crohn-like lymphocytic reaction

Similar to CLR, peritumoural lymphocytic reaction involves an increased lymphocytic infiltrate along the advancing tumour edge. The difference between the two features is that while CLR refers to distinct lymphoid aggregates, a peritumoural lymphocytic reaction refers to band-like ribbons of cells forming an inflammatory mantle along the advancing edge of a tumour (Figure 1.7). This mantle is made up predominantly of lymphocytes, but can include a variety of immune cells including eosinophils, neutrophils, and plasma cells [36]. While mentioned in MacCarty's 1931 article [78], described above, it was in 1986 that Jass [97] first defined this feature and assessed its utility as a prognostic feature in multivariate analysis. In his study of rectal cancer, peritumoural lymphocytic reaction was an independent positive prognostic factor. Graham and Appelman's 1990 paper which described CLR found that there was an

association with peritumoural reaction [102]. While only five percent of tumours without CLR had a peritumoural reaction, 39 percent of those with a strong CLR had peritumoural reaction. This was statistically significant, with *P* value less than 0.01.

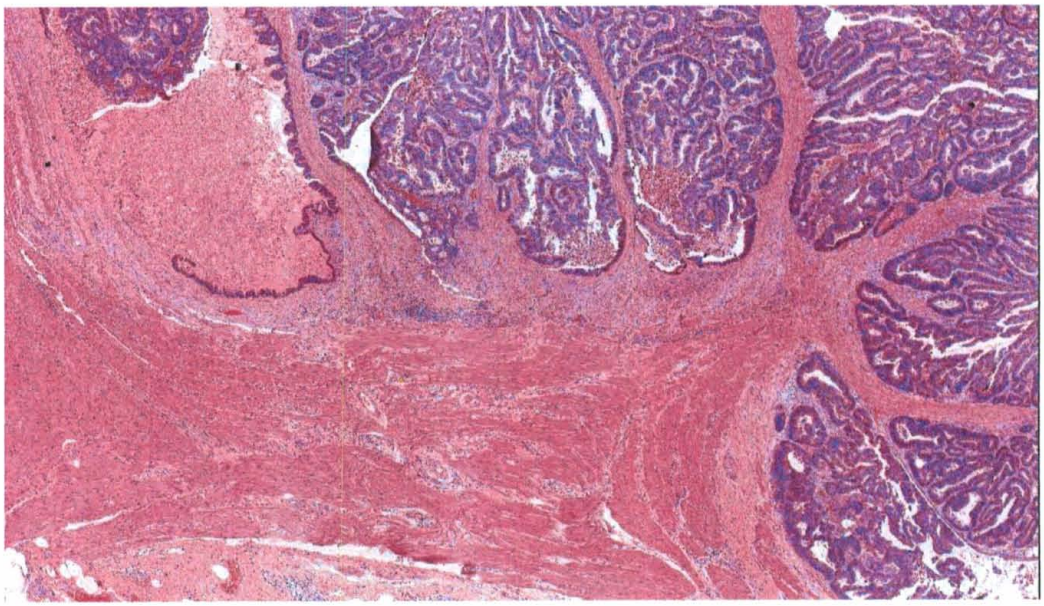


Figure 1.7. Peritumoural lymphocytic reaction

The host immunological response to CRC which has recently received the most attention as a prognostic marker is tumour-infiltrating lymphocytes (TILs). While the term TILs used to be associated with all forms of immune response in and around a tumour, the definition appears to be changing with time, and now often includes only lymphocytes found among the tumour cells, infiltrating the neoplastic epithelium (Figure 1.8). This is the definition of TILs used in this thesis. These intraepithelial TILs are closely associated with microsatellite instability and medullary tumour histology [79].

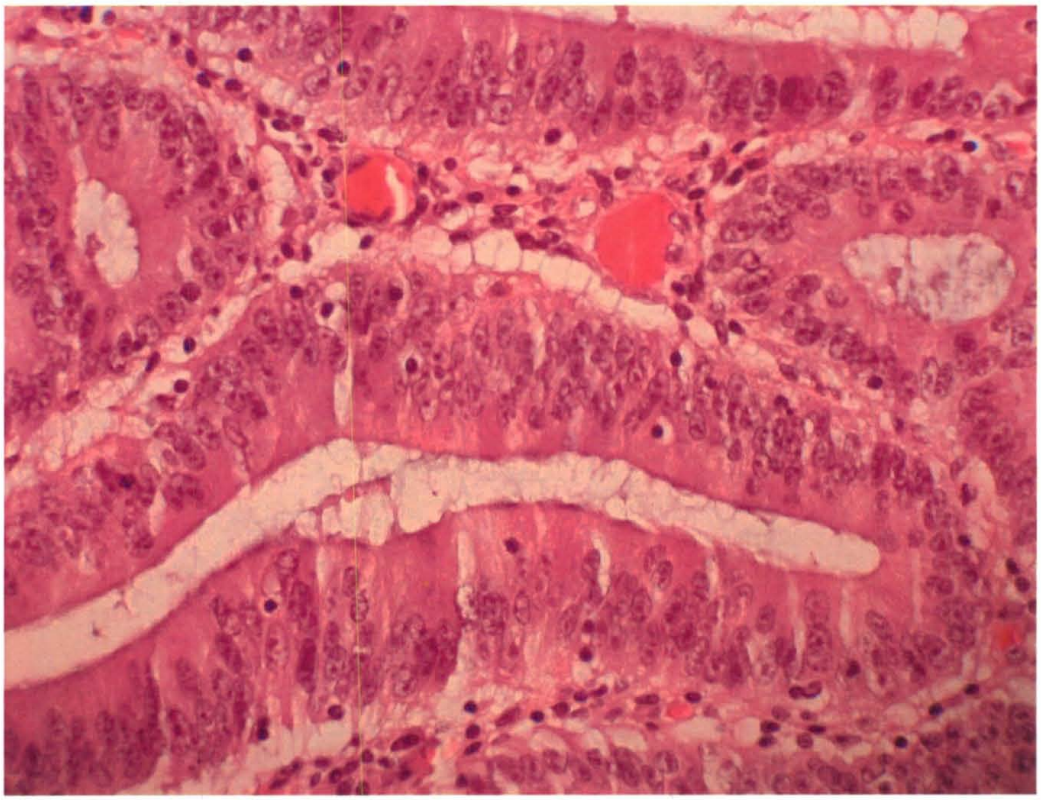


Figure 1.8. Tumour infiltrating lymphocytes

While many credit Ropponen and colleagues [98] for being the first to determine the positive prognostic value of TILs, they actually used the older definition, including all forms of infiltrating lymphocytes, not just those that infiltrate the epithelium. More recent studies have focused on intraepithelial TILs [89,95,101,104,105]. The presence of these infiltrating T cells is strongly correlated with another prognostic indicator, MSI status [72,105,106]. As stated above, the T cells infiltrating MSI-H tumours are often specific for frameshift-derived neopeptides that result from the deficiency in mismatch repair. MSI-H tumours also tend to have Crohn-like lymphocytic reaction and other forms of immune response [72,105,106].

One study has also reported on the independent prognostic significance of eosinophil and mast cell infiltrations in colorectal cancer. In their 1999 article [107], Nielsen and colleagues reported that both features were independent, positive prognostic markers.

1.5.2.2 Molecular Genetic and Protein Markers

While many markers have been identified as having prognostic or predictive utility for colorectal cancer, very few have been robust or informative enough to be incorporated into clinical practice. The complexity of the disease has led many researchers to consider the potential increased utility of considering multiple markers in concert, expanding upon the Dukes' and TNM classification methods to further refine patient prognostication as well as guide more individualized therapies. In 1987, Jass and Morson first proposed the concurrent consideration of multiple markers for these reasons [36]. They suggested including a number of pathological variables into a multivariate logistic regression analysis, with patient survival as the endpoint. Use of these statistical methods identified the variables with independent prognostic power, and provided the ability to use the regression coefficients to weight these variables in an algorithm for assigning prognostic risk. In their analysis, the variables with the strongest independent association with prognosis were: limitation of growth to the bowel wall, number of lymph nodes with metastasis, character of the invasive margin, and presence or absence of conspicuous peritumoural lymphocytic infiltrate. Based on the regression scores for these

variables, Jass and Morson assigned weighted scores, which they added together to generate a total score for each tumour. The total scores could be translated into one of four clinically relevant prognostic groups.

Similar statistical methods have been employed with different endpoints to generate algorithms to predict MSI status [105,106], and to predict carriers of germline MMR mutations [108-110]. Newer technologies, including microarrays, have allowed for the use of multiple genetic and molecular markers for classification and prognostication, rather than pathological and clinical markers. In 1999, Alon and colleagues published the first study using a genome-wide oligonucleotide array to assess colorectal cancer [111]. Using clustering analysis, they demonstrated that the expression profiles of colorectal tumours were distinguishable from those of normal colon tissue.

Numerous gene expression profiling studies of CRC have since been carried out; they can be classified into three distinct types: those studying the carcinogenic process, prognosis prediction, and the prediction of response to therapies [112]. Despite the work done and the numerous studies publishing positive results, there has been very little reproducibility between studies, and no clinically useful gene expression profile has been identified to date [112]. A meta-analysis of such studies, published in 2006 [113], identified a total of 13 genes which were differentially expressed in CRC with good versus poor prognosis, and reported by at least two independent studies. There were, however, discrepancies for six of the 13 genes between the two original studies with regards to the direction of the differential expression (up- or down-regulation) associated with poor prognosis. Most of the 13 differentially expressed genes fell into the categories of either

signal transduction or transcription. A more recent review [112] has used pathway-based bioinformatic enrichment tools to further explore the commonality between published gene-expression profiling studies on prognosis of colorectal cancer. They identified the oxidative phosphorylation chain, extracellular matrix interaction, cell proliferation, and apoptosis as those most consistently overrepresented pathways reported to be differentially expressed in association with CRC prognosis.

The microarray technology has been translated into a tissue-based format – tissue microarrays (TMAs) – which can contain hundreds or thousands of tiny cores of tissue from different tumours, allowing them to be assessed for expression of proteins in a high-throughput manner through the use of immunohistochemistry. They have been used to sub-classify CRC and other cancers based on protein-expression profiles. TMAs are discussed further in chapter 3.

While much research has been done on determining prognostic and predictive factors for CRC, research has also been done on determining the causes and predisposing factors to developing CRC. These are discussed in the following section.

1.6 Aetiology of Colorectal Cancer

1.6.1 Environmental Factors

Increasing age is a top risk factor for colorectal cancer; approximately 95% of new cases and deaths occur after 50 years of age [21]. An accumulation of DNA mutations are required, in the same cell, for malignant progression to occur. Such mutations occur due to intrinsic mechanisms, for example during DNA replication, and through extrinsic mechanisms, such as through exposure to mutagens found in foods or the environment. While most such mutations will be corrected by DNA repair systems, some will persist, and these will accumulate with age. Once a single cell accumulates adequate mutations in proto-oncogenes (which encourage growth) and tumour-suppressor genes (which limit growth), the cell can undergo malignant transformation, and cancer can develop.

A high incidence of colorectal cancer has been associated with a sedentary lifestyle and the consumption of a Western-style diet, which is characterized by a high intake of animal fats, red meat, processed meats, and refined sugars and grains, with low intake of vegetables, fruits, and fibre.

A recent review [114] reported that of the 52 identified studies of physical exercise and colorectal cancer risk, 37 found a statistically significant inverse correlation. The authors concluded from the evidence accumulated in their review, that physical activity

was associated with a 25% decrease in risk of developing colonic cancer [114]. This association did not extend to rectal or other gastrointestinal cancers.

An association between colorectal cancer risk and obesity has long been recognized, although the details of the association have not always been straightforward. A large 2007 Canadian study by Campbell and colleagues [115] clarified some of the intricacies. They confirmed that there was an increased risk of CRC in obese men, but that this risk did not extend to women. They used the WHO definition of obesity, of a body mass index (BMI) of at least 30 kg/m^2, and found that men with such a BMI were almost twice as likely to develop CRC than men with a lower BMI. This increased risk applied to both sporadic and hereditary CRCs. While obesity was not associated with an increased risk of CRC in women in this study, in a later study they and others, have found such an association, although it was not as strong as that seen for men [116]. Campbell and colleagues did, however, find an increased risk associated with height in women. Females taller than 1.75 m had more than double the risk of CRC compared to shorter women. The increased risk did not apply to tall men [115]. In their later study [116], Campbell and colleagues further refined the risks associated with obesity, and concluded that this risk did not extend to MSI-H CRCs.

Consumption of red meats has been associated with an increased risk of developing CRC. The association is strongest with a high frequency of consumption, especially of red meats cooked at high temperatures for prolonged periods of time.

In 1969, while working in Africa, Denis Burkitt published a hypothesis [117] based on his observation that CRC was relatively rare in the African population. He surmised

that the paucity of CRC was due to their high fibre diet, or what he termed their "low residue diet" [117]. Dietary fibre is reported to have a number of effects on the human digestive system, such as increasing the volume of fecal waste and adsorbing some of its contents, thereby diluting and removing some fecal carcinogens. It can also alter bile acid metabolism, lower colonic pH, increase production of short-chain fatty acids, and modulate fecal transit time [118]. High intake of dietary fibre, especially vegetable, fruit, and grain fibre, has been associated with a lower risk of developing CRC in some studies, sometimes by as much as 50% [119-122]; however, the protective effect has not been reported consistently [123-125].

In 1995, Giovannucci published his hypothesis that linked many of the seemingly unrelated factors that increase the risk of developing CRC [126]. He proposed that lack of physical activity, increased abdominal obesity, high intake of red and processed meats, saturated fats, and refined grains and sugars all led to an increase in insulin levels in the blood. He argued that hyperinsulinemia may lead to increased cellular proliferation and decreased apoptosis, thereby increasing the susceptibility for developing CRC. Giovannucci's hypothesis [126] has been supported by the finding that diabetes mellitus is a risk factor for CRC, especially when the patient is on insulin therapy, or has poor glycemic control [127-130].

Other risk factors for CRC include cigarette smoking, and alcohol consumption. A large study of male health professionals found a two-fold risk of colon cancer among men who consumed more than two alcoholic drinks per day, when compared to those who consumed less than 0.25 drinks per day [131]. Another large study from many of the

same authors examined the risk of developing colorectal adenomas in both males and females [132]. They identified an increased relative risk of 1.84 for females, and 1.64 for males among those who consumed more than 2 drinks per day, compared with non-drinkers. In both studies [131,132], the authors attributed the role of alcohol as a methyl-group antagonist which reduces the availability of methyl groups, thereby leading to global hypomethylation, potentially altering the expression of proto-oncogenes, a concept discovered in rat colonic mucosa in the late 1990's [133]. While the earlier studies looked at colon cancer as one disease entity, a more recent study stratified the disease by MSI status and tumour location [134]. The authors found that a high intake of alcohol was associated with CRC overall, but was more strongly associated with rectal than colon cancer, and with MSI-low (MSI-L) rather than MSI-H tumours [134].

The same, recent study also assessed the CRC risk from cigarette smoking. The authors identified an association between pack-years of smoking and rectal cancer, with an odds ratio of 1.85 for those with a greater than 40 pack-year smoking history, when compared with nonsmokers [134]. They also identified an increased risk of MSI-H tumours with increased duration of smoking; those with greater than 30 pack-years were almost twice as likely as nonsmokers to develop MSI-H tumours [134]. The association between smoking and CRC is limited to tumours which have a phenotype including high levels of CpG island methylation (CIMP-high), which account for the sporadic MSI-H tumours, as well as some microsatellite stable (MSS) tumours, and is often associated with mutations in the *BRAF* gene [135,136]. The CIMP pathway of CRC is described in further detail in section 1.7.3.

The association between smoking and CRC has been controversial, since the majority of early studies found no association. Most studies since 1970, however, have found a positive association between the two. The reason for the discrepancy is thought to be the long lag time between exposure and disease (reviewed in [137]). A recent meta-analysis [137] examined four dose-response variables, and found a significant association with each: daily cigarette consumption (with a 38% increased risk per increment of 40 cigarettes per day), duration of smoking history (with a 20% increased risk per increment of 40 years of smoking), pack-years (with a 51% increased per 60 pack-years), and age of initiation (with a 4% decrease in risk per 10 years delayed onset of smoking) [137].

Some medications and therapies have been associated with a decrease in CRC risk; these include aspirin, COX-2 inhibitors, and other non-steroidal anti-inflammatories (NSAIDs), as well as calcium, vitamin D, and estrogen. A recent meta-analysis of four randomized, placebo-controlled trials identified a protective effect from aspirin [138]. The trials all assessed the secondary prevention of colorectal adenomas in patients with a personal history of these lesions. The meta-analysis identified a 17% reduced risk of adenomas among aspirin users, and a 28% decreased risk of developing advanced adenomas for the same group [138]. While aspirin can also be used for the prevention of cardiovascular events, long term use of aspirin is associated with risks of hemorrhagic stroke and gastrointestinal bleeding [139], and recent reports have advocated for more research into the benefits and potential harms of using aspirin for long term chemoprevention [140,141].

COX-2 selective inhibitors, including celecoxib and rofecoxib, can also prevent CRC. COX-2 selective inhibitors generally have fewer gastrointestinal (GI) side-effects than do non-selective NSAIDs, including less GI ulceration and bleeding than is associated with aspirin. In clinical trials intended to assess the potential of these drugs to prevent recurrences in individuals with a history of adenomas, celecoxib was associated with a dose-dependent reduction in recurrence of 33% to 45% [142]. Unfortunately, the benefits do not outweigh the risks of adverse cardiovascular events for these drugs, and they are not recommended for routine prevention of CRC [142]. Risks of adverse cardiovascular events have also been associated with use of other non-aspirin, non-COX2-inhibitor NSAIDs, including sulindac, in the chemoprevention of CRC [143]. The risks of cardiovascular events were highest in patients who were already considered to be at high risk of having such events at baseline.

There are many possible mechanisms of action of NSAIDs in CRC chemoprevention. The anti-inflammatory effects of the drugs may play a role, as well as the inhibition of COX-2, which is overexpressed in colorectal adenomas and carcinomas [144], and is indirectly involved in decreasing immunity and apoptosis, and increasing proliferation, migration, invasion, and angiogenesis [125]. Aspirin also slows down the cell cycle, by activating a replication checkpoint [145]. This may result in improved DNA replication fidelity, and thereby also result in maintenance of genomic stability, preventing cancer initiation.

The high fat nature of the Western diet may promote CRC by increasing the levels of free ionized fatty acids and toxic bile acids into the lumen of the colon. These are

irritating to the lumen, can lead to an increased rate of cellular proliferation for regeneration and repair, and subsequently play a role in cancer initiation [146]. Ingestion of calcium may help counteract the negative effects of the fatty acids and bile acids by binding to these compounds and converting them into insoluble soaps [147].

Calcium and vitamin D are metabolically inter-related in the normal, non-cancerous bowel. In randomized clinical trials, calcium and vitamin D have been found to increase both apoptosis and epithelial cell differentiation [148,149]. In a large pooled analysis of 10 cohort studies including more than 500 000 individuals with a total of almost 5000 incident cases of CRC, total, dietary, and supplemental calcium were all found to decrease the risk of distal colon and rectal cancers, but not proximal colon cancers [150]. There was a 22% risk reduction in the group with the highest calcium intake, when compared to the group with the lowest intake. Another study also reported that distal CRC risk was inversely proportional to calcium intake [151]. These authors also noted that intake of calcium beyond 700 mg/day provided little additional risk reduction, suggesting a threshold effect; that intake of calcium beyond that level would not impart a further benefit.

In the large study described above, the inverse association found between calcium intake and CRC was statistically significant only among those in the group with the highest vitamin D intake [150]. Others have also reported similar findings, that the protective effect of calcium was only significant with high intake of vitamin D, and conversely, that vitamin D intake was only protective with high calcium intake [152].

Women who receive estrogen therapy, either in the form of oral contraceptive pills or post-menopausal hormone replacement therapy, have a reduced risk of developing colorectal adenomas and CRC [125,153]. This decreased risk may be due to a protective effect of estrogens, or to a decrease in production of growth factors, including insulin-like growth factor 1 [154]. Prior to menopause, women are somewhat protected from CRC due in part to their intrinsic hormone levels, and this protection can be extended beyond menopause with the use of hormone replacement therapy. A study of 59 000 women found a 35% decrease in risk of CRC among those taking post-menopausal hormone replacement therapy [155]. While protective against CRC, post-menopausal hormone replacement therapy is associated with other risks, including an increased risk of breast cancer and adverse cardiovascular events [156], and is generally not recommended for routine chemoprevention of CRC.

While for the most part, factors imparting either risk of or protection from CRC were discussed above in general terms, there are some differences in the risk profiles of colon and rectal cancers. When compared in multivariate analyses in one large study [157], colon cancer risk factors included those described above, such as inadequate physical activity, consumption of processed and red meats, alcohol consumption, and others. However, none of these remained significant at the multivariate level for rectal cancer. Risk factors significantly associated with rectal cancer included only age and sex. In this study, even family history was found to be not significantly associated with risk of developing rectal cancer [157].

Many of the risk factors for colon cancer can be avoided through diet and lifestyle changes. A study of almost 48 000 men sought to determine the proportion of colon cancer risk that could be preventable [158]. The authors considered six risk factors: obesity, physical activity, alcohol consumption, early adulthood cigarette smoking, red meat consumption, and low intake of folic acid. After adjusting for age and family history, they found that from one third to more than one half of colon cancers might be avoidable, if the individuals with the highest risk profiles were to change their exposures to be similar to those with the lower risk profiles [158].

Of course, while some risk factors are avoidable, others are not. The following section discusses inflammation, a major risk factor for CRC that is not as easily modifiable as those described in this section.

1.6.2 Predisposing Conditions – Inflammatory Bowel Disease

Chronic inflammation of the gastrointestinal mucosa is a risk factor for malignancy. The two forms of chronic inflammatory bowel disease (IBD), Crohn Disease and Ulcerative Colitis (UC), are associated with an increased risk of epithelial dysplasia and carcinoma [159,160]. Together they account for 1% to 2% of all CRC [161]. IBD is among the top three high risk conditions for developing CRC, along with Lynch Syndrome and FAP [162]. The IBD syndromes are both characterized by chronic, relapsing inflammation of the large bowel, with mucosal damage. Both Crohn disease and UC are systemic disorders, with some patients developing extraintestinal manifestations,

including migratory polyarthritis, sacroiliitis, ankylosing spondylitis, and hepatic manifestations including primary sclerosing cholangitis [19,163]. While the aetiologies are unknown, both disorders are considered to be autoimmune.

Crohn disease can involve any level of the gastrointestinal tract, although the terminal ileum of the small intestine and the colon are the most common sites. It is characterized by sharply demarcated segments of involved bowel, with normal intervening bowel, known as skip lesions. The involved bowel segments are affected by an inflammatory process that typically encompasses the full thickness of the bowel wall. With time, this inflammation causes mucosal damage, granulomas, and fissuring of the epithelial surface, which can lead to bowel adhesions and fistulas [19].

Ulcerative colitis differs from Crohn disease in that it is limited to the colon, extending proximally from the rectum in a continuous manner. The chronic inflammation of UC is generally limited to the mucosal and submucosal layers of the bowel wall, and leads to ulcers and pseudopolyps, but not granulomas or fissures [19].

The risk of carcinoma in IBD increases with longer duration of disease, extent of anatomic involvement, severity of inflammation, presence of primary sclerosing cholangitis, and a family history of colorectal cancer [162,163]. While almost all reports agree that there is an increased risk of developing CRC in IBD [19,161-169], the extent of that risk is controversial. Early studies likely overestimated risk, due to the collection of data from tertiary gastroenterology centres, resulting in an overrepresentation of severe cases with extensive disease. Study design and geographic location also contribute to the discrepancies in published risk estimates. Authors of a large meta-analysis published in

2001 [164] attempted to clarify the extent of the risk in UC. They examined 116 studies, which included a total of 54 478 patients, who had a total of 1698 CRCs. They concluded that the cumulative probability of developing CRC was 2% by 10 years after diagnosis, 8% by 20 years, and 18% by 30 years [164]. They also noted variability in UC-related CRC incidence based on geographical location, with higher rates observed in the USA and UK than in Scandinavia. Rates were also lower in Hungary than they were in Western European and North American studies [163].

Studies included in the meta-analysis employed various study designs and patient recruitment methods, including referral-based, hospital-based, and population-based studies[164]. More recent, population-based, studies have reported lower cancer risks. While this may be related to study design, it may also be due in part to more successful screening and prevention strategies, including the regular intake of anti-inflammatory drugs such as NSAIDs, and prophylactic colectomy for patients with severe disease including dysplasia.

A Canadian population-based case-control study published in 2000 [165] reported cancer risk in IBD patients compared to a control group in terms of the incidence rate ratios (IRR). They could not comment on the impact of disease duration to cancer risk, since they had a maximum possible patient follow-up time of only 14 years. They reported a significantly increased incidence rate of colon carcinoma in both Crohn disease (IRR 2.64; 95% confidence interval [CI] 1.69-4.12) and UC (IRR 2.75; 95% CI 1.91-3.17), as well as a significantly increased incidence rate of rectal carcinoma in UC (1.90; 95% CI 1.05-14.9), when compared to the non-IBD group.

A Danish population-based study published in 2004 [169] followed more than 1000 UC patients for up to 36 years. They reported a cumulative risk of CRC of only 0.4% by 10 years after diagnosis, 1.1% by 20 years, and 2.1% by 30 years. A Hungarian population-based study published in 2006 [163] followed more than 700 UC patients for up to 30 years. They reported a higher cumulative risk of CRC than the Danish study, of 0.6% by 10 years, 5.4% by 20 years, and 7.5% by 30 years, although these numbers are much lower than those of the 2001 meta-analysis. By comparison, the worldwide cumulative incidence of CRC in the general population by age 64 years reaches a maximum of 2.2% for males and 1.7% in females from Australia and New Zealand, and a minimum of 0.2% for males and 0.3% for females in middle Africa [170].

Clinical features of IBD-related CRC are different than for sporadic CRC. IBD-related CRC occurs in patients at a younger age than those in the general population. The large meta-analysis reported a mean age of cancer diagnosis of 43.2 years in IBD patients [164]. The Hungarian study reported CRC diagnosis at a mean age of 51 years in UC patients, as compared to almost 15 years older for sporadic CRC in that country [163]. IBD-related CRC also has a higher mortality rate than sporadic CRC, with more than half of cancer patients dying from their disease [161,167,168]. CRC accounts for one in six deaths in IBD patients [161]. IBD-related CRCs are more likely to be proximally located within the colon, have a higher rate of synchronous cancers, and have a higher proportion of mucinous and signet ring histology than sporadic CRCs [162].

IBD-related CRCs tend not to progress through the traditional adenoma to carcinoma pathway (described in more detail in section 1.7.1), but instead progress from

flat and nonpolypoid dysplasia [159,162]. The chronic inflammation of IBD leads to the production of pro-inflammatory cytokines that are capable of inducing mutations in oncogenes and tumour suppressor genes, including those commonly mutated in sporadic CRC, such as *APC*, *p53* and *KRAS*. Although many of the same genes and pathways are involved in both IBD-related and sporadic CRC, the timing and sequence of events is quite different. While p53 inactivation occurs late in the adenoma to carcinoma pathway of sporadic CRC, it occurs very early in the dysplasia to carcinoma sequence of IBD. Conversely, APC inactivation occurs early in the adenoma to carcinoma pathway, but late in the dysplasia to carcinoma pathway [162,167,168].

Some features of other CRC pathways, including microsatellite instability, and CpG island methylation, are often seen in IBD-related CRC. Again, the timing and sequence of the key changes in these pathways are also different [167], as are the patterns of instability and methylation [171,172]. Microsatellite instability appears to be more related to the initiation of cancer in IBD patients than to the progression of it, which it is for patients without IBD [172,173]. CpG island methylation is associated with high levels of chronic inflammation. This is likely due to the increased cell turnover that occurs with inflammation, and has been described as premature aging of colorectal epithelial cells [174].

1.7 Pathways to Colorectal Cancer

Colorectal cancer develops along one of three major pathways. They are: chromosomal instability (CIN), microsatellite instability (MSI), and the CpG island methylator phenotype (CIMP). Lynch Syndrome cancers develop along the MSI pathway, but the most common pathway for sporadic CRC carcinogenesis is the CIN pathway.

The traditional theory of clonal evolution proposes that any cell is capable of developing into a cancer, through the stepwise accumulation of genetic mutations which confer a growth advantage [175]. Natural selection increases the frequency of advantageous variants. Once a cell possesses a growth advantage over adjacent cells, neoplastic proliferation can occur. As this proliferation proceeds, other mutations can arise and genetic instability can result. In other cases, it is the genetic instability that precedes the increased rate of proliferation [176]. Most variants will be eliminated from the growing tumour, however, those that confer a further selective advantage will persist. Eventually this process may lead to malignant transformation of the neoplasm, through the acquisition of the capacity to invade and metastasize [175]. The genetic and epigenetic changes that occur during this process can vary from tumour to tumour. Thus, colorectal cancer is not a single homogeneous disease, but instead it is a heterogeneous grouping of diseases. However, some key proteins and pathways are involved in the carcinogenesis of many CRCs.

The earliest morphological precursors of CRC are the aberrant crypt foci (ACF). They can be observed macroscopically with the use of a magnifying endoscope on colonic mucosa stained with methylene blue, as well as microscopically. They appear as large crypts with thick epithelium and reduced mucin content, and typically occur in clusters [25,177]. Most ACF will regress, however, some will progress and grow to become larger lesions [178]. The histopathology of ACF is variable, representing the variety of polyps for which they can be precursors. While many of these benign polyps will not progress to cancer, some will. Different types of polyps can be found within the colon; some are more likely to become malignant than others, and some are associated with specific pathways of carcinogenesis, and with specific CRC syndromes.

Traditionally, two main types of colorectal polyps were recognized: adenomatous and hyperplastic. Adenomatous polyps, or adenomas, are neoplastic and show evidence of dysplasia. They are considered precursor lesions for CRC, although most will not develop into an invasive carcinoma. A higher proportion of adenomas will progress to carcinoma, and they will do so at a faster rate of progression, in the setting of increased familial risk than they will in the general risk population [179]. Adenomas can have tubular, villous, or tubulovillous histopatholgy; they can be elevated, flat, or depressed [25]. Elevated polyps can be pedunculated, on long stalks or just slightly raised over the mucosal surface. Sessile polyps are those without stalks [25]. Although adenomas of any size or subtype can be associated with any degree of dysplasia (and even invasive carcinoma), the greatest malignant risk is associated with severe dysplasia, often of large, usually sessile adenomas with villous morphology [19].

Until recently, hyperplastic polyps were generally not considered neoplastic; by definition they were not dysplastic and did not progress to carcinoma. They make up the majority of epithelial polyps of the colon, are generally small, asymptomatic, and can be found in more than half of all adults 60 years of age or older [19].

In recent years, hyperplastic polyps have been reclassified; it is now recognized that they include a heterogeneous group of lesions, some with a characteristic serrated histopathology [180]. In 1999, Jass and colleagues reported a subtype of colorectal cancer which arose from serrated polyps [181,182]. While the traditional, non-neoplastic hyperplastic polyps are by far the most common, representing 80% to 90% of all serrated polyps, the other subtypes can lead to malignancy. They are the traditional serrated adenomas and the sessile serrated adenomas. The traditional serrated adenomas are the rarest of the serrated polyps, representing only 0.7% of all identified polyps in one study [183]. They were characterized in 1990, and were first thought to be a variant of adenomatous polyps [184]. The sessile serrated adenomas represent approximately 9% of all colorectal polyps , and 22% of all serrated polyps [183]. They tend to be right-sided, large, flat, and poorly circumscribed.

Some CRCs can develop in the absence of polyps, such as in chronic inflammatory bowel disease. It is difficult to confirm whether or not some other adenocarcinomas develop from a benign precursor lesion, since even small adenocarcinomas can sometimes completely overgrow such a lesion. The prolonged time interval required for a precursor lesion such as a polyp to develop into an invasive carcinoma provides an opportunity for identification and removal of the lesion, thereby

preventing the development of cancer. A lesion is considered to be a carcinoma with invasive potential once it has spread through the muscularis mucosa and into the submucosa (Figure 1.3.). Methods of screening for colorectal polyps and carcinomas are discussed in section 1.9.

1.7.1 Chromosomal Instability

Chromosomal instability (CIN) is the most common form of genetic instability. Approximately 60% of colorectal cancers exhibit CIN [185], although this number was originally thought to be much higher [186]. This is the most common pathway for the development of sporadic CRC, and is also the pathway of CRC development in Familial Adenomatous Polyposis (FAP), a hereditary CRC syndrome discussed further in section 1.8.2.1, and *MUTYH* polyposis, which is discussed in section 1.8.2.2.

CRCs with CIN exhibit defects in chromosomal segregation, which result in changes in chromosomal copy number (aneuploidy) and structure [187]. Subchromosomal amplifications and a high frequency of loss of heterozygosity (LOH) are common [188]. The rate of chromosomal loss or gain is at least 10^{-2} per generation, which equates to approximately the gain or loss of a chromosome in at least one in every five cell divisions [187]. Coupled with the abnormal karyotype, colorectal tumours with CIN also accumulate a characteristic set of mutations in key oncogenes and tumour suppressor genes which drive tumour progression and malignant transformation.

The CIN pathway is initiated by mutations in the *APC* gene. Located on the long arm of chromosome 5, the *APC* gene encodes a protein (APC) which has multiple roles within the cell, including roles in cell migration, cell adhesion, chromosome segregation, and apoptosis [189]. Truncating mutations of *APC*, as seen in FAP, contribute to chromosomal instability through disruption of mitotic spindle stability and impaired chromosomal segregation, although the exact mechanism is unknown [190]. Other mechanisms have been proposed as potential contributors to CIN. They include telomere dysfunction, impaired DNA damage response, and defects in the mitotic spindle assembly checkpoint (reviewed in [188]), although the causes of CIN are still not completely understood. The progression from normal colonic epithelium to adenoma to carcinoma is associated with an increasing level of CIN.

Another major role of APC is that of tumour suppressor, through antagonism of the Wingless Int (Wnt) signalling pathway [190-193]. APC forms a complex with other proteins including axis inhibitor 1 (AXIN 1), casein kinase 1 (CK1), glycogen synthase kinase 3 (GSK3), and others. This complex, known as the β-catenin destruction complex, binds to β-catenin, ultimately leading to its degradation. Biallelic inactivation of *APC* with loss of functional APC protein leads to the inappropriate stabilization of β-catenin. Such stabilization leads to accumulation of β-catenin within the cytoplasm and its subsequent transfer into the nucleus. There, β-catenin acts as a transcriptional co-activator with the T-cell factor/Lymphoid enhancer factor (TCF/LEF) transcription factor families, and persistently activates the transcription of many different target genes. These Wnt pathway target genes are involved in many different cellular processes, many of

which are involved in cancer progression [194], including proto-oncoprotein c-Myc [195], and cell cycle regulator cyclin D1 [196]. More details on this subject are given in section 3.3.4.1. Activation of the Wnt signalling pathway occurs most commonly through loss of the *APC* gene [190], but also, rarely, through mutations in the gene encoding β-catenin which render the protein resistant to degradation [197].

Biallelic inactivating mutations of *APC* result in dysplastic aberrant crypt foci and disorganized, dysplastic epithelium. Subsequent nuclear accumulation of β-catenin and activation of the Wnt pathway are important for continuing this process, but many more steps are required for the development of cancer. Progression from normal to dysplastic colonic epithelium, to small then larger adenoma, and eventually to carcinoma and metastasis is a multistep process involving the accumulation of genetic and epigenetic alterations which inactivate tumour suppressor genes and activate proto-oncogenes (Figure 1.9.). This multistep model of colorectal carcinogenesis was first proposed in 1990 by Fearon and Vogelstein [186], and has since been refined with the incorporation of recent discoveries [188,198]. Throughout this process lesions grow in size, transform to dysplastic epithelium with subsequent higher grades, and acquire more of a villous morphology [186]. Development of carcinoma takes an average of 8 to 12 years [194], although this is accelerated in affected members of families with hereditary CRC syndromes [199].

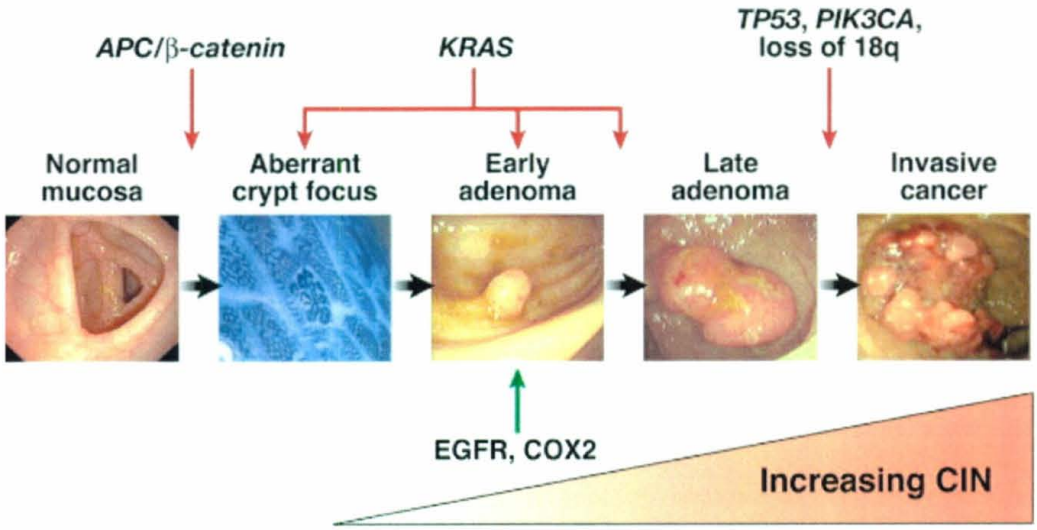


Figure 1.9. Genetic model for colorectal carcinogenesis
The red arrows indicate the approximate timing of the various mutations acquired during tumour progression. Reprinted with permission from Elsevier: *Gastroenterology*, volume 138, pages 2059-2072, copyright 2010 [188]

Mutations of *KRAS* are found in 38% to 50% of CRCs [186,200]. They occur early in the neoplastic process, and lead to cellular proliferation in the growing adenoma. Single oncogenic mutations lead to constitutive activation of downstream signalling, including of the mitogen-activated protein kinase (MAPK) signalling pathway, phosphatidylinositol 3-kinase (PI3K), and NF-κB. Through activation of these pathways and others, mutations in *KRAS* play a role in regulating multiple cellular functions, including cellular proliferation, differentiation, and cell survival.

Allelic loss of 18q is associated with the progression from early to late adenoma, and occurs later than *KRAS* activation. This loss is identified in up to 50% of late adenomas and 70% of CRCs [201]. While the *DCC* (deleted in colorectal carcinomas) gene was initially identified as a candidate tumour suppressor within this region [202], the role of

DCC in colorectal carcinogenesis has since been questioned [203,204]. Other potential candidate genes in the 18q region include *SMAD2* and *SMAD4*, which are mediators in the transforming growth factor-β (TGF-β) pathway, and involved in several important biological processes including cell growth, differentiation, and apoptosis [188,205]. Another potential candidate in the 18q region is *Cables*, a linker protein that increases tyrosine phosphorylation of cyclin-dependent kinases, thereby inhibiting cell growth [206]. It is lost in 65% of CRCs, through a combination of allelic loss of 18q and promoter hypermethylation [207].

The most common allelic loss in CRC involves loss of the *TP53* gene on chromosome 17p [201,208]. Allelic loss of this region is identified in 75% of CRCs, but in a much smaller percentage of adenomas, and is therefore considered a key step in the malignant transformation from colorectal adenoma to carcinoma [201]. The p53 protein, encoded by the *TP53* gene, plays important roles in maintaining cellular integrity. It is a tumour suppressor and regulates growth in two main ways. p53 regulates progression through the cell cycle, halting progression when damage is detected. When the damage cannot be repaired, p53 initiates cell death, through apoptosis.

The progression from normal epithelium to adenoma and eventually carcinoma is not a simple process. The genetic alterations described above are not the only ones involved in progression along the CIN pathway, however, they are the most frequently reported and play key roles in tumourigenesis. Researchers have determined that the progression to invasive carcinoma requires a minimum of five genetic alterations [186]. Sequencing the CRC genome has identified an average of 80 to 90 mutated genes per

tumour, of which 11 to 15 are considered drivers of tumourigenesis, and the others non-causative passengers [209,210]. The most important signalling pathways involved in driving colorectal carcinogenesis include the Wnt, TGF-β, EGFR, MAPK, and PI3K pathways [211]. The acquisition of metastatic potential requires even more genetic changes. While the overall profile of genetic alterations helps define the phenotype of a tumour, the order in which these alterations occur is important for tumour development. Other mutations, such as of *TP53* or *KRAS*, are not capable of initiating tumour growth if they occur before loss of *APC* [212,213].

1.7.2 Microsatellite Instability

Microsatellite instability (MSI) is another form of genetic instability and arises differently than CIN. While the instability of the CIN pathway affects whole chromosomes or parts of chromosomes, the instability of MSI affects DNA sequences. Unlike CIN CRCs which tend to be aneuploid, MSI CRCs have a diploid DNA content. MSI results from defects in the mismatch repair (MMR) pathway, which is responsible for correcting errors in DNA replication, ensuring the fidelity of meiotic recombination, and participating in the initial steps of cell cycle checkpoints and apoptosis in response to DNA damage [214]. Approximately 10% to 15% of all CRCs have high levels of MSI (MSI-H), as a result of MMR deficiency. About one quarter of these are due to germline mutations as seen in Lynch Syndrome, the remainder are due to somatic inactivation of MMR genes, most commonly by hypermethylation of the *MLH1* promoter [215,216]. The

remaining wild-type allele is inactivated through somatic mutation, LOH, or promoter methylation [217]. Somatic inactivating mutations of MMR genes are generally not seen in sporadic CRCs. The MMR system improves the fidelity of DNA replication by about 100 to 1000 fold [218]. Tumours with deficient MMR systems can be identified due to their unrepaired DNA strand slippage in microsatellite sequences, or MSI. MSI has also been identified in a number of non-colorectal solid tumours, most commonly in some endometrial, lung, and gastric cancers [219,220].

In the early 1990s, Manuel Perucho and colleagues were among many researchers searching the genome for novel tumour suppressor genes. Perucho's group used an arbitrarily primed polymerase chain reaction (PCR) method to generate DNA fingerprints from colorectal tumours and adjacent normal colonic mucosa [221]. When comparing these fingerprints they noted both loss and gain of genetic material in some tumours. They further characterized these changes, and reported that they were due to deletion mutations in simple repeats sequences within the tumour genome [222]. They correctly hypothesized that these changes were due to a defective replication factor within the tumour, and that this reflected a previously undescribed form of carcinogenesis which was associated with a unique phenotype. Tumours with these deletions were negatively correlated with *p53* and *KRAS* mutations and with metastasis, but were positively correlated with poorly differentiated tumours and with a proximal tumour location [222].

At around the same time that Perucho's group reported their discovery, Stephen Thibodeau's research laboratory identified the same phenomenon [223]. Thibodeau's group studied dinucleotide repeat sequences located in the chromosomal regions often seen

deleted in CRC. By comparing normal and tumour DNA, they noted a somatic instability at these microsatellite sequences in the DNA of some tumours. They coined the phrase "microsatellite instability" (which they abbreviated MIN) to describe their finding, and reported that it was significantly associated with proximal tumour location and increased patient survival [223]. Like Perucho's group, Thibodeau and colleagues concluded that this instability was the result of a novel mechanism of carcinogenesis.

A third research group, lead by Burt Vogelstein and Albert de la Chapelle, identified the same type of alterations in repeat sequences at the same time, and published alongside Thibodeau's group in Science in 1993 [60]. They compared the molecular features of familial colon tumours with sporadic tumours, and identified widespread alterations of short repeat sequences in most of the familial cancers, as well as 13% of sporadic cancers. They recognized that this finding was likely due to replication errors, and so applied this name to the phenotype, with the abbreviation RER.

Earlier research of mutated bacteria and yeast had identified a similar mutator phenotype of increased replication errors that was due to defects in DNA mismatch repair (MMR) [224,225] (reviewed in [226]). In *Escherichia coli*, the MMR system includes three MMR proteins: MutS, MutL, and MutH, which recognize base pair mismatches and small insertion and deletion mispairs, and initiate their repair. The MMR genes and protein products were named "mut" because of the mutator phenotype that results when they are deficient. Recognition of the similarities between this and the human cancer phenotypes led to the association of microsatellite instability in cancer with defects in the human DNA MMR system.

Three key papers published together in Cell in December 1993 made the connection between the *E. coli* and human cancer hypermutability phenotypes [227], and mapped the first causative gene in Lynch Syndrome, which was a homolog of the *E. coli* *mutS* gene [228,229]. Its protein product binds to mismatched DNA base pairs, and is required for initiation of the DNA MMR system [230]. The human homolog was named for the homologous gene in yeast, *MSH2*, for *mutS* homolog 2 [228]. Since then, mutations have been found in the human homologs of other mismatch repair genes, including *MSH6* [231], which is another homolog of the bacterial *mutS* gene, as well as *MLH1* (*mutL* homolog 1) [232,233] and *PMS2* (post-meiotic segregation protein 2) [234], which are homologs of the bacterial *mutL* gene. Germline mutations in these genes predispose carriers to Lynch Syndrome, the most common form of inherited colorectal cancer, discussed further in section 1.8.1.1.

The DNA mismatch repair system has been conserved throughout evolution, with strong similarities in *E. coli*, yeast, human, and other species. It functions to repair errors in newly synthesized DNA, including insertion/deletion loops (IDLs) and base-base mismatches. IDLs can occur due to misalignment during the replication of short tandem repeat sequences such as microsatellites. This is also known as strand slippage, and results in a different number of nucleotides in the template and newly synthesized DNA, causing a heteroduplex. A number of genes contain microsatellites within their coding regions, including some growth factors and their receptors, as well as some genes involved in apoptosis [235]. IDLs cause frameshift mutations in microsatellites, which result in the production of truncated, inactive proteins. Mutations of microsatellites

located in non-coding regulatory regions may also contribute to altered gene expression [217]. Together, the altered expression of these target genes may drive carcinogenesis through positive selection.

In the human system, MMR proteins act in heterodimer pairs for mismatch recognition and the initiation of repair. MSH2 is central to all mismatch recognition [236]. It binds with other MutS homologs MSH6 and MSH3, to form MutSα and MutSβ heterodimers, respectively. These protein complexes are ATPases, and have different, but overlapping mispair-recognition specificity. MutSα preferentially recognizes base-base mismatches and small IDLs, whereas MutSβ preferentially recognizes some small as well as larger IDLs containing up to 16 extra nucleotides on one strand [236,237]. MutSα is the dominant MutS complex in human MMR, and represents 80% to 90% of cellular MSH2 [214]. Upon recognizing a DNA mismatch, the MutS homolog complex binds to the mismatch and recruits MutL homolog complex [238].

MLH1 is the key human homolog of MutL. It forms heterodimers with other MutL homologs including PMS2, PMS1, and MLH3, to form MutLα, MutLβ, and MutLγ, respectively. MutLα is the primary MutL homolog in human MMR, and accounts for approximately 90% of cellular MLH1 [214]. It acts to mediate the interactions between the MutS homolog complex and the proteins involved in excision and repair of the mismatch. In yeast, MutLβ acts to suppress mutagenesis; its function is unknown in humans, however, it may act as a backup for MutLα in MMR. MutLγ may also have a backup role in MMR, and is involved in meiotic recombination [238].

Once the MutS and MutL complexes have bound together to form a ternary complex, they undergo ATP-dependent conformational changes, resulting in the formation of a sliding clamp around the DNA. The ternary structure then diffuses away from the mismatch site in either a 5' or 3' direction. The conformational changes also allow the complex to interact with the other proteins required for the repair process. These proteins include proliferating cell nuclear antigen (PCNA), replication protein A (RPA), exonuclease I (ExoI), replication factor C (RFC), and DNA polymerase δ.

The ternary complex diffuses away from the mismatch until it encounters a DNA strand break. ExoI then degrades the strand between the break and the mismatch, until the mismatch has been removed. The single stranded DNA that remains after degradation is stabilized through binding of RPA. DNA polymerase δ then repairs the DNA, with the help of PCNA and other replication factors. The final nick is then sealed by DNA ligase I, completing the repair process. The details of the MMR process differ slightly depending on the direction in which the complex diffuses away from the mismatch, as demonstrated in Figure 1.10.

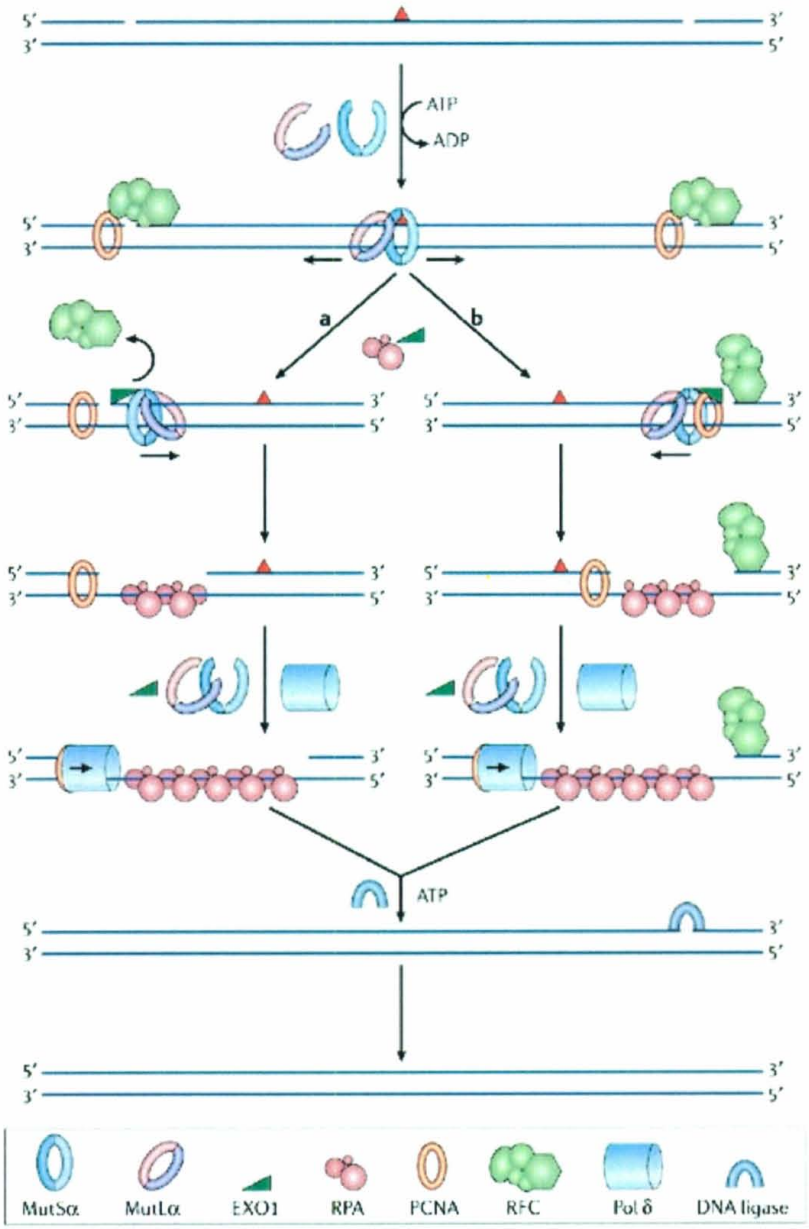


Figure 1.10. The human mismatch repair system
Reprinted by permission from Macmillan Publishers Ltd: *Molecular Cell Biology*, volume 7, pages 335-346, copyright 2006 [238]

In 1998, the National Cancer Institute (NCI) published the outcome of a workshop on microsatellite instability, in which international guidelines for the determination of MSI status in cancer were defined [61]. The authors recommended that tumour and normal DNA should be compared for a reference panel of five microsatellites: two mononucleotide repeats (BAT25 and BAT26), and three dinucleotide repeats (D2S123, D5S346, and D17S250). Designations of high and low levels of MSI (MSI-H and MSI-L respectively), as well as microsatellite stable (MSS) could be assigned to each tumour based on outcome of testing. Instability at two or more of the five microsatellites would be MSI-H, whereas no instability would be MSS. In cases where instability was observed at one of the five microsatellites, a designation of MSI-L would be assigned, and the DNA tested with a second set of five defined microsatellites in an effort to better characterize the tumour.

Since the publication of the NCI reference panel, the methods used to identify MSI have been refined. Researchers have concluded that mononucleotide repeats are easier to interpret and lead to fewer false positives than dinucleotide repeats [239,240]. As well, some mononucleotide repeats are quasimonomorphic, meaning that the use of normal DNA for comparison purposes is no longer required [239-241]. A pentaplex polymerase chain reaction (PCR) of five mononucleotide microsatellites was proposed in 2002, which does not require the testing of normal DNA [240]. It is commercially available, and has a higher sensitivity and specificity for identifying MSI than the NCI reference panel [242].

Mismatch repair deficiencies can also be identified by immunohistochemical (IHC) analysis, by comparing expression of MMR protein products in tumour and non-tumour specimens [6]. IHC analysis has the advantage over MSI testing that it identifies the specific MMR protein which is deficient, thereby better enabling mutation testing for patients with Lynch Syndrome. It also identifies deficiencies in MSH6 that can be missed by MSI testing, as MSH6 deficiencies do not always lead to MSI-H tumours [243-245]. As well, most pathology laboratories are equipped to perform IHC testing, while fewer are equipped for MSI testing. However, IHC can be more expensive and time consuming. Some pathogenic mutations may leave the IHC epitope intact, resulting in misleading IHC results. As well, the interpretation of IHC results can be more subjective; inter-observer variation can be a concern, especially when staining for MMR proteins in MMR deficient tumours [246].

Knowledge of MSI status has value beyond aiding in the identification of Lynch Syndrome; patients with MSI-H tumours have a better prognosis than those with MSS tumours [61,63,64,66-69]. As well, due to the involvement of the MMR system in the cellular response to DNA damage, MSI-H tumours respond differently to some standard chemotherapeutic agents [70,71] (discussed in more detail in section 1.10).

The improved prognosis of patients with MSI-H tumours is not completely understood, and paradoxically, many histological features associated with these tumours are generally considered to indicate a more aggressive phenotype (for example mucinous and poorly differentiated histology). However, some mechanisms that may contribute to the improved prognosis include: an increased anti-tumour immune response [247], the

diploid nature of the tumours [248,249], and a decreased tumour viability due to the accumulation of mutations in genes required for viability [250,251]. The increased host immune response to MSI-H tumours was previously discussed in section 1.5.2.1. Briefly, the microsatellite frameshift mutations create novel peptides, which the host immune system recognizes as foreign and mounts an antigen-specific CD8+ T cell anti-tumour response, which selectively kills tumour cells [247,252].

Although most studies agree that MSI is a positive prognostic marker for colorectal cancer, there have been some controversies and mixed findings. This may be due in part to the proportion of rectal cancers included in different studies. Most MSI-H tumours are located in the proximal colon, and little work has been published to characterize distal tumours. One large study by Samowitz and colleagues [253], examined almost 1000 rectal cancers. They reported a worse prognosis for rectal cancer patients with MSI-H tumours, when compared to those with MSS tumours. This remained significant even after adjustment for patient age and tumour stage at diagnosis.

Discrepancies in prognostic significance may also result from the bimodal age distribution of patients with MSI-H CRCs. Patients with Lynch Syndrome, who make up approximately one quarter of those with MSI-H tumours, tend to be younger at diagnosis than average for CRC. For this young group, MSI-H is a strong, positive prognostic indicator [65]. The other group of patients with MSI-H tumours are those without Lynch Syndrome. They have sporadic tumours which are MSI-H due to hypermethylation of the *MLH1* promoter region. These patients tend, on average, to be older than 70 years old at diagnosis. Due to their older age, there can be an expected higher rate of comorbidities,

which could confound outcome data [254]. Although the tumours of this later group are MSI-H and they share many features with the younger, Lynch Syndrome group, the pathway of carcinogenesis is different. This is discussed in more detail in section 1.7.3., CpG Island Methylator Phenoptype. While methylation is an epigenetic event, a heritable predisposition to it can occur. This phenomenon, as related to MSI-H CRCs, is discussed in the following section.

1.7.2.1 Heritability of epigenetic MMR silencing

Epigenetic silencing of MMR genes can have a heritable component. A study from early 2007 [255] demonstrated that heritable variation in the *MLH1* gene can contribute to the epigenetic silencing of *MLH1* seen frequency in sporadic colorectal and endometrial cancers. Chen *et. al.* carried out a case-control study of women with endometrial cancer; the cases had hypermethylation of the *MLH1* promoter region in their tumours, whereas the controls did not. They typed eight *MLH1* SNPs in the normal DNA of these women and found significant association between the abnormal methylation and the rs1800734 SNP, located in a CpG island at -93 from the translation start site. Their validation studies, which also included a cohort with CRC, found that the odds ratio associated with the AA or AG versus GG genotype of this SNP was 1.61, with a 95% confidence interval of 1.20 to 2.16 [255]. While the methylation itself is not heritable, the *MLH1* variant leading to an increased risk of abnormal methylation is heritable.

While hypermethylation of the of *MLH1* promoter is an accepted mechanism of MMR gene silencing leading to CRC and other cancers, epigenetic silencing of the *MSH2* gene is not as well accepted. However, in 2006 Chan *et. al.* [256] reported on a Chinese family with clinically diagnosed Lynch syndrome, with germline epigenetic silencing of *MSH2* as the MMR defect underlying their disease. Affected family members had colorectal tumours with high levels of microsatellite instability that were lacking expression of the MSH2 protein on immunohistochemistry. However, while germline mutations were not found, germline allele-specific hypermethylation of the *MSH2* promoter region was identified, although in a mosaic pattern.

The phenomenon of heritable epimutation of *MSH2* in Lynch syndrome first described by Chan *et. al.* [256] has recently been revisited and explained [257,258]. Chan's group in China and another research group in the Netherlands collaborated to publish a paper in early 2009 [257] that further characterized the original Chinese family and described five other families with similar findings. All had the clinical appearance of Lynch syndrome, and the tumours were MSI-H, indicating MMR deficiency. The tumours also showed lack of expression of MSH2 by immunohistochemistry, yet there were no detectable *MSH2* mutations, or other MMR gene mutations. All affected individuals had mosaic, allele-specific *MSH2* promoter methylation. They found that affected individuals from all six families had deletions of the most 3' exon of *TACSTD1*, the gene located immediately upstream of the *MSH2* gene, just 17kb away from exon 1. These deletions all involved the polyadenylation site of *TACSTD1*, but left the *MSH2* promoter region intact. The deletions segregated with disease within the families, and

also segregated with the methylated *MSH2* allele and the expression of fusion transcripts that included parts of both the *TACSTD1* and *MSH2* genes. The authors suggested that the deletion of the *TACSTD1* polyadenylation site leads to transcription of the gene continuing past its normal end-point and reading-through into *MSH2*. The read-through transcription leads to methylation of the *MSH2* promoter through an unknown mechanism, resulting in inactivation of the *MSH2* gene. The methylation is seen in a mosaic pattern because it is dependent upon the expression of *TACSTD1*, and does not occur in tissues where *TACSTD1* is not expressed. It is not the methylation itself that is heritable, but instead the deletion, which ultimately leads to methylation [257].

A second paper [258] describing the same type of 3' *TACSTD1* deletions in Lynch families was published close to the same time. Kovacs *et. al.* started with a cohort of 49 Lynch families, and performed multiplex ligation-dependent probe amplification (MLPA) looking for deletions in the 27 families for whom they were unable to detect a disease-causing MMR mutation. They found 3' *TACSTD1* deletions in five of these families, representing 10% of their total cohort [258], suggesting that this mutation type could be underestimated in Lynch syndrome. As with the previous paper, they observed that the tumours in these families were MSI-H and MSH2 deficient. The authors did not test the affected individuals for *MSH2* promoter methylation, but instead suggested that the *MSH2* inactivation was due to transcriptional interference; that the read-through of the *TACSTD1* gene into the *MSH2* gene and the production of fusion transcripts led to a non-functional MSH2 protein product.

Kovacs group also observed what they termed a "restrained phenotype" among these deletion carriers [258]. While families with *MSH2* mutations have been reported to be at higher risk of extracolonic cancers than *MLH1* mutation carriers [259], families with *MSH2* inactivation caused by 3' *TACSTD1* deletions are less likely to have these cancers. They attribute this to the tissue-specific expression pattern of *TACSTD1*. The gene is expressed predominantly in the colon and rectum, resulting in these particular tissues having a larger percentage of cells with *MSH2* inactivation. While other Lynch-involved organs, such as the endometrium, do have *TACSTD1* expression, it is at a lower level, resulting in fewer cells with *MSH2* inactivation, and a decreased likelihood of MMR deficient cancer development from this mechanism [258].

1.7.3 CpG Island Methylator Phenotype

As discussed in section 1.7.2, the majority of colorectal tumours with MSI-H do not result from Lynch Syndrome, with germline mismatch repair mutations, but instead are sporadic, associated with epigenetic hypermethylation of the *MLH1* promoter region and subsequent silencing of that gene. This form of hypermethylation is not limited to the *MLH1* promoter, nor is it found only in MSI-H tumours [185,260-262]. The hypermethylation phenotype in CRCs is associated with female sex, advanced age at diagnosis, proximal tumour location, high tumour grade, and mucinous histology [263].

CpG islands are short sequences of DNA with high concentrations of the CG dinucleotide, and are found in the 5' region of more than half of all human genes [264].

CpG refers to a cytosine adjacent to a guanine in the linear DNA structure, separated by a phosphate group, and not to CG base-pairing. The cytosine can be methylated to form 5-methylcytosine, resulting in repression of transcription of the associated gene. Such epigenetic changes are self-propagating – passed on during DNA replication, and they are potentially reversible, unlike genetic changes [265]. This form of epigenetic gene silencing is associated with normal physiological occurrences including X chromosome inactivation[266], genomic imprinting [267], and aging [268]. Aberrant *de novo* hypermethylation of CpG islands has been associated with genetic diseases including Fragile X Syndrome [269,270], as well as some cancers [268,271-273].

In cancers, the hypermethylation and subsequent silencing of tumour suppressor genes is postulated to impart a growth advantage, thereby initiating tumour formation [273]. Another mechanism of tumourigenesis is evident in colorectal cancer, as mentioned above. Hypermethylation of the *MLH1* promoter region silences this DNA mismatch repair gene, thereby setting the course of tumour development through the MSI pathway [274]. This methylation is not unique to the genes mentioned, but instead reflects a methylator phenotype, termed CpG island methylator phenotype (CIMP), whereby there is widespread methylation of multiple CpG islands. The *MLH1* promoter is not always hypermethylated in CIMP CRCs, but of sporadic CRC it is only CIMP CRCs that are also MSI-H. Other genes commonly silenced in this way in CIMP CRCs include: the cell cycle regulator and tumour suppressor gene *CDKN2A*, which expresses the p16 protein [273-275]; the estrogen receptor gene, *ER* [268]; angiogenesis inhibitor thrombospondin-1

(THBS1) [274,275]; and *O^6-methylguanine DNA methyltransferase*, the gene which encodes the MGMT DNA repair protein [276,277].

The mechanisms which cause the increased methylation in CIMP are not well understood. CpG island methylation increases with age, and it has been proposed that if such methylation inactivates specific genes involved in tumour suppression, a growth advantage can result, and a neoplasm can develop [268,274]. As well, aberrant methylation has been associated with exposure to specific environmental carcinogens [278]. While traditionally considered to affect individual CpG islands and specific genes, more recent research has suggested that epigenetic gene silencing in CIMP can be coordinated across large chromosomal regions [279]. Such regions can be up to one megabase (Mb) in size, and may be controlled in a similar manner to the long-range epigenetic regulation seen in X-chromosome inactivation and imprinted gene clusters [279]. Such long-range gene silencing in cancer could be considered somewhat analogous to loss of heterozygosity from genetic deletion of chromosomal regions, which is frequently observed in many cancers.

CIMP CRCs are epidemiologically associated with patient smoking [135]. This finding has been independent of MSI status, and there is a statistically significant dose-response relationship in terms of the smoking history and the degree of hypermethylation. An association between smoking and aberrant methylation has previously been observed in sputum, bronchial epithelium, and non-small cell lung cancers [280-284], although the mechanism is unknown. It has been proposed that smoking may induce the aberrant methylation phenotype, or possibly *BRAF* mutations, in CRCs [135].

Another theory of the mechanism by which epigenetic changes cause cancer is known as the epigenetic progenitor model [285]. Hypermethylation of CpG islands is only one epigenetic alteration that can be seen in cancers; others can include both global and gene-specific DNA hypomethylation, and hypoacetylation of chromatin. The epigenetic progenitor model of cancer suggests that carcinogenesis occurs through three general steps. The first step involves the polyclonal epigenetic disruption of stem cells or progenitor cells within a given tissue. This leads to the generation of a population of epigenetically altered progenitor cells. The second step in this process involves the acquisition of an initiating mutation within the subpopulation of epigenetically altered progenitor cells. This mutation is traditionally considered to be the first step in tumour development, and disrupts the gatekeeper gene associated with a specific cancer type. For example, for CRC the associated gatekeeper gene is considered to be *APC*. For some types of leukemia this step is not a genetic mutation, but instead a specific chromosomal rearrangement. The third step in this model of cancer involves genetic and epigenetic plasticity, which leads to increased tumour evolution.[285]

Mutations leading to the ability to invade and metastasize are not a part of this model, because these characteristics are inherent to the progenitor cells from which the cancers are proposed to arise [285]. The epigenetic progenitor model of cancer is supported by the finding that the sites of methylation changes in cancer are those that vary normally in tissue-specific differentiation [286]. As well, alterations in methylation can be observed in aberrant crypt foci, indicating that this occurs very early in the development of these cancers [287].

The existence of the CpG island methylator phenotype as a distinct subset of CRC was controversial, as many researchers had concluded that CpG methylation occurs at a continuous frequency spectrum [288,289]. All tumours, and even normal colorectal tissue, have some DNA methylation. This CpG methylation increases with age, and a normal distribution of methylation has been found when studying cohorts of CRC [288,289]. However, different studies used different panels of CpG islands as markers. Since aberrant CpG methylation increases with age, and it is possible that only a subset of CpG islands are involved in CIMP, which itself only affects a subset of tumours, there are many confounders to complicate interpretation of data.

In 2006, Weisenberger and colleagues published a study which strongly supported the existence of CIMP, and suggested an improved marker panel for classifying CIMP tumours [290]. They first carried out a screen of all 195 methylation markers they had available to their lab, using MethylLight technology, a high-throughput assay to measure DNA methylation. They compared the methylation patterns of these markers among ten pairings of normal colon and tumour DNA, and removed all markers that did not show differential methylation. They next used the 92 markers that they found to have tumour-specific methylation on 48 independent colorectal tumours, and performed unsupervised hierarchical clustering analysis on the methylation data. Although some methylation was observed in all tumours, this work identified a distinct cluster group which had much higher levels of methylation. As reported earlier by other groups [261,291-293], the CIMP tumours were significantly associated with mutations in the *BRAF* gene, as well as with proximal location within the colon. All of the tumours in their cohort that fit the CIMP

cluster group had mutations in either the *BRAF* or *KRAS* genes, and the authors suggested that there may be a subgroup within the CIMP tumours, which could be identified through these gene mutations. The authors determined which five markers best identified the CIMP tumours, and suggested that these markers be used in future work, to classify CIMP [290].

CIMP has been observed in some hyperplastic polyps, especially in the proximal colon [294-296]. While traditionally, hyperplastic polyps were considered to be non-neoplastic and without malignant potential, it is now recognized that they are not a single type of lesion, but a complex family of lesions. A subset of them have now been implicated as the precursor lesion of CIMP CRCs [294-296]. Most hyperplastic polyps are small, round lesions located in the distal colon, however, others are located in the proximal colon and tend to be somewhat larger; these are the ones most commonly implicated in CIMP, and have been reclassified. One feature that all of these lesions share is a serrated architecture on histological examination, and they have therefore been collectively termed serrated polyps [297]. The term serrated refers to the sawtooth-like infolding of the surface of the epithelium within the crypt, and is considered to be the result of decreased apoptosis [298].

In 1990, Longacre and Fenoglio-Preiser reported their findings of a group of polyps with features of both traditional hyperplastic polyps and adenomatous polyps [184]. They identified these not as mixed polyps, but as a single entity in which the glands had a serrated structure. They named these serrated adenomas.

In 1996, Torlakovic and Snover studied the polyps of patients with Hyperplastic Polyposis Syndrome (HPS) [299], a condition in which patients have multiple hyperplastic polyps throughout their colon and rectum (further discussed in section 1.8.2.3). These lesions were considered to be non-cancerous, however, Torlakovic and Snover described four patients who developed associated adenocarcinomas. They compared their polyps to classic hyperplastic polyps, and noted some histological differences. They concluded that these patients had serrated adenomas.

The lesions currently called hyperplastic polyps are the small, innocuous lesions most often found in the distal colon and rectum. They are the most common, representing about 80% to 90% of all serrated lesions, and are found in more than 10% of asymptomatic individuals [287]. Other serrated lesions can be sessile or polypoid, and include admixed polyps, serrated adenomas, sessile serrated adenomas and serrated adenocarcinomas [287,297]. Sessile serrated adenomas are generally considered the precursor lesion for serrated adenocarcinomas – the term given to CIMP CRCs; and the pathway of carcinogenesis has been termed the serrated pathway [287]. These are often large and located in the proximal colon. They make up approximately 9% of all colorectal polyps, and 22% of all serrated polyps [183].

Admixed polyps include polyps with histological features of different types of polyps; they have both hyperplastic and dysplastic components. The dysplastic component may include traditional adenoma or serrated adenoma [300]. Usually there is a predominant histological type, and the other mixed features are focal, however, the proportions can vary [301]. Initially, researchers considered admixed polyps to be the result

of a collision between two adjacent polyps of different histological type [302]. More recently, researchers have suggested that admixed polyps may represent a progression from one type of polyp into another, such as a sessile serrated polyp developing an area of dysplasia [184].

Two serrated pathways have been proposed, one that leads to right-sided MSI-H CRCs, and the other that leads to left sided MSS or MSI-L CRCs [287,300,303]. Although the pathways have not been clearly defined, both serrated pathways are thought to initiate through the methylator phenotype, and begin with serrated lesions. Both pathways involve activation of the mitogen-activated protein kinase (MAPK) signal transduction pathway, but through mutations in different members of the pathway. Mutations of the *BRAF* and *KRAS* genes are thought to exert equivalent effects on tumourigenesis, are mutated at a similar point during tumourigenesis, and are mutually exclusive; CIMP tumours will have either *BRAF* or *KRAS* mutations, but not both [185,261,292,304-306]. The patterns of methylation are different, and can be somewhat predicted through knowledge of which of the two genes is mutated [305].

The *BRAF* gene is a member of the *RAF* family of kinases that mediate cellular responses to growth signals via this pathway, and the p.V600E *BRAF* mutation has been closely associated with the CIMP phenotype [260,261,290,292]. This mutation likely mimics phosphorylation of the active site, by placing a negatively charged residue adjacent to the site of regulatory phosphorylation [307]. The constitutive activation observed as a consequence of this mutation promotes cellular proliferation and inhibits apoptosis – a cardinal feature of serrated lesions. Serrated CRCs with the p.V600E *BRAF* mutation

tend to be proximally located, have high levels of methylation (CIMP-high), and can be either MSI-H or MSS, depending on whether or not the *MLH1* promoter region becomes methylated. They also tend to have infiltrating lymphocytes, poor histological grade, and a mucinous morphology [261].

The *BRAF* p.V600E mutation is also associated with a poor patient prognosis [260,306,308]. This poor prognosis only applies to patients whose tumours are MSS; MSI-H tumours are associated with a good prognosis, regardless of *BRAF* mutation status [260,308]. In MSS, *BRAF* mutation positive tumours, the poor prognosis was statistically significant in both univariate and multivariate analysis adjusting for patient age, tumour stage, and tumour site [260]. One study reported that for MSS CRCs, increasing levels of methylation were significantly associated with decreasing survival [309].

The alternative serrated pathway is thought to be initiated by mutations in the *KRAS* gene[310], a member of the *RAS* family of GTPases. Like *BRAF*, *KRAS* also signals through the MAPK pathway, and *KRAS* activation also results in cellular proliferation and inhibition of apoptosis. Rather than methylation of *MLH1*, the *KRAS*-mutated CIMP CRCs are associated with methylation of the DNA repair gene O^6-methylguanine DNA methyltransferase (*MGMT*). Loss of *MGMT* is thought to increase the number of methylG:T mismatches, and overwhelm the DNA repair systems [185,277]. These tumours generally exhibit a moderate level of methylation (CIMP-low), are MSS or MSI-L, and are located distally [185,287,305,311]. While *KRAS*-mutated CRCs usually arise from serrated lesions, they can also arise from traditional adenomas though the chromosomal instability

pathway [287,312]. *KRAS* mutation in CRC is generally not considered an indicator of patient prognosis [262,306].

Screening and treatment for CIMP CRCs should be somewhat different from non-CIMP CRCs. CIMP CRCs are disproportionately represented in interval cancers – those cancers arising within five years of a complete colonoscopy [313]. This could be due either to an increased rate of development of CIMP CRCs, in which the early lesions progress to cancer within the time frame between regular colonoscopic examinations, or because the lesions are not identified on colonoscopy. There is evidence to support both of these possibilities. CIMP CRCs which are MSI-H are likely to progress quickly since they have a defective mismatch repair system similar to Lynch Syndrome CRCs [314], and in one study, interval cancers were four times more likely than non-interval cancers to be MSI-H [315]. As well, the flat subtle nature of the sessile serrated adenomas which precede CIMP CRCs can be difficult to identify on colonoscopic examination, especially if there is adherent stool or mucin, and such lesions can often be missed [287,314,316]. Even when these lesions are identified and biopsied, a pathological diagnosis of hyperplastic polyp may result in a false sense of security in the clinician regarding the malignant potential of the lesion [316].

Chemotherapeutic treatment of CIMP CRCs may be approached differently than non-CIMP CRCs, since they progress through different pathways, and therefore may respond differently to different drugs. An example of this is anti-epidermal growth factor receptor (EGFR) monoclonal antibodies, which are given as the first line of treatment for metastatic CRC. These drugs bind to the EGFR protein and inhibit its downstream

signalling. When not inhibited, EGFR triggers two main signalling pathways including the MAPK signal transduction pathway, which both *BRAF* and *KRAS* act through. Metastatic CRCs that harbour *KRAS* or *BRAF* mutations do not respond to these drugs [317,318].

The fact that CIMP CRCs arise through increased methylation leads to novel therapeutic options. Unlike genetic variants, epigenetic alterations are potentially reversible. A 2009 study by Issa and Kantarjian provided proof of principle for epigenetic therapy [319]. They reported that two nucleoside inhibitors of DNA methylation, azacitidine and decitabine, are successfully used clinically to treat myelodysplastic syndrome. The development of epigenetic-modifying drugs and the adaptation of these drugs to solid tumours may become important in the treatment of CIMP CRCs [320]. *BRAF* inhibitors, including sorafenib, have been shown to reduce signalling through the MAPK pathway, and have clinical efficacy against some tumour types, including melanoma, thyroid, hepatocellular and renal cell cancers [321]. Sorafenib has shown anti-tumour properties in mouse models of CRC [322], but has not yet been used clinically for CRC.

1.7.4 Colorectal Cancer Stem Cells

Although the traditional clonal theory of carcinogenesis suggests that any cell can initiate a tumour through the accumulation of multiple mutations which impart the necessary tumourigenic characteristics [175], this has been challenged by the cancer stem cell theory. The stem cell model of carcinogenesis proposes that only a select few cells,

the stem cells, are capable of initiating tumour formation, sustaining tumour growth, and promoting tumour progression [323,324]. Stem cells are suspects in cancer initiation, as they possess some of the characteristics deemed necessary for this role [325]. Unlike differentiated colonic epithelial cells, which experience relatively quick turnover, colonic stem cells possess longevity, which enables the accumulation of mutations. They also have the capacity to self-renew, which allows for the maintenance and propagation of these mutations. By definition, stem cells are undifferentiated, are capable of both self-renewal and proliferation, and have the capacity to develop into different cell types and regenerate tissue after injury [326]. Colonic stem cells reside at the base of colonic crypts, in niches that maintain the homeostasis required for their proper function; for this, the Wnt pathway, especially β-catenin, is considered to be of central importance [23,327]. There are an estimated five to ten stem cells per crypt, although the exact number is unknown as there is a lack of definitive stem cell markers [326].

According to the stem cell model, cancers are organized in a hierarchical manner, in which cancer stem cells can differentiate and form a heterogenous tumour through epigenetic changes. The bulk of tumour cells lose their stem cell properties through these epigenetic changes and are therefore nontumourigenic. It is the rare cells that maintain their stem cell properties that contribute to tumour progression [328]. Support for the cancer stem cell theory is strongest in some forms of leukemia [329,330], brain cancer [331], breast cancer [332], and colon cancer [324,333].

In 2007, O'Brien and colleagues reported that only one in $5.7x10^4$ human colon cancer cells was capable of initiating cancer when transplanted into the renal capsule of

immunodeficient NOD/SCID mice [324]. They found that there was a more than 200 fold enrichment of these cancer-initiating cells among tumour cells expressing the surface antigen CD133. The bulk of tumour cells were not capable of initiating a new tumour upon transplantation. Ricci-Vitani and colleagues reported similar findings in the same 2007 issue of Nature [333]. They reported that only transplanted $CD133^+$ tumour cells were capable of initiating a new tumour; $CD133^-$ cells were not. These new tumours were identical to the original tumours in both morphology and antigenic pattern [324,333].

Both O'Brien's and Ricci-Vitani's research groups discussed the potential implications of their findings for clinical management of CRC [324,333]. Chemotherapeutics are generally designed to interfere with the replication of highly proliferative cells; therefore, colon cancer stem cells, which may be slower growing, may be spared [334]. The highly proliferative cells which make up the bulk of the tumour are less likely to be tumourigenic. To help prevent both local and distant recurrence of cancer, therapeutics should target the cells capable of initiating the recurrences, the cancer stem cells. Such therapy may include agents which induce differentiation of the stem cells, along with immune therapies, and cytotoxic chemotherapies [334].

Colon cancer stem cells have the capacity for virtually unlimited expansion, both *in vitro* and *in vivo* [333]. The xenographs developed from these cells are identical to the primary tumours, and therefore have potential for use in developing efficacious drugs, and in individualizing patient therapy to better match the specific biology of their own tumour [324,333]. Cancer cell lines are not as reliable for these purposes, as they do not share all of the same features as the primary tumours [324].

While the evidence to support the theory of cancer stem cells is strong, it is not universally accepted, and the theory cannot be applied to all cancers. One concern is with the identification of cancer stem cells. While markers have been identified, none are fully specific for cancer stem cells. This renders analysis of the stem cells to be somewhat difficult. Without reliable, specific markers for cancer stem cells, it is impossible to fully distinguish between tumourigenic and nontumourigenic cells. It is important to be able to accurately identify these groups of cells, in order to determine what types of differences exist between them. Epigenetic changes would support the cancer stem cell model, and a combination of epigenetic and genetic changes would support the clonal model; but such differences may be also be due in part to cellular environment. The mere identification of some tumour cells which are more tumourigenic than others does not prove the cancer stem cell theory, but instead is entirely consistent with the traditional, clonal theory of cancer development.

Other concerns with the cancer stem cell theory arise from the use of animal models to test the tumourigenicity of human cancer cells. Due to different cellular environments and immune mechanisms, cells which are tumourigenic in immunocompromised mouse models may not be so in humans. As well, there is the potential to underestimate the frequency of cells with tumourigenic properties. [328,334]. In a 2008 study of melanoma, the authors reported that only one in one million human melanoma cells was capable of initiating a tumour in recipient NOD/SCID mice [335]. The identification of these melanoma stem cells was supported by the identification of a marker for these cells, ABCB5, as well as by evidence that anti-ABCB5 antibodies had

tumour-inhibitory effects [335]. A second study, published by a different research group later that year, reported that more than one quarter of human melanoma cells could initiate tumours when using modified xenotransplantation conditions [336]. By using a more highly immunocompromised mouse strain and optimizing transplant conditions, they were able to increase the identification of tumourigenic cells by many orders of magnitude. Further tests by the same group revealed that the tumourigenic cells were phenotypically heterogeneous and showed little evidence of hierarchical organization. They concluded that while the evidence supporting the stem cell model was strong for some cancers, this was not the case for melanoma [336].

The stem cell and clonal models of cancer are not necessarily mutually exclusive; a cancer which follows the stem cell model will undergo clonal expansion as it acquires epigenetic changes which confer a growth advantage. Similarly, a cancer which develops according to the traditional clonal theory may acquire stem-like properties, which then lead to further tumour progression. Tumours with stem cell qualities are reported to have a more aggressive phenotype, and worse prognosis [337].

Some evidence supports the theory that migrating stem cells are responsible for metastasis [338]. The migrating cancer stem cell theory suggests that it is cancer stem cells, likely located at the advancing edge of the tumour, that acquire the ability to migrate. They can then dissociate from the tumour, disseminate throughout the body, and initiate the growth of metastases at distant sites [338,339]. While, in theory, any tumour cell could acquire the ability to dissociate and migrate, stem cells are considered to be the only tumour cells capable of metastasis, as they would have the capacity to initiate a new

tumour deposit which recapitulates the morphologic and antigenic phenotype of the primary tumour. This could explain why some patients are found to have circulating tumour cells, yet do not go on to develop clinically identifiable metastases [323].

Cells capable of initiating metastasis must be able to undergo transitions from an epithelial to migratory phenotype to both invade nearby tissues and disseminate from the primary tumour, then reverse this transition, to revert back to an epithelial phenotype to initiate the growth of the metastasis at a new site within the body [339]. Such epithelial-mesenchymal (EMT) and mesenchymal-epithelial transitions (MET) are required for normal embryonic development. Signalling of the Wnt pathway is involved in the induction of EMT in the embryo, and aberrant activation of this pathway, with accumulation of β-catenin, can activate this process in tumour cells [338]. The Wnt pathway, and specifically β-catenin, is also involved in cell-cell adhesion, stem cell formation, and stem cell maintenance [338,339].

Somatic mutations in the *APC* gene, found in the great majority of CRCs, leads to an accumulation of β-catenin, and also to tumour budding [340]. It is an example of EMT within the tumour. These tumour cells have lost both glandular differentiation and cellular adhesion, and their identification is recognized as an independent negative prognostic factor in CRC [80]. This feature is significantly associated with the presence of *APC* mutations, and is rarely observed in MSI-H tumours, which lack such mutations. This may partially account for the improved prognosis in patients with MSI-H tumours [340].

1.8 Hereditary and Familial Colorectal Cancer Syndromes

Hereditary and familial colorectal cancer syndromes can be divided into two general categories: those with and without a phenotype of polyposis. Polyposis is simply defined by the identification of numerous polyps along the mucosa of the colon.

1.8.1 Non-Polyposis Syndromes:

1.8.1.1 Lynch Syndrome

Lynch Syndrome is the most common hereditary form of CRC, and accounts for up to 3% of all cases [4-6]. It is defined by the presence of germline mutations in mismatch repair (MMR) genes, and cancers develop through the MSI pathway, described in section 1.7.2. Predisposing mutations are passed on from parent to offspring in an autosomal dominant pattern of inheritance. Since CRC has a cumulative lifetime risk of 5% in the general population, it is estimated that as many as 1 in 660 people have germline MMR mutations [4], making it one of the most common of the heritable syndromes associated with high levels of morbidity and mortality. As discussed in the introduction, it was originally termed Cancer Family Syndrome.

The original 1971 criteria for Cancer Family Syndrome (CFS) were less stringent than current criteria for Lynch Syndrome. The current criteria have evolved over time as technology has allowed for a shift away from clinical definitions and towards

stratification of disease based on molecular findings. CFS was defined as an increased incidence of adenocarcinoma within a family, inherited in an autosomal dominant pattern, where colon and endometrium were the anatomic sites most commonly involved. Early age of onset of cancer and increased frequency of multiple primary cancers were also observed [14]. A current diagnosis of Lynch Syndrome requires the identification of a pathogenic germline mutation in one of the genes involved in DNA mismatch repair. The clinical picture remains much the same, however, other extracolonic cancers are now recognized as part of the syndrome, and certain molecular and histopathological features are also associated with it. The cardinal features of Lynch syndrome are listed in Table 1.10.

Table 1.10. Cardinal features of Lynch syndrome
(from [341,342])

Familial clustering of colorectal and/or endometrial cancer with an autosomal dominant inheritance pattern
Associated cancers: cancer of the stomach, ovary, ureter/renal pelvis, brain, small bowel, hepatobiliary tract, pancreas, and skin (sebaceous tumours)
Development of cancer at an early age (average 45 years of age for Lynch syndrome versus 65 years for sporadic CRC)
Development of multiple cancers
Features of colorectal cancer: • predilection for proximal colon (70% arise proximal to the splenic flexure) • improved survival when controlled for age and stage • multiple colorectal cancers (25 % to 30% have a second primary CRC within ten years from resection of the first, unless surgical treatment involved a subtotal colectomy) • increased proportion of tumours with mucinous and signet ring morphology, poorly differentiated tumours, and tumours with marked host-lymphocytic infiltration and lymphoid aggregation at the tumour margin
Features of the colorectal adenoma: • the numbers vary from one to a few • increased proportion of adenomas with a villous growth pattern • increased proportion of adenomas with a high degree of dysplasia

• rapid progression from adenoma to carcinoma (2 to 3 years versus 8 to 10 years in the general population)
High frequency of MSI (MSI-H)
Immunohistochemistry: loss of MLH1, MSH2, MSH6, or PMS2 protein expression
Germline mutation in MMR genes (*MSH2*, *MLH1*, *MSH6*, and *PMS2*)

Adapted from [341,342]

In 1991 the name "Cancer Family Syndrome" was deemed no longer practical, and "Hereditary Non-Polyposis Colorectal Cancer" (HNPCC) and "Lynch Syndrome" became the names of choice [343], although many researchers had already adopted these new terms. The word "non-polyposis" was used to differentiate HNPCC from the other known hereditary form of CRC, Familial Adenomatous Polyposis (FAP) (further discussed in section 1.8.2.1.), in which there can be hundreds or thousands of adenomatous polyps lining the colorectal mucosa, some of which will eventually progress to adenocarcinoma unless a prophylactic colectomy is performed. However, the term non-polyposis could be misleading, since it is not unusual for affected individuals to have a small number of polyps. Lynch Syndrome could be divided into Lynch I and Lynch II, to distinguish between the two variants of the syndrome recognized at the time: that which manifested only colorectal cancer and that which included extracolonic cancers, respectively. This distinction could generally only be made in families large enough to have multiple affected individuals.

The decision to officially use these new names instead of CFS came at a gathering of leading experts in the field of hereditary colorectal cancer in Amsterdam in August of 1990 [343]. The participants of this meeting recognized the lack of adequate studies on the epidemiology and natural history of the syndrome, and started the first large-scale

collaboration to address these issues. They called themselves the International Collaborative Group on Hereditary Non-Polyposis Colorectal Cancer (ICG-HNPCC). It was important that the families used in the research of HNPCC/Lynch Syndrome actually had the syndrome, and not a different form of familial or hereditary CRC. For this reason, strict criteria were agreed upon that all families must meet in order to be included in research carried out by the ICG-HNPCC; these criteria have become known as the Amsterdam Criteria I (ACI) (Table 1.11.). They were based on a strong family history and young age of diagnosis. Although they have been used as such, these criteria were not intended to be used clinically for the diagnosis of Lynch Syndrome as they were not sensitive enough for clinical application, but instead they were highly specific, which was necessary to identify families for use in research to better define the syndrome, and from which more sensitive clinical criteria could later be developed [341,343].

Table 1.11. Amsterdam Criteria I
(from [343])

There should be at least 3 relatives with CRC; all the following criteria should be present:
• One should be a first-degree relative of the other 2
• At least 2 successive generations should be affected
• At least 1 CRC should be diagnosed before age 50
• Familial adenomatous polyposis should be excluded
• Tumours should be verified by pathological examination

The ACI were revised by the ICG-HNPCC in 1999 to take into account extracolonic cancers associated with the syndrome; cancers of the endometrium, small bowel, ureter, and renal pelvis were included [341]. These revised criteria have become known as the Amsterdam Criteria II (ACII) (Table 1.12.). These particular extracolonic

cancers were included in the criteria because they are the most specific and carry the highest relative risk [341], however, other cancers are also associated with Lynch Syndrome. While these other cancers were not included in the criteria they were included in the ICG definition of the syndrome, which was published in the same article as ACII [341] (included in Table 1.10). This definition also includes features of associated colorectal adenomas, and some molecular and histopathological features of the tumours. These tumours often show loss of expression of the protein product of one of the mismatch repair genes: *MLH1*, *MSH2* or *MSH6* on immunohisochemistry. They also tend to be MSI-H.

Table 1.12. Amsterdam Criteria II
(from [341])

There should be at least 3 relatives with an HNPCC-associated cancer (CRC, cancer of the endometrium, small bowel, ureter, or renal pelvis): • One should be a first-degree relative of the other 2 • At least 2 consecutive generations should be affected • At least 1 should be diagnosed before age 50 • Familial adenomatous polyposis should be excluded in the CRC case(s) if any • Tumours should be verified by pathological examination

Since the Amsterdam Criteria were not intended for clinical use, it was necessary to develop criteria for clinical diagnosis of Lynch Syndrome. The strict nature of the Amsterdam Criteria meant that 20% of families in which germline mutations had been identified in the MMR genes responsible for the syndrome did not meet the criteria [344], and would be missed if the criteria were applied in a clinical diagnostic setting. In 1996 the National Cancer Institute hosted a workshop in Bethesda, Maryland with the purpose of "clarifying the role of genetics in the pathology of HNPCC"[344]. They examined the

histopathology of colorectal cancers, hoping to find features that would be unique to HNPCC/Lynch Syndrome tumours. The workshop participants concluded that while some histopathologic features were more common in colorectal cancers from patients with the syndrome than in those with sporadic CRC, no feature was specific enough to be used to distinguish the two. They discussed other ways of trying to identify the potential Lynch Syndrome patients that would be missed by the strict Amsterdam Criteria; these discussions lead to the development of the less stringent Bethesda guidelines (Table 1.13.) [344].

Table 1.13. Bethesda Guidelines
(from [344])

1. Individuals with cancer in families that meet the Amsterdam Criteria
2. Individuals with two HNPCC-related cancers, including synchronous and metachronous colorectal cancers or associated extracolonic cancers*
3. Individuals with colorectal cancer and a first-degree relative with colorectal cancer and/or HNPCC-related extracolonic cancer and/or a colorectal adenoma; one of the cancers diagnosed at age <45 y, and the adenoma diagnosed at age <40 y
4. Individuals with colorectal cancer or endometrial cancer diagnosed at age <45 y
5. Individuals with right-sided colorectal cancer with an undifferentiated pattern (solid/cribiform) on histopathology diagnosed at age <45 y†
6. Individuals with signet-ring-cell-type colorectal cancer diagnosed at <45 y‡
7. Individuals with adenomas diagnosed at age <40 y

*Endometrial, ovarian, gastric, hepatobiliary, or small-bowel cancer or transitional cell carcinoma of the renal pelvis or ureter.
†Solid/cribiform defined as poorly differentiated or undifferentiated carcinoma composed of irregular, solid sheets of large eosinophilic cells and containing small gland-like spaces.
‡Composed of >50% signet ring cells.

The Bethesda guidelines were not intended to diagnose Lynch Syndrome, but instead to triage patients in the clinical setting; to identify those at risk of Lynch

Syndrome so that they could undergo further testing which could ultimately make the diagnosis. They recognized that more than 90% of Lynch Syndrome tumours had high levels of microsatellite instability (MSI-H), a molecular feature resulting from MMR deficiency. The Bethesda guidelines were intended to identify tumours at risk of being MSI-H, so that they could go on for testing for microsatellite instability (MSI) status. A status of MSI-H is not, however, enough to diagnose Lynch Syndrome in the patient, since the majority of MSI-H tumours are sporadic. They proposed that those tumours that were MSI-H undergo further testing for germline mutations in the MMR genes. The Bethesda guidelines were published in the interval between the two versions of the Amsterdam Criteria, and they expanded upon the AC I to include the extracolonic cancers later included in the ACII, as well as other family history and age-related criteria, and criteria that included the histopathological features associated with Lynch Syndrome tumours.

There are many reasons why it is important to identify cases of Lynch Syndrome in a clinical setting. A patient with Lynch Syndrome is predisposed to the cancers associated with the syndrome, and therefore needs to undergo regular clinical screening to identify any further colorectal cancers as well as any of the associated extracolonic cancers. Screening by colonoscopy can identify adenomas and remove them before they become malignant. Identifying a cancer at an early stage decreases associated morbidity and mortality. Clinical screening and associated genetic counselling also needs to be initiated for at-risk family members to identify Lynch Syndrome-related cancers and increase survival. Without the initial diagnosis of Lynch Syndrome and subsequent genetic follow-up, the family members would often not know the extent of their risk, and

would not undergo all of the necessary screening. Lynch Syndrome CRCs are more often located in the proximal colon, and therefore screening strategies must account for this. For example, complete colonoscopy is preferred over sigmoidoscopy, which would only evaluate the distal portion of the bowel. Screening must also start earlier than for the general population, since the age of onset of CRC is generally 15 to 20 years younger for Lynch Syndrome than it is for sporadic CRC [341,342].

A diagnosis of Lynch Syndrome also influences cancer treatment. Cancers with mismatch repair deficiency undergo a different biological pathogenesis than MMR-proficient tumours and so behave differently. MMR-deficient tumours tend to react differently to chemotherapeutic agents; importantly, they can be resistant to fluorouracil [70,71,344], which is the basis of standard first-line adjuvant chemotherapy [345,346]. As well, if a patient is known to have Lynch Syndrome when a diagnosis of CRC is made, a different, more extensive surgical approach may be taken to minimize the risk of developing a subsequent CRC [347]. Approximately 35% of patients will develop a synchronous or metachronous CRC [341]. A diagnosis of Lynch Syndrome can also influence patient prognosis, as CRCs in this context have a higher overall survival rate than sporadic CRCs, even when matched for stage and sex [342].

While MSI status determination can help identify Lynch Syndrome, it is not practical to test all colorectal cancers. It would be expensive and time consuming to test all cases, especially considering the low pick-up rate. The Bethesda guidelines were intended to identify cases that were more likely to be MSI-H, to reduce the number of tumours that would have to be tested. The authors of the Bethesda guidelines expected

that its application would lead to MSI testing of only 15% to 20% of all CRC cases [344], yet still identify virtually all cases of Lynch Syndrome. These numbers were later shown to be optimistic, as the reported sensitivity was 96% and specificity only 27% for identifying MSI-H tumours in high risk populations [348]. While most cases of Lynch Syndrome were identified, many more than the estimated 15% to 20% of all CRC cases had to undergo MSI testing.

It is interesting to note that in the original Bethesda guidelines (Table 1.13.), the two criteria which included histopathologic features were written in such a way that the histological features themselves would never alone be the reason for meeting the criteria to go on for MSI testing. Criteria 5 and 6 include various histopathologic features common in Lynch Syndrome tumours, yet they impose an age restriction. Only tumours with these histopathological features which were diagnosed before age 45 years should go on for testing. These criteria are made redundant by criterion 4, which states that all CRCs diagnosed younger than age 45 years should go on for testing. While the authors recognized the utility of the histopathological features they did not effectively include them in the guidelines.

Six years after the National Cancer Institute meeting in Bethesda, a follow-up meeting was held to update and improve the original Bethesda guidelines. The outcome was the publication of the Revised Bethesda guidelines (RBG) in 2004 (Table 1.14.)[72]. The RBG differ from the original in that there are fewer criteria and overall they are less stringent, to increase sensitivity. The first criterion in the original guidelines was for families to meet the Amsterdam criteria. The equivalent criterion in the revised

guidelines includes second degree relatives as well as first degree, includes Lynch Syndrome-related cancers instead of just CRC, and removes the age restriction of one cancer having to be diagnosed younger than age 50 years. Similarly, the age restrictions were revised for other criteria, to now include cancers diagnosed before age 50 years instead of age 45 years, and for cancers with histology suggestive of Lynch Syndrome diagnosed before age 60 years instead of age 45 years. The RBG also aimed to be more specific than the original Bethesda guidelines, in that the criterion for the presence of adenomas was removed, and the definition of Lynch Syndrome-suggestive histology was changed to include mucinous tumours and those with a host-lymphocytic response.

Table 1.14 Revised Bethesda Guidelines
(from [72])

1. Colorectal cancer diagnosed in a patient who is less than 50 years of age.
2. Presence of synchronous, metachronous colorectal, or other HNPCC-associated tumours,* regardless of age.
3. Colorectal cancer with the MSI-H† histology‡ diagnosed in a patient who is less than 60 years of age.§
4. Colorectal cancer diagnosed in one or more first-degree relatives with an HNPCC-related tumour, with one of the cancers being diagnosed under age 50 years.
5. Colorectal cancer diagnosed in two or more first- or second-degree relatives with HNPCC-related tumours, regardless of age.

*Hereditary nonpolyposis colorectal cancer (HNPCC)-related tumours include colorectal, endometrial, stomach, ovarian, pancreas, ureter and renal pelvis, biliary tract, and brain (usually glioblastoma as seen in Turcot syndrome) tumours, sebaceous gland adenomas and keratoacanthomas in Muir-Torre syndrome, and carcinoma of the small bowel.
†MSI-H = microsatellite instability-high in tumours refers to changes in two or more of the five National Cancer Institute-recommended panels of microsatellite markers.
‡Presence of tumour infiltrating lymphocytes, Crohn's-like lymphocytic reaction, mucinous/signet-ring differentiation, or medullary growth pattern.
§There was no consensus among the Workshop participants on whether to include the age criteria; participants voted to keep less than 60 years of age in the guidelines.

The RBG have not been as clinically useful as hoped for the identification of MSI-H tumours and Lynch Syndrome. As a screening test it is important to have a high sensitivity, in other words, to miss very few, if any of the tumours it strives to identify. It is also helpful to have a high specificity, meaning that very few tumours would be identified for further testing. Ideally, a minimum number of tumours would be identified by the RBG for further testing, and a maximum of mutation-carriers would be identified. The RBG have published sensitivities for identifying MMR mutation-carrying tumours in the range of 50% [349] to 95% [109], with most publications reporting close to 75% [6,350,351]. While the RBG misses an average of one in four tumours with germline MMR mutations, it does so even though from 26% to 42% of tumours are identified to go on for further testing [349,351,352]. This is discussed further in section 4.1.1.

1.8.1.2 Familial Colorectal Cancer Type X

As the definition of Lynch Syndrome changed from a clinical description to a molecular diagnosis based on the finding of a hereditary deficiency in the DNA MMR system, a dichotomy began to emerge. MMR defects could only be identified in 60% of families who met the strict clinical Amsterdam Criteria I [353,354]. A large 2005 study by Lindor and colleagues [353] examined the cancer risks of ACI families, and compared the cancer rates in those families with a MMR deficiency to those without such a deficiency. After removing from analysis the triad of patients used to define ACI status in each family, they discovered a significant difference in the risks and types of cancers in the

two groups. Members of families with MMR-defects had a standardized incidence ratio (SIR) of 6.1 for CRC, whereas families without a MMR defect had a SIR of only 2.3 for CRC. In other words, individuals in families with MMR defects had a much higher risk of developing CRC.[353]

There was also an observed difference in the risk of extracolonic cancers. Families with MMR defects had significantly increased risk of developing the extracolonic cancers associated with Lynch Syndrome, including cancers of the endometrium, stomach, urinary tract, ovary, and small intestine. Families without MMR defects did not have an increased risk of these or any other extracolonic cancer. Another difference between the families was the average age of diagnosis of CRC. The average age of diagnosis was 48.7 years for those with MMR defects, whereas those without these defects were diagnosed much older, at an average of 60.7 years of age.[353]

Lindor and colleagues concluded from their findings that families who meet the ACI criteria for Lynch Syndrome, but who did not have germline MMR defects, did not have Lynch Syndrome, but instead had a less severe form of familial cancer. They suggested that the name Lynch Syndrome should apply only to those families with germline MMR defects, and not to those families without such defects; for them they proposed the name Familial Colorectal Cancer Type X (FCCTX). They stressed the importance of their findings for counselling and screening. While families who meet ACI and have MMR defects should follow with the regular screening protocol for Lynch Syndrome, FCCTX families without MMR defects do not need screening for extracolonic cancers, nor do they need to start colonoscopic screening at such an early age.

Prophylactic colectomy and hysterectomy are also not warranted in this group, while they are often recommended in the context of Lynch Syndrome.[353]

In 2009, Lindor published an update and review of the literature of FCCTX [355]. The gap between Lynch Syndrome and FCCTX had widened, with more differences recognized, including the propensity for left-sided, distal CRCs in FCCTX [354,356,357], and the evidence to suggest that there was a slower rate of malignant transformation in FCCTX than in Lynch Syndrome [356]. They summarized the differences between these two syndromes in table format, adapted and included here as Table 1.15.

Table 1.15. Comparison of two syndromes in which families fulfill Amsterdam Criteria I: Lynch syndrome and Familial Colorectal Cancer Type X

	Lynch Syndrome	**FCCTX**
Colorectal		
Cancer risk	Very high	Modestly increased
Age of onset	~45 yr	50s-60s
Usual location	Proximal colon	Distal colon
Polyps	Few	More
Malignant transformation	Rapid	Less rapid
Other cancers		
Endometrial risk	Very high risk	Risk not significantly increased
Other cancer sites	Many others	None known
Germline MMR mutations	Mutations found	No mutations found
CRC tumour testing	MSI-H	MSS (by definition)
CRC tumour IHC	Loss of MMR protein expression	Normal MMR expression

Reprinted with permission from Elsevier: *Surgical Oncology Clinics of North America*, volume 18, pages 637-645, copyright 2009 [355]

There are reports which show that FCCTX tumours have distinct molecular features from both Lynch Syndrome and sporadic CRC, suggesting that there are novel genes and pathways of carcinogenesis associated with FCCTX [358,359]. Abdel-Rahman

and colleagues [358] reported that the Wnt signalling pathway was activated less frequently in FCCTX tumours, as there were significantly fewer mutations in the *CTNNB1* gene, and its protein product, β-catenin, was less often aberrantly localized in the nucleus in FCCTX tumours when compared to Lynch Syndrome tumours. This decrease in aberrant β-catenin localization in FCCTX tumours was also significantly different from sporadic colorectal tumours. Significant differences were also seen between FCCTX and both Lynch Syndrome and sporadic CRC when considering alterations in p53, and the presence of chromosomal instability. Although not statistically significant, they also found a lower rate of *KRAS* mutations in FCCTX tumours than in either Lynch Syndrome or sporadic tumours. [358]

Similarly, Sánchez-de-Abajo and colleagues [359] reported molecular differences between FCCTX CRCs and other colorectal tumours. While the rate of *KRAS* mutations in FCCTX tumours was similar to both sporadic tumours and those associated with Lynch Syndrome, the type of mutations and their location within the gene much more closely resembled those seen in sporadic tumours, especially MSS sporadic tumours. Sánchez-de-Abajo and colleagues also reported that activation of the Wnt signalling pathway was not common in FCCTX tumours . The rate of *APC* mutations was significantly lower than that of sporadic tumours, and somewhat lower than that of Lynch Syndrome tumours. [359]

These differences between Lynch Syndrome and FCCTX strongly suggest different pathways of carcinogenesis for the two syndromes. However, a single specific pathway of carcinogenesis in FCCTX has not been identified, and is unlikely to exist.

Lindor and colleagues stated that FCCTX is likely due to a heterogenous grouping of causes. They proposed three possible explanations for the strong familial clustering in FCCTX: (1) cancer aggregation by chance alone, (2) some aggregation due to shared lifestyle factors, and (3) some yet-to-be-defined genetic syndromes [353,355]. It is possible that some patients labelled as having FCCTX actually have Lynch Syndrome, with a mutation in the *MSH6* gene, since mutations in this gene do not always cause microsatellite instability. There may also be examples of phenocopies within Lynch Syndrome families. If the individual whose tumour was tested for MSI was a sporadic phenocopy within a family with Lynch Syndrome, a diagnosis of FCCTX may be applied incorrectly. The remaining cases of FCCTX are not likely to be due to a high risk monogenic trait, as such a loci has not been identified; a polygenic model is much more likely. Studies are underway to try to identify some of the genes that may be involved. It has been estimated that approximately one third of the familial risk of colorectal cancer is not accounted for by mutations in the known genes [360].

1.8.2 Polyposis Syndromes

1.8.2.1 Familial Adenomatous Polyposis

Germline mutations in the adenomatous polyposis coli (*APC*) tumour suppressor gene are associated with familial adenomatous polyposis (FAP), an autosomal dominant

disorder characterized by hundreds or thousands of colorectal adenomatous polyps. These adenomas usually have a tubular histology, rarely tubulovillous or villous, and typically carpet the mucosal surface. They occur throughout the colon and rectum, but tend to occur with the greatest density distally, within the sigmoid colon and rectum. An exception to this is found in about one third of FAP families, who have the greatest density in the proximal colon, sometimes with rectal sparing. The distribution of associated carcinomas follows that of the adenomas. FAP is the second most common inherited CRC syndrome, after Lynch Syndrome, and has a prevalence of approximately one in 10,000 individuals [361]. It represents less than 1% of all cases of CRC [25].

Although FAP is an inherited disorder, up to half of all patients are solitary cases representing new mutations [25]. The polyps often start to develop in early adolescence, and increase in number and size with age. Although the majority of these polyps will not progress to adenocarcinoma, the potential for progression coupled with the large number of polyps virtually ensures the development of adenocarcinoma if the colon is not prophylacticly removed. Even after colectomy, regular screening of the remaining rectum and ileal pouch is warranted. The lifetime risk of CRC associated with FAP is 100%, with the average age of CRC diagnosis at 39 years. Approximately 7% of FAP patients will develop CRC by age 21, and 95% by age 50 [361].

Extracolonic manifestations are frequently encountered in FAP, and can assist in a clinical diagnosis, as they are not generally seen in the other polyposis syndromes. Polyps of the upper GI system are common in FAP. Gastric adenomas and fundic gland polyps are common, but rarely progress to adenocarcinoma. Adenomas of the duodenum

are seen in almost all patients with FAP, and these also have a tendency to progress to carcinoma [362]. Patients with FAP have a lifetime risk of up to 10% of developing duodenal cancer [362], making it the second most common malignancy in this syndrome, and one of the leading causes of death for patients who have undergone prophylactic proctocolectomy [363].

Risk of hepatoblastoma and carcinoma of the biliary tree are increased in individuals with FAP, as are risk of dental abnormalities and endostoses of the mandible. Mesodermal fibromatosis, known as desmoids tumours, can arise in retroperitoneal tissue or the abdominal wall, often after trauma such as previous surgery at the same site. Desmoid tumours can expand to surround and cause damage to nearby structures including major blood vessels, ureters, and the intestines. Patients with FAP also commonly have epidermal cysts, often on the face or dorsal surface of the hands. Up to 80% of patients have multiple patches of congenital hypertrophy of retinal pigment epithelium (CHRPE), which are generally asymptomatic.[25]

The diagnostic criteria of FAP, as defined by the World Health Organization [25], are provided in Table 1.16.

Table 1.16. Clinical diagnostic criteria for Familial Adenomatous Polyposis (from [25])

1. 100 or more colorectal adenomas, or
2. germline mutation of the APC gene, or
3. family history of FAP and at least one of the following: epidermoid cysts, osteomas, desmoids tumour.

Subtypes of FAP have been defined based on clinical findings. Attenuated FAP (aFAP) is a less severe form of the syndrome, with patients diagnosed later in life, and with fewer adenomas than with traditional FAP. Diagnosis tends to be in patients older than 40 or 50 years of age, and the number of adenomas is generally more than ten, but less than 100. The adenomas tend to be quite small, rarely reaching 1cm in diameter until adulthood. They can be so small and few in numbers that they are difficult to identify on rigid sigmoidoscipy [25]. Contrary to traditional FAP, the adenomas found in aFAP tend to be located proximal to the splenic flexure and may be flat rather than polypoid [364-367]. While gastric fundus polyps are seen in aFAP, ocular CHRPE lesions and desmoids tumours are generally not observed [364,368]. Due to the smaller number of adenomatous polyps and the tendency for them to be right-sided, there is considerable clinical overlap between aFAP and Lynch Syndrome. There is also clinical overlap between aFAP and MUTYH-associated polyposis (MUTYH), described further in the following section, 1.8.2.2. Biallelic *MUTYH* mutations have been found in more than 25% of aFAP patients for whom germline *APC* mutations could not be identified [369].

Gardner Syndrome is a variant of FAP associated with colorectal adenomas and carcinoma as well as manifestation of the extracolonic features described above, especially epidermoid cysts, osteomas, dental anomalies, and desmoids tumours [25]. Gardner Syndrome was traditionally considered to be a unique disease entity, however, it is now recognized as a variant of FAP. While many patients with FAP develop extracolonic manifestations, in Gardner Syndrome the extracolonic manifestations are more prominent, and tend to develop years before the onset of colonic adenomas [370].

A third variant of FAP is Turcot Syndrome. Turcot Syndrome is defined by the presence of multiple colorectal adenomas and associated risk of CRC, along with a primary cancer of the central nervous system, most commonly medulloblastoma. A 1995 study by Hamilton and colleagues [371] identified two genetic causes of Turcot Syndrome: germline mutations in the *APC* gene, or in MMR genes associated with Lynch Syndrome. The phenotypes are somewhat different based on the different germline defects, and so they are considered different disease subtypes. Turcot Syndrome associated with germline mutations in the *APC* gene is therefore also known as brain tumour-polyposis syndrome 2 [372].

The first clue in the identification of the gene responsible for FAP was published in 1986 [373]. Herrera and colleagues reported on a male patient with suspected Gardner Syndrome who had an interstitial deletion of 5q [373]. Five years later the *APC* gene was associated with FAP in three concurrent publications by two independent research groups [374-376]. The *APC* gene and its functions have previously been described here in section 1.7.1. Briefly, it is a multi-domain, 2843 reside protein which resides in the cytoplasm. The protein product, APC, plays in integral role in tumour suppression by antagonizing the Wnt signalling pathway through its interactions with the β-catenin protein. Binding of β-catenin by APC ultimately leads to degradation of β-catenin, thereby preventing the aberrant activation of Wnt. APC is also involved in cell migration, cell adhesion, chromosome segregation, and apoptosis [189].

More than 98% of *APC* mutations are either frameshift or nonsense mutations that lead to a truncated protein product [377]. There have been numerous reports of genotype-

phenotype correlations, in which the location of the *APC* mutation is associated with a specific clinical manifestation of disease. Ocular CHRPE may be a presentation of disease only if the truncating mutation is located between codons 463 and 1387 [378]. Other extracolonic manifestations including desmoids tumours and mandibular lesions are common when the truncating mutation lies between codons 1403 and 1578, however, patients with mutations in this region do not generally develop CHRPE [379]. The attenuated form of FAP is associated with mutations found in two clusters within the first third of the gene, as well as the entire last third of the gene. More specifically, there are three regions associated with aFAP, they are: (1) exons 3 and 4, (2) within a region of exon 9 which can be alternatively spliced, and (3) the 3' end of the gene, beyond codon 1595 [367,369]. Patients with aFAP-associated mutations in the third region tend to have desmoids tumours, but do not have CHRPE [377]. A mutational hotspot exists between codons 1250 and 1464, and has been termed the mutation cluster region [380]. The functional domains of the APC protein, along with the regions of genotype-phenotype associations are illustrated in Figure 1.11.

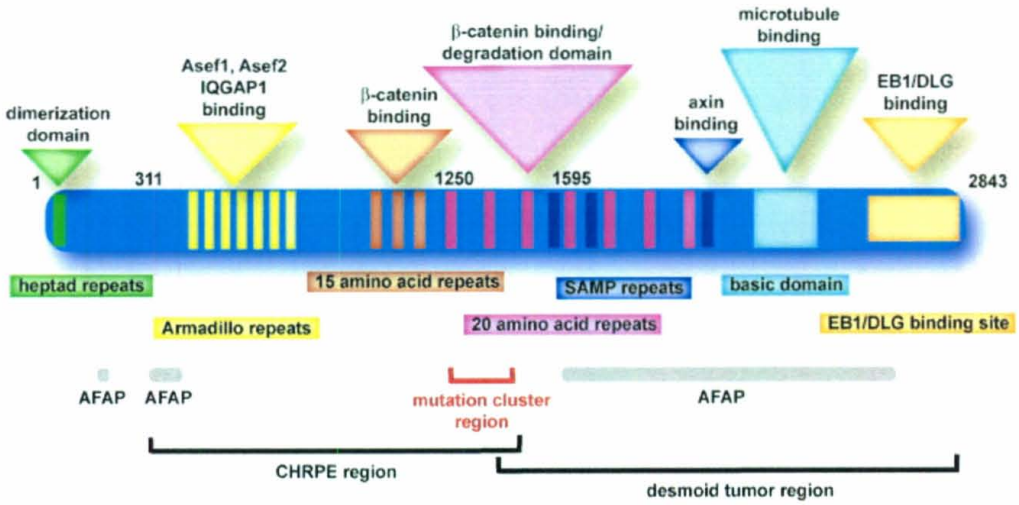


Figure 1.11. The functional domains of the APC protein and regions associated with genotype-phenotype correlations.
Reprinted by permission from Elsevier: *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, volume 693, pages 32-45, copyright 2010 [369]

While it can be useful to apply genotype-phenotype correlations to the clinical management of patients, these correlations cannot be relied on completely. The increased risk of desmoids tumours associated with mutations in the 3' end of the gene can guide a delay in surgical intervention, mutations outside of the CHRPE region would indicate no need for ophthalmological screening, and regions associated with early manifestation of disease should be associated with early screening. However, along with the allelic heterogeneity found in FAP, there is also considerable clinical heterogeneity. There are many examples of clinical heterogeneity even within family members with the same germline mutation [213]. For example, only a small number of people within a FAP kindred will develop the more rare manifestations, including thyroid cancer, hepatoblastoma, and medulloblastoma [213,371]. As well, some patients will develop features of Gardner

Syndrome, including mandibular osteomas and desmoids tumours, while other patients with the same mutation will not [375]. There is also variability within families considering age of development of adenomas and carcinomas, as well as number of adenomas.

1.8.2.2 *MUTYH* Associated Polyposis

MutY homolog (MUTYH) is a protein involved in the base excision (BER) pathway, which functions primarily to repair oxidative damage. Germline mutations in the *MUTYH* gene have been associated with an increased risk of adenomatous polyps and CRC [381-385], which progress through the CIN pathway of carcinogenesis. This disorder is known as *MUTYH*-Associated Polyposis (MAP), and inheritance generally follows an autosomal recessive pattern. The predominant lesions are small, mildly dysplastic tubular and tubulovillous adenomas, and hyperplastic polyps are also identified [383]. Due to the presence of multiple adenomas and sometimes multiple hyperplastic polyps, MAP can be difficult to differentiate clinically from FAP, attenuated FAP, and Hyperplastic Polyposis Syndrome (HPS – described in the following section) [383-386].

The association between CRC and the *MUTYH* gene was first described in 2002 [381]. Al-Tassan and colleagues reported a British family with three siblings affected with multiple colorectal adenomas and carcinomas, but who lacked a germline mutation in the *APC* gene associated with FAP. They did note, however, that the tumours of these siblings harboured a number of somatic inactivating mutations in APC. These mutations

included a significantly higher proportion of G:C to T:A transversions than would normally be expected. This type of transversion is characteristic of oxidative damage, so Al-Tassan and colleagues studied three key genes involved in oxidative repair – *OGG1*, *MUTYH*, and *MTH*. They identified biallelic germline *MUTYH* mutations in the affected siblings [381].

The same group subsequently tested 21 more unrelated British patients who had multiple adenomas, and identified biallelic *MUTYH* mutations in seven (six of whom had CRC) [382]. Other groups soon started to report similar findings. A 2003 publication by Sieber and colleagues [383] reported their findings of *MUTYH* mutations in two cohorts of patients: those clinically diagnosed with multiple colorectal adenomas, and those who had classical FAP, but without germline *APC* mutations. They identified some patients with biallelic *MUTYH* mutations in both groups. Approximately one third of patients with 15 to 100 adenomas harboured biallelic *MUTYH* mutations, as well as approximately 8% of patients with classic FAP. Sieber's group recommended that *MUTYH* testing be undertaken in any patients with 15 or more adenomas, especially if no germline *APC* mutation was identified and the family history was compatible with autosomal recessive inheritance [383]. Sieber's group also reported the existence of extracolonic manifestations, in the form of duodenal polyposis, in some patients with biallelic *MUTYH* mutations.

The autosomal recessive inheritance pattern reported initially has since been challenged, and appears to be more complex. In 2006, Jenkins and colleagues from the Ontario Familial Colorectal Cancer Registry (OFCCR) conducted a family-based study whereby they estimated the CRC risk of 300 first-degree relatives of 39 patients with

CRC who were carriers of either monoallelic or biallelic *MUTYH* mutations [387]. They reported that while biallelic carriers had a 50-fold increase in risk of CRC, monoallelic carriers also had an increase in risk, although it was much smaller, at 3-fold over the population level. This suggested a co-dominant model of inheritance rather than a recessive model. This means that homozygotes have a higher risk than heterozygotes, and that they, in turn, have a higher risk than wild type.

Jenkins and colleagues also included a meta-analysis in their publication [387]. They examined seven previous studies which had reported CRC risk for carriers of monoallelic *MUTYH* mutations. The data from the meta-analysis suggested a 40% increase in risk of CRC for monoallelic mutations carriers. A larger meta-analysis has since been published [388], which examined a total of 20 565 cases and 15 524 controls from a total of nine both published and unpublished studies. They concluded that while there is an increased risk of cancer in mono-allelic mutation carriers, the risk is marginally significant when including all *MUTYH* mutations, but stronger, at approximately 35% when including only the common Y179C mutation.

More work is needed to definitively establish risks for *MUTYH*-associated polyposis, not only in terms of biallelic and monoallelic mutation carriers, but also in terms of any differences in risk that may be due to the sex and age of patients.

1.8.2.3 Hyperplastic Polyposis Syndrome

Hyperplastic Polyposis Syndrome (HPS) is a model for the serrated pathway of carcinogenesis, in the same way that FAP is a model for the traditional CIN, adenoma to carcinoma pathway of CRC. It was first described by Williams and colleagues in 1980 [389], who reported seven patients with sufficient hyperplastic polyps that they had been incorrectly diagnosed with FAP. They reported this condition to be more common in young males, and that it was not associated with an increased risk of CRC. HPS has since been redefined in the World Health Organization Classification of Tumours by Jass and Burt [25] using the criteria outlined in Table 1.17. While the definition refers only to hyperplastic polyps, many types of serrated lesions are identified in patients with HPS, as well as some traditional adenomas [386,390-392]. The phenotype overlaps somewhat with *MUTYH* polyposis [393,394].

Table 1.17. Definition of Hyperplastic Polyposis
(from [25])

1. At least five histologically diagnosed hyperplastic polyps proximal to the sigmoid colon, of which two are greater than 10 mm in diameter, or
2. Any number of hyperplastic polyps occurring proximal to the sigmoid colon in an individual who has a first-degree relative with hyperplastic polyposis, or
3. Greater than 30* hyperplastic poylps of any size but distributed throughout the colon.

*Some publications use 20 instead of 30 polyps in this criterion [314,386,390,392].

While this rare syndrome was not initially thought to confer a risk of colorectal cancer, it is now recognized that it imparts a substantial risk [391]. These cancers are

associated with somatic *BRAF* p.V600E mutations and CIMP, and tend to be located in the proximal colon, have mucinous histology and poor differentiation[260,261,311]. They can also be associated with a family history of CRC [183,260,395]. While these tumours can be either MSI-H, MSI-L or MSS, colorectal cancer in HPS is more often MSS than MSI [386]. Also, patients with MSS CRCs harbouring the *BRAF* p.V600E mutation had the highest rates of positive family history in one large North American study [260]. Unlike general, sporadic CIMP CRCs, those occurring in the context of HPS do not have an increased association with female sex, nor with increased age [263,396].

Familial aggregation of HPS has more recently been called by other names, including Familial Serrated Neoplasia and Jass Syndrome, after the late pathologist who first defined the syndrome [397]. The mode of inheritance has not yet been completely determined, however, based on the diagnosis of HPS in sibships of consanguineous families [386], as well as in identical twins, it has been proposed that an autosomal recessive or co-dominant mode is the most likely [396]. While HPS is relatively rare, individuals carrying one putative co-dominant allele would be much more common. Such individuals may have a less severe phenotype, but with an increased risk over the general population of developing serrated polyps, and possibly CRC. It has been proposed that these individuals may account for a portion of CRCs which are thought to be sporadic [396]. Genome-wide linkage analysis has recently been utilized to identify a putative locus for HPS at 2q32.2-q33.3 [397].

1.8.3 Hamartomatous Polyposis Syndromes

There is a rare group of hereditary disorders characterized by hamartomatous polyps of the colon and rectum. Hamartomatous polyps are masses of normal, mature cells, indigenous to the area in which they are found, which are growing at the same rate as surrounding tissues, but tend to be organized in a different way. While these lesions are inherently benign, some of these syndromes are associated with an increased risk of malignancy, both of the colon and rectum as well as extracolonic sites. Together they account for less than one percent of all hereditary colorectal cancers [398]. They include Juvenile Polyposis Syndrome, PTEN Hamartoma Syndrome (which includes both Cowden Syndrome and Bannayan-Riley-Ruvalcaba Syndrome), as well as Peutz-Jeghers Syndrome. They are described in the following sections.

1.8.3.1 Juvenile Polyposis Syndrome

Juvenile Polyposis Syndrome can be inherited as an autosomal dominant trait, although up to half of all patients do not have a family history of the syndrome [399]. Patients typically develop between 50 and 200 polyps throughout the colon and rectum, as well as polyps in the small intestine and stomach [25]. The polyps were traditionally referred to as hamartomatous, however, they are now recognized to impart some risk of malignancy and are generally referred to as juvenile polyps. The majority of polyps are small, however, up to 20% of polyps can be large, up to 5cm in diameter, with

multilobulated heads [400]. The lifetime risk of CRC is up to 40%, and the risk of upper gastrointestinal carcinoma is up to 15% [401]. There is also an increased risk of pancreatic cancer. Juvenile polyposis is also associated with a number of extracolonic manifestations, including congenital anomalies, which are seen in up to 15% of patients [399]. These can include anomalies of the heart, central nervous system, gastrointestinal tract, and genitourinary system [399].

Juvenile polyposis is rare, approximately ten-fold less common than FAP [25]. Two-thirds of patients are diagnosed within the first two decades of life, with a mean age at diagnosis of 18 years [399]. The most severe form of juvenile polyposis presents in infancy, and is associated with diarrhea, anemia, protein-losing enteropathy, bleeding, and rectal prolaspse. Patients with this form of juvenile polyposis rarely survive past two years of age.

Mutations in the *SMAD4* tumour suppressor gene have been implicated in some cases [402,403]; *SMAD4* is located on chromosome 18q21.1 [404], a region often deleted on CRC. Other genes implicated in this syndrome include *PTEN* and *BMPR1α* [405], all of which are members of the transforming growth factor β superfamily of proteins.

1.8.3.2 Cowden Syndrome

Cowden Syndrome is not associated with a clear increase in CRC risk, however, non-malignant hamartomatous polyps of the gastrointestinal tract are a common manifestation. This syndrome is best characterized by cutaneous lesions which occur in

99% of patients, including trichilemmoma, which is a benign tumour of the hair follicle infundibulum, and mucocutaneous papules [406]. Patients can also develop hamartomas of the breast, thyroid, skin, and central nervous system. There is an increased risk of breast, uterus, and non-medullary thyroid cancer [25,405], with breast cancer risk in affected women reported to be as high as 50%, and thyroid and endometrial cancer risks each of 10% [398]. It is an autosomal dominant disorder caused by mutations in the *PTEN* gene, with mutations occurring in all exons except 1, 4, and 9 [398].

1.8.3.3 Bannayan-Riley-Ruvalcaba Syndrome

Like Cowden syndrome, Bannayan-Riley-Ruvalcaba syndrome (BRRS) is also caused by mutations in the *PTEN* gene, and inheritance is in an autosomal dominant pattern. Mutations causing BRRS have been found in all exons with the exception of 1, 4, and 9, similar to Cowden Syndrome, however, in BRRS they are preferentially located in exons 6 and 7 of the *PTEN* gene, and can include balanced translocations and deletions [406]. The two syndromes are collectively called PTEN Hamartoma Tumour Syndrome, and families have been identified which exhibit both syndromes, suggesting that they may be different manifestations of one single syndrome [406]. Both syndromes include hamartomas of the colon and rectum, however, are best characterized by extracolonic manifestations. Similar to Cowden syndrome, BRRS is not associated with a clear increase in CRC risk. The most common manifestations of BRRS include macrocephaly, developmental delays, pigmented speckling of the penis, and multiple lipomas [398,407].

Hamartomatous polyps of the intestine occur in approximately 45% of patients [407]. Due to the rarity of the syndrome and paucity of publications describing it, the risks of cancer are unclear; however, cases of breast and endometrial cancers have been reported [406].

1.8.3.4 Peutz-Jeghers Syndrome

Like the other hamartomatous polyp syndromes, Peutz-Jeghers Syndrome is inherited in an autosomal dominant manner. It is characterized by mucocutaneous melanin pigmentation and hamartomatous polyps throughout the gastrointestinal tract, but predominating in the small intestine. The polyps tend to be pedunculated with a long stalk, range from 0.5cm to 5cm in diameter, and have a characteristic histological appearance with a central core of smooth muscle that has tree-like branching, which separates the glands into multiple lobules [25]. Pigmentation is seen in the vermillion border of the lips in more than 95% of patients, and in the buccal mucosa of 80%; it also occurs on the hands, feet, genitals, and around the eyes and nose [398]. Most patients are diagnosed in the first to third decade of life, often presenting with abdominal pain due to intussusception, but also presenting with anemia and related symptoms from gastrointestinal blood loss [398]. There is a nine- to 13-fold increased risk of developing cancer over the general population. This increased risk includes both gastrointestinal and non-gastrointestinal cancers [405], such as colorectal, gastric, small intestinal, esophageal, and pancreatic cancers, as well as cancers of the ovary, uterine cervix, testis, and breast

[25,398]. Mutations in the *LKB1* gene, also known as *STK11*, are responsible for up to 80% of cases [408].

1.8.4 Other Syndromes and Variants Associated with Colorectal Cancer

There are numerous other syndromes, genes, and loci associated with an increased risk of developing CRC. While I will not describe all of them here, I will briefly mention some. CRC is listed as a possible manifestation of Familial Breast and Ovarian Cancer Syndrome. Individuals with germline mutations in the ovarian cancer cluster region of exon 11 of the *BRCA2* gene have a relative risk of 3.4 of developing CRC, compared to the normal population [409]. Specific mutations in the *CHEK2* protein kinase gene are associated with an increased risk of CRC [410], as are mutations in *AXIN2*, a gene whose protein product plays a role in the Wnt signalling pathway [411]. Muir-Torre is a syndrome associated with both sebaceous skin tumours and internal malignancies including CRC. It is a phenotypic variant of Lynch Syndrome, with causative mutations identified in the MMR genes [412,413]. Large population-based studies have found and association between specific variants at 9p24 and 8q24 and an increased risk of developing CRC [414,415]. Recently, linkage to 9q22 has been confirmed in familial CRC [416], and both inactivating germline and somatic mutations have been identified in the *GALNT12* gene [417], which resides at this locus. These studies demonstrate that common variants may play an important role in the risk of CRC in the general population.

With so many syndromes associated with CRC, as well as variants in other genes and loci associated with an increased risk of CRC, it becomes obvious how CRC is as common as it is, and it leads to the conclusion that screening for this relatively common, yet heterogeneous disease must extend beyond members of families with recognized CRC syndromes. Clinical screening is especially important, given that early recognition and treatment can lead to a substantially better patient outcome [418].

1.9 Clinical Screening for Colorectal Cancer

Early detection of colorectal cancer results in improved patient prognosis, with decreased morbidity and mortality [418]. A number of different screening methods are available, which vary by sensitivity and specificity of identifying polyps and CRC, and also by their invasiveness, availability, the time required to perform the test, and the types of qualifications required of the health care professional who administers the test.

1.9.1 Digital Rectal Exam

The simplest, quickest, and least expensive method for CRC screening is the digital rectal exam. Bowel cleansing is not required, and the procedure is quick and usually painless. A physician inserts a lubricated, gloved finger into the patient's rectum, and feels for masses or abnormal areas. The digital rectal exam is often incorporated into

a routine adult physical examination by general practitioners, and is recommended as an annual test for men over the age of 50, where it is also used to detect abnormalities in the prostate. Of course, the test is limited to the reach of the physician, and can only identify masses of the distal rectum or anus.

1.9.2 Faecal Occult Blood Test

The faecal occult blood test (FOBT) is a quick, simple test that can be performed by a general practitioner during a routine visit, or even by the patient themselves in their own home. FOBT can detect blood in the stool, through smearing a small amount of stool onto a test card. There are currently two general types of FOBT that use different methods to identify the haemoglobin in blood. The more commonly used guaiac FOBTs detect the haeme in haemoglobin. The test is based on the ability of haeme to release oxygen from hydrogen peroxide, which then reacts with the guaiac to form a blue dye, indicating a positive test result. The other type of FOBT is immunochemical. It uses monoclonal or polyclonal antibodies raised against the globin in haemoglobin, and can detect intact or partially degraded haemoglobin [419].

The sensitivity and specificity of FOBT in identifying CRC is limited by numerous possible causes of false negative and false positive results. False negative results can occur if the patient has taken high doses of vitamin C. False positive results can result from the presence of blood from other sources, including upper GI lesions, upper GI bleeding as a result of aspirin or nonsteroidal anti-inflammatory drug use, other

lower GI lesions including haemorrhoids, and from dietary sources. Haeme ingested in the diet through the consumption of red meat can lead to a false positive test result, and plant peroxidises found in certain fruits and vegetables can lead to a false positive result when using the guaiac-based FOBT. Consumption of iron supplements can also lead to false positive results. For these reasons, there are restrictions imposed upon patients for the three to five days prior to a test, which include abstaining from aspirin and other NSAIDs, vitamin C, iron supplements, meat and fish, as well as some raw vegetables [420].

Advantages of FOBT include that it is relatively noninvasive when compared to other CRC screening methods, and stool samples can be collected by the patient at home. This leads to greater patient compliance and use of the test, and therefore more patients being screened for CRC than if such a noninvasive test were not available [420]. As well, thorough cleansing of the colon to remove faecal material is not required, and there is no risk of bowel perforation. Some disadvantages of FOBT include the high false positive and false negative results, which result in the failure to detect most polyps and some cancers [421], as well as the dietary restrictions required. Additional procedures including colonoscopy are required in the case of a positive FOBT result.

Publications from 2005 [420] and 2007 [419] reviewed articles reporting the accuracy of FOBTs. Both reviews reported finding great variations, based on the commercial test used, the number of tests carried out per patient, the type of study (cohort or case-control), the type of FOBT (guaiac or immunochemical), and the clinical screening test chosen as gold-standard for comparison purposes. Sensitivity of FOBTs can be increased with serial testing, but is reported to range between 9% and 65% [421]. Specificity of

FOBTs is generally relatively high, ranging from 90% to 98% [420]. However, these high values may be misleading, since only about one in ten patients with a positive FOBT has a detectable CRC [420].

1.9.3 Faecal DNA Testing

This test is similar to the FOBT, in that a stool sample is collected and used for analysis. Rather than testing for blood, however, faecal DNA testing involves using tests to identify abnormal DNA from a carcinoma, which may be shed into the stool. The advantages and disadvantages of the test are similar to those of the FOBT. The samples may be collected by the patient in their home, it does not require bowel cleansing, and there is no risk to injury of the colon. However, unlike guaiac FOBT there are no dietary or medication restrictions prior to the test. If an abnormality is identified, follow-up with a more invasive test, such as colonoscopy is often required.

This is still a new test, it is more expensive than other stool tests, and is not widely available. It is not included in the current Canadian guidelines for colorectal screening [422,423]. One published estimate of the efficacy of the test is that it has a similar sensitivity and specificity to guaiac FOBT, but costs approximately 80 times more [424]. When compared to another commercially available guaiac FOBT, faecel DNA testing had a similar specificity, but a much higher sensitivity, although it still missed approximately half of the invasive carcinomas in the study cohort [425]. With improving DNA analysis technology and decreasing cost, it could become a more practical test [424]; some

researchers have suggested it be used in conjunction with other stool-based detection methods, such as immunochemical FOBT, to increase overall sensitivity [426].

1.9.4 Sigmoidoscopy

Flexible sigmoidoscopy enables visualization of the rectum and distal colon by rectal insertion of a lighted tube, a sigmoidoscope. It can be performed by general practitioners, but is often performed by gastroenterologists or surgeons. If polyps or other suspicious lesions are identified during the procedure they can be biopsied (and sometimes removed) for pathological examination, however, such findings often lead to referral for full colonoscopy to examine the remainder of the colon. Flexible sigmoidoscopy is a relatively quick procedure, with few complications, although there is a small risk of bowel tearing or perforation. It produces less discomfort for the patient than full colonoscopy, requires a less thorough bowel cleansing, and generally does not require sedation.

Screening by sigmoidoscopy has a three times higher detection rate for advanced neoplasia when compared to FOBT [427]. In a case-control study, having had a screening sigmoidoscopy was associated with reducing the risk of developing a CRC to half that of individuals who had not had the screening procedure [428]. However, many advanced neoplasias are still missed by this method, especially in women [429]. Authors of a 2005 study performed colonoscopy on almost 1500 women who had negative FOBTs and no family history of CRC [429]. They then age-matched these women with men from another

study, who also had negative FOBTs and no family history. They compared yield of flexible sigmoidoscopy, which they defined as the proportion of patients with advanced colorectal neoplasias that would have had their lesions identified by sigmoidoscopy. This included patients with advanced distal lesions that would have been directly identified by sigmoidoscopy, as well as those with advanced proximal lesions along with small adenomas in the distal colon, that would have lead to a full colonoscopy during which the proximal lesions would have been identified. They reported that the yield of flexible sigmoidoscopy was 35.2% for women, and 66.3% for men. In other words, 65% of advanced neoplasias in women would have been missed, along with 34% of these in men. The difference between the sexes was statistically significant with $P < 0.001$ [429].

1.9.5 Colonoscopy

Performed by gastroenterologists or surgeons, a colonoscopy is currently the gold standard for clinical CRC screening [421]. Similar to sigmoidoscopy, it involves the rectal insertion of a lighted tube, a colonoscope, for visualization of the bowel lumen and the biopsy or removal of identified polyps or other lesions. Depending on the characteristics of the lesion, surgical excision of that portion of the bowel may be required. The colonoscope is much longer than the sigmoidoscope, enabling visualization of the entire colon and into the distal end of the small intestine. The procedure often lasts 30 minutes or more, and sedation is often used to aid in patient comfort.

As with many sigmoidoscopes, colonoscopes have a small video camera on the end which is connected to a display monitor. For optimal visualization of the luminal wall during colonoscopy, patients are required to complete a thorough bowel preparation prior to the procedure. This involves cleansing the bowel through the implementation of dietary restrictions, the use of strong laxatives, and sometimes also an enema on the day of the procedure. Many patients report the bowel preparation to be the most unpleasant part of the test [421]. Since air is delivered into the colon to aid visualization, many patients experience gas pains and cramping after the test. Patients can also experience some mild bleeding. Heavy or persistent bleeding can result from biopsy or polypectomy in some cases, and on rare occasion can indicate a bowel perforation, which may sometimes require surgical repair.

The sensitivity and specificity of colonoscopy in the detection of polyps and carcinomas are high relative to other screening methods, but are reduced by an incomplete bowel preparation. Visualisation of the entire colon is not possible in about 5% of tests, due to either incomplete bowel preparation or a bowel which is technically difficult to navigate with the colonoscope [423]. Although some small polyps and even cancers can be missed, it is considered to be the gold standard for colorectal screening, and as such there is no other test against which to compare it and accurately determine its detection rate for neoplasia.

1.9.6 Double Contrast Barium Enema

In preparation for a double contrast barium enema (DCBE), a barium solution is administered by enema into the rectum, and the rectum is filled with air. A radiology technician takes a series of plain x-rays of the abdomen, for which the barium and air aid in outlining the colon and rectum. The test requires a full bowel preparation for optimal visualisation. Advantages of this test include that complications are rare, patient sedation is not required, and the entirety of the colon and rectum can be visualised. Disadvantages include the required bowel preparation and the inability of physicians to biopsy or remove polyps or other lesions during the test. If such lesions are identified, they are often assessed by follow-up colonoscopy, and then surgically removed.

In a study comparing the results of colonoscopy and DCBE in 580 patients who underwent both procedures, colonoscopy was determined to be the superior screening test [430]. The detection rate of DCBE for identifying adenomas was 39% that of colonoscopy. This was significantly related to the size of the adenomas, with only one third of the smallest adenomas (up to 5mm diameter) identified by DCBE, and approximately one half of larger ones identified. The detection rate of colorectal lesions by DCBE was also significantly influenced by their location within the bowel. DCBE was least effective in identifying lesions in the sigmoid colon, due to the difficulties in confidently interpreting x-rays of the region in patients with diverticulosis and redundancy. Only 35% of lesions identified in the sigmoid colon by colonoscopy were also identified DCBE. Detection rates in the ascending colon and cecum were also low, with less than half of lesions

identified by colonoscopy also identified by DCBE. This was most commonly due to faecal residue or poor mucosal coating, which again obscured x-ray visualisation of the bowel wall. In the remaining segments of bowel, the detection rates of DCBE were approximately two thirds those of colonoscopy [430].

1.9.7 Virtual Colonoscopy

Virtual colonoscopy is also known as computed tomographic colonography. It refers to the use of a traditional computed tomography (CT) scanner to take images of the entire rectum and colon. Integrated computer software then assembles these images into both two-dimensional pictures and a three-dimensional endoluminal "fly-through" view of the bowel, which allows both anterograde and retrograde examination, and resembles the view seen in traditional endoscopic colonoscopy [421,431]. A radiologist can examine the complimentary two- and three-dimensional images for signs of polyps, cancers, or other lesions.

Although much less invasive than traditional colonoscopy, virtual colonoscopy still requires a full bowel preparation to remove faecal matter. Patients are also often asked to consume a contrast solution to tag any residual stool or fluid, to aid in interpretation of images. A flexible tube is inserted into the rectum prior to the imaging, and is used to inflate the bowel with air. This distention of the bowel is necessary to enable visualization of the luminal wall. While there is a risk of bowel over-inflation and perforation, it is smaller than the risk of perforation in traditional colonoscopy. Most

patients will report some bloating and cramping from the air introduced into the colon. Generally, virtual colonoscopy is tolerated quite well by patients. The procedure does not require pain medication or sedation.

The utility of virtual colonoscopy at detecting colorectal neoplasia is equivalent to traditional colonoscopy. Lesions on the luminal wall can be visualised as clearly as they can with traditional colonoscopy [421,431]. Virtual colonoscopy may provide superior visualisation in some locations, such as within 10 cm of the ano-rectal verge and along the back-side of folds or other blind spots [431]. While it is comparable to traditional colonoscopy at identifying most colorectal lesions, there is evidence that it may be slightly inferior at identifying small polyps [431].

Virtual colonoscopy may be an excellent alternative to traditional colonoscopy for frail, elderly patients, or for other patients that may not be able to tolerate traditional colonoscopy, including those with certain comorbidities, or who are on anticoagulation therapy. It is also an alternative for patients whose colons are technically difficult to navigate with a colonoscope, for whom examination of the entire colon is not possible using traditional methods.

One disadvantage of virtual colonoscopy is that polyps and other identified lesions cannot be biopsied or removed; in cases where lesions are identified, a traditional colonoscopy or surgical procedure would need to be arranged. Another disadvantage of this procedure is the risk of radiation exposure to the patient. While the radiation dose from one test is considered to be low, the accumulated exposure from multiple repeated

scans may be significant, and could put patients at a slightly increased risk of developing cancer [431].

1.9.8 Current Canadian Clinical Screening Recommendations

Current guidelines for clinical screening for colorectal cancer vary by patient age and family risk. In 2001 the Canadian Task Force on Preventive Health Care published recommendations for colorectal cancer screening [422]. Three years later the Canadian Association of Gastroenterology and the Canadian Digestive Health Foundation published recommendations which were similar, yet much more detailed [423]. Both sets of recommendations were stratified by the risk of the patient developing CRC. This included differing recommendations for individuals based both on age and on family history of CRC. Colonoscopy, the gold standard test for identifying colorectal lesions, was generally not recommended for general population screening, since it is expensive, invasive, is associated with low patient compliance, and incurs risks to the patient which may not be justified by the benefits of the test in this low-risk group. As well, the lack of infrastructure and access to specialised gastroenterology care necessary for this test limit its use for population-based screening. Just as colonoscopy may not be suitable for population-based screening, cheaper, more accessible, and less invasive tests such as FOBT are not sensitive enough for clinical screening within a high-risk population [423].

Current recommendations are to begin CRC clinical screening at age 50 years for people of average population risk – defined as those who are asymptomatic and have a

negative family history of CRC. Individuals with only one first degree relative with CRC, who was older than 60 years old at diagnosis, and those with history of CRC only in second and third degree relatives are also advised to follow screening for people of average population risk. Individuals who are members of a kindred with a known genetic predisposition for CRC, for whom the genetic mutation has been identified, can also follow average-risk screening, if they are found not to harbour the predisposing mutation in their family. The screening method used can be chosen by the individual and their physician, based on patient preference, evidence, and availability of resources [423]. Five screening options are listed: (1) FOBT every two years, (2) flexible sigmoidoscopy every five years, (3) flexible sigmoidoscopy combined with FOBT every five years, (4) DCBE every five years, or (5) colonoscopy every 10 years [423].

Individuals at higher risk of developing CRC are defined as those with a strong family history of CRC or who have had polyps identified. For most of these individuals, regular colonoscopy is generally recommended. For individuals with Lynch Syndrome, colonoscopy is recommended at a frequency of every one to two years, starting at age 20, or ten years younger than the earliest case in their family, whichever comes first [423]. Individuals who are not members of a recognized Lynch Syndrome kindred, but who do have a family history of CRC which includes at least one first degree relative with a history of CRC diagnosed younger than age 60, or two or more first degree relatives with CRC or polyps diagnosed at any age, are advised to undergo colonoscopy every five years. The screening should begin ten years earlier than the youngest case in the family, or no later than age 40 years.

Clinical screening recommendations for members of FAP kindreds differs, to reflect the different phenotype generally associated with that syndrome. Since polyps can develop at a young age, initiation of screening is recommended at age 10 to 12 years. Because polyps have traditionally been found in the rectum as well as the rest of the colon, annual screening with sigmoidoscopy is considered to be appropriate [423]. Although these are the recommendations, they may not be appropriate for all FAP families. Many Newfoundland FAP kindreds include individuals with an attenuated phenotype who have fewer polyps, in some cases only located in the proximal colon. For some individuals polyps have been found even younger than the age given above, while other individuals have remained polyp-free into their 40s (personal communication, Dr. Jane Green). For these families screening must start earlier, and should include colonoscopy. Recommendations for families with attenuated FAP includes annual colonoscopic screening beginning in the early to mid teens [423].

Individuals of any age who have a positive family history of CRC and are found to have polyps, or had an incomplete examination, should have follow-up colonoscopy after a shorter interval than normally recommended, based on clinical judgement [423]. Others who should have more frequent screening include individuals with a personal history of CRC, and those with a personal history of inflammatory bowel disease or primary sclerosing cholangitis, as they are also at an increased risk of developing CRC over the general population [423].

Regular screening is important for early detection of neoplasms. Identified polyps can be removed, thereby preventing their possible development into carcinoma, and

carcinomas identified early have a much greater chance for cure than those identified at a more advanced stage. Despite the publication of clinical screening recommendations, adherence to these recommendations is poor [432]. In a recent study of almost 13 000 individuals from four Canadian provinces, all of average risk of developing CRC, only 17.6% had a history of CRC screening which followed the recommendations. This number increased to 23.5% when considering those that had a history of some screening, but whose screening was not up to date. Reasons for the low levels of screening were not presented [432].

1.10 Colorectal Cancer Treatment, Microsatellite Instability Status, and Chemotherapy Response

While surgical resection can be curative for many colorectal tumours, many others go on to develop distant or local recurrences. Patients considered to be at increased risk of tumour recurrence are often offered adjuvant therapy to improve the likelihood of a good outcome. This therapy can be in the form of radiation, chemotherapy, or immunotherapy, and is administered before (neo-adjuvant) or after (adjuvant) surgical resection with intent to cure [433]. These therapies are cytotoxic, and therefore the potential benefits must outweigh the risks of treatment for the individual patient.

The current clinical practice guidelines for the use of adjuvant therapy were developed at a National Institutes of Health (NIH; National Cancer Institute, United

States) Consensus Development Conference in 1990, and published later that year [433]. The guidelines recommend against administering adjuvant therapy for stage I and IV cancers. Stage I tumours by definition have minimal invasion, and therefore tend to be the least likely to recur. For patients with stage I tumours the risk-to-benefit ratio of the therapy is unfavourable. Patients with stage IV tumours already have distant metastases at diagnosis, and therefore any therapy given would generally be for palliative reasons.

Stage II cancers invade into the bowel wall past the submucosa, but have not spread to lymph nodes or distant organs. Stage III cancers have spread to lymph nodes, but not to distant organs. These are the cancer stages for which adjuvant therapy is considered. The guidelines [433] advise assessing colon and rectal cancers differently for adjuvant therapy. Rectal cancers are more likely to recur, and therefore therapy is more aggressive. Stage II and III rectal cancers are commonly treated with neo-adjuvant or adjuvant radiation therapy, in conjunction with adjuvant chemotherapy. The recommendation for colon cancers is to treat all stage III with adjuvant chemotherapy, but only some stage II. The NIH guidelines list features that indicate an increased risk of disease recurrence, and recommend that only stage II colon cancers with one or more of these features receive adjuvant chemotherapy. The listed features include: aneuploid tumour DNA, high pre-operative CEA level, a high portion of cells in S phase of the cell cycle, loss of heterozygosity (LOH) at 17p or 18q, and mucinous, signet ring, or poorly differentiated histology [433]. Interestingly, the histological features listed are those associated with high levels of microsatellite instability, although MSI status was not mentioned in the guidelines.

Updated recommendations were published in 2004 by both the American Society of Clinical Oncology (ASCO) [434], and the Cancer Care Ontario Program in Evidence-Based Care (CCOPEBC) [435] for the use of adjuvant chemotherapy in the treatment of stage II colon cancers. ASCO and CCOPEBC collaborated to perform a review of available literature from randomized controlled trials and meta-analyses comparing adjuvant chemotherapy to observation after surgical resection for stage II colon cancer. The most commonly used chemotherapeutic agent is 5-fluorouracil (5-FU), given with or without folinic acid, irinotecan, or oxaliplatin. The ASCO and CCOPEBC reviews [434,435] concluded that while there was no compelling evidence to support the standard use of adjuvant chemotherapy in stage II colon cancers, there did appear to be a small survival benefit that may not be considered clinically significant. Both reviews recommended that patients with stage II colon cancer should discuss the risks and benefits of therapy with their oncologists, and if they have characteristics of increased risk of recurrence or there was an inadequate lymph node harvest at surgery they should be encouraged to seek out active clinical trials for adjuvant therapy. Two large, recent studies [436,437] had similar conclusions.

Concordance with the current clinical practice guidelines was recently assessed for the Newfoundland and Ontario populations [438]. Among the 274 incident cases of colon cancer diagnosed in the province of Newfoundland and Labrador in 1999 and 2000, 130 were included in this study, of which 55 patients had stage II disease. Of these, 41 (75%) had at least one high-risk feature, defined as: "clinical obstruction or tumour perforation at presentation, T4 lesion, poor differentiation, lymphatic invasion, perineural

invasion, vascular invasion or mucin production" [438]. Of those, 29 (71%) were referred to medical oncology, and 18 (62%) received adjuvant chemotherapy. Of the 14 patients with low-risk stage II colon cancers, approximately the same percentage were referred to oncology (11 cases, 73%), but only 3 (27%) received chemotherapy. In the Ontario cohort, 116 patients had stage II disease, of which just less than half were considered to be high-risk. Two thirds of the high-risk patients were referred to oncology, as were one third of low-risk patients.

The authors found that the guidelines were followed for stage I and III colon cancer, as none of the stage I cancers received chemotherapy, and the vast majority of stage III cancers did, in both provinces [438]. They concluded, however, that the guidelines were not universally followed in the treatment of stage II colon cancers. While the patients with high-risk stage II cancers were significantly more likely to receive adjuvant chemotherapy than those with low-risk cancer, the authors found that factors not included in the guidelines were important in decision-making regarding therapy. The biggest example of this was age; patients younger than 50 years of age at diagnosis were more likely to be treated with chemotherapy than were older patients with the same level of recurrence risk [438]. The authors recognized the potential significance of this, that younger patients are more likely to have tumours with high levels of microsatellite instability.

Does MSI status affect response to chemotherapy? MSI is not mentioned in any of the guidelines [433-435]. The NIH guidelines [433] list histology features more common in MSI-H tumours as features which warrant high-risk status for disease recurrence, and therefore recommend adjuvant chemotherapy for stage II colon cancer with these

features. The review of colon cancer in Newfoundland and Ontario [438] found young age at diagnosis to be a factor used in clinical decision-making in favour of treating stage II colon cancer with adjuvant chemotherapy, although age is not mentioned in the guidelines, and again leads to an increase in patients with MSI-H tumours receiving adjuvant chemotherapy.

Tumours with high levels of microsatellite instability have a DNA mismatch repair deficiency; thus, they respond differently to certain types of DNA damage than microsatellite stable tumours. Since many standard chemotherapeutic agents function by damaging DNA, these two groups of tumours can be expected to respond differently to chemotherapy. The damage induced by some agents is recognized by the intact MMR system, but cannot be repaired, so cell cycle arrest and cell death result. Meanwhile, other chemotherapeutic agents induce DNA damage that leads to cytotoxicity independently of the MMR system [439].

For example, the methylating agent N-methyl-N'-nitro-N-nitrosoguanidine (MNNG) can produce methylated purines, including 0^6-methylguanine, which are incorporated into newly synthesized DNA. An intact MMR system can recognize the adduct due to its mispairing, usually with a T, on the new strand. Since an appropriate match cannot be found for 0^6-methylguanine the cell cycle will arrest in the G2 phase, resulting in cell death. However, in the absence of a proficient MMR system, the damage is not recognized and processed in the same way, and the cell survives, resulting in tolerance to the chemotherapy agent [440]. Similarly, 6-thioguanine (6-TG) is an antimetabolite chemotherapeutic agent that becomes incorporated into newly synthesized

DNA, but cannot form a stable pair with any other base. The resultant mispairing is recognized by the MMR system and cannot be repaired, leading to cell cycle arrest at the G2/M checkpoint. Defective MMR leads to tolerance of this agent as well [441].

Platinum agents used in adjuvant cancer therapy form inter-strand-DNA cross-links, resulting in DNA adducts [442,443]. The adducts of some platinum agents are recognized by the MMR system, while others are not. Consequently, drugs such as cisplatin, whose adducts are recognized by the MMR system, are tolerated by MMR deficient tumours (although this drug is not used in treatment for CRC), while MMR proficient tumours can be sensitive to them. Other platinum drugs, such as oxaliplatin, produce DNA adducts that are not recognized by the MMR system, so the mismatch repair status of the tumour is not relevant to the efficacy of the agent [443]. Oxaliplatin is commonly used today in colorectal cancer chemotherapy as an adjunct to 5-fluorouracil and leucovorin. Together, these agents are used in the FOLFOX protocol, which was described in 2000 as a first-line palliative therapy for advanced CRC [444], then applied to the adjuvant setting four years later [345], where it proved more efficacious than the other therapies at the time, and is still used today in adjuvant chemotherapy for CRC.

While the sensitivity of a tumour to the oxaliplatin component of the FOLFOX protocol is not dependent on MMR status, the tumour sensitivity to 5-FU is. 5-FU is the main component of the protocol, and was the basis of virtually all first-line adjuvant chemotherapy for CRC described in a large meta-analysis from 2009 [346]. A more recent trial has identified a benefit in both overall and disease-free survival when oxaliplatin was added to 5-FU-based adjuvant therapy in patients with stage III CRC [437].

5-FU is a member of the antimetabolite class of chemotherapeutic agents. *In vitro* studies [70,71] have clearly demonstrated a reduced response by MSI-H tumours to fluorouracil-based agents. 5-FU acts on the cell in many ways. It is incorporated into RNA, which interferes with RNA processing [445]. 5-FU also prevents DNA synthesis, through inhibition of thymidylate synthetase (TS). 5-FU is converted to 5-fluoro-2'-deoxyuridylate (FdUMP) in the cell, which irreversibly binds TS, an enzyme required for *de novo* synthesis of thymine, a nucleotide required for DNA synthesis [445].

While these functions were traditionally believed to be the cause of the cytotoxicity of 5-FU, they do not explain the relationship with MMR proficiency. This relationship is likely related to the observation that metabolites of 5-FU can be incorporated into DNA [446]. The MMR complex hMutSα recognizes and binds 5-FU-modified DNA with greater affinity than for complementary DNA or DNA containing a C/T mismatch [71]. The specific mechanism by which the MMR recognition of 5-FU-modified DNA leads to cytotoxicity is still unknown. Unlike MNNG and 6-TG, which lead to cell cycle arrest in the G2 phase, there have been no cell cycle deficiencies observed in 5-FU-treated cells [70].

Even though the mechanism is unknown, it is recognized that 5-FU treatment does not hinder the growth of MMR-deficient cell lines, while it does significantly hinder the growth of MMR-proficient cell lines [70]. One study [447] looked at MMR deficiency caused by *hMLH1* promoter methylation, and found that when the promoter was demethylated MMR function was restored, and the cells became sensitive to 5-FU treatment. A similar

study [448] has also been done which demonstrated the same effect of *hMLH1* promoter demethylation when using other MMR-dependent drugs.

Clinical studies of the impact of MMR status on 5-FU therapy have not been as clear-cut as the *in vitro* studies. Early studies [64,449-451] showed a survival benefit for patients with MSI-H CRCs who received 5-FU-based adjuvant therapy. However, these studies had inherent biases resulting from the use of small or non-randomized cohorts. One study [64] had very small numbers; they looked at 19 patients treated with 5-FU, all of which were diagnosed younger than age 40 years, and of which eight had MSI-H tumours, and 11 had MSS tumours. Another study [449] had a non-randomized cohort, with potential for bias due to the difference in age of patients in the two compared cohorts. The patients that received adjuvant treatment were an average of 13 years younger than those that did not receive treatment. Another study [452] did not control for patients not receiving treatment. All patients in their cohort received 5-FU, and they compared the outcomes of those with and without microsatellite instability. Their conclusion of better outcomes for patients with MSI-H tumours receiving 5-FU over those with 5-FU-treated MSS tumours did not account for the known confounder of better overall survival of patients with MSI-H tumours [64,65].

The first study to use a large, randomized cohort which adequately limited biases was published in 2003 by Ribic *et. al.* [453]. They assessed MSI status in 570 CRC tissue specimens from patients with stage II and III cancer. Their study was retrospective, but used specimens from patients that were enrolled in multicentre, prospective, randomized, controlled trials of fluorouracil-based adjuvant chemotherapy. They concluded that

adjuvant chemotherapy provided no survival benefit for patients with MSI-H CRCs. These patients instead demonstrated a slight survival disadvantage when treated with adjuvant chemotherapy, when compared to patients with MSI-H tumours that did not receive adjuvant therapy. By contrast, there was an overall survival benefit for patients with MSS and MSI-L tumours who received chemotherapy when compared to those who did not. This observation held true in multivariate analysis, when also considering the stage and grade of the tumours.

More recent studies have had similar findings to those of Ribic. A 2009 prospective study by Jover *et. al.* [62] evaluated the efficacy of 5-FU-based adjuvant chemotherapy on a cohort of 505 CRC patients with stage II or III disease. Approximately half of the patients received adjuvant treatment; mismatch repair deficiency was identified in about 10% of tumours. They found that adjuvant chemotherapy both improved survival and increased time to tumour recurrence only in patients with MMR-proficient tumours. Patients with MMR-deficient tumours did not benefit from adjuvant chemotherapy. These differences remained significant in multivariate analysis controlling for age, sex, and tumour stage. The authors concluded that MSI was not only a prognostic marker in CRC, but also a valuable predictive marker, and should be used in decision-making for patient treatment [62].

A meta-analysis, also from 2009 [346], looked at seven studies representing 3690 CRC patients. The results were similar to those found by Jover *et. al.* Patients with MMR-proficient, MSS tumours showed a clear benefit from adjuvant chemotherapy in both tumour-free survival and overall survival, while patients with MSI-H tumours

showed no benefit from the treatment. The authors of this meta-analysis also concluded that MSI status should be used as a predictive marker in deciding CRC treatment [346].

While it is becoming evident that MSI-H tumours do not respond to 5-FU-based chemotherapy, testing for MSI status is not a part of current clinical guidelines, which are still based almost entirely on tumour stage. The use of MSI and other molecular markers as predictive tools for guiding cancer therapies has the potential to make substantial clinical impact. The standard of care includes 5-fluorouracil-based chemotherapy, however it has only a 20% to 30% tumour response rate [439]. MMR deficiencies in the tumours of patients with Lynch Syndrome and those with MSI-H sporadic tumours account for some of the resistance, but not all. Some tumours that are initially sensitive to chemotherapy can develop a resistance to it. Some chemotherapy agents increase the rate of spontaneous mutations, which can lead to an acquired MMR deficiency in some cells. With continued treatment, these cells develop a growth advantage and can develop tolerance to the drug [442,454]. This same acquired MMR deficiency can occur in non-tumour cells in other tissues as well, as chemotherapeutic agents don't act exclusively on tumour DNA. These acquired MMR deficiencies can lead to secondary cancers; most commonly leukemias [442].

For the tumours that are tolerant to 5-FU there are other chemotherapeutic agents available. Some of these agents act independently of the MMR system, so are equally effective on both MSI-H and MSS tumours, for example the taxane class of drugs [455]. There are others, however, that do act through the MMR system, but are more sensitive to MSI-H tumours than they are to MSS tumours. Such drugs include the topoisomerase

poison CPT-11 (Ironotecan) [451]. Knowledge of the mismatch repair status of a tumour can help tailor the adjuvant treatment plan for optimum patient outcome. At the present time, most clinical laboratories do not test for the MSI status of CRCs, as it is expensive and time-consuming.

1.11 Introduction to Thesis Work

The location and anatomy of the colon and rectum provides the ability to identify, collect, and examine lesions of various stages of development, from the smallest aberrant crypt foci, through polyps of various types and sizes, to large, metastatic or obstructive adenocarcinomas. The study of the histopathology, genetics, and molecular profiles of these various lesions has led to a tremendous knowledge base of the different pathways involved in carcinogenesis, which have relevance not only to the colon and rectum, but also to other sites in the body. The use of this data, together with clinical and demographic information gleaned from the patients and their medical records, has led to a better understanding of the prognostically relevant features of the disease, and has begun to allow for the individual tailoring of some therapies. Further research will lead to an even better understanding of the heterogeneous nature of colorectal cancer, and from that will come an improved understanding of how the differences seen at the histopathological, molecular, and genetic levels can further refine patient prognostication and treatment. Improved patient outcome is, of course, one of the ultimate goals of colorectal cancer research.

In order to improve patient outcome in a specific geographic region, it is helpful to first study the scope and type of disease in that area. The population structure and level of genetic heterogeneity can play a role in the types of disease seen within a population, as can shared cultural practices, which may include a common diet. The province of Newfoundland and Labrador has a unique population structure, resulting from the way in which the area was originally settled, as well as geographic isolation, religious segregation, and traditionally large family pedigrees.

Newfoundland has a young founder population, which is less than 20 generations old [456]. The vast majority of current residents are descendants of a very small group of original founders. The first Europeans came here as transient, seasonal workers in the 1600s [457]. They were fishermen from small regions in southwest England and southeast Ireland, and came in the summer months to catch cod for European markets. The current population of approximately 510 000 people are almost exclusively of English or Irish decent, and can trace their ancestry back to those same two small regions in England and Ireland that once supplied the seasonal fisherman [456,457].

The English settlers were of Protestant faith, while the Irish were Catholic. These two religious groups generally did not intermix for marriage, leading to religious segregation. The original immigrants settled into geographically isolated communities along the coast, to enable fishing. This geographic isolation, both of the small coastal communities from each other, and of the island of Newfoundland from the mainland, further contributed to genetic isolation. There has been very little migration into these communities, although out-migration has recently been high [456]. Traditionally, family

sizes were large with couples having many children, and family members remained close to the core family, settling in the same, or nearby communities. The result is a province consisting of a young founder population living in multiple genetic isolates. This genetic isolation and homogeneity is still present in many small coastal communities [456,458], despite the fact that improved transportation infrastructure has reduced the impact of geographic isolation in recent years.

The founder population and genetic isolates have resulted in an overabundance of some genetic disorders [456]. While some genetic traits are over-represented in some areas of the province, others are more rare. A founder effect has been observed in Newfoundland for a number of different traits, including: Bardet-Biedl Syndrome [459,460], hemophilia A [461], multiple endocrine neoplasia type 1 [462], hereditary spastic ataxia [463], and hereditary diffuse gastric cancer [464], among others. This population has been instrumental in identifying the genetic defects responsible for many of these diseases. It is important to note, however, that although Newfoundland has a unique population structure, in a study of 12 founder populations, Newfoundland was found to have the greatest generalizability to other Caucasian populations [465]. This indicates that Newfoundland not only has a valuable population structure for genetic studies, but that the results of these studies are not unique to this population, but can be applied to other populations as well.

There has been a long, fruitful history of colorectal cancer research in Newfoundland. Large families with hereditary CRC have been clinically followed for decades, and most family members have been generous and eager to participant in

genetics research. As briefly mentioned previously, a large family from Newfoundland was instrumental in the mapping of a locus on chromosome 2 associated with CRC risk [18], as well as with the later cloning of the *MSH2* gene at that locus [228]. That family, Family C, has a germline splice site mutation, c.942+3A>T (p.Val265_Gln314del), which results in an in-frame deletion of exon 5. This Family C mutation has since been identified in more than 12 other individually ascertained Newfoundland Lynch Syndrome families (many of which have since been linked to a common founder), and accounts for approximately one third of all individuals with known mutations predisposing to CRC in the province [466].

A second Lynch Syndrome founder mutation is also quite predominant in Newfoundland, c.1277-?_1386+?del (p.Lys427GlyfsX4). It has been identified in five families, and haplotype analysis strongly suggests that there is a common ancestor for at least four of these families [467]. While the specific breakpoints are not yet known, this mutation results in deletion of *MSH2* exon 8, and subsequent frameshift and early truncation of the protein product [466].

A third founder mutation predisposing to CRC in this province is associated with FAP. A single base pair change at a splice site in the *APC* gene resulting in deletion of exon 4 is the cause of attenuated FAP in at least five Newfoundland families, four of which have been traced to a common ancestor [468]. The mutation is a single base pair change at the splice-acceptor site of intron 3, and causes deletion of exon 4; c.221-1G>A (p.Glu74GlyfsX29) [466,468].

The founder population of Newfoundland, along with traditionally large family sizes, has contributed to a higher than average proportion of hereditary and familial CRC in Newfoundland [466,469]. The typical Newfoundland diet with high intake of fat and processed meats may also contribute to the high rates of CRC here. The Canadian Cancer Statistics reports that Newfoundland has higher rates of CRC than any other province in Canada [21]. The incidence rates in this province are among the highest in the world, second only to a small prefecture in Japan (Dr. John McLaughlin, personal correspondence). The high burden of disease in this province warrants further CRC research. We have the potential to build upon the long history of research and patient cooperation to further improve risk assessment, organize risk-tailored access to screening, and thereby improve cancer prevention. We can also learn more about the genetic, clinical, and pathological features of disease and their relationships to patient prognosis and drug response, to enable a more personalized approach to patient management, and ultimately improve patient outcome.

With the local need and opportunities for CRC research, in 1999 an interdisciplinary team of clinicians and researchers embarked upon a multi-year collaborative project with a team in Ontario. The work reported in this thesis was carried out in conjunction with that project. The collaborative study, the Canadian Institutes of Health Research (CIHR) Interdisciplinary Health Research Team (IHRT), provided funding for the development of the Newfoundland Colorectal Cancer Registry (NFCCR). The NFCCR was modeled after the existing Ontario Familial Colorectal Cancer Registry (OFCCR) [470], and intended to provide a resource for numerous research studies.

Different components of the project included: molecular genetics work studying the prevalence of MMR mutations, as well as the role of polymorphic genes, and the prognostic significance of microsatellite instability; clinical genetics work evaluating the risk-profiles of families, as well as the risks to family members; the development of a risk-assessment tool for primary care physicians; establishment of a more efficient pathology reporting system; as well as the psychosocial and behavioural impact of genetic counseling and testing; and more. The project was also intended to support the training of students within the various disciplines.

In Newfoundland, all pathology-confirmed incident cases of CRC diagnosed in the province between 1999 and 2003 were identified from the provincial tumour registry, and efforts were made to contact all patients or proxies [469]. Of 1150 tumours that met these inclusion criteria, 730 consented to take part in the research study. Patients or proxies were asked to fill out a family history questionnaire, as well as risk factor questionnaires pertaining to diet and epidemiological data. They were also asked to provide a blood sample, and access to surgical pathology specimens and medical records. Both population-based as well as family-based controls were also collected.

The OFCCR was established two years prior to the NFCCR, as one of six North American and Australian institutions participating in an international collaborative study, the Colorectal Cancer Family Registry (Colon CFR) [471]. The Colon CFR was established by the US National Cancer Institute in 1997 to support interdisciplinary studies of CRC. As collaborators with the OFCCR, the Newfoundland team has been included in the international Colon CFR as an affiliate site.

My graduate research programme began as the Newfoundland team was investigating the possibility of undertaking a large-scale CRC research initiative, just before the NFCCR was established. The first project in my programme was a pilot study for the NFCCR. It is presented in Chapter 2 of this thesis. I used the provincial cancer registry as well as searches of the pathology reporting database to identify all incident cases of CRC on the Avalon peninsula of Newfoundland for the years 1997 and 1998. I carried out microsatellite instability analysis and immunohistochemistry for mismatch repair proteins on this population-based cohort, and correlated the results with the limited available family history data. The main purpose of this study was to provide a molecular characterization of CRC in the province and estimate the contribution of hereditary cancer to the burden of disease.

After completion of the pilot study I commenced two other projects within the CIHR IHRT. Both of these studies used various aspects of tumour pathology and histology to sub-classify colorectal tumours for the purpose of improved prognostication and prediction of drug response. The first of these studies used tissue microarray (TMA) technology with immunohistochemistry to assess expression of a number of proteins within the tumour tissue. All available specimens from the pilot study were included, along with a portion of cases from the NFCCR. I selected and marked appropriate areas of tumour blocks to use as donor regions for removing tissue cores for TMA construction. With the help of pathologists I was involved in scoring all stained cores for all proteins studied. I also collected and cleaned all demographic, clinical, pathological, treatment, and follow-up data for each patient, along with vital status and cause of death, when appropriate. I then directed statistical analysis, and wrote the resulting manuscripts. The

aim of this study was to assess the prognostic and predictive value of the expression of the selected proteins in isolation as well as together in protein expression profiles, and to evaluate the utility of TMA and unsupervised hierarchical clustering analysis to identify prognostically relevant sub-groups of CRC. This work is presented in Chapter 3. It was intended to be a pilot project for a larger TMA study that was planned for the CIHR IHRT.

The remaining project was a review of the histology of all tumours included in the NFCCR, which is presented in Chapter 4. Together with the help of a pathologist, I evaluated representative formalin-fixed, paraffin-embedded tissue sections mounted on glass microscope slides and stained with haematoxylin and eosin, from the 710 primary tumours for which such blocks were available. We assigned scores for a number of histological features. I then correlated these scores with molecular data from the tumour DNA and clinical data from the patients. These scores were used to develop an algorithm for the prediction of MSI status, which we validated using the cohort from the original pilot project.

1.12 Thesis Goal

The overall goal of this programme was to further characterize colorectal cancer in the province of Newfoundland and Labrador, to better understand the etiology of this disease and to identify sub-groups of the disease with prognostic and predictive utility. Many methods were used to characterize the tumours, including molecular methods (Chapter 2), protein expression profiles (Chapter 3), and histological features (Chapter 4).

Chapter 2 – The Epidemiology and Molecular Characterization of Colorectal Cancer in Eastern Newfoundland

2.1 Introduction

When we started this project in the year 2000, Newfoundland had among the highest rates of CRC in the country, as it still does. According to the 2000 Canadian Cancer Statistics, Newfoundland males had the highest age-standardized incidence of colorectal cancer in the country, while females had the second highest rate, estimated at 77 and 46 per 100,000 respectively [472]. These high CRC incidence rates were not due to a high overall cancer burden in the province, since for the same time period Newfoundland males had the second lowest and females the lowest overall cancer incidence rates in the country [472]. Although we knew that Newfoundland had high levels of CRC, the contribution of hereditary cancer to the burden of CRC in the province was unknown. A main objective of this project was to determine that contribution.

In 2000, the known hereditary colorectal cancer syndromes were Lynch Syndrome and FAP (discussed in sections 1.8.1.1 and 1.8.2.1 respectively). While polyposis syndromes had the obvious clinical characteristic of multiple polyps, Lynch Syndrome did not have a distinctive clinical feature that could be used to easily distinguish it from sporadic CRC [473]. Germline mismatch repair (MMR) deficiency was the molecular characteristic used to identify Lynch Syndrome, but it was not routinely used in clinical

practice. MMR deficiency could be identified by three methods: microsatellite instability (MSI) analysis, immunohistochemistry (IHC) to test for expression of mismatch repair proteins, and sequencing to find mutations in mismatch repair genes. Gene sequencing was not feasible in a population-based clinical setting, and was similarly unfeasible for this study, so we focused on the other two methods. At the time, we hypothesized that MSI and IHC would closely correlate with each other, and would correlate somewhat with family history of CRC. While virtually all Lynch Syndrome colorectal cancers have microsatellite instability, 13% of sporadic CRCs also have this feature, but these patients generally do not have a family history of CRC [60,474].

Planning was underway in the year 2000 to apply for an interdisciplinary health research team (IHRT) grant from the Canadian Institutes of Health Research (CIHR) for CRC research in the province. This application was approved as a collaborative project between the provinces of Newfoundland and Ontario, as described briefly in Section 1.11. The project described here was a pilot study for some aspects of the CIHR-IHRT CRC, including patient identification, patient recruitment, and molecular characterization of tumours.

This was a population-based, retrospective study. We identified and examined all available incident cases of colorectal cancer on the Avalon Peninsula of the province for a two-year period, from January 1, 1997 to December 31, 1998 (N=294). The population of the Avalon Peninsula represented approximately half the total population of the province. We examined all tumours for microsatellite status and immunohistochemistry for the major MMR proteins: MLH1, MSH2, and MSH6. A family history study was

conducted on a small subset of these patients (n=79), to examine the correlation of family history with molecular findings.

2.2 Objectives

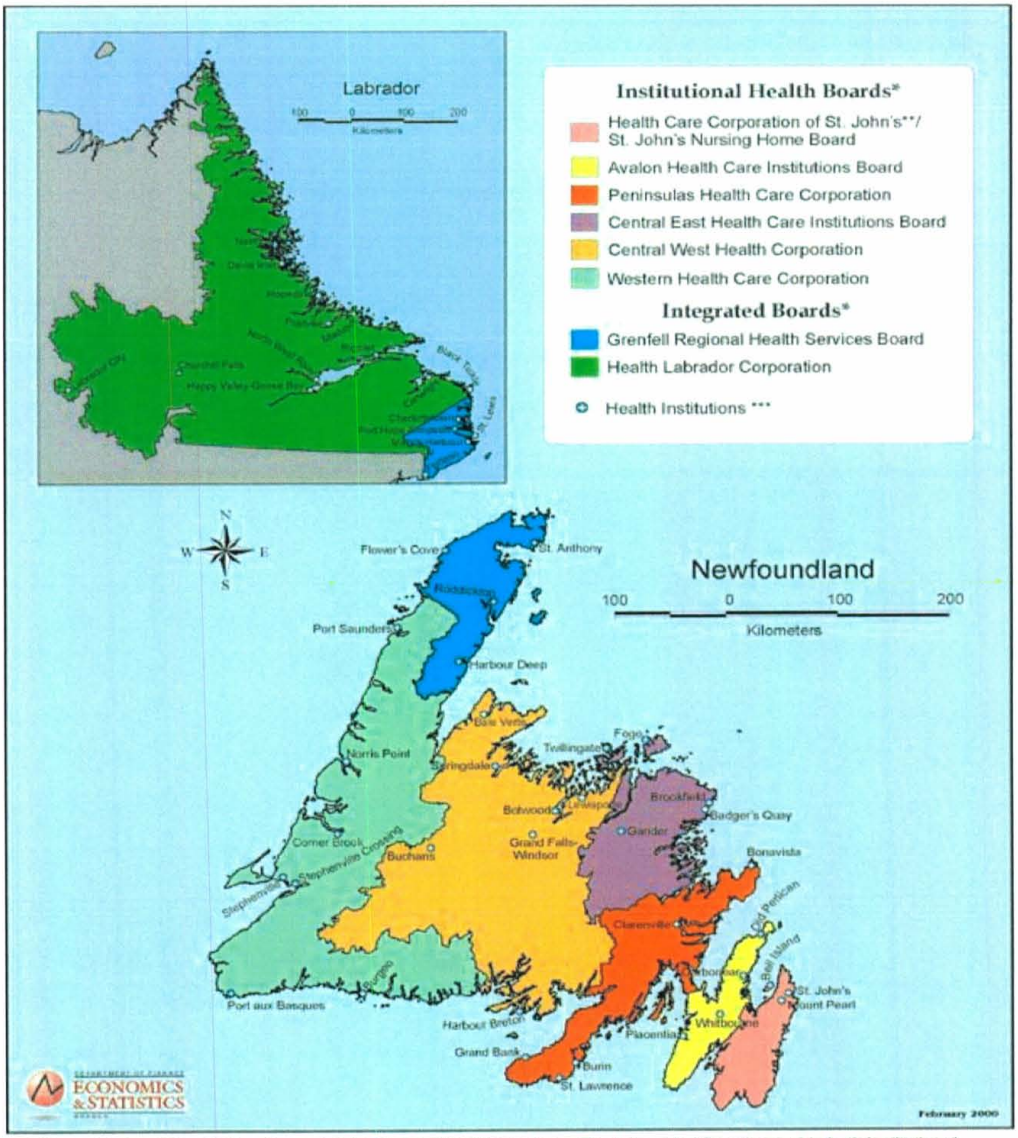
The main objective of this study was to determine the contribution of hereditary disease to the colorectal cancer burden of Newfoundland by analysis of both the molecular and family history components of the study cohort. We also sought to characterize the molecular features of colorectal cancer in the province, in a population-based cohort of patients. We were most interested in examining the molecular features observed when there is a deficiency in the DNA mismatch repair system, such as in Lynch Syndrome. This was done by examining the tumours and normal colonic epithelium of all eligible cases for microsatellite instability and immunohistochemical expression of three key mismatch repair proteins, as described below.

2.3 Molecular Study

2.3.1 Methods

2.3.1.1 Identification of Eligible Cases

As co-principal investigator of this study, my role was to identify all incident cases of colorectal cancer on the Avalon Peninsula of Newfoundland for the two-year period of interest, locate the surgical specimens, perform the molecular and histological analyses, and carry out any associated statistical analyses. At the time, the province of Newfoundland and Labrador was divided into eight healthcare regions. Each region was managed by a different Health Board, as illustrated in Figure 2.1. The Avalon Peninsula was represented by two of the Health Boards: the Healthcare Corporation of St. John's, and the Avalon Healthcare Institutions Board. Ethics approval for this study was obtained from each of these Health Boards, as well as from the Human Investigations Committee (HIC) of Memorial University of Newfoundland and Labrador, HIC#00.105.



*The Institutional Health Boards provide institutional health services. The Integrated Boards provide both institutional, and health and community services.
**The Health Care Corporation of St. John's has a tertiary care mandate for the entire province.
***Locations of Hospitals, Nursing Homes, Community Health Centres, or Community Clinics administered by the Boards.
Note: The Newfoundland Cancer Treatment and Research Foundation offers services throughout the province.

Figure 2.1. Institutional Health Boards of Newfoundland in 2000
Our region of interest was the Avalon Peninsula, which encompassed Health Boards 1 and 2 (in pink and yellow), managed by the Health Care Corporation of St. John's, and the Avalon Health Care Institutions Board.

Eligible patients were identified using the Provincial Cancer Registry at the Newfoundland Cancer Treatment and Research Foundation (NCTRF), inpatient discharge summaries accessed on the electronic medical records system MEDITECH, and pathology reports from the four hospitals in the region. The four hospitals included the Health Sciences Centre, St. Clare's Mercy Hospital, and the Grace General Hospital, all located in the city of St. John's, as well as the Carbonear General Hospital, located in the town of Carbonear. Fiona Curtis (F.C.), a Master's student in clinical epidemiology, searched the NCTRF and MEDITECH records, while I searched pathology records. I obtained pathology reports from each of the four hospitals with the help of their departmental secretaries, searching their pathology record system for terms including: colorectal, colon, rectum, bowel, and large intestine, for the two year period of interest. I then went through each report to remove any ineligible cases.

The patients included in the study were sorted by age into two cohorts for molecular analysis: those diagnosed up to age 74, and those diagnosed at age 75 years or older. The molecular analysis on the older patients was carried out by undergraduate honours student Payal Sipahimalani (P.S.), while I carried out the analysis for the younger patients. F.C. carried out a family history study on the subset of patients diagnosed younger than age 70 years. The correlation of the family history and molecular studies is presented in section 2.5.5.3.

2.3.1.2 Inclusion Criteria

Eligible cases were patients diagnosed with a primary colorectal cancer in 1997 or 1998, who lived on the Avalon Peninsula of Newfoundland. All patients were included in the molecular component of the study, regardless of age at diagnosis. Patients with carcinoma arising in a polyp were eligible and included in the study, as long as there was invasion of the carcinoma into the stalk.

2.3.1.3 Exclusion Criteria

Cases were excluded from the study if the CRC identified in our time period was the recurrence of an earlier CRC, diagnosed before January 1, 1997, or if the CRC was a metastasis from a distant primary. Patients with carcinoid tumours, and those who had known FAP were also excluded, as were those with carcinoma *in situ* and mucosal carcinoma. Patients who did not have their tumours surgically resected, or for whom the resected surgical specimens were unavailable were excluded from the molecular study, but still included in the family history study, as long as they met the age criteria for that study. F.C. used MEDITECH to confirm the home address of each patient to ensure that they lived in the Avalon Peninsula of the province. Those that lived outside of this region were excluded from the study.

2.3.1.4 Selection of Surgical Specimens

Once we had a complete list of eligible cases, I searched for the haematoxylin and eosin (H&E) -stained slides of the pathology specimens removed at surgical resection. These surgically-removed tissues were fixed in formalin and embedded in paraffin. I reviewed every slide for each of the tumours diagnosed up to age 74 years with pathologist Dr. Desmond Robb. With his help I chose the most appropriate surgical specimen block to represent normal colonic tissue, tumour tissue, and mixed normal and tumour. We selected the blocks of normal and tumour tissue based on the number of viable cells, and the lack of contaminating cells seen on the slide. The normal blocks could have no tumour cells, and the tumour blocks were chosen to have the fewest number of normal colonic epithelial cells, while not including large areas of necrotic tumour. The normal block and tumour block were used for DNA extraction for MSI analysis. The mixed blocks had both normal, non-malignant, colonic mucosal cells, as well as an area of viable tumour cells. An equal balance of normal and tumour cells was preferred, although not always available. The mixed blocks were used for IHC staining, with the normal, non-cancerous cells providing a convenient control for the tumour when examining protein expression. If a suitable mixed block was not available, then sections from the separate normal and tumour blocks were used together for IHC, mounted next to each other on the same glass slide, if size allowed. Once we selected which blocks would be used, I located and collected each of the chosen blocks from hospital storage. This

same process was used by P.S., also under the guidance of Dr. Desmond Robb, for selection of specimens for the patients diagnosed age 75 years and older.

2.3.1.5 MSI Analysis

Five 10-micron thick sections were cut from each of the formalin-fixed, paraffin-embedded normal and tumour specimen blocks, using a microtome. The tissue was de-waxed with two washes of toluene and then rehydrated in alcohol. DNA was extracted using the Qiagen DNeasy Tissue Kit, as per the manufacturer's instructions (Qiagen Inc., Mississauga, ON, Canada).

Both the extracted normal and tumour DNA were amplified by polymerase chain reaction (PCR), using primers for each of the five microsatellite markers in the panel recommended for microsatellite instability analysis by the National Cancer Institute (NCI) [61] (primer sequences are included in Appendix A, and recipes for all solutions described below are included in Appendix B). These included the mononucleotide markers BAT25 and BAT26, and the dinucleotide markers D5S346, D2S123, and D17S250. Tumours with instability at two or more of the five markers were considered to have a high level of microsatellite instability (MSI-H). Tumours with no instability were considered stable (MSS). Tumours with instability at one marker were considered to have a low level of microsatellite instability (MSI-L). MSI-L tumours were investigated with a second panel of five microsatellite markers, to determine if they could be re-classified as MSI-H. The second set of markers included mononucleotide BAT40,

and dinucleotides D17S787, D18S58, D20S100, and D7S519, and was recommended by the NCI for this purpose [61]. To be re-classified as MSI-H a tumour must be instable for at least thirty percent of all tested loci.

The patients diagnosed older than age 74 years were analyzed by P.S. using a minimized protocol. MSI analysis for these patients was done using only one microsatellite marker. The mononucleotide BAT26 was reported by Hoang and colleagues [241] to be sufficient for MSI analysis, and this was supported by others [475]. As well, this marker was shown to have 100% concordance with MSI-status in the younger cohort, so was used without the other markers for the older cases.

For each PCR amplification, one microlitre of the extracted DNA was included in a 25 microliter reaction volume along with: 1x reaction buffer, 1x Qiagen Q solution (Qiagen Inc., Mississauga, ON, Canada), 0.2 millimolar of each of the four deoxyribonucleotide triphosphates (dNTPs), 1.5 millimolar magnesium chloride, 0.20 units of Qiagen HotStar Taq DNA polymerase (Qiagen Inc., Mississauga, ON, Canada), and 0.6 micromolar of each of the forward and reverse primers. The reactions were run in a thermocycler, starting with a 15 minute denaturation at 95 degrees Celsius, followed by 36 cycles of 45 second denaturation at 94 degrees, 45 second annealing at 57 degrees, and 60 seconds elongation at 72 degrees. The reactions were finished by a seven minute elongation period at 72 degrees, and then stored at four degrees.

Individual PCR products from matched normal and tumour tissues were electrophoresed side-by-side on six percent polyacrylamide gels, with 7 molar urea concentration. The flat (shorter) glass plate of the gel apparatus was treated with Bind

Silane (g-methacryloxypropyltrimethoxysilane) to ensure the gel would stick, to enable staining. Once the gel was loaded and allowed to polymerize, the gel apparatus was run to preheat the gel to 50 degrees before loading PCR products. A formamide-based loading dye was added to individual PCR products, and this was denatured at 95 degrees for 2 minutes, then placed on ice prior to loading into the gel. Matched normal and tumour DNA were run side-by-side to enable comparison of microsatellites.

Once the electrophoresis was complete, the apparatus was disassembled with the polyacrylamide gel remaining bound to the shorter glass plate. The DNA in the gel was stained with silver nitrate by submerging the plate in a series of solutions, using the protocol described by Promega in their SILVER SEQUENCE™ DNA Sequencing System Technical Manual (Promega Corporation, Madison, U.S.A.).

Stained microsatellite DNA was interpreted independently by two people, Dr. Ban Younghusband and I, and any discrepancies in interpretation were discussed and resolved, or analysis was repeated. Instability at any marker was defined as the gain of new alleles in the tumour DNA when compared to the matched normal DNA from the same patient. Loss of heterozygosity (LOH) was noted if alleles were lost in the tumour DNA when compared to normal.

2.3.1.6 Immunohistochemical Analysis

Immunohistochemical (IHC) analyses were performed on all available tumours using antibodies against human mismatch repair proteins MLH1, MSH2, and MSH6. Five micron-thick sections were cut from the formalin-fixed, paraffin-embedded tissue blocks, and mounted on glass microscope slides treated with Histogrip (Zymed Laboratories Inc, U.S.A.). The Histogrip was used to improve tissue adherence to the slides, thereby decreasing tissue loss during the staining process.

The tissue was deparaffinized with two five-minute immersions in toluene, then rehydrated in a series of immersions in alcohol of decreasing concentration. Endogenous peroxidases were blocked by a ten-minute incubation with three percent hydrogen peroxide, after which the slides were rinsed in running tap water for ten minutes. Antigen retrieval was carried out by immersing the slides in ten millimolar citrate buffer (pH 6.0), then heating at full temperature and pressure in a standard, household pressure cooker for nine minutes.

One of two different immunohistochemistry kits was then used for the remaining stages of staining. IHC for MSH2 was performed with the DAKO ENVISION System (DAKO Corporation, California, U.S.A.), while MLH1 and MSH6 IHC were performed using the DAKO LSAB+ System, HRP kit.

For MSH2, after antigen retrieval, the cooled slides were treated in a humidifying chamber for twenty minutes with five percent goat serum in water, to limit nonspecific antibody binding. Anti-MSH2 polyclonal antibody (Oncogene Research Products, Massachusetts, U.S.A.) was diluted to a one in thirty concentration with one percent bovine serum albumin in phosphate buffered saline (PBS). The excess goat serum was removed, and antibody applied to each tissue section. The slides were then incubated in a

humidifying chamber for 24 hours at four degrees Celsius. The slides were then rinsed with PBS, then covered in ENVISION (DAKO Corporation, California, U.S.A.) and placed in the humidifying chamber for 30 minutes. ENVISION is comprised of a secondary antibody, biotin, streptavidin, and horseradish peroxidase. The slides were then rinsed again with PBS, and covered with 3,3'-diaminobenzidine tetrahydrochloride (DAB), and left for seven minutes, before being rinsed again with PBS, then running tap water. The tissue was then counter-stained with a three minute immersion in haematoxylin. Excess counterstain was removed by rinsing in water, followed by a quick immersion in one percent acid alcohol. After another rinse in water the slides were put in a solution of Scott's tap water until the counterstain turned blue, which took approximately one minute. The slides were then rinsed again in water, and dehydrated in a series of immersions in alcohol of increasing concentration. The slides were then treated with toluene, mounted with a drop of Permount (Fisher Scientific, New Jersey, U.S.A.), and cover-slipped.

The IHC protocols for the anti-MLH1 and anti-MSH6 antibodies were different only in the IHC kit used. After antigen retrieval the cooled slides were treated for twenty minutes in a humidifying chamber with DAKO Universal Blocker to reduce non-specific antibody binding. Excess blocker was removed, and the antibody applied. The slides were then incubated in a humidifying chamber for 24 hours at four degrees. The anti-MLH1 mouse anti-human monoclonal antibody (BD Biosciences, Ontario, Canada) was used at a one in fifty dilution. The anti-MSH6 mouse anti-human monoclonal antibody (BD Biosciences) was used at a one in seventy-five dilution. The slides were then rinsed with PBS, then covered with a biotinylated secondary rabbit anti-mouse antibody from

the DAKO kit, called Link Antibody, and incubated in the humidifying chamber for 30 minutes. After another rinse with PBS, the DAKO Streptavidin Peroxidase was applied, and again the slides were left in the humidifying chamber for 30 minutes. The slides were again rinsed in PBS, then treated with DAB for seven minutes. From here the procedure to finish the staining process was as described above for MSH2.

2.3.1.7 Interpretation of IHC staining

Slides stained by IHC for expression of MLH1, MSH2, and MSH6 were scored by pathologist Dr. Desmond Robb and myself. Stained slides for the older patients were scored by Dr. Robb and P.S.. For each of these MMR proteins, expression was reported when there was non-equivocal staining observed in the nuclei of the tumour cells. Both the intensity and frequency of the staining were recorded. The intensity of staining was recorded on a four-point scale as follows: 0 = no stain; 1 = weak stain; 2 = moderate stain; and 3 = strong stain. The frequency of staining was recorded on a five-point scale as follows: 0 = all tumour cells were negative; 1 = 1%-10% positive; 2 = 11%-50% positive; 3 = 51% to 75% positive; and 4 = greater than 75% of tumour cells were positive. Normal colonic epithelial cells and stromal lymphocytes were used as positive internal controls. If they did not show protein expression the stain was repeated. If, after repetition, a clear stain was not observed in the normal control cells, then a result was not recorded for that tumour.

2.3.1.8 Statistical Analysis

The proportion of MSI-H tumours, those with MMR deficiency by IHC analysis, and patient sex and age at diagnosis were compared using Pearson's Chi square tests. Fisher's exact test was used instead when comparison required a two-by-two table and the count in any one cell was less than 5. Significance values were 2-tailed, and all variables were considered significant if $P < 0.05$. Chi square tests were performed using SPSS for Windows, version 15.0 (SPSS, Chicago, U.S.A.).

2.3.2 Results

2.3.2.1 Study Participants – Molecular Project

Two hundred and four cases of colorectal cancer were identified by the NCTRF as being diagnosed under the age of 75 years on the Avalon Peninsula during 1997 and 1998. An additional 67 cases, not included on the NCTRF list, were identified through pathology reports. Of the total 271 cases, 79 had to be excluded, leaving 192 in the study. Half of the excluded cases (40) lived outside the Avalon Peninsula, and nine were recurrent CRCs. The others were excluded for the various reasons listed as exclusion criteria, the most common being the non-availability of the surgical specimens.

Although the hospital discharge records were used for the identification of cases, they did not identify any additional patients, and so were not used for the older cohort. The NCTRF and pathology reports were used to identify the patients diagnosed older than age 75 years. I searched both of these, and identified a total of 127 cases. Again, a number of cases were excluded from the study. Fourteen cases identified by the NCTRF were not surgically resected, so no tumour tissue was available for analysis. Pathology specimens could not be located for four cases. Two more cases identified by the NCTRF had to be excluded, one because the patient was diagnosed younger than age 75 (and had appropriately been included in the younger cohort of this study), and the other because the diagnosis date fell outside the time period of interest. On examination of pathology slides another three cases were excluded from the study, as two had non-invasive carcinoma in polyps, and one did not have cancer of the colon. Ultimately, 104 cases were found to be appropriate and were included in this cohort of older patients.

2.3.2.2 Molecular Analysis

Of the 192 cases diagnosed younger than 75 years of age, microsatellite analysis could be carried out on 191. There were no appropriate blocks for DNA extraction in the remaining case, so it was included only in the immunohistochemistry analysis, where it was found to express all three tested MMR proteins. One of the 104 cases in the older cohort had the same limitation, and was also included only in the IHC analysis, where it too was found to express all three MMR proteins tested.

Of the 191 cases included in the microsatellite analysis for the younger cohort 26 (13.6%) were MSI-H, and nine (4.7%) were MSI-L. The remaining 156 (81.7%) were MSS. Examples illustrating MSI at the markers tested is given in Figure 2.2. After testing for the second set of five microsatellite markers, all nine MSI-L cases remained MSI-L. Each of these tumours was stable at each of the five markers (with the exception of one case, which showed loss of heterozygosity at two of the five markers). For the remainder of the analysis the MSI-L cases were included as stable, along with the MSS cases. Nine tumours showed LOH at one of the three dinucleotide markers, all of these were MSS.

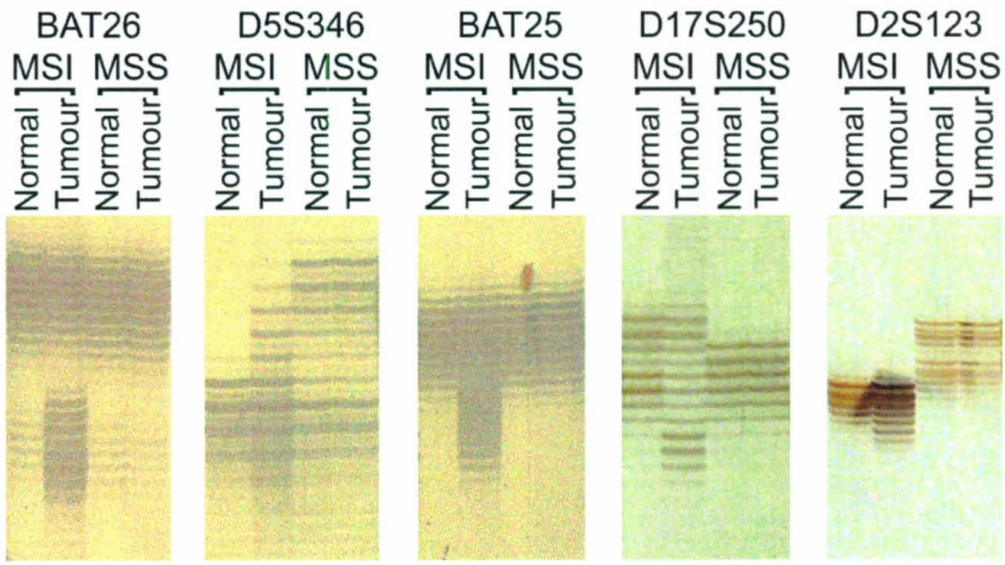


Figure 2.2. Microsatellite instability analysis
New smaller (BAT26, BAT25, D17S250 and D2S123) or larger alleles (D5S346) seen in the tumour when compared to normal DNA from the same case, indicating MSI, and no change seen between normal and tumour DNA, indicating MSS.

The rate of microsatellite instability found in the younger cohort was comparable to the older cohort, in which 13 (12.6%) of the 103 cases were MSI-H. Since only one marker was tested for this cohort MSI-L could not be determined, so the remaining 90 cases (87.4%) were considered MSS. Overall, for the entire population 39 (13.3%) of 294 cases were MSI-H.

Twenty-six (13.5%) of the 192 cases in the younger cohort had a deficiency of one or more of the MMR proteins by IHC. Twenty-two of these cases were MSI-H, while the other four were MSS. Of the 22 IHC deficient MSI-H tumours, 13 had loss of expression of MLH1 only. Two were deficient for MSH6, another two were deficient in MSH2, and five were deficient in both MSH2 and MSH6. The four MSS tumours showing IHC deficiency all had loss of MSH6, and one of these also had loss of MLH1. Examples of IHC results for MSH2 and MLH1 are given in Figure 2.3.

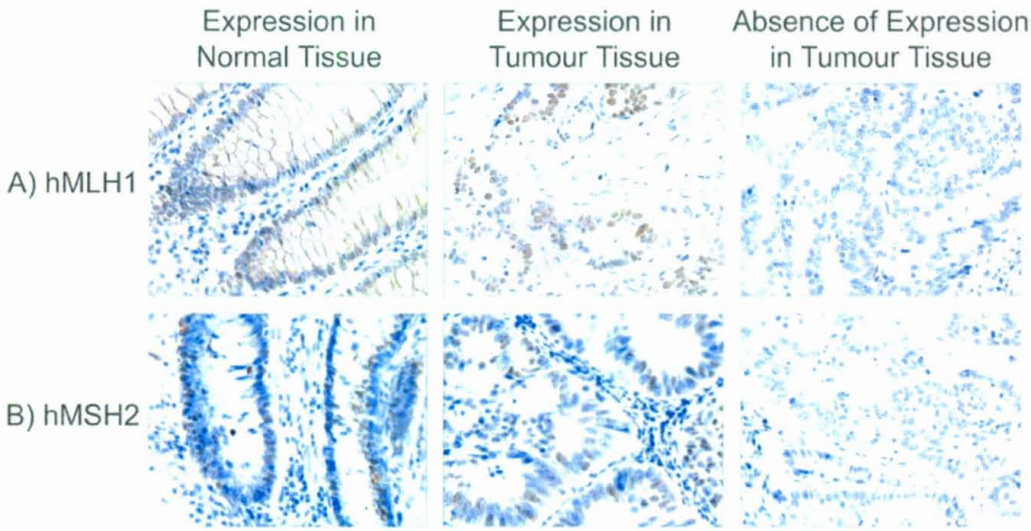


Figure 2.3. Immunohistochemistry of mismatch repair proteins MSH2 and MLH1 A brown stain in the tumour cell nuclei indicates expression of protein.

In the older cohort, 12 (11.5%) of the 104 tumours had a deficiency in expression of one or more of the MMR proteins. Nine of these were in MSI-H tumours, of which eight were deficient in MLH1 and one was deficient in both MSH2 and MSH6. The other three were in MSS tumours, of which two were deficient in MLH1 and the other in MSH6. For the entire population including both cohorts, 38 (12.9%) of the 294 tumours had loss of at least one of the three mismatch repair proteins.

Although both deficiency of a MMR protein by IHC and high levels of microsatellite instability were each observed in approximately 13% of the overall population, there was not complete correlation between these two indicators of mismatch repair deficiency. Eight (20.5%) of the 39 MSI-H tumours showed expression of all three MMR proteins, while the other 31 were deficient in at least one. Similarly, of the 38 tumours deficient in protein expression by IHC, seven (18.4%) were MSS. The IHC and

MSI correlated, indicating MMR deficiency for 31 (10.5%) of the 294 tumours, and another 15 (5.1%) were either MSI-H or IHC deficient, but not both. These combine for a total of 46 tumours, or 15.6% of all tumours, with at least one molecular feature of mismatch repair deficiency. The correlations of MSI and IHC data are given in Table 2.1.

Table 2.1. Correlation of microsatellite instability and immunohistochemistry results A) Patients diagnosed younger than age 75 years, B) patients diagnosed age 75 years or older, C) all patients

A

Age <75 yrs	MSI	MSS	Totals
IHC deficient	21	4	25
IHC intact	5	161	166
Totals	26	165	191

B

Age >=75 yrs	MSI	MSS	Totals
IHC deficient	9	3	12
IHC intact	4	87	91
Totals	13	90	103

C

All ages combined	MSI	MSS	Totals
IHC deficient	30	7	37
IHC intact	9	248	257
Totals	39	255	294

The data for both cohorts was combined and categorized by age, sex, MSI status, and IHC results. The resulting table is presented as Table 2.2. Deficiency of each of the three mismatch repair proteins by immunohistochemistry was significantly associated with MSI-H status, all with *P*-values less than 0.001 (Table 2.3).

Table 2.2. Description of entire population, categorized by age, sex, microsatellite instability status and immunohistochemistry results

Age (years)	Total	Female (%)	MSI-H (% for age)	MSI & Female (% of MSI-H)	MSI-H & IHC deficient				MSI-H & IHC intact	MSS & IHC deficient	Total IHC deficient	Total MSI-H and/or IHC deficient (% for age)
					MSH2	MLH1	MSH6	MSH2 &6				
0-29	1	1 (100.0)	0 (0)	0 (0)	0	0	0	0	0	0	0	0 (0)
30-39	8	8 (100.0)	3 (37.5)	3 (100.0)	1	0	1	1	0	0	3	3 (37.5)
40-49	20	7 (35.0)	4 (20.0)	2 (50.0)	1	2	0	1	0	0	4	4 (20.0)
50-59	64	26 (40.6)	9 (14.1)	4 (44.4)	0	4	1	2	2	1	8	10 (15.6)
60-69	72	35 (48.6)	8 (11.1)	6 (75.0)	0	5	0	1	2	2	8	10 (13.9)
70-79	75	31 (41.3)	8 (10.7)	5 (62.5)	0	4	0	0	4	1	5	9 (12.0)
80-89	50	28 (56.0)	6 (12.0)	6 (100.0)	0	4	0	1	1	3	8	9 (18.0)
90+	4	3 (75.0)	1 (25.0)	1 (100.0)	0	1	0	0	0	0	1	1 (25.0)
Total	294	139 (47.3)	39 (13.3)	27 (69.2)	2	20	2	6	9	7	37	46 (15.6)

Table 2.3. Association of microsatellite instability with patient age, sex, and tumour immunohistochemistry results

	MSI-H (n=39)	MSS (n=255)	*P*-value
Female	27	112	0.003*
Age <50 years	7	22	0.083
MSH2 deficient	8	0	<0.001*
MLH1 deficient	20	3	<0.001*
MSH6 deficient	8	5	<0.001*

* statistically significant, $P < 0.05$

Of the 294 patients 139 (47.3%) were female. Although females represented just less than half of the total population, they were over-represented among those with MSI-H tumours. Twenty-seven (69.2%) of the 39 MSI-H tumours were from female patients. This is a statistically significant difference, with *P*-value 0.003, as seen in Table 2.3.

Patients were assigned to one of six categories based on age at the time of diagnosis: younger than age 40 years, 40 to 49, 50 to 59, 60 to 69, 70 to 79, and 80 years or older. Figure 2.4 shows the distribution of cases by age at diagnosis, sex, and MSI status. Figure 2.5 illustrates the trend that was observed when looking at the proportion of MSI-H tumours within each of the age categories, both for the whole population, and by sex. For each of the six age categories the percent of women with MSI-H tumours was greater than the percent of men with MSI-H tumours. The difference was most dramatic for the oldest age category, with the smallest difference in the 50 to 59 year age group. The only age category for which the difference in MSI status was statistically significant between the two sexes was for those diagnosed older than age 80 years. In this cohort 7 of 31 females had MSI-H tumours, whereas none of the 23 males had MSI-H tumours. This is significant with a *P*-value of 0.003.

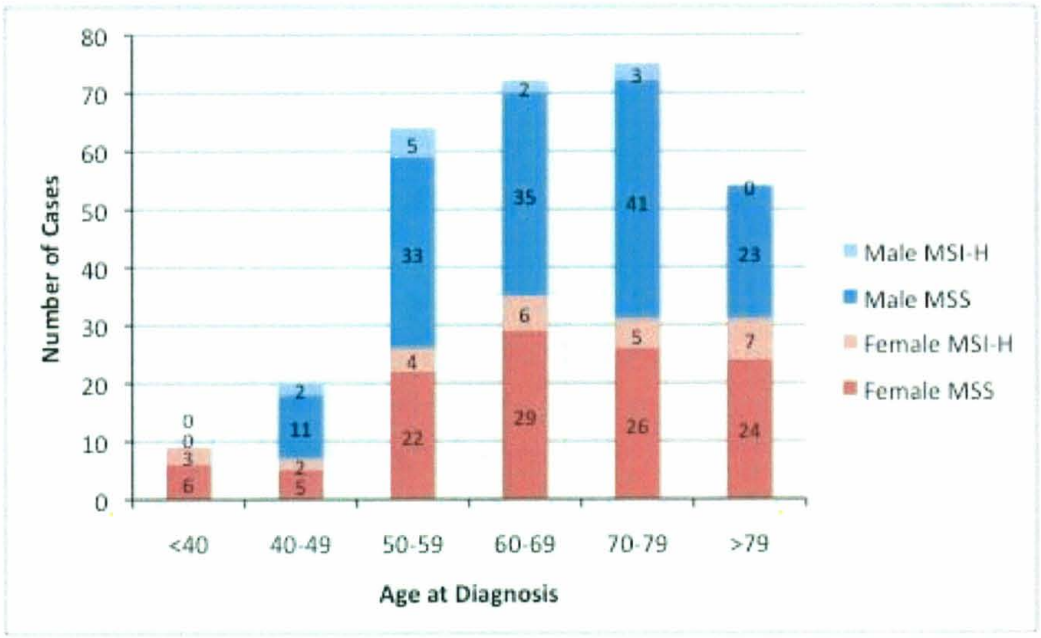


Figure 2.4. Distribution of population by age at diagnosis, sex, and microsatellite instability analysis

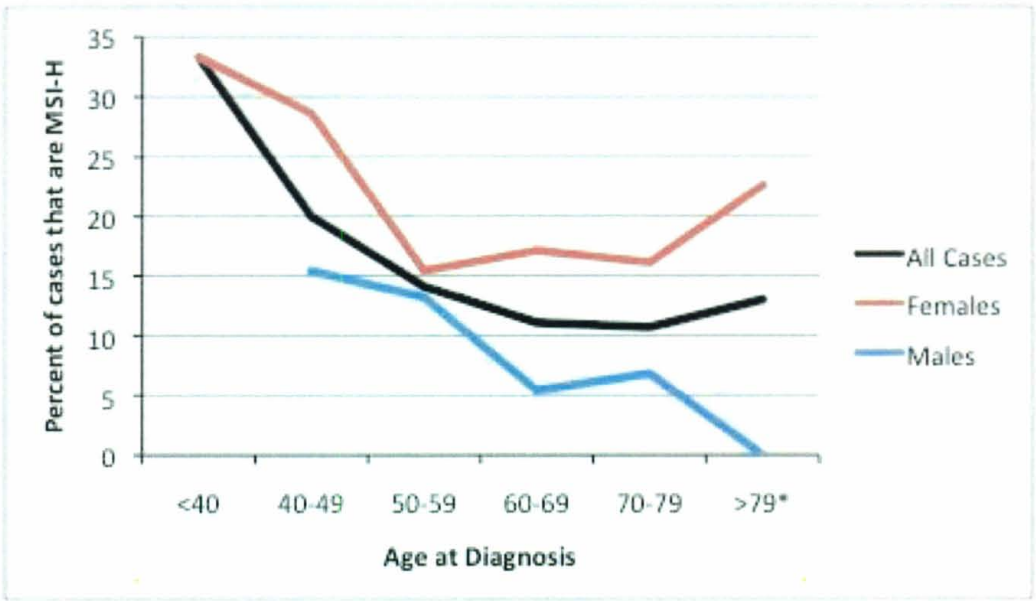


Figure 2.5. Age and sex distribution of all cases with high levels of microsatellite instability
The raw numbers used to compute the above percentages were given in Figure 2.4.
* Diagnosis at age greater than 79 years is the only age category for which there is a statistically significant difference in the proportion of MSI-H tumours between the sexes, P=0.016.

When sex was excluded from this analysis, age at diagnosis was not significantly associated with MSI status. This was tested for four different methods of categorizing age, although only one of these is shown in Table 2.3. The age at diagnosis was divided into two categories, younger than age 50 years and 50 or older, giving a P-value of 0.083. Age was also divided into younger than age 60 years and 60 and older, giving a P-value of 0.197. When divided into three age categories (younger than 50, 50 to 74, and 75 years and older), the P-value was 0.187. Finally, when divided into five age categories (younger than 50 years, 50 to 59, 60 to 69, 70 to 79, and 80 years and older) the P-value was 0.509. All methods of categorization showed no significant association between age

at onset and MSI status. Although not statistically significant, Figure 2.5 shows a trend towards an increased proportion of MSI-H tumours seen in the youngest age groups, those patients diagnosed younger than age 50 years, as well as a small increase in the oldest age group, those diagnosed at age 80 years or older.

2.3.3 Discussion

We have characterized the molecular features of the incident colorectal tumours diagnosed on the Avalon Peninsula of Newfoundland in 1997 and 1998. Of the 294 patients included in the study, we found that the tumours of 31 (10.5%) had both high levels of microsatellite instability and loss of expression of at least one mismatch repair protein, thereby indicating a defect in the mismatch repair system. Another 15 tumours (5.1%) had one or the other of these features, but not both, for a total of 46 tumours (15.6%) with features of a mismatch repair defect. This is as expected, since approximately 13% of sporadic tumours have MMR deficiency [60,474], as do the vast majority of Lynch Syndrome cases, which account for between 1% and 3% of all CRC [4,6].

In 2005, Hampel and colleagues [6] studied more than 1000 population-based incident CRCs from Ohio, U.S.A. They had similar results to those of this study for microsatellite instability, identifying this feature in 12.7% of tumours, whereas we identified it in 13.3%. However, they had better correlation between the MSI and IHC results. In their study, absence of staining in IHC analysis was observed in 123 of the 132

MSI-H tumours, resulting in a sensitivity of 93.2%. In our study however, IHC identified only 31 of the 39 MSI-H tumours, for a sensitivity of 79.5%.

Our MSI and IHC results did not correlate for 15 tumours, as mentioned above. Eight tumours were MSI-H, but had intact staining for all three MMR proteins by IHC. These could perhaps be explained by an epitope-sparing mutation in one of the MMR genes, or a mutation in a MMR gene that was not examined. They could also be explained by certain mutations in the *MLH1* or *MSH6* genes, as described below. Another seven tumours were MSS, but had loss of expression of one or more proteins by IHC. These discrepancies could also be explained in part by observations from other researchers that some *MLH1* and *MSH6* mutations can lead to variable MSI and IHC results [243,244,476-478].

Seven of the eight MSI-H tumours that were not identified by IHC in our study had a weak level of staining for the MLH1 protein by IHC. One of these seven also had a weak level of staining for MSH6, and another had a weak level of staining for all three proteins. The eighth tumour had a strong level of staining for all three proteins tested. Staining intensity can be affected by many factors, including the quality of tissue fixation, antigen retrieval, antibody concentration, incubation time, and staining procedure. However, it is also possible that some of these tumours could have a mutation in *MLH1*. The seven MSI tumours with intact IHC were not the only ones with a weak score for MLH1; forty-nine (18.4%) of the 267 tumours with a positive score for MLH1 had a weak-intensity stain (data not shown). However, there is a statistically significant difference in the proportion of weakly stained tumours when comparing these seven of eight MSI-H tumours, to the 42 of 259 remaining positively scored tumours, with a *P*-

value less than 0.001. This indicates that the weak MLH1 staining is likely associated with the discrepancy seen between the MSI and IHC results for these cases.

A paper published in 2005 by Mangold and colleagues [476] demonstrated that weakly positive staining for MLH1 can be observed in tumours from some *MLH1* mutation carriers. In their study of 82 MSI tumours from known MMR mutation carriers, IHC deficiency identified 100% (38 of 38) of tumours from *MSH2* mutation carriers, but only 66% (29 of 44) of tumours from *MLH1* mutations carriers.

An earlier study by Raevaara and colleagues [477], from 2003, examined the tumours from individuals in a family with a pathogenic *MLH1* mutation. They found that not only did IHC results for MLH1 vary from tumour to tumour, but MSI status did as well. The protein was not detected by IHC in two tumours, but was, at a low intensity in a third. MSI results indicated that some tumours were MSI-L, while others were MSI-H. The mutation segregating in that family was *MLH1* del616, a single amino acid deletion that the authors reported to be one of the most widespread recurring mutations in Lynch Syndrome.

While Raevaara's paper described the inconsistencies seen in MSI and IHC results from one particular non-truncating mutation [477], a 2009 paper from Zighelboim and colleagues [478] described a truncating mutation in the *MLH1* gene. This pathogenic germline deletion of exons 14 and 15 did not affect the epitope for the antibody used in analysis, and resulted in normal expression of MLH1 as assessed by IHC.

Given that the above data demonstrates that some pathogenic *MLH1* mutations lead to MSI-H tumours with intact normal or weak staining by IHC, it is possible that this

could explain at least some of the discrepancy seen in the MSI-H tumours with intact IHC results from our study. It is also possible that some of these individuals had a germline mutation in the *MSH6* gene. *MSH6* mutations could also explain some of the discrepancies seen in the seven MSS tumours that were deficient for one or more proteins by IHC. Five of these tumours were deficient in MSH6, while the other two were deficient in MLH1.

A 1999 paper by Ku and colleagues [243] demonstrated that *MSH6* mutations do not always lead to microsatellite instability. They described two colon cancer cell lines in which *MSH6* mutations did not result in MSI.

Berends and colleagues reported in a 2002 paper [244] that tumours from patients with *MSH6* germline mutations had mixed results for both MSI status and IHC. They studied six endometrial and 22 colorectal tumours from 25 patients with germline mutations in *MSH6*. IHC demonstrated a loss of MSH6 expression in most (but not all) of the tumours of patients with truncating mutations, however, most of those with missense mutations had intact staining for MSH6. Fourteen (54%) of the tumours were MSI-L, while the others were MSI-H.

An unexpected finding in this study was the paucity of young males with CRC, especially of those with MSI-H CRC. A total of 9 females were diagnosed younger than age 40 years, but there were no males in this age group (Figure 2.4). Looking at all of those diagnosed younger than age 50 years brings the number closer to parity, with 16 females, and 13 males, however, within this group there are more females than males with MSI-H tumours, as five (31.3%) of the 16 tumours from female patients were MSI-

H, as compared to only two (15.4%) of the 13 tumours from males. We had expected to find higher numbers of males since males have a higher overall incidence of CRC than females [472]. We had expected this discrepancy would be exaggerated in the young MSI-H patients, since Lynch Syndrome affects more young patients, and more males than females with Lynch Syndrome get CRC [479]. In unpublished data from the interdisciplinary study for which this was a pilot, there were more males than females with CRC, and many more males with MSI-H CRC in those patients diagnosed younger than 50 years of age. In that study there were 95 patients diagnosed with CRC younger than age 50 years, of which 62 (65.3%) were males. In that cohort 16 tumours were MSI-H, of which 15 (93.8%) were from male patients (Dr. Roger Green, personal communication). It is unclear why the pilot study has discrepant results, however, the numbers were small, so even a small variation could have a large effect on observed results.

While the relatively high number of females with MSI-H tumours within the youngest age category was unexpected, the high number of females with MSI-H tumours seen in the oldest age category was expected. We found a significantly higher proportion of females than males with MSI-H CRC in the oldest age group, those diagnosed at age 80 years or older, with *P*-value 0.016. Most of the older-onset MSI-H tumours reported by others have a mutation in the *BRAF* gene [480,481]. In these cases the microsatellite instability is due to hypermethylation of the *MLH1* promoter region, and not to a germline MMR mutation, as would be seen in Lynch Syndrome [261,291,292]. CRCs with these

features are associated with older females [308,481], and could explain the trend we found in this study.

A weakness in this study was the use of only one microsatellite marker for the assessment of MSI status in the cohort of patients diagnosed age 75 years and older. While the evidence at the time supported the use of only BAT26 [241,475], more recent studies have demonstrated that it should not be used alone for screening for MSI status, especially for the purpose of identifying Lynch Syndrome. The BAT26 microsatellite is located immediately downstream of exon five in the *MSH2* gene, and deletions within the *MSH2* gene are a frequent cause of Lynch Syndrome [482]. Large intragenic deletions of *MSH2* account for up to 15% of known MMR mutations, and half of these span the BAT26 microsatellite region [483-485].

Pastrello and colleagues published their findings in 2006 [486] that some MSI-H tumours appear to be stable at BAT26, due to deletions within the *MSH2* gene. They studied ten Lynch Syndrome families with large intragenic MSH2 deletions, encompassing exon five and intron five, including the BAT26 microsatellite region. Thirteen (68%) of the 19 MSI-H tumours tested appeared to be stable at the BAT26 marker. They determined this to be due to the absence of target DNA within the tumour cells, resulting in PCR amplification of only contaminating normal cells from within the tumour. Jaskowski and colleagues supported these findings with their own publication the following year [485].

Since it was not an objective of this study to determine if either MSI or IHC was the superior test for identifying Lynch Syndrome, and because neither MSI analysis nor

IHC identified 100% of the tumours with molecular features of MMR deficiency in our study, we cannot recommend one method to be used exclusively to screen for Lynch Syndrome in a population-based setting. Family history also plays an important role in identifying Lynch Syndrome, and will be discussed further in the following section.

2.4 Family History Study

2.4.1 Methods

2.4.1.1 Identification of eligible cases

The family history study included the subset of patients from the molecular study that were diagnosed between the ages of 20 and 69 years, inclusive. Availability of surgical specimens was not a requirement for inclusion in the family study. With the exceptions of age and specimen availability, inclusion and exclusion criteria were the same as those used for the molecular study (described in sections 2.3.1.2, and 2.3.1.3).

2.4.1.2 Collection of Family History Data

We sent a standardized letter to the attending surgeon, or family doctor when necessary, of patients identified as eligible for the study. The letter was signed by Dr. William Pollett, the surgeon involved in the study, and introduced the project and the research team (Appendix C). The physicians made primary contact with the patients, or with next of kin if patients were deceased, and provided them with a letter introducing the study and inviting them to participate (Appendix D). The patients could then opt into or decline participation in the study. Once oral consent was obtained, a consent form was sent to the patients and proxies, which explained the study in detail (Appendices E and F). After written, informed consent was obtained, an interview was arranged with F.C. for the purpose of collecting a detailed family history. At the time of interview, patients also signed the Newfoundland and Labrador Medical Genetics Program DNA consent form (Appendix G), and a form for release of medical information (Appendix H).

Information was collected at interview to construct a three-generation pedigree, which included details about types of cancer within the family, age at diagnosis, age at death, cause of death, and other related medical information. Forms for release of medical information were provided to patients to give to family members who had been diagnosed with cancer. Medical records were reviewed for accuracy whenever possible (Appendix I).

2.4.1.3 Classification of Risk

The pedigrees were evaluated to determine the risk of Lynch Syndrome in the proband and their family. Risk was classified as high, intermediate or low. High risk families met at least one of the following three criteria: Amsterdam Criteria I (Figure 1.11.), Amsterdam Criteria II (Figure 1.12.), or Age and Cancer Modified Amsterdam Criteria (ACMAC). ACMAC follows the definition of the ACI, but includes the tumour spectrum of ACII and a modified age of onset. Rather than 50 years, the youngest individual must be 60 years or younger at diagnosis to meet ACMAC. Intermediate risk families did not meet the criteria for high risk, but met at least one of the following seven criteria: (1) proband and two relatives with any of the Lynch Syndrome-associated cancers (as defined by ACII), with two of the three being first degree relatives; (2) proband and any one relative with a Lynch-associated cancer diagnosed younger than age 35 years; (3) proband and relative both diagnosed with CRC younger than age 50 years; (4) proband diagnosed younger than age 35 years; (5) proband with multiple primary CRCs; (6) proband with another Lynch-associated cancer; and (7) proband meets at least one of the following pathologic criteria: multiple primary CRCs, at least five adenomas, inflammatory bowel disease, active chronic colitis, colitis-associated neoplasia, or other cancer. Low risk families met neither the high nor intermediate risk criteria.

2.4.1.4 Follow-up

A follow-up letter was sent to each of the probands that participated in this part of the study. This letter explained study outcomes, and included a personalized explanation of their family history risk assessment. Families with a high or intermediate risk of having a hereditary form of CRC were referred to a clinical geneticist at the Provincial Genetics Program for clinical follow-up and management. Results of molecular analyses were not released to patients, since they were considered to be of a research and not clinical nature.

2.4.2 Results

Using the NCTRF Cancer Registry, we identified 178 CRC cases diagnosed between ages 20 and 69 years that met our inclusion criteria. On further analysis, eight of those patients had to be removed from the study because they met one or more exclusion criteria. Four were removed because they were inappropriately diagnosed, two were not primary CRCs, but instead recurrences, one appeared in the registry twice, and one lived outside the included geographic region. The search of MEDITECH records identified another six patients, and the pathology reports identified another eight cases not included in the NCTRF Cancer Registry. These cases were reported to the NCTRF, and were subsequently registered. Of the 14 cases identified by these other sources, five were

excluded from our study; four were inappropriately diagnosed, and one could not be properly identified. The remaining nine were included in our study, for a total of 179 eligible cases in this cohort.

Of the 179 eligible patients, contact was made with 158 patients or proxies. Of these, 52 declined to participate, 106 verbally agreed to participate, but 27 did not sign a consent form to participate in the study. The remaining 79 consented to participate, and were interviewed. Two participating patients were first cousins, so their family risk was considered the same, and evaluated together. This resulted in a family history assessment of 78 families, for which molecular data was collected for 72. Results of the correlation of these data were published in a peer-reviewed journal. The manuscript is included in the following section.

2.5 Manuscript - "High Frequencies of Hereditary Colorectal Cancer in Newfoundland Likely Involves Novel Susceptibility Genes"

The population-based study described above (sections 2.1 to 2.4) was published along with data from a study of CRC patients identified through the Provincial Medical Genetics Program. The main finding described in the manuscript is that the correlation between molecular and family history data was not as strong as expected. Many families meeting high risk criteria for hereditary CRC had tumours lacking the molecular features which define the known hereditary CRC syndromes. The conclusion presented was that

there were likely strong and novel genetic causes of hereditary CRC yet unidentified, which were responsible for a large proportion of CRC in the Newfoundland population.

A version of this manuscript was published in October 2005, in *Clinical Cancer Research*, volume 11, issue 19, pages 6853 to 6861. Supplemental data is included in Appendix J. My role in this publication was in patient ascertainment and molecular analysis for the population-based cohort, as described in the preceding sections.

2.5.1 Authors and Affiliations

Michael O. Woods[1], Angela J. Hyde[1], Fiona K. Curtis[2], Susan Stuckless[2], Jane S. Green[1], Aaron F. Pollett[3], J. Desmond Robb[4], Roger C. Green[1], Marina E. Croitoru[3], Amanda Careen[1], Jason A.W. Chaulk[1], Jegan Jegathesan[3], John R. McLaughlin[3], Steven S. Gallinger[3], H. Banfield Younghusband[1], Bharati V. Bapat[3*] and Patrick S. Parfrey[2*]

1. Discipline of Genetics, Memorial University of Newfoundland, St. John's, Newfoundland, Canada
2. Clinical Epidemiology Unit, Memorial University of Newfoundland, St. John's, Newfoundland, Canada
3. Samuel Lunenfeld Research Institute, Mount Sinai Hospital, Toronto, Ontario, Canada
   Toronto, Ontario, Canada
4. Department of Pathology, Eastern Health, St. John's, Newfoundland, Canada

*These authors contributed equally to this work.

2.5.2 Summary

Background: Newfoundland has one of the highest rates of colorectal cancer (CRC) in North America. The most common hereditary form of CRC is Lynch Syndrome, caused by mutations in genes involved in mismatch repair (MMR). Our purpose was to determine the proportion of hereditary CRC and to determine the genetic basis of disease in both population and clinically referred cohorts from Newfoundland.

Methods: Seventy-eight CRC patients were accrued over a two year period from the Avalon Peninsula of Newfoundland. We also examined 31 Lynch Syndrome-like families which had been referred to the Provincial Medical Genetics Program. Tumors from probands were tested by immunohistochemistry for deficiencies in MLH1, MSH2 and MSH6 proteins; and tested for DNA microsatellite instability. Mutation analyses of *MLH1*, *MSH2* and *MSH6* were undertaken by direct sequencing and an assay to detect deletions, amplifications and rearrangements in *MSH2* and *MLH1*.

Results: We identified eight population-based families which fulfill the Amsterdam I or II criteria (ACI or ACII), four (50%) of which appear to have hereditary cancer not attributable to the most commonly mutated MMR genes. Also, in 16 of 21 (76%) referred families fulfilling ACI or ACII, no mutations were found in the three most commonly altered MMR genes, and tumor analyses corroborated these findings.

Conclusions: It appears that strong and novel genetic causes of hereditary CRC are responsible for a high proportion of CRC in this population. Conditions are suitable for the identification of these genes by linkage studies of large Newfoundland cancer families.

2.5.3 Introduction

The most strongly associated risk factor for colorectal cancer (CRC) is family history, with genetic susceptibility factors estimated to play a role in up to 35% of cases [487-489]. Known hereditary forms of CRC, including Lynch Syndrome and familial adenomatous polyposis (FAP), comprise less than 3% of all CRC [473,490]. Therefore, there remains a large proportion of familial clusters of CRC which have no identifiable hereditary cause [491-493]. These include families that are considered high risk kindreds according to the Amsterdam criteria established by the International Collaborative Group on HNPCC (ICG-HNPCC) [341,343]. The most common inherited form of CRC is Lynch Syndrome (HNPCC). This is an autosomal dominant cancer susceptibility syndrome characterized by a young age of onset of often right-sided colon cancer [494]. Additionally, tumors may develop in a variety of other sites, including the extra-colonic gastrointestinal tract, genitourinary tract, endometrium, ovaries and brain [494]. Lynch Syndrome accounts for between <1% and 5% of all CRC cases, depending on the population studied and the methods used to determine Lynch Syndrome status [473,494-496]. Under the Amsterdam I criteria (ACI) [343], the following are required for a diagnosis of Lynch Syndrome: (1) three

or more relatives with CRC, one of whom is a first degree relative of the other two; (2) colorectal cancer occurring in at least two generations; and (3) one or more CRC diagnosed before the age of 50 years (Familial Adenomatous Polyposis (FAP) should be excluded). Subsequently, due to concern that these criteria were too restrictive, they were expanded to include cancers of the endometrium, small bowel, ureter and renal pelvis [341]. These are known as Amsterdam II criteria (ACII).

Mutations in three mismatch repair (MMR) genes account for greater than 95% of the known mutations causing Lynch syndrome: *MSH2* [228,229], *MLH1* [232,233], and *MSH6* [497]. A number of other MMR genes have been associated with Lynch syndrome; however, their roles are not well established. Few kindreds have been identified with germline mutations in *PMS1* [234,498] or *PMS2* [234,499,500], and there is no convincing evidence for involvement of *MLH3* [501-503] or *MSH3* in high risk CRC families.

The Canadian province of Newfoundland and Labrador, with a current population of about 510 000, is considered a collection of genetic isolates [504,505]. Approximately 60% of the population lives in communities of fewer than 2 500 inhabitants and 41% in communities of less than 1 000 people [458]. Such a population structure and a willingness to participate in research studies, has made Newfoundland and Labrador an invaluable resource for the study of genetic disorders. In fact, a large Newfoundland CRC kindred was used to identify the importance of *MSH2* in Lynch Syndrome [228].

In the current study, we used a population-based study of CRC patients diagnosed in 1997 and 1998 in order to determine the proportion of hereditary CRC cases on the Avalon Peninsula of Newfoundland (Figure 2.6), and to determine the genetic basis of

disease in hereditary cases. We found a large number of families which appear to have a familial form of colorectal cancer which cannot be explained by mutations in MMR genes. We then analysed families referred to the Provincial Medical Genetics Program to determine if there are additional high risk families in Newfoundland that have hereditary CRC not associated with MMR deficiencies. Our results suggest there are probable novel genetic causes of hereditary CRC in this distinct population and that conditions are suitable for the identification of these factors by linkage studies of large Newfoundland cancer families.

2.5.4 Materials and Methods

2.5.4.1 Population-Based Probands

The study included all pathologically confirmed (International Classification of Diseases 9th edition codes 153 and 154; ICD-0, 1963) incident cases of CRC in patients between 20 and 69 years of age, diagnosed between January 1, 1997 and December 31, 1998 and were resident on the Avalon Peninsula at the time of diagnosis. Cases were identified using the provincial cancer registry at the Newfoundland Cancer Treatment and Research Foundation. A review of pathology reports and hospital databases confirmed that all cases had been reported to the tumor registry.

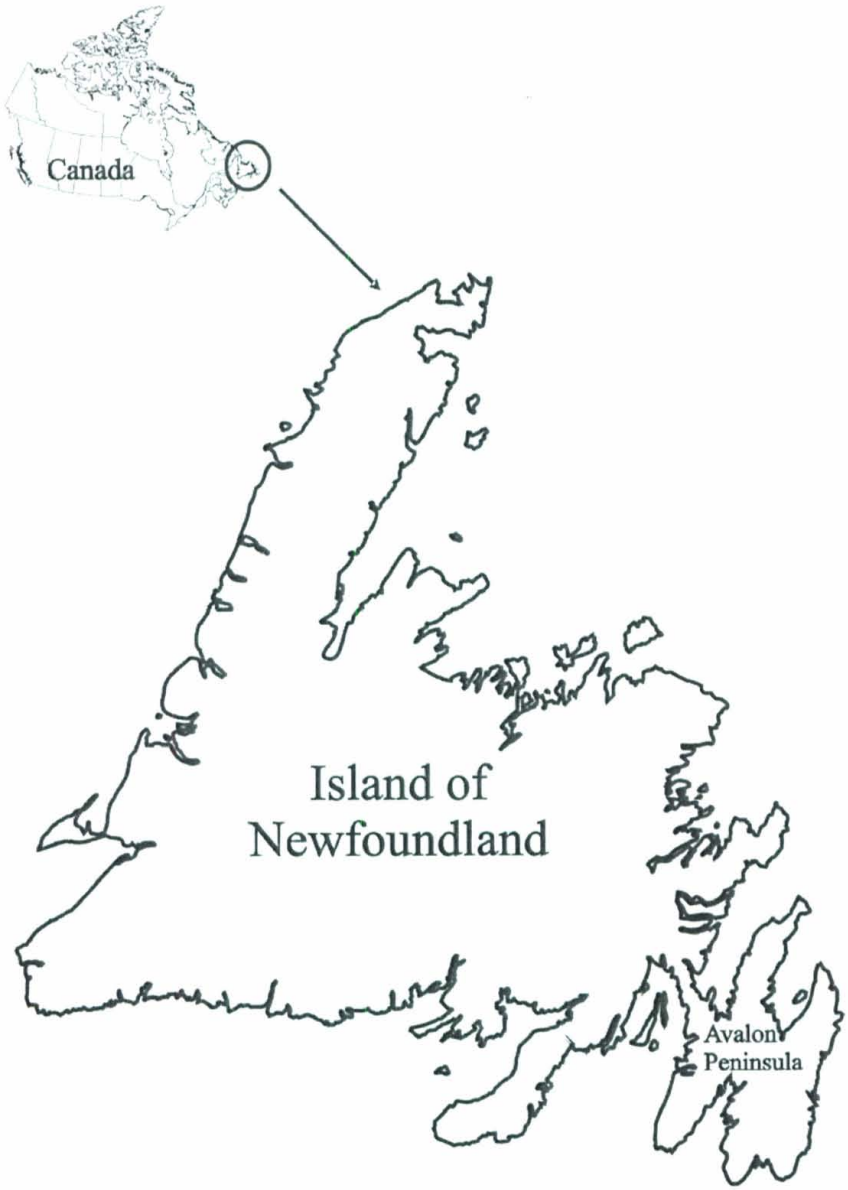


Figure 2.6. The island of Newfoundland
Part of the Province of Newfoundland and Labrador, the island of Newfoundland lies off of the east coast of Canada. The population-based probands were ascertained from the Avalon Peninsula where over 50% of the province's population resides. The origins of the families referred to the Provincial Medical Genetics Program are spread throughout the island.

One hundred and seventy nine cases met the inclusion criteria. Contact was made with 158 eligible cases/proxies (88%) of which 106 (67%) agreed to participate in the study. The final number of subjects who provided family histories was 79 (44% of eligible cases).

2.5.4.2 Families Referred to the Genetics Clinic

Over the last 20 years, more than 100 families have been referred to the Provincial Medical Genetics Program for assessment of possible hereditary CRC. Thirty-one families having the greatest family risk of HNPCC and with DNA samples available were examined in this study. All referred families had been previously screened and did not carry the most common Newfoundland founder HNPCC mutation (*MSH2* – c.942+3A>T) [506]. Informed consent was obtained from all subjects or an appropriate proxy. Ethics approval was granted by the Human Investigations Committee (HIC) of the Faculty of Medicine, Memorial University of Newfoundland, the Health Care Corporation of St. John's and the Avalon Peninsula Health Board.

2.5.4.3 MSI and Immunohistochemistry

For the population-based probands, matched tumor and normal tissue from formalin fixed, paraffin-embedded blocks were compared for MSI using at least five

microsatellite markers. Markers used were from the National Cancer Institute panel [61]. Individual PCR reactions for each of the markers were electrophoresed on 6% polyacrylamide gels and silver stained. Each case was designated as either MSI-high (MSI-H; ≥ 30% markers unstable), MSI-low (MSI-L; < 30% markers unstable) or MS stable (MSS; no unstable markers) [61,507]. When an MSI-L result was obtained, a second set of markers including the mononucleotide *BAT40* and four dinucleotide markers *D17S787*, *D18S58*, *D20S100* and *D7S519*, were evaluated as above.

Colorectal tumors from the referred group underwent similar MSI analysis as above with the exception that microsatellite markers were radioactively labeled and visualized by autoradiography [508]. Only *BAT-25* and *BAT-26* were analyzed in the referred group because in the population-based probands there was 98.6% (70/71) concordance between overall MSI-H and MSS status, and the MSI status of these two mononucleotide microsatellites.

For the immunohistochemical (IHC) analysis of MLH1, MSH2 and MSH6, formalin-fixed, paraffin-embedded tissues were sectioned at 4 μm, deparaffinized, and rehydrated with xylene and alcohol. The slides underwent either pressure cooker or microwave antigen retrieval (10 mmol/L citrate buffer, pH 6.0; 3 min at 115°C in microMED T/T Mega (Hacker Instruments & Industries, Inc., Fairfield, NJ)). Non-specific binding was blocked by 20% Protein Blocker with Avidin (Signet Laboratories, Inc., Dedham, MA). The slides were washed with Tris-buffered saline. The sections were then incubated with mouse antibodies against MLH1 (1:40; G168-728, PharMingen, San Diego, CA), MSH2 (1:100; FE 11, Oncogene Research Products, Cambridge, MA),

or MSH6 (1:100; 44, BD Transduction Laboratories, Mississauga, ON, Canada) for 1 hr. The antibodies were detected with the avidin-biotin complex method. 3,3`-Diaminobenzidine tetrahydrochloride was used as the chromogen, and haematoxylin was used as a counterstain.

2.5.4.4 Mutation Detection

DNA from probands and, when available, additional family members, was prepared from whole blood using a simple salting-out method [509]. Alterations of *MLH1*, *MSH2* and *MSH6* were determined by direct sequencing of all 45 exons and intron/exon boundaries. DNA from all clinic-based patients, and those population-based probands who had tumors deficient in an MMR protein, was subjected to direct sequencing. Automated sequencing was performed on an ABI 377 DNA Sequencer (Applied Biosystems, Foster City, CA). Sequence information of the coding region was derived from RefSeq NM_000249.2 (*MLH1*), NM_000251.1 (*MSH2*) and NM_000179.1 (*MSH6*). Primer sequences and intronic nucleotide information were derived from genomic sequences from NCBI - AC011816.17 (*MLH1*), AC079775.6 (*MSH2*), AC006509.15 (*MSH6*). Primer sequences are available from the authors upon request.

Exon deletions in *MSH2* and *MLH1* were detected by Multiplex Ligation-dependent Probe Amplification (MLPA) using DNA from all clinic-based probands and those population-based probands whose tumors were deficient in MLH1 or MSH2 [510].

MLPA, using the Lynch Syndrome probes (kit # SALSA P003), was conducted and analyzed according to the protocol provided by MRC-Holland (Amsterdam, Holland). All deletions identified in probands by MLPA were confirmed in other affected relatives by MLPA and/or reverse transcription PCR from lymphocyte RNA.

All sequence variant nomenclature conforms to the recommendations found on the Human Genome Variation Society website (www.hgvs.org), updated as of February 27th, 2005.

2.5.4.5 Detection of *MLH1* Promoter Methylation

We used bisulfite modification coupled with methylation specific PCR (MSP) to determine methylation status of region C of the *MLH1* promoter [511]. One µg of template DNA was treated as outlined in the ChemiCon CpGenome Bisulfite Modification Kit (ChemiCon International, Temecula, CA). Primers and conditions for MSP were obtained from previously published work [512] and were examined by gel electophoresis. Fully methylated DNA (ChemiCon International, Temecula, CA) and normal blood DNA were used as controls.

2.5.4.6 Statistical Analyses

Time to first cancer was analysed using the Kaplan Meier time to event analysis in first and second degree family members of probands. Low risk families were used as the reference group for comparison between risk groups using Cox regression analysis. Probands were excluded from these analyses. In low risk families, both parents and siblings of the parents were included, whereas in higher risk families only the parent and siblings on the transmitting side were included.

2.5.5 Results

2.5.5.1 Family History of Population-Based Probands

Three-generation family histories were collected from all 79 study participants (see Table 2.4 for proband characteristics). Tumors and blood samples were obtained from 74 patients. Two probands are first cousins and so were considered to have the same familial risk. Therefore, family histories of a total of 78 families were obtained. Six families fulfilled ACI and two families ACII. However, tissue and blood samples were not available from one ACI family. An additional twenty-seven probands fulfilled at least one of the revised Bethesda criteria [72].

Table 2.4. Characteristics of population-based probands (n=78) and referred probands (n=31)

		Population-based probands	Clinic-based probands
Mean Age (yrs) colorectal cancer diagnosed		56.1	49.5
Median Age (yrs)		57	44
Age Range		36-69	33-76
Male		45 (57%)	13 (41.9%)
Female		34 (43%)	18 (58.1%)
Deceased		11 (13.9%)	6 (19.4%)
Tumor site:	Proximal	26 (32.9%)	8 (25.8%)
	Distal	45 (57%)	22 (71.0%)
	Not specified	8 (10.1%)	1 (3.2%)
Other Lynch cancers:	stomach	1	0
	endometrial	2	0
	metachronous colorectal	7	4
	synchronous colorectal	1	3
Other non-Lynch cancers		6*	6†

*two of the cancers were in a single proband
† four of the cancers were in two probands

Due to the observation that many families appeared to be at relatively high risk for hereditary cancer but did not fulfill the criterion that at least one family member had developed cancer before 50 years of age, we modified the age and cancer criteria of the Amsterdam classifications, creating the Age and Cancer Modified Amsterdam Criteria (ACMAC): (1) three or more relatives with CRC or a Lynch-associated cancer (according to the revised Bethesda criteria), one of which is a first-degree relative of the other two; (2) CRC or a Lynch-associated cancer occurring in at least two generations; and (3) one or more colorectal or Lynch-associated cancers diagnosed ≤ 60 years of age. Eleven additional families fulfilled the ACMAC classification. The thirty-two probands/families which did not meet any of these criteria were classified as low risk.

2.5.5.2 Time to Cancer in Population-Based Families

The cumulative probability of developing colorectal cancer or any Lynch Syndrome-associated cancer (as defined by the revised Bethesda criteria) in family members at 50% risk was stratified by risk classification (Supplemental Tables S2.1 and S2.2, Appendix J). In ACI and ACII families, over 50% of family members had developed a Lynch Syndrome-associated cancer by age 70 years. In families defined by ACMAC, the time to first cancer curve was parallel to that of members from ACI and ACII families, except that cancer developed about 10 years later in life (Figure 2.7a), with 26% developing the cancer by age 70 years. The observed relative risk of any Lynch-related cancer in ACMAC families (9.9 times that of low risk families), suggests a genetic predisposition for cancer in these families. Members of families defined by the revised Bethesda guidelines (excluding those defined by ACMAC) had a 3.9 fold increased risk of developing any Lynch-related cancer compared to the low risk families, with 14% developing the cancer by age 70 years. Forty-seven percent of ACI and ACII families and 18% of ACMAC families developed CRC by age 70 years. The risk of developing CRC was 31 times greater for ACI and ACII families and 9.9 times greater for ACMAC families compared to low risk families (Figure 2.7b). For those families defined by the revised Bethesda guidelines, the relative risk was 3.8 (compared to low risk families), with 9.3% developing CRC by age 70 years.

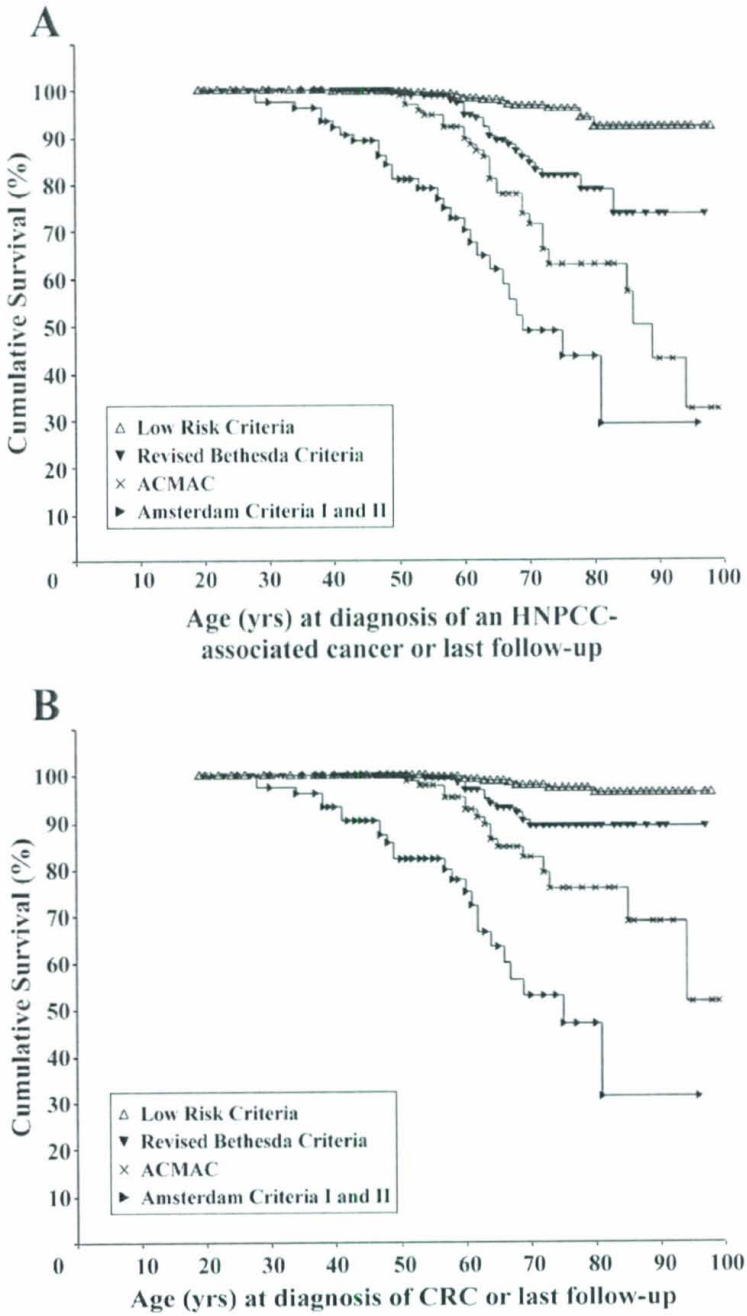


Figure 2.7. Kaplan Meier survival curves of probands of families from the four different familial risk classifications
A) shows time to CRC or last follow-up
B) shows time to any Lynch-associated cancer or last follow-up

2.5.5.3 Immunohistochemical and Microsatellite Instability Analyses in the Population-Based Probands

IHC and MSI analyses were completed for 71 cases from 78 families (Figure 2.8). Fifteen proband tumors (21.1%) had deficient expression of at least one of the three proteins (Table 2.5). All of the tumors deficient in MSH2, MLH1 and both MSH2 and MSH6 were MSI-H. Two of the four MSH6 deficient tumors were MSI-H and two were MSS.

Fifty-five cases (77.4%) had intact staining of all three proteins in their tumors and were MSS; and one tumor was MSI-H and intact for the three proteins. Interestingly, only two of the five tested families (40%) fulfilling the Amsterdam I criteria had probands with tumors that were deficient in an MMR protein. As well, a proband from an ACII kindred had an IHC intact and MSS tumor. CRC tumors from seven ACMAC probands (63.6%) were MSS and expressed all three MMR proteins.

Of the tumors from 22 probands who fulfilled revised Bethesda guidelines, five (22.7%) had tumors which were MSI-H (Table 2.5). The remaining 17 probands (77.3%) had tumors which were MSS and expressed all three MMR proteins.

Five of 32 tumors (15.6%) from probands of low risk families were MSI-H and/or IHC deficient (Table 2.5). The remaining 27 probands (84.4%) from low risk families had tumors which expressed all proteins and were MSS.

2.5.5.4 Mutation Analyses of Population-based Probands

MLPA was implemented as the first screen for mutations in *MLH1* and *MSH2* in the population-based cases. A deletion of exon 8 (c.1277-?_1386+?del (p.Gly426_Gln462>GlyfsX5)) in *MSH2* was identified in a single kindred using this technique (Table 2.5 and Figure 2.8). Subsequent DNA sequencing identified a single proband (study ID 122) carrying a previously identified founder mutation [506] affecting the splice donor site 3' of exon 5 in *MSH2* (c.942+3A>T). One proband (00MG1598; Table 2.5) who had a tumor which was deficient in both MSH2 and MSH6, did not have an identifiable pathogenic alteration in *MSH2*, but did have a novel missense mutation in *MSH6* (c.3415G>A (p.Gly1139Ser)). In four probands with MSH6 deficient tumors, a causative alteration was not identified. A mutation was also not identified in the seven probands with tumors deficient in MLH1. Thus, definitive cancer causing mutations were found in only two of the 16 (12.5%) probands screened for sequence alterations.

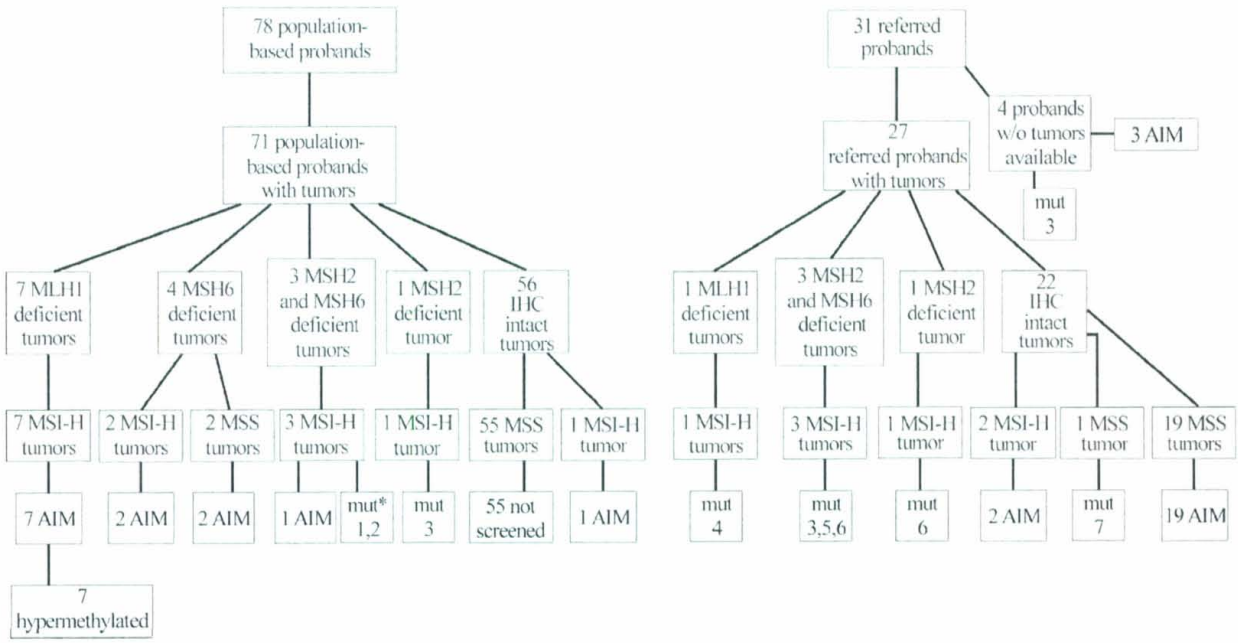


Figure 2.8. Summary of the immunohistochemistry (IHC), microsatellite instability (MSI), methylation, and mutation analyses in population-based and referred probands
*mut 1 = G1139S (*MSH6*); mut 2 = V265_Q314del (*MSH2*); mut 3 = G426_Q462del>GfsX5 (*MSH2*); mut 4 = V49E (*MLH1*); mut 5 = N596S (*MSH2*); mut 6 = I216_T934del (*MSH2*); mut 7 = S144I (*MSH6*); AIM = Absence of Identifiable Mutation.

Table 2.5. Mismatch repair analyses of the population-based probands with MSI-H and/or MMR deficient tumours; and those in Amsterdam criteria I or II families with tumours available

Proband	Age*	Clinical Criteria†	MSI Status‡	IHC MLH1§	IHC MSH2	IHC MSH6	Mutation‖
7	57	ACI	S	+	+	-	AIM
12	38	ACI	H	+	-	+	c.1277-?_1386+?del p.Gly426_Gln462del>GlyfsX5 (MSH2)
19	51	ACI	S	+	+	+	NT
35	49	ACI	S	+	+	+	NT
110	62	ACI	S	+	+	+	NT
99	68	ACII	S	+	+	+	NT
122	36	ACII	H	+	-	-	c.942+3A>T p.Val265_Gln314del (MSH2)
17	57	ACMAC	H	+	+	-	AIM
14	58	ACMAC	H	-	+	+	AIM
50	64	ACMAC	H	-	+	+	AIM
78	51	ACMAC	H	-	+	+	AIM
45	43	rBethesda	H	+	-	-	c.3415G>A p.Gly1139Ser (MSH6)
162	68	rBethesda	H	-	+	+	AIM
10	51	rBethesda	H	ND	+	+	AIM
152	56	rBethesda	H	-	+	+	AIM
180	38	rBethesda	H	+	+	-	AIM
135	51	LR	H	+	-	-	AIM
108	55	LR	H	-	+	+	AIM
63	68	LR	H	+	+	+	NT
119	64	LR	S	+	+	-	AIM
148	66	LR	H	-	+	+	AIM

*Age in years at which proband was diagnosed with the CRC tumor that was used for IHC and MSI analyses.

†Amsterdam I = proband is from a kindred which fulfills the Amsterdam I criteria;
Amsterdam II = proband is from a kindred which fulfills the Amsterdam II criteria;
rBethesda = proband fulfills at least one of the revised Bethesda criteria;
ACMAC = proband is from a kindred which fulfills the ACMAC criteria;
LR = proband is from a family which does not fulfill any of the above criteria

‡S = all markers tested were stable; H = ≥30 of markers tested were unstable

§+ = protein expression was intact in tumor; - = protein expression was absent in tumor;
ND = not determined due to ambiguity in the interpretation of the staining

‖NT = not tested; AIM = Absence of Identifiable Mutation

2.5.5.5 Methylation Analysis of *MLH1* Promoter

Seven tumors from population-based probands which were MSI-H, IHC deficient and with no mutation identified in *MLH1* were tested to determine if their tumors showed methylation of *MLH1*. In all seven tumors, hypermethylation of the *MLH1* promoter was evident (Figure 2.8).

2.5.5.6 Description of the Referred Families

Thirty-one probands from families referred to the Provincial Medical Genetics Program were selected for HNPCC testing, 13 fulfilled the Amsterdam I criteria, eight fulfilled the Amsterdam II criteria, one was categorized as ACMAC, and nine satisfied the revised Bethesda criteria (see Tables 2.4 and 2.6).

2.5.5.7 Immunohistochemistry and Microsatellite Instability Analyses of the Referred Probands

Of the 27 CRC tumors available from probands referred to the genetics clinic, five (18.5%) were deficient in at least one MMR protein and were MSI-H (Figure 2.8 and Table 2.6). However, there were two tumors (proband #s 11566 and 13450) which were intact for all MMR proteins tested but MSI-H. For two of the four probands (#s 2155 and

10763) for which there were no tumors available, a CRC tumor from a blood relative was obtained for IHC and MSI testing. In both cases the tumor was intact for all three proteins and the DNA was microsatellite stable.

2.5.5.8 Mutation Analyses of Referred Cases

Using MLPA, two deletions were identified in four referred probands. A deletion of exon 8 of *MSH2* was found in three probands, and a deletion of exons 4 through 16 inclusive (c.646-?_2802+?del (p.Ile216_Thr934)) was identified in a single proband. To our knowledge, this 13 exon deletion has not previously been documented. No exonic deletions or amplifications were found in *MLH1*. When tumors were available from probands who had an exonic deletion, they were MSI-H and deficient in MSH2 or both MSH2 and MSH6 (Table 2.6).

Regardless of MLPA, IHC, or MSI results, all probands underwent direct sequencing of exons and intron/exon boundaries to determine potential disease-causing mutations, to catalog known and novel polymorphisms and to identify potential disease modifying alleles. In the two tumors, which were MSI-H but intact for all tested MMR proteins, analysis of genomic DNA was negative for mutations in the three MMR genes sequenced (Table 2.6). Another proband (#10737) carried a c.1787A>G (p.Asn596Ser) mutation, presumably causing the tumor to be MSH2 deficient and MSI-H.

Surprisingly, 11 of 17 probands (64.7%) from ACI and ACII families, with cancers which were MSS and intact for the MMR proteins, had no identifiable mutations

in *MLH1*, *MSH2* or *MSH6*. In addition, tumors from 9 of 10 probands (90%) fulfilling the ACMAC or revised Bethesda criteria were MSS and intact for all proteins tested. However, in one of these probands (#2275) a p.Ser144Ile substitution (c.431G>T) in *MSH6* was detected, which segregates with the disease in this family.

A number of other genomic variants were identified, some of which have not been previously reported and may represent alterations unique to the Newfoundland population (Supplemental Table S2.3).

Table 2.6. Mismatch repair analyses of thirty-one referred CRC patients

Proband	Family	Age*	Clinical Criteria†	MSI Status‡	IHC MLH1§	IHC MSH2	IHC MSH6	Mutation‖
2152	10	43	AC I	S	+	+	+	AIM
2155**	76	56	AC I	n/a	n/a	n/a	n/a	AIM
11111	R0014	64	AC I	H	+	-	-	c.1277-?_1386+?del p.Gly426_Gln462del>GlyfsX5 (MSH2)
11825	R0139	42	AC I	S	+	+	+	AIM
11816	R0203	39	AC I	S	+	+	+	AIM
12668	R0245	44	AC I	S	+	+	+	AIM
13017	R0294	43	AC I	S	+	+	+	AIM
12536	R0339	62	AC I	H	-	+	+	c.146T>A p.Val49Glu (MLH1)
12501	R0386	76	AC I	S	+	+	+	AIM
12463	R0458	38	AC I	n/a	n/a	n/a	n/a	AIM
12571	R0366	43	AC I	S	+	+	+	AIM
13450	R0665	57	AC I	H	+	+	+	AIM
334	11	43	AC I	H	+	-	+	c.646-?_2802+?del p.Ile216_Thr934del (MSH2)
10719	R0004	33	AC II	n/a	n/a	n/a	n/a	c.1277-?_1386+?del p.Gly426_Gln462del>GlyfsX5 (MSH2)
10763**	R0163	51	AC II	n/a	n/a	n/a	n/a	AIM
11566	R0183	42	AC II	H	+	+	+	AIM
12859	R0336	44	AC II	S	+	+	+	AIM

12858	R0457	49	AC II	H	+	-	-	c.1277-?_1386+?del p.Gly426_Gln462del>GlyfsX5 (MSH2)
12870	R0512	47	AC II	S	+	+	+	AIM
10283	R0006	43	AC II	S	+	+	+	AIM
13603	R0653	63	AC II	S	+	+	+	AIM
10838	R0026	64	ACMAC	S	+	+	+	AIM
2275	51	38	rBethesda	S	+	+	+	c.431G>T p.Ser144Ile (MSH6)
13359	R0525	43	rBethesda	S	+	+	+	AIM
13544	R0624	72	rBethesda	S	+	+	+	AIM
11308	R0152	61	rBethesda	S	+	+	+	AIM
10737	R0156	44	rBethesda	H	+	-	-	c.1787A>G p.Asn596Ser (MSH2)
11799	R0234	49	rBethesda	S	+	+	+	AIM
11988	R0262	60	rBethesda	S	+	+	+	AIM
12398	R0311	47	rBethesda	S	+	+	+	AIM
12173	R0314	39	rBethesda	S	+	+	+	AIM

*Age in years at which proband was diagnosed with the CRC tumor that was used for IHC and MSI analyses.

†ACI = proband is from a kindred which fulfills the Amsterdam I criteria;
ACII = proband is from a kindred which fulfills the Amsterdam II criteria;
rBethesda = probands fulfills at least one of the revised Bethesda criteria;
ACMAC = proband is from a kindred which fulfills the ACMAC criteria

‡S = both BAT-25 and BAT-26 were stable; H = both BAT-25 and BAT-26 were unstable; n/a = tumor from proband was not available.

§+ = protein expression was intact in tumor; - = protein expression was absent in tumor; n/a = proband's tumor was not available

|| AIM = Absence of Identifiable Mutation

** CRC tumor from proband's relative was MSS and IHC intact for all proteins

2.5.6 Discussion

Our initial population-based study of CRC cases less than 70 years of age suggests there is a high incidence of CRC due to hereditary factors on the Avalon Peninsula of Newfoundland. Of 71 CRC tumors, 15 (21.1%) were deficient in at least one of the MMR proteins tested by IHC. Since epigenetic silencing of *MLH1* [216,511,513] appears responsible for the seven tumors which are deficient in MLH1, there remain eight (11.3%) cases with tumors deficient in MSH2 and/or MSH6 - indicating a known causative hereditary component.

Additionally, there are many families from the population-based probands which have a strong history of Lynch syndrome-related cancers not attributable to MMR deficiency and with MSS tumors. Tumors from four of the seven (57.1%) available probands of Amsterdam criteria I and II families, and 17 of 22 (77.3%) tumors from probands fulfilling the revised Bethesda criteria, do not have an MMR deficiency in the proteins analyzed. Many of these families had a large number of Lynch-related cancers segregating in a pattern consistent with an autosomal dominant mode of inheritance and failed to meet Amsterdam Criteria because of late onset of the cancers. The Age and Cancer Modified Amsterdam Criteria (ACMAC) incorporates the principles of ACII [341] but with an age cut-off of less than 60 years instead of 50 years, and includes all the cancers outlined in the revised Bethesda criteria [72]. In this group of eleven families, the risk of CRC or any other Lynch Syndrome cancer was substantially higher than that in low risk families, or those fulfilling revised Bethesda criteria. Also, tumors from four

probands (36.4%) of ACMAC families were deficient in an MMR protein and MSI-H. Tumors from the remaining seven (63.6%) were MSS and intact for all proteins tested. Taken together, these findings indicate that the majority of families with a strong familial predisposition for Lynch-associated cancers may have mutations in novel susceptibility genes. The likelihood that mutations in other MMR genes such as *PMS1*, *PMS2*, *MLH3* or *MSH3* are causing susceptibility to cancer in these families is low. Very few families have been identified with MMR defects caused by heritable mutations in these genes [234,500,501,514]. However, the involvement of these loci cannot be dismissed as we have not screened for mutations in these genes.

Due to the high frequency of Lynch Syndrome-like families without MMR deficiencies identified in the population-based cohort, we investigated families referred to the Provincial Medical Genetics Program. In total, four families fulfilling ACI or ACII criteria segregated an exon 8 deletion in *MSH2*, causing a frameshift and a premature stop codon in exon 9 (p.Gly426_Gln462>GlyfsX5). A deletion of this same region was identified previously; however, the exact location of the breakpoints could not be determined as they fell within a 45 bp *AluY* consensus sequence, which is present in both intron 7 and intron 8 of *MSH2* [515]. We were able to map the breakpoints to this same 45 bp region in our probands (data not shown), suggesting this alteration has a common etiology and could be a mutational hotspot, as there have been other published reports of the deletion of exon 8 in Lynch Syndrome families [508,516]. Genotyping our families with microsatellite markers is underway to determine, by haplotype analysis, if this is a founder mutation or represents mutations that have occurred independently due to a

common mechanism. As well, a multi-exonic deletion of *MSH2*, spanning exons 4-16 inclusive (p.Ile216_Thr934del), segregates in the very large Family 11. This kindred includes more than 80 mutation carriers which, like the families segregating the exon 8 deletion, would be good candidates for genotype-phenotype studies.

As in other studies which screen MMR genes for genomic alterations in Lynch Syndrome families, we have identified a number of missense mutations of unknown significance. In one proband with an MSI-H tumor deficient in MSH2 and MSH6 (proband #10737), a p.Asn596Ser substitution was found which occurs in the MSH3/MSH6 interaction domain of MSH2. Although both asparagine and serine are uncharged polar amino acids and such a change is tolerated according to the Sorting Intolerant From Tolerant (SIFT) algorithm [517], there has been documentation that this is a functionally important amino acid [518-521]. Also, it was absent from 192 control chromosomes. Thus, there is some evidence to suggest that this is the causative alteration in our family. Unfortunately, there is no other DNA available from this family to determine the segregation pattern of the variant.

In another proband (#45) with a tumor that is MSI-H and deficient in MSH2 and MSH6, there is no identifiable mutation in *MSH2* but there is a p.Gly1139Ser alteration in *MSH6*. This variant is found in the P-loop of the ATP-binding domain required for hydrolysis of ATP [522]. SIFT analysis indicated that any substitution of this amino acid is predicted to affect protein function. In *S. cerevisiae*, an alteration of the homologous residue (p.Gly987Asp) caused a decrease in ATP hydrolysis and the failure of mismatch repair, but the MSH2-MSH6 complex showed greater binding affinity to the mispair *in*

vitro [522]. Therefore, the MSH6-MSH2 complex may be sequestered onto the heteroduplex DNA, without mending the mispair, and then targeted for degradation, thus causing the deficiency of both the MSH2 and MSH6 proteins. Unfortunately, there were no other family members available for testing, however, this alteration was not observed in 192 control chromosomes.

Another proband (#2275), fulfilling the revised Bethesda criteria was shown to be heterozygous for p.Ser144Ile in *MSH6*. This substitution is a change from an uncharged polar to a nonpolar residue and occurs in a conserved amino acid. This missense variant has previously been identified in four families [244,245,523], none of which fulfilled ACI or ACII. One study also reported functional analysis using an *S. cerevisiae* model system [523]. They determined that this change, which they did not identify in 199 normal control chromosomes, caused the loss of MSH6 function observed in their assay. We also did not observe this change in any of 192 control chromosomes. Additionally, our proband has two siblings with the same missense alteration who have had tubular adenomas at the ages of 38 and 53 years. Also, two unaffected sibs without adenomas over the age of 53 years, undergoing regular colonoscopy screening, do not carry the alteration. The above evidence is consistent with this variant being the cause of CRC in our proband.

A final proband (#12536) was heterozygous for p.Val49Glu in *MLH1* and had a CRC which was MLH1 deficient and MSI-H. Additionally, his nephew, who was diagnosed with CRC at age 49, was MSI-H, MLH1 deficient, and had the same substitution. Two other cancer patients from the family also have this alteration. The Val49 residue falls within the ATPase domain of MLH1 and is a nonpolar, hydrophilic

amino acid while Glu49 is polar and hydrophobic. Recently, functional studies of the ATPase domain have been performed on yeast-human hybrid constructs indicating that this change impairs MMR [524]. Additionally, no substitutions were observed at this amino acid in 192 of our control chromosomes.

In the four population-based probands with an MSH6 deficient tumor no mutations were identified. The *MSH6* alteration may include deletions too large to be detected by sequencing. It was not possible to perform MLPA for *MSH6*, due to the unavailability of the assay. Also, if mutations were located deep within intronic regions as cryptic splice sites, or considerably upstream of the translation start site, their detection would not be possible.

The dearth of identifiable mutations in the referred families is noteworthy. In 16 of the 21 families (76.2%) fulfilling Amsterdam criteria, a mutation in *MLH1*, *MSH2* and *MSH6* could not be identified. Other investigators have also noted that a large fraction of high risk families have no detectable mutations in the most commonly mutated MMR genes [525,526]. Such observations could be due to inadequate mutation detection methods. Also, mutations in other CRC predisposition genes could be responsible for disease in our families. It has been shown that attenuated FAP and MYH-associated phenotypes may mimic Lynch Syndrome in that there are very few polyps present and tumors are MSS [527,528]. Therefore, defects in these genes cannot be excluded solely by phenotypic analysis.

However, when comparing probands with an MMR deficiency, to those without, there was a significant difference between the colorectal tumor sites of each group (Table

2.7). Distal tumors were much more prevalent in non-MMR deficient probands than in those who had an MMR defect. As well, unlike the MMR deficient probands, there were no other cancer types detected in the non-MMR deficient probands. As has been suggested previously, there appears to be important clinical differences between those families with MMR defects and those without [353,529]. We support the suggestion that there be another designation ("Familial Colorectal Cancer Type X") to describe those families without an MMR defect, suggesting there is an unexplained cancer etiology [353]. In addition, some recent studies have analyzed the molecular features of families not associated with MMR defects, and have indicated that novel pathways toward cancer may be involved [358,530].

Table 2.7. Characteristics of probands with (n=24) and without (n=15) mismatch repair deficiencies

	Probands with an MMR deficiency*	High-risk probands† without an MMR deficiency	
Mean Age (yrs) colorectal cancer diagnosed	52.3	50.5	
Median Age (yrs)	54	45	
Age Range	33-68	39-76	
Male	11 (45.8%)	7 (46.7%)	
Female	13 (54.2%)	8 (53.3%)	
Tumor site:			
Proximal	13 (54.2%)	1 (6.7%)	
Distal	9 (37.5%)	14 (93.3%)	P=0.001‡
Not specified	2 (8.3%)	0 (0%)	
Other Lynch cancers§ in proband	8 (in 7 probands)	0	
Other non-Lynch cancers in proband	3	0	

*Any proband who carried a germline MMR mutation and/or had a tumor which was IHC deficient and/or was MSI-H was considered to be MMR deficient.
†Includes only probands in families fulfilling either Amsterdam criteria I or II
‡Obtained from a Chi-squared test (X^2=10.4)
§Cancers included in the revised Bethesda criteria other than CRC

As a result of our detailed analysis for detecting MMR defects in the families and patients presented here, and the fact that the majority of families fulfilling the Amsterdam criteria had an absence of MMR mutations, we propose that there are mutations in other cancer-causing genes segregating in this population. Some of these families are large, multiplex kindreds which have been followed for over 20 years, and are feasible for linkage studies to identify novel gene(s) causing a Lynch Syndrome-like disease.

2.5.7 Acknowledgments

We especially thank the families who participated in this study for their time and patience. Also, we thank Cathy Searle and Michelle Simms for their technical assistance on the immunohistochemical and MSI analyses; and Hong Cheng for help in sequencing *MLH1*, *MSH2* and *MSH6*.

Chapter 3 – Tissue Microarray Screening for Prognostic Markers in Colorectal Cancer

3.1 Introduction

3.1.1 History and Description of Tissue Microarray

Immunohistochemistry (IHC) can be used to assess the level and pattern of protein expression in different tissues and cell types, and the cellular location of expression. In 1987 Wan *et. al.*[531] developed a method for performing immunohistochemistry on multiple tissue samples simultaneously on the same microscopic slide, thereby saving time and reducing the quantity of tissue, antibodies, and reagents used when compared to using whole sections.

It was 11 years later that this concept became popular, when Kononen *et. al.* [532] developed a technique that could more easily produce quality composite blocks, termed tissue microarrays (TMAs). Kononen's method of constructing TMAs was based on the use of a custom-made device that could take 0.6 mm diameter, 3 mm to 4 mm deep tissue cores from donor blocks and precisely insert them into new recipient paraffin blocks. The size of the cores allowed for preservation of the histological information of the tissue, yet was small enough to minimize damage to the donor block and allowed for up to 1000

tissue cores to be incorporated into one 45 mm by 20 mm recipient paraffin block [532]. Once a TMA was prepared in this way, sections could be cut using a microtome, carefully mounted onto standard microscope slides, and used for further studies. Because the cores were located at known coordinates within the block they could be linked back to the source, and any data resulting from analysis of the TMA could be studied in conjunction with known clinical or pathologic features of that case.

Kononen's method [532] is still used today, although commercial TMA-building devices are now used to construct the composite blocks. TMA blocks are often divided into quadrants, or sectors, to help the user maintain position during scoring. As well, cores of tissue with obviously different morphology and origin from the tissue of study can be included in TMAs in an asymmetrical nature to help maintain orientation when scoring. This tissue can also be chosen to act as an internal control, since it is included in the same block as the tissue of interest, and therefore will be exposed to exactly the same conditions.

3.1.2 Advantages of using Tissue Microarrays Versus Whole Sections

There are many advantages of using TMAs instead of whole sections for both research and clinical applications. Automated image analysis of TMAs can produce scores on a continuous scale of staining intensity, which reflects the continuous scale of biological expression of protein [533]. Automated image analysis cannot easily be performed on whole sections. These scores hold more analytical power than semi-

quantitative scores based on arbitrary cut-offs, which are at the limit of what can be achieved through manual scoring. When analysed in conjunction with clinical data, this information can also be used to determine standardized, biologically meaningful cut-off points to be used instead of the arbitrary cut-offs traditionally used in scoring. Such biological cut-offs are likely to be more powerful for identifying sub-groups of tumours associated with different clinical outcomes.

The most obvious advantages of the use of TMAs are the savings of time and money. There is a huge savings of reagents, antibodies, and tissue used, as well as of time for the technologists and pathologists who cut, stain and score them. With whole section analysis the pathologist must replace the slide on the microscope stage between each case, refocus, and scan the entire slide to find the area of interest, then determine the score. For TMA analysis the specimens are all on one slide, so no slide replacement or refocusing is necessary. Instead of scanning the slide for the area of interest it has been preselected, and the area limited to a reasonable size for scoring.

The removed tissue cores are so small that they do not cause much damage to the donor block, which allows multiple cores to be removed, and therefore multiple recipient TMA blocks to be composed from the same specimens, further increasing the number of proteins that can be studied. The small size of the extracted tissue cores is particularly advantageous when dealing with small specimens.

These savings are so great that they enable much larger studies to be carried out, including larger cohort sizes and a greater number of proteins evaluated, than could ever be reasonably done by whole section analysis. Increased cohort sizes result in greater

statistical power for producing reliable results. It also enables the study of proteins differentially expressed in only a small percentage of cases [533].

The format of TMAs, in which multiple tissue samples are stained on the same slide in parallel, reduces a lot of the batch-to-batch variability that is inherent in IHC. There are many potential sources of variability in IHC, including: the time and temperature of antigen retrieval, concentrations of reagents and antibodies, incubation times with antibodies, washing conditions, and temperatures [534]. All of these variables are identical when the specimens are all processed together on a single slide. Because all of the specimens are stained at the same time, under exactly the same conditions with TMAs, there is a standardization of protocol that is difficult to attain when staining multiple whole sections in different batches [535]. Another advantage is that the inclusion of control tissues within the TMA block results in identical staining conditions for test and control specimens.

3.1.3 Considerations and Disadvantages of Using Tissue Microarrays

A major concern with using TMAs is whether or not the small tissue cores, also called histospots, are representative of the whole specimen. This is of particular relevance when studying heterogeneous tissue, and assessing either tissue morphology or the IHC staining pattern of a biomarker that exhibits tissue-specific heterogeneous expression.

In 2000, Camp and colleagues published a validation study [536] of the then-new tissue microarray technology. They used 38 invasive breast tumours, and stained both whole sections and 0.6 mm histospots for expression of the three proteins commonly tested in breast cancer: ER (estrogen receptor), PR (progesterone receptor) and HER2 (human epidermal growth factor receptor 2), using IHC. They compared the staining results obtained from using two to ten histospots with those from the full-face sections, and found that the results from just two histospots were comparable to the whole section results for more than 95 percent of cases [536].

A year later Torhorst and colleagues published a similar article [537], in which they also examined the validity of TMA for breast tumour IHC. They studied 553 breast cancers, and compared the expression of ER, PR and p53, using 0.6 mm TMA histospots and whole sections. They found that one single histospot was sufficient to give comparable results to whole sections for 95 percent of cases for ER, up to 81 percent for PR, and up to 74 percent for p53 [537]. These correlations could be increased by increasing the number of histospots examined; the TMA and whole section scores for PR were virtually the same when they examined three histospots. They concluded from their findings that tissue heterogeneity did not decrease the power of TMA results to predict clinical endpoints. Increasing the number of cores used per case can help in two ways: it can help identify heterogeneity within a tumour, and it can validate antibody reproducibility [533,537]. When using more than one core per specimen to identify heterogeneity the cores should be taken from different areas of the specimen. Torhorst's group included four cores per specimen in their study [537]; one from a central region of the

tumour and three from peripheral regions. They did not make any recommendations as to the ideal number of cores to use; however, from their findings it is suggested that one core is adequate for homogeneous tissues and antibodies with homogeneous expression patterns, and three cores are sufficient to account for heterogeneity.

Although most researchers seem to agree that TMA can be a valid alternative to whole section analysis, some negative results have been published in this regard. Jensen and colleagues [538] reported that 20 percent of the ovarian serous tumours of low malignant potential that they studied showed loss of one or more mismatch repair (MMR) proteins by TMA IHC, yet subsequent whole section IHC showed intact expression of all MMR proteins for all informative tumours. As this was not the focus of their study they did not offer an explanation for the discrepancies.

Although traditional paraffin tissue blocks and TMAs are similar in many ways, and both contain tissue and paraffin, TMAs require a substantially greater level of technical skill to produce and cut [533]. TMA construction requires the proper cutting, handling, and embedding of hundreds of small tissue cores, while traditional blocks usually include just one, larger sample. Each requires the use of a microtome; however, when cutting a TMA it is important to minimize waste of the block with initial trimming. It is also more important that the block be cut completely flat, and not at an angle. Finally, it is necessary that the sections contain tissue from all cores, and that none of the cores are lost during the transfer of the section to a glass microscope slide. Any stretching, tearing or folding of the section can result in loss of orientation during scoring, or the loss of interpretable cores.

While the ability to study large cohorts using archival tissue in a retrospective study is an advantage of TMAs, it can also cause some problems. Such studies often include cases diagnosed over many years. Over time there can be changes in tissue fixation and processing methods, which can lead to variations in antigen preservation, staining intensity, and interpretability for the different tissue cores in the TMA. As well, the age of the tissue specimens can have an impact on the level of expression of some antigens. Camp and colleagues [536] reported that most proteins retain their antigenicity for more than 60 years. One exception they observed was that breast tumours from the early 1930s had significantly less positive staining for the ER protein than did more recently diagnosed breast tumours tested in the same TMA block. They recommended the development of TMAs containing tumours diagnosed over different decades, for the use of validating antibodies for archival tissues of different eras before those tissues are used in retrospective studies.

The use of tissue microarrays can lead to an abundance of data. The efficiency of the method enables the staining of hundreds of specimens with a large number of antibodies, and the ability to link the immunohistochemical scores with clinical and pathological data. While this is a great advantage of the method, the large volume of interconnected data needs to be managed appropriately. Software has been developed to aid in data management and analysis. This technology has allowed researchers to use digital imaging technology to capture and store images of stained histospots. Stained tissue microarrays on conventional glass microscope slides are scanned at high resolution; the resulting images can be viewed on any computer with appropriate software. These images can be retrieved quickly and linked to other data from the same case using

software such as Stainfinder [539]. These images can also be shared online with collaborating researchers located at other institutions. Image annotations can be made and shared with other users, for example to point out areas of interest on an image for teaching purposes or to facilitate discussion with colleagues. The application of imaging technology can facilitate standardized interpretations of immunostains and lead to more reproducible results.

Whether using a traditional microscope or digital imaging software, manually scoring hundreds or thousands of histospots can be a tedious, time-consuming task, and the scoring can be somewhat subjective. The human eye is also subject to fatigue, and biases in interpretation associated with the colour and size of adjacent images, or the context in which the scoring takes place [534]. The important step of scoring can impede the high-throughput potential of TMAs. Because TMAs are constructed with tissue chosen by pathologists as being representative of the tissue of interest, some argue that it is not necessary for a pathologist to validate the tissue again while interpreting each immunohistochemical stain. For this reason it is considered possible for the scoring of TMAs to be automated, and a number of companies have developed software for automated TMA scoring, such as AQUA (HistoRx, New Haven, CT) [533].

3.1.4 Use of Tissue Microarrays in the Discovery of Biomarkers

An important use of TMAs is in the identification and development of cancer biomarkers. Cancer biomarkers can have many clinical applications, including use as

tools for diagnosing and staging disease, and for assessing the extent of disease [540]. A potential role for TMAs is in the development of tissue-based cancer biomarkers for predicting disease outcome and response to therapy. While traditional biomarkers rely on the measurement of one protein or one characteristic, the complex nature of cancer requires a less simplistic approach for optimal sensitivity and specificity. The protein products of many different gene types can be involved in any given cancer, including genes involved in cell death, proliferation, differentiation, and DNA repair. Tumour protein expression can be increased or decreased compared to equivalent normal tissue, or can be present or absent in one tumour while not in another tumour, for any number of reasons. These reasons include a change in gene or chromosomal copy number within the tumour, loss of DNA or chromosomal material, genetic mutation, and epigenetic modification. Given the genetic complexity of cancer it is unlikely that any one protein product will hold all the answers in regards to prognosis and optimal disease treatment; rather a profile of multiple biomarkers will likely have higher sensitivity and specificity[541,542]. In this case, sensitivity is the proportion of patients who will have a specific outcome, for whom this is correctly predicted by their biomarker profile; in other words, the percentage of true positives correctly identified. Specificity represents the percentage of true negatives correctly identified. TMAs can be used to develop a set of multiple biomarkers for a certain cancer type, for which the pattern of protein expression of any specific tumour can be used to predict clinical and therapeutic outcome.

TMAs can be used to increase the speed and efficiency of biomarker identification, development, and validation. The expression of hundreds of proteins can

be assessed by TMA using large cohorts to help identify candidate markers. Those markers can then be validated using TMAs, and exchanged between research groups for further validation. Finally, the use of TMAs can facilitate large prospective screening trials, in part by providing a standard control for use in different laboratories, or within the same laboratory over time. [533]

The proteins identified as biomarkers through this process can also contribute to the understanding of disease pathways and drug response. Tumour biomarker profiles may identify groups of patients who have different subtypes of the disease. These subtypes may be clinically and histologically indistinguishable, however, they may respond differently to therapy. Identification of such subtypes of disease should help develop more targeted therapies for individual tumours, and guide drug selection to maximize clinical response while minimizing adverse events. [540]

Once such tumour subtypes are identified, TMAs can be used in clinical practice for adjuvant testing. Some tumour types are routinely stained with a panel of antibodies to better classify the tumour type and guide therapy. An example of this is the testing of all breast cancers for ER, PR, and HER2 status. Assembling batches of breast tumours into TMAs for these analyses instead of using whole sections of each tumour would save time and reagents. One concern with using TMAs for this is whether or not the cores are representative of the whole tumour. It has been suggested [543] that TMA cores may actually be superior to whole sections, since tumours which exhibit just weak, focal staining would more likely be scored as negative by TMA analysis, which for some proteins could be a more clinically appropriate score. This needs to be studied more

thoroughly before it can be applied to clinical practice, to ensure that the focally weak staining tumours behave clinically more like those which are negative than those which are positive.

3.1.5 Introduction to our Study

While it requires more technical skill to create, process, and analyse tissue microarrays, and there are concerns regarding tissue heterogeneity, overall TMA is a valuable technique, and greatly increases the potential uses of immunohistochemistry. By taking care with the technical aspects of TMA, and using software packages to help with data storage and analysis, TMA can be a powerful tool, with many potential roles in research and clinical medicine, including the identification of new cancer biomarkers.

We used tissue microarrays to study the expression of potential biomarkers for colorectal cancer. Antibodies were chosen empirically, based on biological relevance to colorectal cancer and recent publications indicating an association with prognosis. Also considered was the availability of the antibodies, and the suitability for use on formalin-fixed, paraffin-embedded archival tissue specimens. A total of 12 antibodies were included in the study. They are described in section 3.3.

A total of 397 tumours were included in the construction of the tissue microarrays. They include 286 of the 294 tumours from the pilot study (described in chapter 2). Specimens were not available from the remaining eight cases at the time of TMA construction. To increase study numbers and boost statistical power, another 111 tumours

were included from the CIHR-IHRT study, described in section 1.11. These 111 cases were chosen for the timely availability of pathology specimens, and were found to be representative of the study population, as described in the manuscript included in section 3.4.

A total of four TMA blocks were constructed from formalin-fixed, paraffin embedded surgical pathology specimens, using two 0.6 mm donor cores from each tumour. The cores were arranged asymmetrically to ease scoring and ensure proper alignment of sections. Sections of 4 μm thickness were cut from the TMA blocks, mounted on microscope slides, and individually stained with 12 different antibodies. Figure 3.1 shows a section from one of these TMA blocks, stained with haematoxylin and eosin, to illustrate the layout and appearance.

For the tumours included in this study, I collected pathology data, patient demographic and baseline clinical data, as well as patient treatment information, clinical follow-up, vital status, and cause of death, where appropriate. Data was collected from original pathology reports, the provincial medical information database, MEDITECH, doctors reports, clinical imaging reports, hospital charts, Cancer Clinic patient charts, the Newfoundland Cancer Registry, the Newfoundland and Labrador Centre for Health Information, and the Statistics Canada Annual Mortality Files.

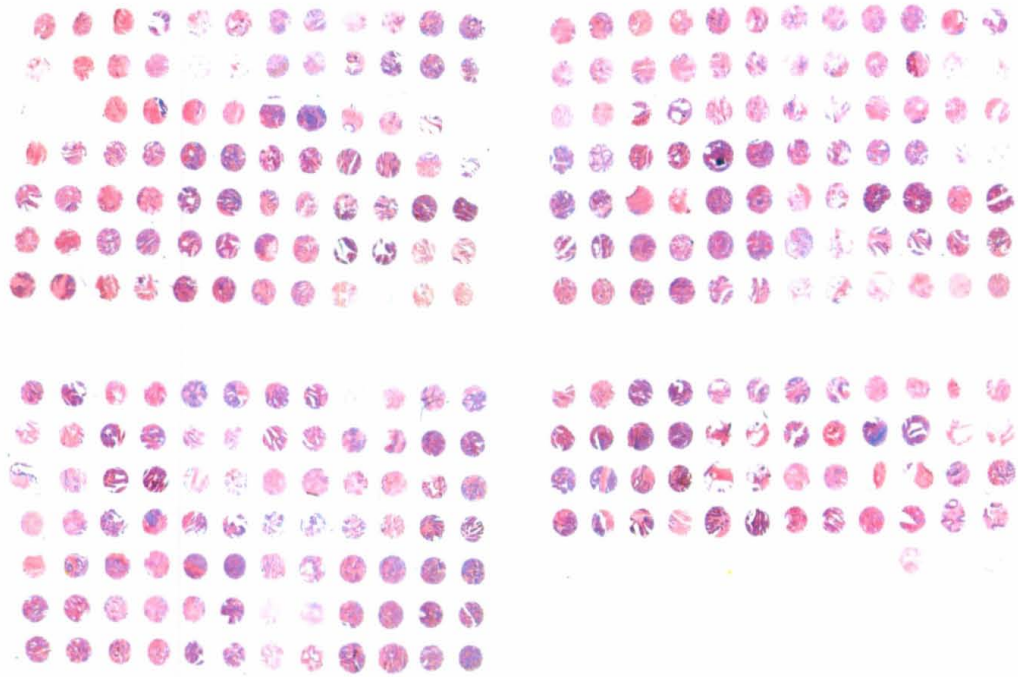


Figure 3.1. Section of a TMA block stained with haematoxylin and eosin

3.2 Objectives

The primary objective of this study was to identify proteins, or profiles of proteins, whose expression was of clinical prognostic or predictive value in colorectal cancer. The ultimate, long-term objective was to develop tools to aid in clinical management and to guide therapeutic decision making for colorectal cancer, leading to increased survival in patients with CRC.

3.3 Antibodies Used

A total of 12 antibodies were used in this study. They include markers of cell proliferation Ki67 and PCNA; tumour suppressor gene and cell cycle regulator p53; other cell cycle-related proteins p16, p21, p27, cyclin D1; cell adhesion proteins β-catenin, and E-cadherin; as well as structural proteins lamin A/C; DNA helicase MCM7; and cell signalling molecule EGFR. Each is described in the following sections, along with their relevance to colorectal cancer.

3.3.1 Proteins Related to the Cell Cycle – p53, p16, p21, p27, and Cyclin D1

3.3.1.1 Brief Introduction to the Cell Cycle

The cycle of DNA replication and cell division is fundamental for growth, as well as for the regeneration and renewal of most cells and tissues in the human body. The cell cycle is tightly regulated to ensure the integrity of the replicated DNA, and to maintain appropriate limits on cellular proliferation [544,545].

The cell cycle consists of a temporal sequence of four main phases. G1 is the first gap phase, which separates the previous cell division from the onset of DNA replication. During this phase, diploid, 2n cells grow and accumulate the proteins necessary for replication. DNA is also checked for errors, which are repaired before continuation

through the cycle. During the S (synthesis) phase, nuclear DNA is replicated, resulting in cells with 4n DNA content. This is followed by a second gap phase, G2, during which the cells again increase in size, and prepare to divide. M phase ensues, during which the cells undergo mitosis and cytokinesis. Mitosis is the multistep process of nuclear division, during which the duplicated genomic material separates and becomes encased in separate nuclear membranes, thereby creating two nuclei. Cytokinesis begins during the late stages of mitosis. It is the process of cellular division, whereby the cytoplasm, organelles and nuclei are divided into two new daughter cells; each encased in their own plasma membrane, and each identical to the other and to the parent cell. Some cells can also enter a state of quiescence (G0 phase), whereby they are not actively cycling.

Cells can enter the G1 phase either from the quiescent state of the G0 phase, or from a previous cell cycle. Whether a cell in G1 continues through the cell cycle, or switches to the G0 state, is dependent on extracellular factors including growth factors, growth inhibitors, and spatial cues [546,547]. Such factors include mitogens, which trigger signal transduction pathways and encourage cell division. Once a cell progresses past the restriction point, late in the G1 phase, it is committed to enter the S phase, without the requirement of any further mitogenic stimulation [546]. Once a cell has past the restriction point, it is no longer sensitive to these extracellular stimuli; progression beyond this point relies on intracellular controls.

Progression through the phases of the cell cycle is regulated by checkpoints [545]. The cell cycle will not progress past them unless certain prerequisites have been met. One such prerequisite is the proper repair of any DNA damage. Such repair is important

at the G1/S checkpoint to prevent the perpetuation of DNA errors to the daughter cells, and at the G2/M checkpoint to prevent the loss of genomic material that would occur if mitosis were to proceed with unrepaired DNA breaks. If such repair were not possible, the cell would undergo controlled death through apoptosis. If deleterious mutations exist in proteins involved in these regulatory checkpoints, the cell cycle may continue through them inappropriately, potentially leading to genomic or chromosomal instability, and increased susceptibility to DNA damaging agents [545,548]. Unregulated cellular proliferation and carcinogenesis can ensue [175]. The G1/S and G2/M checkpoints are commonly defective in cancer [549].

Cyclins and cyclin-dependent kinases (CDKs) are the key regulatory proteins of the cell cycle, and are required for all major transitions within the cycle [549]. CDKs require phosphorylation or association with a member of the cyclin family of proteins to be functionally active as a kinase. Each is activated at a specific point during the cycle. They act by phosphorylating their substrates on either a serine or threonine, and so are known as serine/threonine kinases. CDKs and cyclins are positive regulators of the cell cycle, inducing progression through the cycle. The major negative regulators are cyclin-dependent kinase inhibitors (CDKIs). For the sake of simplicity, these pathways will not be described in detail, and very few proteins will be named or described in the following sections.

Five proteins involved in the cell cycle were chosen for inclusion in this study, based on their relevance to colorectal cancer, and the availability of a robust antibody for use in formalin-fixed, paraffin-embedded tissue. Included were: (1) p53, a transcription

factor, tumour suppressor protein, and key regulator in the cell cycle, (2) p16, a member of the INK4 (inhibitor of cyclin dependent kinase 4) family of CDKIs, (3) p21, a mediator of p53 tumour suppression and member of the CIP/KIP (CDK-interacting protein/CDK inhibitory protein) family of CDKIs, (4) p27, another member of the CIP/KIP family of CDKIs, and (5) cyclin D1, a key cyclin involved in the G1/S checkpoint. The role of each of these proteins in the cell cycle, and some evidence of their role in colorectal cancer is described briefly in the following sections.

3.3.1.2 p53

Intracellular levels of p53 increase in response to various forms of cellular stress including DNA damage from ionizing irradiation, hypoxia, oxidative stress, and aberrant growth signals. Depending on the kind of cellular stress, this transcription factor can induce cell cycle arrest, apoptosis, cellular senescence, DNA repair, or changes in metabolism, by regulating expression of its target genes [550]. It plays key roles in regulating cell cycle arrest and programmed cell death (apoptosis). It is mutated in at least half of all colorectal cancers [551,552], and indirectly inactivated in most others [552,553]. Allelic deletions of the chromosomal region later found to contain the *TP53* gene was included in Vogelstein's 1988 report of common alterations which occur during colorectal tumour development through the chromosomal instability pathway[201] (previously described in section 1.7.1). They described the loss of this locus to occur late

in the progression from adenoma to carcinoma, as the loss was identified in 75% of carcinomas, but in a much smaller proportion of adenomas.

As a tumour suppressor, p53 limits growth in two main ways: halting the cell cycle when damage is detected; and initiating cell death through the apoptosis, when repair is not possible. Lee and Bernstein first reported in 1993 [554] that p53 mutations affected cellular response to DNA damage. They found that cells with mutant p53 did not undergo apoptosis when exposed to ionizing irradiation, whereas cells with wild type p53 would.

The pathways of p53 activation due to various types of cell stress and damage, as well as the affects of that activation, whether it be cell cycle arrest, cell death, or other, are controlled through complex pathways involving many different proteins. For example, MYC, a transcription factor and proto-oncogene plays a key role in determining if p53-activation will result in cell cycle arrest or apoptosis [555].

The role of p53 in cell cycle arrest is mediated by another cell cycle protein, p21 [556,557]. DNA damage, for example from ionizing radiation, increases cellular levels of p53. In turn, p53 binds p21-regulatory elements, and activates p21 transcription. As a CDKI, p21 inhibits the function of many CDKs, thereby arresting cell cycle progression [556]. Most notably, p21 inhibits the complexes formed between CDK2 and either of two cyclins, cyclin A or cyclin E. These complexes are required for progression through the G1/S checkpoint, and inhibition by p53 and p21 leads to cell cycle arrest at this checkpoint (reviewed in [558]). In normal cells, p21 is found in a complex of proteins including PCNA [559]. PCNA is a proliferation factor, with dual roles in both DNA

replication and repair (discussed further in section 3.3.2.2). In G1 arrest, p21 inhibits the function of PCNA in DNA replication, but allows its function in PCNA-dependent nucleotide excision repair [560]. Due to the general nature of the CDK inhibition by p21, the roles of p21 and p53 are not limited to the G1/S checkpoint. These proteins have proven to be essential for sustaining arrest at the G2/M checkpoint as well [561].

Rodriques and colleagues reported in 1990 that p53 mutations in colorectal cancer could be identified by overexpression of the protein [551]. This overexpression can be identified through immunohistochemical analysis. Although p53 plays such a prominent role in the development of colorectal cancer, its role as a prognostic or predictive marker has not been well established. While many groups have studied the potential prognostic use of p53 in colorectal cancer, the results have been varied. Some groups have found that high levels of p53 expression were associated with a poorer outcome [562-564], while others have found the opposite, that high expression levels were associated with a better outcome [565,566], and others have found no significant correlations [567-569].

3.3.1.3 p16

Cyclin-dependent kinase inhibitor 2A (*CDKN2A*) is a tumour suppressor gene which acts as a part of the G1/S cell cycle check point, and is upregulated in cellular senescence. It encodes $p16^{INK4A}$, one of the three genes with tumour suppressor functions located within a 35 kilobase region on the short arm of chromosome 9. An alternate reading frame within *CDKN2A* gives rise to $p14^{ARF}$, which is functionally unrelated. The

third protein encoded in this chromosomal region is $p15^{INK4B}$. The $p16^{INK4A}$ and $p15^{INK4B}$ proteins are both members of the INK4 family of CDKIs, which act by specifically binding CDK4 and CDK6. Thus, when p16 levels are elevated, they block these CDKs from associating with the regulatory D cyclins. This prevents phosphorylation of the pRb protein, a key tumour suppressor, and causes cell cycle arrest at the G1/S checkpoint [570].

Cellular senescence is defined as an irreversible state of growth arrest (reviewed in [571]). In this state, cells no longer proliferate, but are still stable and metabolically active. It was first described in relation to the limited lifespan of human fibroblast cells in culture [572,573], and was initially considered to play a role in the normal aging process [573]. In 1961, Hayflick and Moorhead reported that cancer cells did not enter this growth arrest, but could proliferate indefinitely [572]. They hypothesized that cellular senescence was attributable to intrinsic factors of the cell. The three "Hayflick Factors" are: telomere shortening, accumulation of unrepaired DNA damage, and upregulation of the *CDKN2A* (p16/p14) locus [571]. It is now recognized that in addition to the aging process, senescence can also be initiated in response to various types of cellular stress, including oncogenic stimulation. In 1997, Serrano and colleagues described that expression of oncogenic *ras* in primary human and rodent fibroblasts resulted in an upregulation and accumulation of both p53 and p16, and permanent cell cycle arrest in the G1 phase [574]. It has since been shown that p16 is activated in response to other oncogenic and mitogenic stimulation, and is thus involved in protection from cancer [575-579]. Expression of p16 in colorectal cancer is regulated, at least in part, by β-catenin [81]. With increased nuclear accumulation of β-catenin, transcription is initiated in multiple target genes, including *CDKN2A*.

Inactivation of p16 by *de novo* methylation of the *CDKN2A* promoter region CpG island is seen in about 20% of human cancers [273], and 30% to 40% of human colorectal cancers [580,581]. In 2009, Mitomi and colleagues reported that p16 methylation results in transcriptional silencing, and is associated with large tumour size, increased risk of recurrence, and reduced disease-specific survival [582]. Methylation of p16 is associated with the CpG island methylator phenotype (CIMP) in colorectal cancer [583], which was described in section 1.7.3. Payá and colleagues described the use of p16 immunohistochemistry (IHC) to help identify patients with Lynch Syndrome [583]. While tumours negative for MSH2 or MSH6 usually have associated mutations and are positive for Lynch Syndrome, the majority of CRCs that are negative for MLH1 on IHC are so due to methylation of the MLH1 promoter region, as part of the CIMP pathway. Payá and colleagues suggested that those tumours with methylation of the MLH1 promoter region would also have methylation at other loci, including the p16 promoter. Therefore, they concluded that p16 IHC should be used in conjunction with MLH1 IHC, to identify patients with Lynch Syndrome. Those tumours without staining for either MLH1 or p16 were likely due to CIMP, whereas those with absence of MLH1, but presence of p16 were more likely to be due to Lynch Syndrome[583].

Xie and colleagues reported a positive correlation between p16 expression levels and tumour progression [584]. They performed p16 IHC on TMAs including normal colonic mucosa, adenomas, and CRCs with different depths of invasion (pT), as well as lymph nodes and distant metastases from CRC primary tumours. They found that there was a progressive increase in the percentage of tumours which overexpressed p16, with

the lowest levels in normal colonic mucosa (0% overexpressed p16), and highest levels in lymph node metastases (74% overexpressed p16). Levels dropped off in distant metastases, with 33% of these overexpressing p16. They concluded that p16 expression was associated with tumour progression and lymph node metastasis [584].

Despite the recognized role of p16 in colorectal cancer, its role in patient prognosis has yet to be fully determined. A recent study by Shima and colleagues reviewed a number of previous studies that examined the prognostic value of p16 in CRC, and they also examined their own cohort of more than 900 CRC patients [581]. They determined that most previous studies of p16 prognosis did not account for the confounding effects of the association of p16 methylation with the CIMP pathway. Without accounting for this confounder, some studies did report a correlation between p16 methylation and poor prognosis [585], including the study by Mitomi and colleagues, which was described above [582]. Other studies concluded that there was not an independent prognostic effect [586,587]. In their own study, Shima and colleagues adjusted for CIMP status and other potential confounders. They reported that while p16 methylation was associated with shorter overall survival in univariate analysis, it was not independently significant in multivariate analysis when other confounders, including CIMP status, were evaluated [581].

Expression of p16 along the infiltrative front of invasion of CRC has been associated with tumour budding, which is in turn associated with poor prognosis [81]. In a 2009 study by Wassermann and colleagues [81], in multivariate analysis, adjusting for sex,

age, T, N, and M stages, tumour grade, and adjuvant therapy, expression of p16 at the invasive front of the tumour was significantly correlated with poor patient survival.

3.3.1.4 p21

The *CDKN1A* gene encodes the p21 protein. The main role of p21 has already been discussed in section 3.3.1.2, in the context of its induction by p53, leading to cell cycle arrest. It is a member of the CIP/KIP family of CDKIs, and is a broad spectrum inhibitor of CDKs. This protein was first reported by two independent groups, in the same issue of the journal Cell, in 1993 [556,557]. El-Deiry and colleagues described a gene they called *WAF1*, whose expression was induced by wild type, but not mutant p53 [556]. They suggested that this gene could be an important mediator of p53-dependent tumour suppression. Meanwhile, Harper and colleagues described a gene they called CDK-interacting protein 1 (*CIP1*), which encoded a protein that interacted with and inhibited CDKs involved in the G1/S cell cycle checkpoint, leading to cell cycle arrest. They named this 21 kD protein $p21^{CIP1}$, and suggested that it was a potent inhibitor of CDKs, and was involved in cell cycle control [557].

Expression of p21 can also be induced by transcription factors other than p53. In human pancreatic adenocarcinoma cell lines p21 expression has been induced by transforming growth factor beta [588], suggesting that p21 may have other physiologic roles.

Although best known for its role as a cyclin-CDK inhibitor, p21 is also involved in some active cyclin-CDK complexes, and may be required for proper assembly and stabilization of these complexes, and for normal cell cycle progession [589-591].

The interaction of p21 and PCNA was briefly described in section 3.3.1.2. This interaction provides p21 with a regulatory role in both DNA replication and DNA damage repair [559,560]. In general, p21 inhibits various DNA repair pathways by disrupting the interactions between PCNA and DNA repair molecules (reviewed in [592]).

The p21 protein is also involved in cellular senescence and aging. Cellular senescence, a permanent cell cycle arrest, is controlled by two main pathways, one involving p16 and the retinoblastoma protein (pRb), and the other involving p21 and p53. However, p21 can also induce senescence independently from p53 [593]. It has been found in high levels in senescent fibroblasts from patients with premature aging syndromes including Werner Syndrome [594].

Some research groups have identified a correlation between p21 expression in colorectal cancer and patient outcome, while others have found no such correlation. In 1999, Ropponen and colleagues [595] reported that disease-specific survival and recurrence-free survival were both significantly lower in patients whose tumour cells had either a low intensity or a low frequency of p21 expression. Similarly, two years later Holland and colleagues [596] reported that patients had increased survival if their tumours expressed p21 in more than 50% of cells, although this did not reach statistical significance. They attributed this increased survival to the association they identified between p21 expression and the subcellular localization of cyclin D1.

Tornillo and colleagues later reported that in their large study of 1274 mismatch repair proficient CRCs, p21 was not a significant independent prognostic factor [597]. However, when p21 was included with other tumour markers including p27 and either p53 or bcl-2, these multimarker phenotypes were significant prognostic indicators. In 2008, Ioachim also reported that p21 was of no prognostic value of its own [598].

Zlobec and colleagues studied MSI-H CRCs, and reported in 2008 that loss of p21 was an independent adverse prognostic factor [599], which is consistent with the earlier reports discussed above. They also reported that the prognostic utility of p21 was greater when evaluated in a combined phenotype with other markers.

In 2010, Cacina and colleagues reported that certain functional polymorphisms of the *CDKN1A* gene were associated with a higher risk of developing colorectal cancer [600].

3.3.1.5 p27

The *CDKN1B* gene encodes the p27 protein. Like p21, it is a member of the CIP/KIP family of CDKIs, and is a broad-spectrum inhibitor of CDKs. The p27 protein was first reported by Polyak and colleagues in 1994 [601]. They identified it as a protein which linked transforming growth factor beta and cell-cell contact inhibition to cell cycle arrest at G1. They called the 27 kDa protein Cdk inhibitory protein 1 (KIP1), or $p27^{KIP1}$. The same group cloned the $p27^{KIP1}$ gene, *CDKN1B*, later that year, and reported that it shared sequence homology with the $p21^{CIP1}$ gene, *CDKN1A* [602]. Toyoshima and Hunter

reported that p27 associated predominantly with the cyclin D1-CDK4 complex, but also with other cyclin-CDK complexes [603].

As a tumour suppressor, p27 acts by interacting with cyclin E-CDK2 and other cyclin-CDK complexes, thereby blocking activation of these complexes, and inducing cell cycle arrest [601]. However, the results of its interactions with some of these complexes, including the cyclin D-CDK4,6 complexes, can be more complicated. In different growth states of the cell, and with different configurations of phosphorylation of the p27 protein, it can act as a pro-tumorigenic oncoprotein [604].

Downregulation or disregulation of p27 has been identified as a negative prognostic factor in numerous types of human cancers [604]. The oncogenic potential of p27 is not dependent on cyclin-CDK complexes, but instead on translocation from cell nucleus to cytoplasm. It is this misregulation, and not mutations of p27, that are found in different cancers [604,605]. Loss of p27 in the tumour nuclei is correlated with poor outcome in breast cancer [606-608], prostate cancer [609,610], gastric cancer [611], lung cancer [612], and oesophageal cancer [613]. An exception to this was recently discovered, as germline mutations in the p27 gene are responsible for multiple endocrine neoplasia type 4 (MEN4) [614,615], an inherited predisposition to parathyroid and pituitary tumours, in both rats and humans.

As with many other markers, however, the prognostic utility of p27 in colorectal cancer is not clear. Some research groups have reported that low levels of nuclear expression, or cytoplasmic localization of p27 expression is also associated with poor outcome in CRC patients [597,616,617], similar to the findings in other cancers. Other groups

have reported p27 to have no prognostic value in CRC [569,598], while one group reported that altered expression of p27 was associated with improved outcome [618].

3.3.1.6 Cyclin D1

While the other cell cycle proteins described in the preceding sections all have roles in tumour suppression and cell cycle arrest, cyclin D1 does not. As a cyclin, it instead promotes progression through the cell cycle. Cyclin D1, expressed by the *CCND1* gene, accumulates in response to signalling by mitogenic growth factors. It reaches its highest expression levels in mid to late G1, and acts primarily as the regulatory unit of a holoenzyme which initiates the phosphorylation-dependent inactivation of the tumour suppressor protein pRb [619,620], thereby promoting progression through the G1/S cell cycle checkpoint. To do this it must bind one of its CDKs, either CDK4 or CDK6, which can be inhibited through their binding of the INK4 CDKIs, including p16. The cyclin D1-CDK4/6 complex must incorporate a member of the CIP/KIP CDKI family (usually p21 or p27), in a non-inhibitory role. This aids in assembly of the cyclin D1-CDK4/6 complex [590,591], insures nuclear localization of the complex during the G1 phase [621], and sequesters the CDKI, thereby facilitating activation of the cyclin E-CDK2 complex, which is required for further cell cycle progression [601,603].

As a protein involved in the mitogen-dependent progression of the cell cycle, overexpression and accumulation of cyclin D1 can lead to a distinct growth advantage – a

hallmark of malignancy. Such overexpression of cyclin D1 is seen in many cancers; as a result of chromosomal translocations, gene amplification, and deregulation of cyclin D1 protein degradation [622]. Cyclin D1 was initially cloned and recognized as a candidate oncogene, when Motokura and colleagues noticed chromosomal inversions in three independent parathyroid adenomas, which linked the parathyroid hormone gene to a locus at chromosome 11q13 [623]. They initially called it *PRAD1*, for parathyroid adenomatosis, but quickly realized that it had DNA which matched the conserved sequences shared by cyclins, and that it could play roles in the development of various tumours.

As predicted by Motokura and colleagues [623], overexpression and increased accumulation of cyclin D1 in tumour nuclei has since been identified in a number of different cancer types. Chromosomal translocations involving cyclin D1 are found in approximately 70% of mantle cell lymphomas [624], and 15% to 20% of multiple myelomas [625]. Gene amplification is a cause of cyclin D1 accumulation in approximately 40% to 50% of non-small cell lung cancer [626-628], and 30% to 50% of squamous cell carcinomas of the head and neck [629,630]. For some cancers, researchers have reported gene amplification, and also higher levels of cyclin D1 expression which is unrelated to gene amplification [631]. Examples of this are pancreatic cancer, in which gene amplification was reported in 25% of tumours, but nuclear overexpression detected by immunohistochemistry was recorded in 68% of tumours [632], and breast cancer, in which gene amplification is reported in 15% to 30% of tumours, while overexpression is observed in approximately 65% of tumours [633,634]. This overexpression above the level of gene amplification is likely due to deregulation of cyclin D1 protein degradation [622].

In the above-mentioned cancers for which cyclin D1 overexpression has been reported, this has generally been linked to a poor patent prognosis [628,629,632,634].

In colorectal cancer, the predominant mechanism which leads to cyclin D1 overexpression is mutation in the *APC* gene or *WNT* signalling pathway (described previously in section 1.7.1). Mutations in the *APC* gene occur in the common chromosomal instability pathway to colorectal cancer development. Such mutations lead to an accumulation of b-catenin, which in turn activates a family of transcription factors. The *CCND1* gene is responsive to these transcription factors, and overexpression of cyclin D1 ensues [196]. This can be a contributing factor in the initiation of colorectal neoplasia [635].

Cyclin D1 is overexpressed in approximately one third of colorectal tumours [635,636]. The prognostic value of this overexpression is somewhat conflicted, however. Some studies have shown cyclin D1 overexpression to be associated with a poor prognosis [637,638], others fund an association with an improved prognosis [596,636], yet most have shown it to have no independent prognostic value [639-643]. Of these studies, perhaps the most robust and comprehensive was that of Ogino and colleagues, in 2009 [636]. They used a cohort of 602 patients with CRC, and adjusted for other factors which were known to influence prognosis, including certain patient demographics, and molecular characteristics such as MSI status, CpG island methylator phenotype, patient BMI, and mutation status of the *BRAF* and *KRAS* genes. They found cyclin D1 overexpression in CRC to be an independent positive prognostic indicator; patients whose tumours overexpressed cyclin D1 had a lower risk of disease-specific death than those whose

tumours did not overexpress this protein. The adjusted hazard ratio was 0.57, with 95% confidence interval 0.39 to 0.84, and *P* value of 0.0048 [636].

3.3.2 Markers of Cellular Proliferation - Ki-67 and PCNA

3.3.2.1 Ki67

The growth rate of a tumour can be evaluated by two parameters: the growth fraction, which is the percentage of proliferating cells, and the rate of proliferation, which is the time taken to complete the cell cycle [644]. Although the proliferation rate is difficult to assess by immunohistochemistry, antibodies for two proteins have been used repeatedly to assess the growth fraction of tumours, they are Ki-67 and PCNA.

Ki-67 is a marker of cellular proliferation. This protein is only expressed in the nuclei of proliferating cells, during the S, G2, M and sometimes also G1 phases of the cell cycle. Continuously cycling cells entering G1 after mitosis do express Ki-67, but cells entering G1 from G0 do not [645]. Ki-67 is a cell cycle regulating protein. It has some structural elements found in other known cell cycle regulating proteins, such as a forkhead-associated domain [646], more than 200 phosphorylation sites for many different kinases [647], and some nuclear targeting sequences [647], but the exact function of Ki-67 is unknown. Little progress has been made on elucidating the function of the protein, since overall it does not have close structural similarity to any other known proteins [648].

Although the function is unknown, it is known to be required for maintenance of cellular proliferation. This was confirmed when DNA synthesis was inhibited in cell lines by treating them with antisense oligonucleotides specific for Ki-67 cDNA [647]. It has also been determined that it is not involved in DNA repair [649].

In normal colonic mucosa Ki-67 expression is seen in the depths of the crypts, in the proliferative zone [650]. Within colorectal tumours, the percent of cells expressing Ki-67, known as the Ki-67 labelling index (Ki-67 LI), increases with histologic grade, and varies among different histologic types [650]. While Ki-67 is detected in almost all CRC tumours, the reported Ki-67 LIs for CRC range from 33% [651] to 59% [652]. When assessing association with prognosis, a cut-off is often imposed upon Ki-67 LI for the purpose of analysis. Different research groups have used different cut-offs, ranging from 5% [643] to 60% [567], to compare the tumours with Ki-67 expression levels below, with those above, the assigned cut-off. This may explain, in part, why there have been inconsistent conclusions as to the prognostic value of Ki-67 in colorectal cancer.

Ki-67 expression significantly correlates with prognosis in the majority of studies of some other cancer types (reviewed in [653]), including breast cancer, soft tissue tumours, lung cancer, astrocytoma and meningioma. However, there have been inconsistent and conflicting reports of the correlation with prognosis in colorectal cancer [653,654]. Allegra and colleagues published in 2002 [567] that they found no significant correlation between Ki-67 expression and outcome, when they used a cut-off of 60% positive cells. A year later the same group published [562] that there was a significant correlation, with better outcomes for both overall survival and disease-free survival, when at least 40% of tumour

cells expressed Ki-67. Rosati and colleagues used an even lower cut-off, of 10%, in their study [655], but found no correlation. Brown's review [653] refers to ten more studies that were published in the late 1990's, which all found no correlation between Ki-67 expression and prognosis

In 1999 Palmqvist and colleagues published a paper [656] in which they described an association between low levels of Ki-67 expression and poor prognosis. Just months later, Kimura and colleagues published opposing findings [657]. They reported an association between high Ki-67 levels and poor outcome. These results were supported by Canna's group, who reported that increased tumour proliferation, as measured in part by Ki-67 LI, was associated with a systemic inflammatory response in the patient, and poorer disease-specific survival [658]. Anjomshoaa's group had related findings [659], although their measure of proliferative activity did not include Ki-67 levels. They found that higher proliferation levels in colorectal tumours were associated with reduced aggressive behaviour. The reverse of this has been found in other cancer types including breast, whereby higher proliferation levels have been associated with more aggressive behaviour [659].

Some other literature has, however, sided with Palmqvist's group; stating that high levels of Ki-67 expression indicated a better prognosis. Garrity's group found an association between Ki-67 expression in colon cancer and both better overall and disease-free survival, when using a cut-off of 27% [566]. Interestingly, Hilska's group found a similar association using a cut-off of 5%, but only with rectal cancers, not colon cancers

[643]. With such conflicting reports, the value of Ki-67 as a biomarker for colorectal cancer still needs to be studied further.

3.3.2.2 PCNA

Proliferating cell nuclear antigen (PCNA) is another marker of cellular proliferation; it is only expressed in proliferating cells [660]. PCNA levels start to accumulate in the G1 phase of the cell cycle, reach their highest levels during the S phase, then disappear by the end of mitosis [661]. It plays major roles in both DNA synthesis and repair. PCNA is a cyclin; it is involved in controlling progression through the cell cycle, yet it possesses no enzymatic activity. PCNA forms a ring-shaped homotrimer that encircles DNA, and acts as a sliding clamp. It is an important cofactor for the DNA polymerases; tethering them to the DNA strand during replication, thereby greatly increasing their processivity [662]. In normal colonic mucosa PCNA is expressed in approximately 30% of the cells [663], but only those of the proliferative zone in the intermediate and deep levels of the crypts [660,664]. There is little or no expression observed in the upper areas of the crypts and the mucosal surface, as the cells at these levels are differentiated. Within colorectal tumours, a higher percentage of cells express PCNA, but the cells are dispersed throughout the tumour, and not limited to any particular region within it.

PCNA is required for the DNA mismatch repair system. PCNA interacts directly with MLH1, MSH2, MSH6, and MSH3 mismatch repair (MMR) proteins [662,665].

Absence of functional PCNA leads to inactivation of the MMR pathway, and reintroduction of PCNA can rescue it [665]. PCNA is involved not only in the synthesis component of the DNA repair, but also in earlier events in the MMR pathway. It is involved in orienting the repair system to the damaged DNA strand, and recruiting many of the required factors to the repair site [662,665].

PCNA is also required for other DNA repair pathways, including base excision repair, nucleotide excision repair, double strand break repair through homologous recombination, and translesion synthesis [666]. Its function in the error-prone translesion synthesis is also involved in creating mutations in the hypervariable region of immunoglobulin genes [667] .

PCNA has binding regions for many other types of proteins as well, including DNA ligase, topoisomerase, DNA methylatransferase, helicases, and protein kinases. PCNA also binds apoptotic factors, cell cycle regulators, and enzymes involved in different DNA repair pathways. For its many crucial roles in DNA replication, some have termed PCNA "the maestro of the replication fork" [662]. Through its interaction with different proteins PCNA is involved in joining of Okazaki fragments, DNA methylation, chromatin assembly, and cell cycle control [662]. It also plays a role in down-regulating cell apoptosis in response to DNA damage [668].

The use of PCNA as a marker of cell proliferation has been assessed previously for its possible correlation with prognosis in colorectal and other cancers. PCNA staining by immunohistochemistry is scored quantitatively. The number of positively staining cells is divided by the total number of cancer cells in the microscope field, to produce the

PCNA labelling index (PCNA LI). Some studies have looked for correlations using the PCNA LI itself, while others have used a cut-off, typically of 25% of cells. For some types of cancer a correlation has been made between increased proliferative activity and increased malignant potential [669]; however, the results for colorectal cancer have been inconsistent.

Sun and colleagues [670] found no correlation between PCNA expression and prognosis or any clinicopathological variable. Guerra and colleagues [671] also found no significant correlation; however, they did observe a trend of increased survival for patients with a higher proliferation index. On the other hand, both Mayer [672] and Al-Sheneber [669] and colleagues found significant correlations with PCNA and prognosis. Both groups found an inverse relationship between survival times and PCNA LI; patients whose tumours had the highest proliferation indices had the shortest survival times [669,672]. Al-Shebener's group [669] also found a direct correlation with PCNA LI and recurrence rate. Onodera and colleagues reported in 1998 [673] that they found PCNA to have prognostic significance in recurrent rectal cancer. Earlier the same year, the same group published another article [674] in which they found PCNA to have no prognostic value in primary colorectal cancer. In the same article, however, they did find significant correlations between PCNA expression and many clinicopathological factors, including tumour size, depth of invasion, lymphovascular invasion, and histological type [674].

Many groups have looked at PCNA correlations with clinicopathological factors. Georgescu and colleagues [650] found a significant correlation of PCNA LI and tumour grade: the higher the PCNA LI, the more poorly differentiated the tumour cells.

Kanazawa and colleagues [675] also found a correlation with tumour grade, as well as other features including depth of invasion, and lymphovascular invasion. However, Kanazawa's group only found these correlations when evaluating what they termed "polypoid type" cancers. These correlations did not hold true for their other category of CRCs, the "flat type". This group also reported that for both types of CRCs, PCNA-LI was higher in cases with lymph node metastases, when compared to those without. Similarly, Guzinska-Ustymowicz and colleagues [663] also observed a correlation with both PCNA and Ki-67 expression and lymph node metastases, but they only studied moderately differentiated CRCs with a depth of invasion into the subserosal fat.

With the conflicting results obtained by different research groups, the use of PCNA as a prognostic marker for colorectal cancer has yet to be fully defined. Further complicating this situation is the fragile, finicky nature of the immunoreactivity of the PCNA protein. The use of heat to aid in adhering cut tissue sections to glass microscope slides greatly reduces the immunoreactivity of PCNA, as does extended tissue fixation time [660]. Hall and colleagues [660] found that 48 hours of fixation lead to a great reduction in PCNA staining, and that 72 hours of fixation resulted in a complete lack of staining. However, with the multiple, important roles PCNA plays in the cell, it is an intriguing potential biomarker.

Because both Ki-67 and PCNA are markers of cellular proliferation, it could be argued that it is redundant to study them both. In fact for some cancers, such as non-Hodgkin's lymphoma, a linear relationship, with correlation coefficient of 0.91, has been demonstrated between the number of cells expressing PCNA and the number of cells

expressing Ki-67 [660]. The same has not been found for some other cancers types, including breast, in which PCNA and Ki-67 expressions are not correlated [676]. It has been stated that Ki-67 expression more truly represents the growth fraction of tumours [650,660,676]. PCNA LI is consistently higher than KI-67 LI, and may over-represent the growth fraction due to its longer half life, its role in DNA repair and other cellular functions, and perhaps even due to autocrine or paracrine growth factor signalling [650,660]. Both antigens have potential roles as biomarkers in colorectal and other cancers, and it has been recommended that both be used [650].

3.3.3 Cell Adhesion Molecules - β-Catenin and E-Cadherin

3.3.3.1 β-Catenin

Encoded by the *CTNNB1* gene, b-catenin is a proto-oncoprotein. It is a multifunctional protein, with roles in cell adhesion, signalling pathways, and transcription. It is an adherens junction protein, essential in cadherin-mediated cell adhesion [677,678]. The extracellular domains of E-cadherin proteins from adjacent epithelial cells bind to one another in a calcium-dependent manner [679], leading to cell adhesion. The intracellular domains of these E-cadherin proteins bind to b-catenin, which in turn binds to α-catenin, which binds the actin cytoskeleton [680,681]. Cellular adhesion is required for maintaining the multicellular structure and organization of tissues and

organs. It is also important for embryonic development, stem cell formation, cellular differentiation, and motility.

In the normal developing embryo, β-catenin and E-cadherin are involved in the epithelial-mesenchymal transition (EMT). This process involves the conversion of immobile epithelial cells into highly motile fibroblastoid mesenchymal cells, and is required for the formation of the mesoderm during gastrulation and the development of the neural crest and craniofacial system [682]. In a tumour, loss of E-cadherin or nuclear accumulation of β-catenin (through Wnt system activation, described below) can promote an EMT-like de-differentiation of tumour cells, whereby they acquire the ability to migrate. Alterations of the structure or expression of these proteins may also play a role in attachment of these migrating tumour cells to other organs including the liver, resulting in an increased risk of the development of distant metastasis [54].

The second major function of β-catenin is in the nucleus, where it acts as a transcriptional co-activator, activating genes responsive to Wingless Int (Wnt) signalling. This function of β-catenin is considered to be linked to its role in cell adhesion, integrating cellular morphogenesis with gene expression [682,683]. The Wnt/β-catenin signal transduction pathway is important in normal embryonic development and adult homeostasis, as well as in various diseases processes, including congenital malformations and cancer [684]. In the absence of Wnt, β-catenin is located in the cytoplasm, bound by a complex of proteins which is known as the β-catenin destruction complex. It includes axis inhibitor 1 (AXIN1), the adenomatous polyposis coli (APC) tumour suppressor, casein kinase 1 (CK1) and glycogen synthase kinase 3 (GSK3). This complex facilitates

phosphorylation of β-catenin, which targets it for ubiquitination and proteosomal degradation, preventing β-catenin from reaching the nucleus. However, in the presence of Wnt binding to its cell surface receptors, other proteins, including members of the β-catenin destruction complex, are recruited to the receptors. This disables the β-catenin destruction complex, resulting in the stabilization of β-catenin, and its subsequent cytosolic accumulation and nuclear translocation. Within the nucleus β-catenin interacts with members of the T-cell factor/Lymphoid enhancer factor (TCF/LEF) transcription factor families, thereby displacing co-repressors and activating the transcription of many different target genes, depending on cellular context [685]. These target genes are involved in many processes including the regulation of cellular fate, cell cycle progression and proliferation, cellular differentiation, inhibition of apoptosis, and cellular migration [194]. The protein products of these targets include: c-Myc, a transcription factor and proto-oncoprotein [195]; VEGF, a growth factor involved in angiogenesis [686]; cyclin D1, a cyclin involved in cell cycle progression [196]; and p16, a tumour suppressor whose expression in CRCs, especially at the infiltrating edge, is associated with poor outcome [81].

The requirement of APC in the β-catenin destruction complex is significant to colorectal carcinogenesis. As discussed in section 1.7.1, inactivation of the *APC* gene is an early occurrence in the development of the majority of sporadic CRCs, as well as the hereditary CRCs associated with FAP (those that develop through the chromosomal instability pathway)[191-193]. Without a functional APC protein, the destruction complex no longer sequesters β-catenin or targets it for degradation. Through inactivating mutations

of *APC* or stabilizing mutations of β-catenin there is an accumulation of β-catenin within the nucleus, and subsequent constitutive activation of Wnt target genes.

CRC stem cells are characterized by activation of the Wnt signalling pathway, through nuclear accumulation of β-catenin [338,687,688]. In order for a tumour to metastasize, cells capable of self-renewal must dissociate from the primary tumour, then convert back to a non-motile form which once again has the ability to form adherens junctions to new cells and form a solid metastasis. These cells must therefore be capable of an EMT-like transformation, as described above, and also have the self-renewal ability of stem cells. As β-catenin is involved in both stem cell formation and the EMT transition in the developing embryo, aberrant nuclear accumulation of β-catenin can confer these two abilities onto tumour cells, promoting tumour invasion and metastasis [339]. Since β-catenin is considered a marker of CRC stem cells [688,689], tumours with nuclear accumulation of this protein may have a more stem cell-like phenotype, and thereby exhibit more aggressive behaviour, with high risks of tumour progression and decreased patient survival [688].

As expected, nuclear accumulation of β-catenin in CRC is associated with decreased patient survival in most publications [81,688,690,691]. In a 2007 study [690], Lugli and colleagues stratified their cohort of 1420 patients with CRCs into three groups: those that were mismatch repair proficient, those that were deficient for MLH1 expression, and those with presumed Lynch syndrome. They reported a significantly higher proportion of tumours with high nuclear β-catenin expression in the MMR proficient group. They also reported that this increased nuclear expression was significantly associated with increased

lymph node involvement, vascular invasion, and worse survival in the MMR proficient group. In the presumed Lynch syndrome and the MLH1 deficient groups, however, increased nuclear β-catenin expression was not associated with either tumour progression or worse survival [690]. These findings are as expected, since the Lynch-related and MLH1 deficient tumours would likely have developed through the MSI, and CIMP pathways, respectively, and not the APC/β-catenin-dependent chromosomal instability pathway.

3.3.3.2 E-Cadherin

E-cadherin, encoded by the *CDH1* tumour suppressor gene, is a member of the cadherin family of adherens junction proteins, which are found in the cell membrane as transmembrane glycoproteins, and act as calcium-dependent adhesion receptors. Members of the cadherin family are expressed in a tissue-specific pattern, with E-cadherin expressed in epithelial cells, where it regulates cell contact [679]. As described above, E-cadherin has roles in embryonic development, cellular differentiation, cell motility, and more; its suppression can lead to disassociation of cells, resulting in tumour invasion and metastasis [54]. Many of the roles described above for β-catenin also apply to E-cadherin, as they are both essential for adherens junctions, and because suppression of E-cadherin results in more available cytoplasmic β-catenin, which can then be relocated to the nucleus resulting in increased expression of Wnt target genes [690,692,693].

The role of E-cadherin in colorectal cancer is not as well characterized as the role of β-catenin. While some studies have reported that reduced membranous expression of

E-cadherin in CRC is associated with dedifferentiation, progression and metastasis [690,694-696], others have not. Some studies have reported that reduced E-cadherin expression is somewhat associated with tumour dedifferentiation, but not with progression or metastasis [584,697], and others have reported no association with E-cadherin expression and any of these tumour characteristics [698]. Dedifferentiation at the invasive edge of a tumour, or tumour budding, is considered a negative prognostic indicator in CRC. Loss of tumour E-cadherin expression is an independent predictor of tumour budding [83].

In other cancers, expression of the E-cadherin repressor Snail is associated with decreased survival in ovarian cancer patients [699], and loss of membranous expression of E-cadherin is correlated with tumour invasion, lymph node spread and liver metastasis in pancreatic endocrine tumours [693]. Loss of E-cadherin expression was identified in diffuse gastric cancer cell lines [700], as well as in similar patient tumours [701]. These cancers are made up of cells with greatly reduced cell-cell adhesion, and have been described as having a "scattered" phenotype [701]. Germline *CDH1* mutations resulting in loss of E-cadherin expression have since been implicated in autosomal dominant inherited predisposition to diffuse gastric cancer [702], in which asymptomatic mutation carriers have been identified with multifocal disease, and prophylactic gastrectomy has been recommended [703]. An increased risk of lobular breast cancer has been identified in these families [704]. Complete loss of E-cadherin expression has also been reported in most sporadic infiltrative lobular breast cancers [705,706], which have also been described as having a scattered phenotype [706]. Germline *CDH1* mutations have been identified in

families with inherited lobular breast cancer, in some cases with an absence of gastric cancer in the family [707].

3.3.4 Structural Proteins – Lamin A/C

Lamins are type V intermediate filament proteins. They bind together to form a protein meshwork that makes up the main architectural component of the nuclear lamina, which is found underneath the inner membrane of the nuclear envelope [708]. The nuclear lamina provides strength and support for the nuclear envelope, and the lamins are responsible for determining the size, shape, and strength of the nucleus [708,709]. The lamins also play an important role in the organization of the nuclear envelope; they are involved in the anchoring of various elements of the nuclear envelope to their correct positions, and recruiting proteins to the inner nuclear membrane [710-712]. They bind integral membrane proteins, DNA, and chromatin proteins [710,711,713]. In this way they are involved in chromatin organization and structure.

Three genes encode lamin proteins in human somatic cells. They are *LMNA*, *LMNB1* and *LMNB2*. Lamin proteins are divided into two categories. A-type lamins are produced by alternative splicing of *LMNA*, and include lamins A, AΔ10, C, and C2. B-type lamins are encoded by *LMNB1* (lamin B1), and *LMNB2* (lamins B2 and B3). A-type lamins are expressed only in differentiated cells [714], whereas B-type lamins appear to be required by all cells, and play a role in development [715,716]. Lamins C2 and B3 are only expressed in germ cells [717,718].

A-type lamins, lamins A and C, are found not only at the nuclear periphery, but also within the nucleoplasm. There they have been identified in discrete internal foci, in association with DNA and other proteins [719-722], and are likely involved in transcriptional activity, cell cycle progression and differentiation (reviewed in [723,724]). They also make up a veil-like lamin network within the nucleoplasm. The exact function of this network is unknown, but it has been proposed that it provides a scaffold for DNA replication and transcription [725-727]. While B-type lamins have also been identified in the nucleoplasm, their role there has not been as well studied. In the nucleoplasm A-type lamins bind lamin-associated proteins (LAPs). Nuclear lamins and LAPs have chromatin binding domains, indicating a possible direct role in the epigenetic regulation of gene expression through chromatin organization.

Lamins, especially lamin A, are involved in the apoptosis cascade, although they have not proven to be strong immunohistochemical markers for apoptosis [728]. Proteolytic cleavage of lamins results in degradation of the nuclear lamina, and therefore of the nucleus [729,730]. Lamin degradation in apoptosis requires both hyperphosphorylation by protein kinase C-δ, and proteolyic cleavage [731], which is mediated primarily by caspase 6 [732,733].

A key lamin-associated protein in the nucleosome is LAP2α. The LAP2α-lamin A/C complexes are involved in the regulation of gene expression through roles in both transcriptional activity and epigenetic control pathways (reviewed in [723]). These LAP2α-lamin A/C complexes also regulate the retinoblastoma protein (pRb), a tumour suppressor protein involved in some cancers. The LAP2α-lamin A/C complex binds pRb, and

anchors it in the nucleoplasm [734]. This nuclear retention of pRb is required for its proper functioning, which includes involvement in cell cycle control at the G1/S checkpoint, promotion of cellular differentiation, and roles as both a transcriptional activator and repressor (reviewed in [723]).

A-type lamins are also involved in the regulation of the oncogene β-catenin. Lamins A and C are required to anchor emerin to the inner nuclear membrane. Emerin, in turn, is responsible for preventing the accumulation of β-catenin in the nucleus, thereby inhibiting its functions. Without functional lamins A/C and emerin, β-catenin would accumulate in the nucleus, leading to unregulated signalling and auto-stimulatory growth [735].

The key roles played by the A-type lamins in both nuclear structure and transcriptional regulation have led to the conclusion that these lamins are essential for maintenance of somatic cells and tissues. Thus, the term "guardians of the soma" has been proposed [736]. Mutations in A-type lamins cause a variety of diverse degenerative disorders, collectively termed laminopathies. The different laminopathies affect different tissues, mainly of mesenchymal origin, including striated muscle, fat, bone, skin, and neuronal tissues. Symptoms of the various laminopathies can include cardiomyopathies, joint contractures, muscle wasting, altered fat distribution, peripheral neuropathy, and very tight skin, termed restrictive dermopathy. Some also involve premature aging, such as Hutchison-Gildford progeria syndrome.

In 1999, Moss and colleagues reported reduced expression of nuclear lamins A/C and B1 in gastrointestinal neoplasms, sometimes associated with aberrant cytoplasmic

expression [737]. They suggested that altered lamin expression could be a biomarker of gastrointestinal malignancy. In more recent years lamins have been studied for their potential role as biomarkers for a number of cancer types. Since B-type lamins are required for cell survival, they should be expressed in all normal and malignant cells. For this reason B-type lamins have limited function as cancer biomarkers [738]. A-type lamins, however, have more variable expression, and therefore have greater potential as useful cancer biomarkers [738].

A-type lamins have been considered potential cancer prognosis biomarkers for their roles in cellular differentiation, proliferation, and motility (reviewed in [738]). As stated above, A-type lamins are only expressed in differentiated cells. Reduced expression of A-type lamins has been observed in poorly differentiated squamous cell carcinoma [739] and gastric carcinoma [740]. Poor tumour cell differentiation is generally associated with poor prognosis. While little is known about the actual relationship between A-type lamin expression and cancer cell proliferation, it has been proposed that there would be an inverse correlation [738]. A 1997 paper by Broers and colleagues [741] described the expression of lamins in normal human tissues. They found that proliferating epithelial cells had reduced expression of A-type lamins. A 2006 paper by Ivorra and colleagues [742] found evidence supporting this relationship, as overexpression of A-type lamins inhibited cellular proliferation.

The role of A-type lamins in cell motility is perhaps the strongest factor in its potential as a prognosis biomarker. Expression of lamin A in colorectal cancer cell lines has been associated with increased invasive potential of the tumour. In a 2008 paper by

Willis and colleagues [337], lamin A expression led to an increased expression of T-plastin, an actin bundling protein, and reorganization of the actin cytoskeleton of the cell. This in turn led to decreased expression of the cell adhesion molecule E-cadherin, and increased invasive, or metastatic, potential of the tumour cells. In agreement with this, they found lamin A/C to be a significant negative prognostic biomarker for CRC.

Willis and colleagues also determined that lamin A is a probable stem cell marker in CRC [337]. Lamin A is expressed in the location at the base of the colonic crypt, thought to be the stem cell niche, and in the same number of cells (five to ten per crypt), expected to be stem cells [337]. Strengthening this association is the role of lamin mutations in degenerative laminopathies, thought to be due to impaired function and decreased longevity of adult stem cells. As well, a truncated version of lamin A responsible for the laminopathy Hutchison-Gilford progeria syndrome, has been found to play a major role in the regulation of mesenchymal stem cell homeostasis [743].

The stem cell model of CRC development suggests that CRC initiates from mutations in colonic stem cells, and that these stem cells are the only cells capable of metastases [324]. These cells are thought to be crucial for sustaining tumour growth [324], and tumours rich in these cells may have a more aggressive phenotype [744]. Since lamin A appears to be a CRC stem cell marker, it has been suggested that a tumour which expresses lamin A may have a more stem-cell-like phenotype, and thereby exhibit a more aggressive behaviour [744]. Willis' 2008 paper [337] did indeed find a significantly worse prognosis for CRC patients whose tumours expressed lamin A/C, as compared to those whose tumours did not. They found that patients whose tumours expressed lamin A/C

were almost twice as likely to die from cancer related causes (hazard ratio 1.85, 95% confidence interval 1.16-2.27).

3.3.5 DNA Helicase – MCM7

Minichromosome maintenance proteins (MCM) are a family of proteins, originally discovered in yeast, that are required for both the initiation and maintenance of DNA replication [745-749]. They are considered to be the licensing factors, first proposed in 1988 [750], which ensure that only one round of replication occurs per cell cycle [749], thereby coupling DNA replication and cell cycle progression, and maintaining genome integrity. MCM proteins are conserved among all eukaryotes and found in a simplified form in archaea [751,752]. All eukaryotes have six MCM proteins, named 2 through 7, which share similar sequence homology, and each contains an AAA^+ (ATPases associated with various cellular activities) domain [751].

The MCM2-7 proteins form a ring-shaped heterohexamer, with a central pore, which acts as a DNA helicase [753,754], unwinding the double stranded DNA ahead of the replication fork, providing the single stranded DNA template required by DNA polymerases for replication. These ring-shaped complexes are assembled just following mitosis, early in the first growth phase (G_1) of the cell cycle, and are transported to DNA origins of replication in an inactive form as part of a pre-replication complex. The presence of these complexes is essential for the activation of these origins of replication as cells enter the synthesis phase (S) of the cell cycle [745-748] (reviewed in [755-757]).

MCM7 is of particular interest because it directly interacts with many proteins involved in proliferation and cell cycle regulation including pRb [758], papillomavirus E6 protein [759], and cyclin D [760]. MCM7 has been used as a proliferation marker, and has been found to have more utility than Ki67 and PCNA as such a marker in glioblastoma [761], and other tumours [762]. It is upregulated in a number of cancers, including those of the cervix [763,764], prostate [765], head and neck [766], and lung [767], with elevated expression levels generally correlating with aggressive disease.

In a 2006 publication [768], Honeycutt and colleagues demonstrated that MCM7 was more than just a marker of proliferation, but actually contributed to tumorigenesis. They developed transgenic mice with deregulated MCM7 expression in the basal layer of the epidermis. These mice developed skin papillomas quicker than control mice. The papillomas of the test mice also had a much higher rate of conversion to squamous cell carcinomas than those of the control mice; indicating a role for MCM7 in papilloma formation and progression, as well as in conversion to malignancy [768].

The few published articles which assess the relationship between MCM7 and CRC suggest a correlation between high MCM7 expression levels and poor prognosis. In 2008 Nishihara and colleagues [769] reported that MCM7 expression was an independent prognostic factor in their cohort of 202 patients with Dukes' stage B and C cancers. Further, MCM7 positive, but Ki67 negative tumours were significantly correlated with positive lymph node status, development of distant metastasis, and higher clinical stage [769]. The same group of authors published again the following year [770], using a 191 of the 202 patients from their original study. They reported that MCM7, Ki67, and geminin

expression levels should be assessed together for predicting prognosis. In this study they reported that MCM7 was a significant prognostic indicator in univariate analysis, but did not report if it remained significant in multivariate analysis.

In 2010, Pillaire and colleagues [771] assessed MCM7 expression in CRCs from 74 patients with MSS CRCs, who did not receive adjuvant therapy. They considered expression of MCM7 in greater than 72% of tumour cells to be over-expression, and in less than 40% of tumour cells to be low-expression. In their cohort, overall survival was significantly reduced in patients whose tumours over-expressed MCM7. Two years following surgery, only about 40% of patients over-expressing MCM7 were still alive, whereas approximately 90% of patients with low-expressing tumours were still alive [771]. They noted that although MCM7 was used by some as a marker of proliferation, it was not associated with the level of PCNA expression, indicating that MCM7 over-expression was not related to the proliferation status.

3.3.6 Cell Signalling Molecule – EGFR

Epidermal growth factor receptor (EGFR) is a transmembrane protein which acts as a cell surface receptor, and is a member of the ErbB tyrosine kinase receptor family. There are four members of this family: EGFR/ErbB-1/HER1, HER2/neu/ErbB-2, HER3/ErbB-3, and HER4/ErbB-4. These receptors, and the more than 30 recognized ligands that associate with them[772], play a leading role in a complex signal transduction network. These receptors are inactive in isolation; the intracellular tyrosine kinase

domains are activated through the cell-surface binding of certain ligands, mostly growth factors, which induce the formation of receptor homodimers or heterodimers with other members of the ErbB family[772]. Through phosphorylation, these activated receptors then stimulate many signal transduction cascades, including: the RAS-RAF mitogen-activated protein kinase (MAPK) pathway, the phosphatidylinositol 3-kinase (PI3K)-AKT-PTEN pathway, and the phospholipase C pathway [773]. In this complex signalling network, the specific receptor and ligand combination determines the downstream pathways which will become activated, which types of transcriptional control will result, and what the eventual outcome will be [772]. Activation of EGFR, since a growth factor receptor, often results in cellular proliferation, but can also be involved in cellular migration, cellular adhesion, differentiation, dedifferentiation, inhibition of apoptosis, angiogenesis, or other functions [772,774].

While expression of EGFR is required for normal growth and epithelial development in several organs in mammals [775-777], its over-expression is observed in many cancers, and promotes solid tumour growth [778]. In many cancers, including ovarian, cervical, bladder, lung, and oesophageal, EGFR over-expression is often associated with a poor patient outcome, including reduced relapse-free survival time, and reduced overall survival [778]. In other cancers, including colorectal, EGFR over-expression is not considered to be a strong prognostic indicator, as different studies have reported conflicting results [774,778]. While some studies have reported EGFR expression in tumour cells to have no prognostic utility for CRC patients [779,780], others have found that higher levels of expression are associated with a poorer outcome [672,781,782]. Some studies

have found an association with EGFR expression and other negative prognostic indicators, and inferred its prognostic utility [783,784]. The variability in these reports is due in part to the variability in protocols used, including tumour fixation times, antibody assessed, and scoring system used; as well as different populations studied [782].

Since EGFR plays a leading, upstream role in such a large, complex signal transduction network, and its over-expression is observed in many cancers, inhibition of its cellular actions may have therapeutic benefits for patients. Ant-EGFR therapies with two main mechanisms of action have been developed. The first, monoclonal antibodies (mAbs) against EGFR, bind to the extracellular portion of the protein, prevent ligand binding, and thereby inhibit signalling pathways. The second, tyrosine kinase inhibitors (TKIs) bind to the intracellular portion of the protein, and prevent tyrosine kinase activation, thereby inhibiting signalling pathways [785]. To date, only certain mAbs have been approved for clinical use, including Cetuximab and Panitumumab [773,774], and are used in patients with refractory metastatic disease.

While anti-EGFR therapies are available, the tumour expression level of EGFR is not a reliable indicator of drug response [786]. Several important downstream proteins can be activated independently of EGFR, thereby rendering a tumour resistant to anti-EGFR therapy [774]. Approximately 40% of metastatic CRCs are sensitive to anti-EGFR treatment, the remaining 60% do not depend on EGFR for tumour growth [773]. *KRAS* is the most commonly mutated gene downstream of EGFR, with somatic, activating mutations found in approximately 35% to 40% of CRCs [787-790], which result in tumour resistance to anti-EGFR therapy. Mutations in other key proteins downstream of EGFR

appear to be responsible for the remainder of anti-EGFR drug resistance, including: the V600E mutation in *BRAF* [318], which is found in about 10% of CRCs [773]; activating mutations of the *PIK3CA* gene [791,792]; and inactivating mutations of the *PTEN* gene and hypermethylation of its promoter [793].

3.4 Manuscript – "Lamin A/C Expression is a Prognostic Indicator in Stage III Colorectal Cancer and a Predictor for the Development of Distant Metastasis"

Due to the limited number of publications on the utility of lamin A/C as a prognostic marker in CRC, and the biological relevance of this protein to CRC, I was particularly interested in exploring this potential association. Our findings indicate a strong prognostic role for lamin A/C in patients diagnosed with stage III disease. Patients with high levels of expression of this protein in tumour nuclei were more likely to develop distant metastasis and more likely to die of their disease, when compared with patients whose tumours expressed this protein at low levels. These findings were statistically significant, and independent of other known prognostic indicators.

The following manuscript has been submitted for publication. My role in this study was in the ascertainment of patients and the collection of tumours from the majority of cases (those included from the earlier pilot study, described in chapter 2), assisting in

determination of tumour area to be used for TMA construction, and collecting and cleaning all demographic, pathological, and clinical patient data. I also assisted pathologist Dr. Dan Fontaine in scoring the stained TMA slides, directed statistical analysis, and wrote the manuscript.

3.4.1 Authors and Affiliations

Angela Hyde[1], Tyler Wish[2], Dan Fontaine[3], Nikita Makretsov[3], Des Robb[3], Patrick Parfrey[2], Ban Younghusband[1]

1. Discipline of Genetics, Memorial University of Newfoundland, St. John's, Newfoundland, Canada
2. Clinical Epidemiology Unit, Memorial University of Newfoundland, St. John's, Newfoundland, Canada
3. Department of Pathology, Eastern Health, St. John's, Newfoundland, Canada

3.4.2 Summary

Background: Very few colorectal cancer biomarkers have been adopted for clinical use in predicting the prognosis of individual patients. Lamin A/C expression in tumour nuclei has some potential use as a prognostic biomarker.

Methods: We collected demographic, molecular, pathological, and clinical data from 391 patients with colorectal cancer, and developed a tissue microarray with their surgical specimens. We performed immunohistochemistry with antibodies against lamin A/C, and assessed its value as a prognostic marker.

Results: Expression of lamin A/C in >25% of tumour nuclei was significantly associated with advanced clinical stage at diagnosis (P=0.01), and with increased risk of disease-specific death (HR 1.59, P=0.02), and development of distant metastasis (HR 2.53, P=0.006). After multivariate adjustment, the independent risk of developing metastases was 2.5 fold greater for patients whose tumours expressed lamin A/C in >25% of cells compared to those whose tumours expressed lamin A/C in ≤25% of cells (HR 2.55, P=0.006). These results were driven by outcomes in patients with stage III tumours. Patients with stage III tumours and lamin A/C expression in >25% of tumour nuclei had a significant, independent, increased risk of disease-specific death (HR 4.14, P=0.01), and development of distant metastasis (HR 3.42, P=0.03). In stage III patients treated with chemotherapy this biomarker was an independent predictor of death from their disease (P=0.01).

Conclusions: Lamin A/C expression in tumour nuclei is an independent, negative prognostic indicator for CRC patients diagnosed with stage III tumours. It is also a strong, independent predictor of developing distant metastasis.

3.4.3 Introduction

Despite the extensive study of colorectal cancer (CRC), it remains one of the most common and deadly malignancies. CRC represents almost 10% of all new cancer diagnoses and 8% of deaths from cancer worldwide [1,794]. It has the second highest prevalence in the world, after breast cancer, with an estimated 2.8 million people living who have received a diagnosis of CRC in the past five years [2]. Although the general prognosis is good for CRC when compared to some other cancers, there is a paucity of adequate biomarkers to predict individual outcome and response to therapy.

Lamin A/C expression in tumour nuclei has been identified as a potential biomarker for prognosis in CRC [337]. Lamins A and C are A-type lamins, and alternative splice variants of the same *LMNA* gene product. They are found in the nucleus of differentiated somatic cells [714,741]. In the nuclear lamina they are involved in providing strength and support to the overlying nuclear envelope [708,709], as well as recruiting and anchoring various proteins and chromatin, to the inner nuclear membrane [710-712]. Within the nucleoplasm they have been identified in discrete foci in association with both DNA and other proteins [719-722]. They also make up a veil-like lamin network, which may provide a scaffold for DNA replication and transcription [725-727].

Through their roles in providing the structural framework of the nucleus and the organization of genetic material, A-type lamins are important for protecting cells from physical damage, and for maintaining proper chromatin organization and the function of key transcriptional elements. They have been dubbed "guardians of the soma" [736].

Defects in the A-type lamins are associated with numerous laminopathies, a diverse group of inherited diseases that involve premature aging, and degeneration of tissues of mesenchymal origin.

Lamins A and C have been considered potential biomarkers for cancer prognosis because of their involvement in cellular differentiation, proliferation, and motility (reviewed by Foster *et. al.* in 2010 [738]). Briefly, tumour cell de-differentiation is generally associated with poor prognosis, and lamin A/C is a marker of differentiated cells [714,739]. There is reduced expression of A-type lamins in proliferating epithelial cells [741], and overexpression of these lamins can inhibit cellular proliferation [742]. In 2008, Willis and colleagues [337] described the unexpected association between lamin A/C expression and increased cellular motility in cell culture. They suggested that this association was involved with an increased potential for tumour invasion. This theory was not tested on clinical data from CRC patients, but was later supported by further work [795].

Lamin A expression has also been described as a marker for stem cells in CRC [337]. The stem cell model of CRC development suggests that CRC initiates from mutations in colonic stem cells, and that these are the only cells capable of metastasizing [324]. These cells are thought to be crucial for sustaining tumour growth [324], and tumours rich in these cells may have a more aggressive phenotype [744]. Since lamin A appears to be a stem cell marker, it has been suggested that a tumour which expresses lamin A may have a more stem cell-like phenotype, and thereby exhibit a more aggressive behaviour [744]. Willis' 2008 paper [337] did indeed report a significantly worse prognosis for CRC

patients whose tumours expressed lamin A/C, as compared to those whose tumours did not. They found that patients whose tumours expressed lamin A/C were almost twice as likely to die from cancer-related causes (hazard ratio [HR] 1.85, 95% confidence interval [CI] 1.16-2.27). In contrast, a recent publication by Belt and colleagues [796] found that in their population of patients with stage II and stage III colon cancers, low levels of lamin A/C expression were associated with increased risk of disease recurrence and metastasis.

In this study, we demonstrate the utility of lamin A/C expression in tumour nuclei as a biomarker for CRC. Using tissue microarrays (TMAs), we examined the expression of lamin A/C in a cohort of 391 unselected CRCs, and correlated the findings with baseline clinical and pathological data, as well as with patient outcome.

3.4.4 Materials and Methods

3.4.4.1. Surgical specimens

A total of 391 primary colorectal tumours were included in this retrolective project, which were collected from two previous studies. In the first, we studied archival colorectal tumours diagnosed on the Avalon Peninsula of Newfoundland in 1997 and 1998, in a population-based cohort. Of the 296 tumours identified in this study, surgical specimens were available from 280 (a portion of this cohort was previously described [797]). To increase statistical power, 111 tumours were included from a second study (this second study has been previously described [469]). Briefly, all incident colorectal cancer

patients diagnosed younger than 75 years of age, from 1999 to 2003 in the province of Newfoundland and Labrador were contacted. Of the 750 patients who eventually consented to take part in that study, we included only the first 111 for whom tissue specimens were collected and available. These 111 patients were representative of the larger population, as there were no significant differences found in comparison of the two groups for age at diagnosis, sex, clinical stage at baseline, tumour grade, tumour location, and tumour microsatellite instability (MSI) status (data not shown).

3.4.4.2. Collection of Demographic, Pathological, Molecular and Clinical Data

Demographic, clinical, and outcome data were collected from the patients' hospital and Cancer Clinic charts, and the Newfoundland Tumour Registry. Tumour grade was collected from the original pathology reports. Vital status and cause of death was obtained from the Newfoundland and Labrador Centre for Health Information and the Statistics Canada Annual Mortality Files. Tumour MSI status was determined as described previously [469,797].

3.4.4.3. Construction of Tissue Microarray

Formalin-fixed, paraffin-embedded surgical specimens from CRC resection were used to create the TMAs. Two 0.6mm cores were retrieved from each of the 391

tumours, as well as from associated normal colonic mucosa from the 111 cases included from the later study, described above. These were assembled asymmetrically into four TMA blocks. A TMArrayer (Beecher Instruments, Inc, Sun Prairie, USA) was used for construction of the TMAs.

3.4.4.4. Immunohistochemistry

Sections of 4μm thickness were made from the TMA, and mounted on Fisherbrand Superfrost Plus microscope slides (Fisher Scientific Company, Ottawa, Canada). Automated antigen retrieval was carried out using CC1 standard HIER on Ventana Discovery XT (Ventana Medical Systems, Inc, Tucson, USA). Lamin A/C was detected using the JOL2 clone from Abcam, antibody ab40567. Incubation with the primary antibody was for one hour, at a 1:100 dilution. Incubation with Universal secondary antibody was for 32 minutes. Expression was detected with the DAB Map Detection Kit (Ventana, USA).

3.4.4.5. Interpretation of Immunohistochemical Data

The proportion of tumour cell nuclei expressing the lamin A/C antigen was estimated through manual scoring by two of the authors together (D.F. and A.H.), blinded to pathological, clinical, and outcome data. Only the frequency of nuclear staining was

considered. Scores were given on a five point scale, as follows: (0) 0% to 1% of tumour nuclei expressing the antigen; (1) >1% to 25%; (2) >25% to 50%; (3) >50% to 75%; and (4) >75% of tumour cells staining for lamin A/C.

Staining of fibroblast nuclei from tumour stroma was used as an internal positive control. Any tumours for which these stromal cell nuclei did not stain were considered a technical failure and not scored. In situations where one of the two cores from the same tumour was lost or uninterpretable, the score from the remaining core was used. In situations where the two cores from the same tumour were dissimilar, the score representing the core with the highest staining frequency was used.

3.4.4.6. Statistical Analysis

X-Tile software [798] was used to select a single dichotomous cut-point for lamin A/C expression. Kappa statistic was used to evaluate the level of agreement between tumour cores for lamin A/C expression. Univariate and multivariate binary logistic regression models were used to estimate unadjusted and adjusted odds ratios (OR), respectively. The risk of metastasis and disease-specific death was estimated from univariate and multivariate Cox regression models. All P values were two-sided and $P < 0.05$ was considered significant. All analyses were performed with PASW Statistics, version 18 (Chicago, IL, USA).

3.4.5 Results

3.4.5.1. Patient Demographics

Of the 391 patients included in the study, 19.4% (n=76) were diagnosed with stage I disease, 32.5% (n=127) with stage II, 26.1% (n=102) with stage III, 20.5% (n=80) with stage IV, and 1.5% (n=6) had unknown clinical stage at diagnosis. The average age at diagnosis was 64.5 years (range 21 to 91 years). Males represented 56% (n=221) of patients, 46% of tumours (n=180) were right sided, and 11.5% (n=45) were MSI-H.

3.4.5.2. Immunohistochemistry

In normal colonic mucosa the lamin A/C stain was strong in epithelial cell nuclei, stromal fibroblast nuclei, and underlying muscle. There was a tendency for a greater frequency of positive cells at the surface and upper crypts, as opposed to deeper into the crypts (Figure 3.2A). This was as expected, with A-type lamins known to be markers of differentiated cells [714,739]. Within the tumours, many tumour and stromal cells showed an accentuation of staining in the nuclear envelope (Figure 3.2B). In some cells this perinuclear rimming was present while typical nucleoplasmic staining was not apparent. These cells were classified as having a nuclear stain.

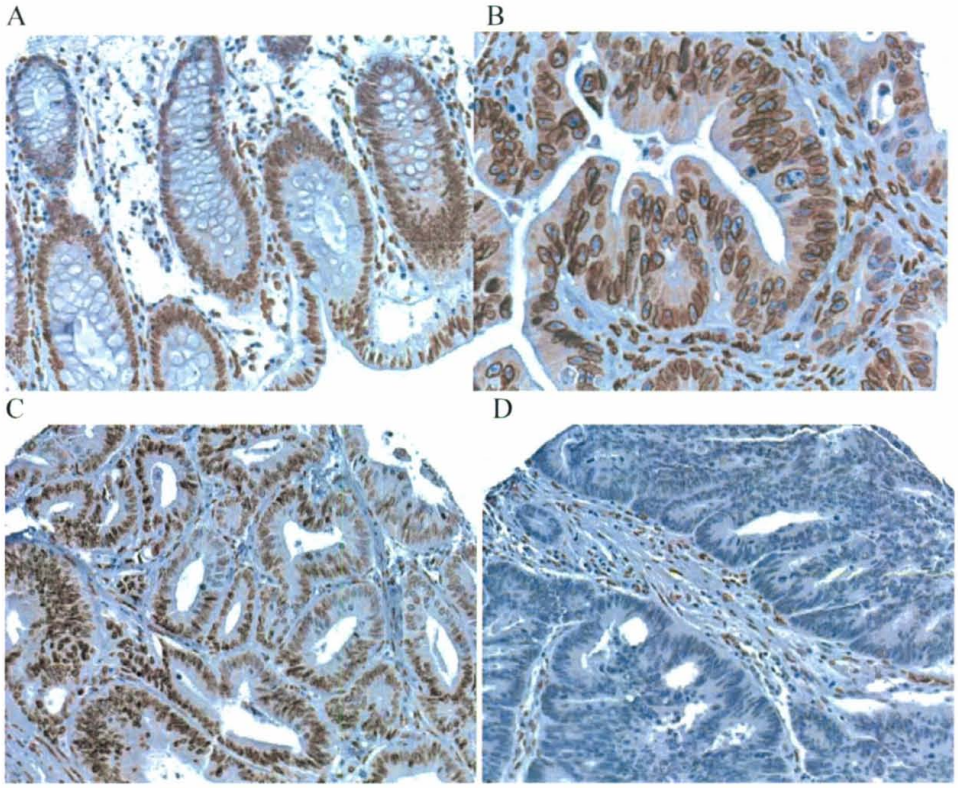


Figure 3.2. Representative images of lamin A/C immunohistochemistry
Images show: (A) increased expression at the luminal surface of normal colonic crypts (original magnification 120x), (B) pattern of perinuclear rimming in colorectal tumour (original magnification 200x), (C) >25% of tumour nuclei stained (original magnification 120x), and (D) ≤25% of tumour nuclei stained, with strong stain visible in nuclei of stromal fibroblasts (original magnification 120x).

The staining of 305 (78.0%) of 391 tumours was interpretable. For 21 cores, representing 12 tumours (3%), the stromal fibroblast nuclei did not stain, representing a technical failure. Nine of these tumours were excluded from further analysis, however, data for the other three tumours could still be interpreted, as only one of the two cores failed. Another 218 cores were either lost during processing or nonrepresentative of tumour, resulting in a complete lack of immunohistochemical (IHC) data for 77 tumours

(19.7%), and data for only one of the two cores for 64 tumours (16.4%). The 77 patients without IHC data had similar rates of disease-specific death as those for whom there was data ($P = 0.233$). When compared for sex, tumour grade, tumour location, clinical stage, lymphatic and perineural invasion, local recurrence and distant metastasis, the only factor that showed a significant difference was tumour grade ($P = 0.02$).

Of the 305 interpretable cases, more than 90% showed some nuclear staining. The distribution of scores is given in Table 3.1. Concordance between duplicate cores for nuclear staining was moderate (kappa value = 0.52, $P = 0.001$).

Table 3.1. Distribution of lamin A/C scores

Score *(% positive cells)*	**Nuclear Stain** *No. (%)*
0 (0%-1%)	30 (7.7)
1 (>1%-25%)	82 (21.0)
2 (>25%-50%)	55 (14.1)
3 (>50%-75%)	65 (16.6)
4 (>75%-100%)	73 (18.7)
X (failure)	86 (22.0)
Total	391 (100)

Figure 3.3A is a Kaplan-Meier graph showing disease-specific survival for all patients, stratified by lamin A/C expression score. While there was no clear, direct correlation of patient outcome with the percent of tumour cells positive for lamin A/C, there was an observed trend, with the lowest expression levels associated with the longest survival. There was no statistical difference between those with 0%-1% expression and those with >1%-25% expression. Evaluation with X-Tile software identified 25% as the

most significant cut-point to dichotomize staining frequency for predicting disease-specific survival. Figure 3.3B is a Kaplan-Meier graph showing disease-specific survival in all patients, using the 25% cut-point. Figures 3.2C and 3.2D show representative tumours with staining in >25% and ≤25% of nuclei, respectively.

High levels of lamin A/C expression, defined as expression in >25% of tumour cell nuclei, was found in 63% of tumours. Table 3.2 describes the demographic and clinicopathological data at diagnosis, according to lamin A/C expression. A significant association was identified between high nuclear expression and advanced clinical stage at diagnosis ($P = 0.01$). High levels of expression were identified in 48% of stage I cancers, 65% of stage II, 69% of stage III, and 70% of stage IV.

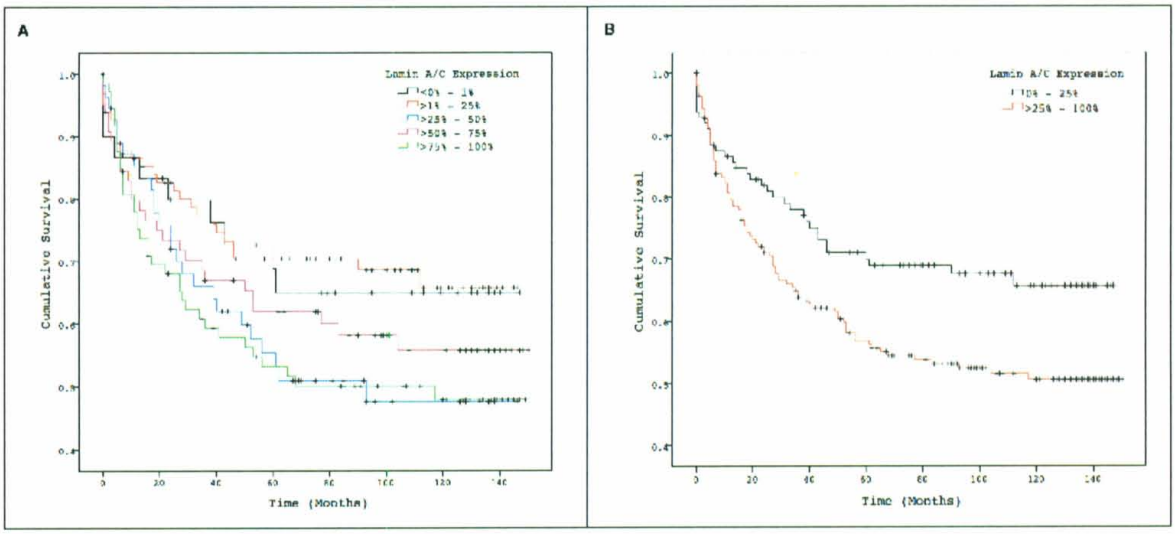


Figure 3.3. Kaplan Meier curves illustrating disease-specific survival for patients stratified by lamin A/C expression A) all lamin A/C scores (B) using cut-off of 25%.

Table 3.2. Description of patients at CRC diagnosis, according to frequency of lamin A/C expression within the tumour nuclei

	Nuclear Lamin A/C Expression		Unadjusted
	≤25%	>25%	
	No. (%)	*No. (%)*	*OR† (95% CI)*
N	112	193	
Age, mean (SD), yrs	66.1 (12.1)	64.7 (12.3)	0.99 (0.97 - 1.01)
P			0.33
Sex			
Male	59 (52.7)	113 (58.5)	1.00
Female	53 (47.3)	80 (41.5)	0.79 (0.49 - 1.26)
P			0.32
Tumour grade			
Well or moderate	102 (91.1)	169 (87.6)	1.00
Poor or undifferentiated	10 (8.9)	23 (12.4)	1.45 (0.67 - 3.15)
P			0.35
Clinical stage			
I	34 (30.4)	31 (16.1)	1.00
II	36 (32.1)	67 (34.7)	2.04 (1.08 - 3.85)
III	23 (20.5)	51 (26.4)	2.43 (1.22 - 4.86)
IV	19 (17.0)	44 (22.8)	2.54 (1.23 - 5.25)
P trend			0.01
Tumour location			
Proximal	50 (44.6)	94 (48.7)	1.00
Distal	62 (55.4)	99 (51.3)	0.85 (0.53 - 1.36)
P			0.49
MSI			
MSS	103 (92.0)	170 (88.1)	1.00
MSI-H	9 (8.0)	23 (11.9)	1.55 (0.69 - 3.48)
P			0.29

† Odds ratios (OR) and 95% confidence intervals (CI) determined from univariate binary logistic regression models.

3.4.5.3. Survival Analysis

Cox regression analysis showed that patients with >25% of tumour cells expressing lamin A/C had a significantly reduced 5 year disease-specific survival (unadjusted HR 1.59, $P = 0.02$) (Table 3.3A). This was no longer significant following adjustment for age, sex, clinical stage, tumour grade, tumour location, and microsatellite instability status, although the trend was similar (adjusted HR 1.46, $P = 0.07$). Similarly, these patients had a worse overall survival (unadjusted HR 1.29, $P = 0.11$); however, the result did not reach statistical significance (Table 3.3A).

Table 3.3. Risk of (A) death and (B) metastasis for colorectal cancer patients stratified by clinical stage and according to lamin A/C expression

A	Patients	Disease-Specific Death			Overall Death		
		Events	Unadjusted	Adjusted‡	Events	Unadjusted	Adjusted‡
	No. (%)	*No.*	*HR† (95% CI)*	*HR (95% CI)*	*No.*	*HR† (95% CI)*	*HR (95% CI)*
Stage I							
≤25%	34 (52.3)	3	1.00		13	1.00	1.00
>25%	31 (47.7)	3	0.93 (0.19 - 4.61)	na*	7	0.52 (0.21 - 1.30)	0.80 (0.28 - 2.29)
P			0.93			0.16	0.68
Stage II							
≤25%	36 (35.0)	9	1.00	1.00	18	1.00	1.00
>25%	67 (65.0)	17	1.02 (0.45 - 2.28)	1.17 (0.52 - 2.65)	31	0.93 (0.52 - 1.67)	1.15 (0.64 - 2.09)
P			0.97	0.71		0.81	0.64
Stage III							
≤25%	23 (31.1)	4	1.00	1.00	7	1.00	1.00
>25%	51 (68.9)	28	3.90 (1.36 - 11.12)	4.14 (1.41 - 12.15)	36	2.91 (1.29 - 6.54)	3.06 (1.32 - 7.13)
P			0.01	0.01		0.01	0.009
Stage IV							
≤25%	19 (30.2)	19	1.00	1.00	19	1.00	1.00
>25%	44 (69.8)	41	1.02 (0.59 - 1.78)	1.33 (0.71 - 2.46)	42	1.05 (0.60 - 1.81)	1.35 (0.73 - 2.51)
P			0.94	0.37		0.88	0.34
Stages I - IV‡‡							
≤25%	112 (36.7)	35	1.00	1.00	57	1.00	1.00
>25%	193 (63.3)	89	1.59 (1.08 - 2.36)	1.46 (0.97 - 2.18)	116	1.29 (0.94 - 1.78)	1.24 (0.90 - 1.72)
P			0.02	0.07		0.11	0.19

B	Patients	Metastasis		
		Events	Unadjusted	Adjusted‡
	No. (%)	*No.*	*HR† (95% CI)*	*HR (95% CI)*
Stage I				
≤25%	34 (52.3)	2	1.00	
>25%	31 (47.7)	3	1.40 (0.23 - 8.42)	na*
P			0.71	
Stage II				
≤25%	36 (35.0)	5	1.00	1.00
>25%	67 (65.0)	17	1.81 (0.67 - 4.92)	2.02 (0.74 - 5.53)
P			0.24	0.17
Stage III				
≤25%	23 (31.1)	4	1.00	1.00
>25%	51 (68.9)	22	3.29 (1.13 - 9.58)	3.60 (1.22 - 10.67)
P			0.03	0.02
Stages I - III‡‡				
≤25%	93 (38.4)	11	1.00	1.00
>25%	149 (61.3)	42	2.60 (1.34 - 5.05)	2.65 (1.35 - 5.18)
P			0.005	0.004

† Hazard ratio (HR) and 95% confidence interval (CI) determined from Cox regression models.

‡ Adjusted for age, sex, clinical stage, tumour grade, tumour location, and microsatellite-instability status.

‡‡ Adjusted model also includes clinical stage as a covariate.

* Model does not converge

3.4.5.4. Subset Analysis of Survival

When stratified by clinical stage at diagnosis it is apparent that, for both disease-specific and overall survival, the adverse trend was driven by outcomes in the stage III patients. Those with stage III disease had a significantly worse survival with low lamin A/C expression levels, but this was not seen in stage I, II, or IV patients (Table 3.3A and Figure 3.4). Figure 3.4 illustrates the observed survival differences for patients whose tumours expressed high versus low levels of lamin A/C. Mantel-Cox Log Rank tests identified a significant disease-specific survival advantage for patients with stage III tumours expressing lamin A/C in ≤25% of tumour cells, as compared to those expressing it in >25% ($P = 0.005$). This was statistically significant for both disease-specific and overall survival, even when adjusted for age, sex, tumour grade, tumour location, and microsatellite instability status. Stage III patients with higher levels of lamin A/C expression were approximately four times more likely to die of their disease than stage III patients with low levels of expression (unadjusted HR 3.90, $P = 0.01$; adjusted HR 4.14, $P = 0.01$). These same patients were about three times more likely to die of any cause during follow-up (unadjusted HR 2.91, $P = 0.01$, adjusted HR 3.06, $P = 0.009$).

When the cohort was further assessed to compare the MSI and MSS tumours, the above trends remained apparent for the MSS cohort, however, there were too few MSI tumours to make any conclusions for this group. In unadjusted analysis of stage III, MSS tumours, patients with high levels of tumour lamin A/C expression were 3.64 times more

likely to die of their disease than those with low levels of expression ($P = 0.016$) (data not shown).

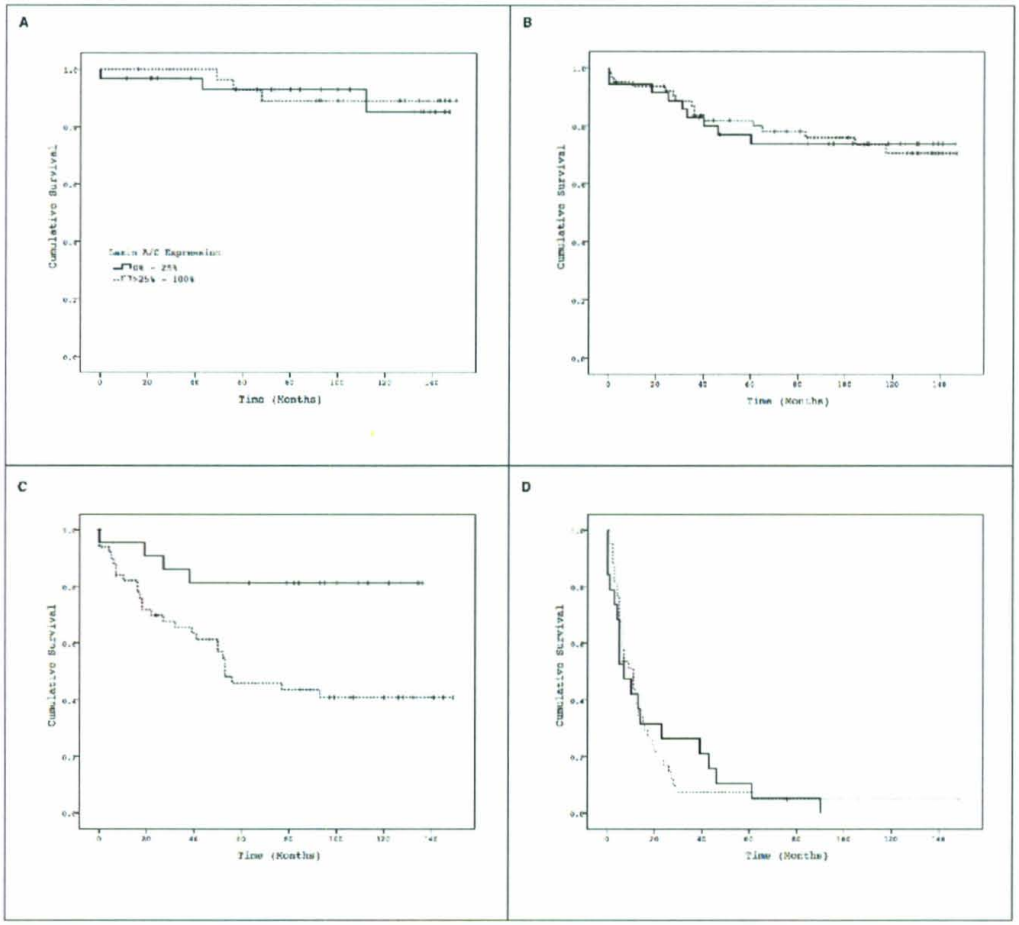


Figure 3.4. Kaplan Meier curves illustrating disease-specific survival for patients based on lamin A/C expression and stratified by clinical stage at diagnosis
(A) patients with stage I disease, (B) stage II, (C) stage III, and (D) stage IV.

3.4.5.5. Risk of Distant Metastasis and Locoregional Recurrence

We tested the whole cohort for a relationship between lamin A/C expression and the two main causes of disease-specific death: distant metastasis and locoregional recurrence. Tumour expression levels of lamin A/C did not affect risk of locoregional disease recurrence (data not shown). However, there was a significantly increased risk of distant metastasis with high levels of lamin A/C expression (unadjusted HR = 2.53, P = 0.006). This risk remained significant after multivariate adjustment for patient age, sex, tumour grade, tumour location, clinical stage at diagnosis, and microsatellite instability status (adjusted HR 2.55, P = 0.006). This equates to a 2.5 fold increase in risk of metastasis in patients whose tumours have a >25% frequency of nuclear expression of lamin A/C, when compared to those with lower levels of expression.

3.4.5.6. Subset Analysis of Risk of Distant Metastasis

As shown in Table 3.3B, the increased risk of distant metastasis, described above, is due to outcomes in patients with stage III disease. Patients with stage III CRC and high expression of lamin A/C are almost 3.5 times more likely to develop metastasis than those with low expression. There are no significant increased risks for the other clinical stages.

We questioned if the risks observed for patients with stage III disease, for both development of distant metastasis and disease-specific death, were due to increased cell

motility, as suggested by Willis *et. al.* [337], or to an altered response to chemotherapeutic agents. If it were due to increased cell motility, we would have expected the increased risk of metastasis and disease-specific death in patients with tumours which had high levels of lamin A/C, diagnosed at all clinical stages. This was not observed.

3.4.5.7. Subset Analysis of Adjuvant Treatment and Survival

Adjuvant chemotherapy is typically only given to patients diagnosed with stage III tumours, and to a select few with stage II tumours. We postulated that the observed risks in patients with stage III tumours expressing higher levels of lamin A/C could have been due to decreased response to adjuvant chemotherapy. Of the 74 stage III tumours, 46 received standard 5-fluorouracil and folinic acid based adjuvant chemotherapy, and three received an alternative chemotherapy. Of the 25 patients who did not receive adjuvant chemotherapy, seven did not survive long enough after surgical resection to be offered further therapy, four were not seen by the Cancer Clinic, eight were not offered chemotherapy for clinical reasons, and six patients refused chemotherapy. Of the 49 who received chemotherapy, 17 had low levels (≤25%) of lamin A/C expression, and 32 had higher levels (>25%). Despite the small numbers, in multivariate Cox regression analysis, patients with stage III CRC who received chemotherapy were significantly more likely to die of their disease if they had higher levels of lamin A/C tumour expression ($P = 0.01$) (Table 3.4A).

Table 3.4B shows the details of a multivariate model which includes lamin A/C expression, treatment, age at diagnosis, patient sex, tumour grade, tumour location, and microsatellite instability status. In the stage III patients in this cohort, lamin A/C expression level was the only significant predictor for both the development of distant metastasis and disease-specific death. While we would expect the treatment to have an effect on these outcomes, patients who did not receive chemotherapy were significantly older, had more severe comorbidities, and died sooner; they were generally censored from analysis before they had the chance to develop metastasis.

Table 3.4. (A) Risk of metastasis and disease-specific death for clinical stage III patients stratified by chemotherapy management and according to lamin A/C expression; (B) details of the multivariate models

A	Stage III Patients	Metastasis			Disease-Specific Death		
		Events	Unadjusted	Adjusted‡	Events	Unadjusted	Adjusted‡
	No. (%)	*No.*	*HR† (95% CI)*	*HR (95% CI)*	*No.*	*HR (95% CI)*	*HR (95% CI)*
Chemotherapy							
≤25%	17 (34.7)	3	1.00	1.00	2	1.00	1.00
>25%	32 (65.3)	16	3.75 (1.09 - 12.93)	3.39 (0.97 - 11.94)	17	5.75 (1.33 - 24.94)	7.01 (1.54 - 31.93)
P			0.04	0.06		0.02	0.01
No chemotherapy							
≤25%	6 (24.0)	1	1.00	1.00	2	1.00	1.00
>25%	19 (76.0)	5	1.73 (0.20 - 14.89)	na‡‡	11	1.74 (0.39 - 7.85)	1.40 (0.25 - 7.71)
P			0.62			0.47	0.70

† Hazard ratios (HR) and 95% confidence intervals (CI) determined from Cox regression models.
‡ Adjusted for age, sex, tumour grade, tumour location, and microsatellite-instability status.
‡‡ Model does not converge

B	Multivariate Model			
	Metastasis		Disease-Specific death	
	HR (95% CI)	*P*	*HR (95% CI)*	*P*
Lamin A/C Expression				
≤25%	1	0.03	1	0.01
>25%	3.36 (1.13 - 10.0)		3.86 (1.31 - 11.42)	
Treatment				
Chemo	1	0.63	1	0.2
No Chemo	1.31 (0.44 - 3.93)		1.80 (0.74 - 4.38)	
Age	1.01 (0.96 - 1.06)	0.66	1.04 (1.00 - 1.09)	0.07
Sex				
Male	1	0.93	1	0.36
Female	0.97 (0.41 - 2.27)		0.70 (0.33 - 1.51)	
Grade				
Well or Moderate	1	0.71	1	0.75
Poor or Undifferentiated	0.79 (0.23 - 2.73)		1.17 (0.46 - 2.97)	
Tumour Site				
Proximal	1	0.21	1	0.09
Distal	1.78 (0.72 - 4.39)		2.05 (0.89 - 4.73)	
MSI				
MSS	1	0.98	1	0.25
MSI			0.40 (0.08 - 1.93)	

3.4.6 Discussion

In this study we have shown that lamin A/C expression in a high number of tumour cell nuclei is a negative prognostic marker for colorectal cancer. While a previous prognosis study of lamin A/C [337] assessed expression as absolute, positive or negative, we determined that 25% of nuclei stained was the most appropriate threshold for dichotomizing staining frequency scores for predicting disease-specific survival.

A weakness of this study is that only 78% of tumours were interpretable. However, even with the reduced cohort size statistically significant findings were in accordance with previously published results from another group. We have validated the previous report by Willis and colleagues [337], who found that lamin A/C expression in tumour nuclei was a prognostic biomarker in colorectal cancer. In their study, Willis *et. al.* [337] reported that patients whose tumours expressed lamin A/C were more likely to die of their disease than patients whose tumours did not express lamin A/C, adjusted for age at diagnosis and sex. Similar findings were observed in our population using a cut-off of 25% of tumour cells expressing lamin A/C, as assessed manually on dual 0.6 mm tissue cores in a tissue microarray. We found that patients whose tumours expressed lamin A/C in at least 25% of cells were significantly more likely to die of causes related to their cancer. When adjusted for patient age at diagnosis, sex, clinical stage, tumour grade, tumour location, and MSI status the trend was the same, however, it no longer reached statistical significance.

A novel finding of this study was the association of lamin A/C expression and distant metastases. The work by Willis and colleagues [337] and Foster and colleagues [795] determined that lamin A expression was associated with increased cellular motility, by performing wound closure scratch tests on cell cultures. While they projected that this would lead to increased invasive potential of tumours, they didn't test this theory with clinical data. We found a clear clinical association between high lamin A/C expression in primary colorectal tumours and increased rates of developing distant metastases. After adjusting for other factors including patient age at diagnosis, sex, tumour grade, tumour location, clinical stage at diagnosis, and MSI status, lamin A/C expression was a strong independent indicator of risk of metastases. Tumours with more than 25% expression of lamin A/C were two and a half times more likely to develop distant metastases than those with lower levels of lamin A/C expression.

Following stratification of our patient population by clinical stage at diagnosis, and found that the impact of lamin A/C expression was driven by outcomes in patients with stage III tumours. While there was no significant difference in disease-specific survival based on lamin A/C expression in tumours of patients with cancers diagnosed at clinical stages I, II, or IV, there was a striking difference for those diagnosed at clinical stage III. Lamin A/C expression was a strong, independent indicator for metastasis and disease-specific survival in stage III patients. While lamin A/C has prognostic utility for the stage III cohort as a whole and for the MSS tumours within it, we were unable to assess its utility in the MSI tumours, due to small sample size.

Since stage III CRC is treated with adjuvant chemotherapy, we stratified the patients by the treatment they received, and assessed the utility of lamin A/C expression

in predicting outcomes. In multivariate models of patients with stage III CRCs, lamin A/C expression in >25% of tumour nuclei was the only statistically significant predictor of metastasis and of disease-specific death.

We conclude that lamin A/C is an independent, negative prognostic indicator for stage III colorectal cancer, and that it is a significant, independent biomarker for the risk of metastasis. Future studies on the predictive and prognostic values of lamin A/C in colorectal cancer are suggested.

3.4.7 Acknowledgements

We would like to thank the Centre for Translational and Applied Genomics at the Provincial Health Services Authority Laboratory of the British Columbia Cancer Agency (Alicja Parker and Julie Lorette) for TMA staining; Susan Ryan and Elizabeth Hunt from the Cancer Care Program, Eastern Health Newfoundland for data collection; Laura Edwards for data entry; the Newfoundland Familial Colorectal Cancer Registry (Project Coordinator Dr. Elizabeth Dicks) for access to the surgical specimens obtained from 1999-2003; and Dr. Roger Green for assistance in editing the manuscript.

3.5 Manuscript – "Prognostically Significant Subgroups of Colorectal Cancer Identified Through Tissue Microarray Biomarker Profiles"

I wrote a second TMA-related manuscript, which will also be submitted for publication. My roles in this work were the same as for the previous manuscript. For this work I included all 12 antibodies used in the TMA study, and assessed the prognostic utility of each protein individually, as well as in biomarker profiles identified by unsupervised hierarchical clustering analysis. A number of these markers were found to have prognostic utility in our population. Clustering analysis identified three cluster groups, which also have prognostic utility.

3.5.1 Authors and Affiliations

Angela Hyde[1], Tyler Wish[2], Dan Fontaine[3], Nikita Makretsov[3], Nina Curcin[3], Des Robb[3], Patrick Parfrey[2], Ban Younghusband[1]

1. Discipline of Genetics, Memorial University of Newfoundland, St. John's, Newfoundland, Canada
2. Clinical Epidemiology Unit, Memorial University of Newfoundland, St. John's, Newfoundland, Canada
3. Department of Pathology, Eastern Health, St. John's, Newfoundland, Canada

3.5.2 Summary

Background: Colorectal cancer is a common, heterogeneous disease, with multiple pathways of development, and variability in features including microsatellite instability status, tumour histology, tumour cytology, and level of host immune response. We sought to determine whether unsupervised hierarchical clustering of multiple immunomarkers could identify prognostically significant groups in colorectal cancer.

Methods: We used immunohistochemistry to stain tissue microarrays containing samples from 391 unselected colorectal tumours with antibodies against 12 different proteins that have a biological relevance to colorectal cancer. These included markers of cellular proliferation, cell cycle regulators, structural proteins, and others. Unsupervised hierarchical clustering analysis of immunostaining scores was used to identify cluster groups.

Results: Three distinct clusters groups were identified, which are associated with different clinical outcomes. The cluster group characterized by overall lower protein expression levels when compared to the other groups, was associated with a statistically significant increased risk of disease-specific death, even when adjusted for age and clinical stage at diagnosis, tumour grade, and MSI status.

Conclusions: We have identified subgroups of colorectal cancer using unsupervised hierarchical clustering analysis of tissue microarray immunostaining data. These subgroups differ for patient prognosis, with lower overall expression of biologically relevant immunostains being associated with adverse survival.

3.5.3. Introduction

Colorectal cancer (CRC) is a common and often deadly form of cancer. It is estimated that one in 14 Canadians will be diagnosed during their lifetime, with the highest incidence rates in the province of Newfoundland and Labrador [21]. CRC comprises approximately 12% of all new cancer diagnoses in men and women in Canada, and 12% of all cancer deaths [21].

Colorectal cancer is a heterogeneous disease. Patients can develop CRC in the context of Mendelian hereditary syndromes, in familial clusters for which no known genetic cause can usually be identified, and most commonly as sporadic occurrences, with no known family history. These tumours can result from different pathways of carcinogenesis, and can involve large-scale chromosomal instability or finer scale microsatellite instability. Histology and cytology can vary greatly from tumour to tumour, as can the level of host immune response. Patient prognosis and treatment are based largely on the Tumour Nodes Metastasis (TNM) classification system [37], which does not consider many of these differences. Not surprisingly, there can be differences in

tumour aggression, response to treatment, and likelihood of recurrence and metastasis, even among tumours categorized into the same TNM stage.

Considering the high population burden of CRC, it is important to have markers of survival and of drug response to guide individual patient prognosis and therapy. Considering the heterogeneity of the disease, it is unlikely that one sole biomarker will be the answer. Instead, looking at the expression profiles of multiple potential biomarkers representing many pathways that are biologically relevant to CRC may be more fruitful, and has been demonstrated using both gene expression profiles [799-801] and protein expression profiles [597,642,802,803]. An important first step in developing clinically relevant prognostic and predictive biomarker profiles is identifying subgroups of tumours that share multimarker profiles. In this study we constructed tissue microarrays (TMAs) from 391 unselected colorectal tumours from the province of Newfoundland and Labrador. We used immunohistochemistry to stain TMA sections for 12 biomarkers with biological relevance to CRC, then performed unsupervised hierarchical clustering analysis on the immunostaining data, and assessed the resultant cluster groups for associations with patient outcome. Our objective was to see if clustering analysis could identify subgroups of CRC that were associated with prognosis.

3.5.4. Materials and Methods

3.5.4.1. Construction of Tissue Microarrays

Dual 0.6mm cores from the formalin-fixed, paraffin-embedded surgical specimens of 391 colorectal carcinomas were assembled into three TMA blocks using a TMArrayer (Beecher Instruments, Inc., Sun Prairie, USA). The tumours were collected from two studies, both of which have been previously described [469,797]. Briefly, in the first study, 296 cases were identified in a population-based study of all incident CRCs diagnosed on the Avalon Peninsula of Newfoundland in 1997 and 1998. Of these, surgical specimens from 280 tumours were available, and used in the construction of the TMAs. In the second study, all incident CRCs diagnosed in patients younger than age 75 years in the province of Newfoundland and Labrador from 1999 to 2003 were identified, and patients or proxies were contacted. Of the 750 patients who eventually consented to the study, we included only the first 111 for whom surgical specimens were collected and available. Normal colonic epithelium from the same patients was used to create a fourth TMA block. This was used to assess the normal staining pattern of the antibodies, and also served as a control. These 111 patients were representative of the larger population, as there were no significant differences found upon comparison of the two groups for age at diagnosis, sex, clinical stage at diagnosis, tumour grade, tumour location, and microsatellite instability status (data not shown).

TMA blocks were assembled asymmetrically into quadrants of cores to aid orientation for ease of scoring. Cores of normal kidney were also included in each TMA block, to aid orientation and to act as controls.

3.5.4.2. Collection of Demographic, Pathological, Molecular and Clinical Data

Demographic and clinical data, including outcomes, were collected from the patients' hospital and Cancer Clinic charts, and the Newfoundland Tumour Registry. Pathology data were collected from the original pathology reports. Vital status and cause of death were obtained from the Newfoundland and Labrador Centre for Health Information and the Statistics Canada Annual Mortality Files. Tumour MSI status was determined as described previously [469,797].

3.5.4.3. Immunohistochemistry

Antibodies were chosen empirically, based on biological relevance to colorectal cancer and recent publications indicating an association with prognosis. Also considered was the availability of the antibodies, and the suitability for use on formalin-fixed, paraffin-embedded archival tissue specimens. A total of 12 antibodies were included in the study: markers of cell proliferation Ki67 and PCNA; tumour suppressor gene and cell cycle regulator p53; other cell cycle-related proteins p16, p21, p27, cyclin D1; cell

adhesion proteins β-catenin, and E-cadherin; as well as structural proteins lamin A/C; DNA helicase MCM7; and cell signalling molecule EGFR. Immunostaining protocols were first optimized on whole sections. Staining protocols are summarized in Table 3.5.

Table 3.5. Antibodies and staining protocols for immunohistochemistry

Marker	Source	Dilution	Pretreatment
β-catenin	BD Biosciences	1:400	CC1 mild HIER on Ventana Discovery
Cyclin D1	Ventana	Pre-dilute	CC1 standard HIER on Ventana Discovery
E-cadherin	Dako	1:25	CC1 standard HIER on Ventana Discovery
EGFR	Ventana	Pre-dilute	Protease 2
Ki67	Ventana	Pre-dilute	CC1 mild HIER on Ventana Benchmark
Lamin A/C	Abcam	1:100	CC1 standard HIER on Ventana Discovery
MCM7	Abcam	1:25	CC1 standard HIER on Ventana Discovery
p16	Mtm Laboratories	Pre-dilute*	CC1† standard on Ventana Benchmark XT
p21	Dako	1:10	CC1 standard HIER‡ on Ventana Discovery
p27	Ventana	Pre-dilute	CC1 mild HIER on Ventana Discovery
p53	Ventana	Pre-dilute	CC1 mild HIER on Ventana Benchmark
PCNA	Ventana	Pre-dilute	none

Note: Manufacturers locations are as follows: Abcam (Cambridge, MA, USA), BD Biosciences (Franklin Lakes, NJ, USA), Dako (Glostrup, Denmark), Mtm Laboratories (Heidelberg, Germany), Ventana (Tucson, AZ, USA)
*Pre-dilute indicates that the antibody was obtained ready to use from the manufacturer, it was pre-diluted for the automated system.
†CC1 refers to a commercial antigen retrieval solution, and is a TRIS EDTA buffer with a pH of 8.0 to 8.5.
‡HIER refers to heat induced epitope retrieval, and is performed on the automated Ventana machines.

3.5.4.4. Interpretation of Immunohistochemical Data

All slides were scored without knowledge of clinical features or patient outcome. Ki67 and PCNA were scored quantitatively (Ki67 by pathologist N.M., and PCNA by

pathologist N.C.), by manually counting the number of positive and negative cells, and calculating the percent positivity. Immunoreactivity of the remaining markers was scored semiquantitatively by pathologist D.F. and student A.H. together, by manually estimating the percentage of positive tumour cells, given pre-described cut-offs. All protein expression scores were based on frequency of positive staining, with the exception of p53, for which we scored the intensiy of the stain. Cut-offs for immunostaining in previous studies have been widely variable. We selected cut-offs to allow evaluation of markers with a four point scoring system, with consideration of those used in previous studies. Scoring methods are summarized in Table 3.6, along with log rank univariate *P*-values for the association of each immunomarker with disease-specific survival.

Raw data scores were entered into an Excel spreadsheet (Microsoft, Redmond, WA, USA), and converted into a format suitable for cluster analysis using TMA-Deconvoluter version 1.10 [539]. Scores from duplicate cores were combined. In the case of discrepancy between cores, the higher value was accepted. Only tumours with immunostaining data for at least 80% of markers were included in clustering analysis.

Table 3.6. Description of scoring methods for immunostains, and association with disease-specific survival

Marker	Staining	Description of Scores				Disease-Specific Survival
		Score 0	Score 1	Score 2	Score 3	*P** (direction of association†)
β-catenin	Membranous	<30%	30% to 50%	N/A	≥50%	0.05 (Positive)
Cyclin D1	Nuclear	<5%	5% to 10%	11% to 50%	>50%	0.25
E-cadherin	Membranous	<1%	1% to 50%	N/A	≥50%	0.01 (Positive)
EGFR	Membranous and/or cytoplasmic	<1%	N/A	N/A	≥1%	0.74
Ki67	Nuclear	≤10%	11% to 30%	N/A	>30%	0.007 (Negative)
Lamin A/C	Nuclear	<25%	25% to 50%	51% to 75%	>75%	0.09 (Negative)
MCM7	Nuclear	<1%	1% to 10%	11% to 30%	>30%	0.003 (Positive)
p16	Nuclear	<1%	1% to 10%	11% to 50%	>50%	0.04 (Positive)
p21	Nuclear	≤5%	6% to 10%	11% to 50%	>50%	0.10 (Positive)
p27	Nuclear	<25%	25% to 50%	N/A	>50%	0.03 (Positive)
p53	Nuclear	Negative	Weak	N/A	Strong	0.34
PCNA	Nuclear	Quartile 1	Quartile 2	Quartile 3	Quartile 4	0.001 (Positive)

* *P* value determined by log-rank test.

† Association between score and survival determined by Kaplan-Meier survival analysis. Positive association means that a greater score was associated with greater survival.

3.5.4.5. Statistical Analysis

Each marker was assessed for its relationship with disease-specific survival using univariate Kaplan-Meier analysis and log rank tests. Unsupervised hierarchical clustering analysis using the average linkage algorithm and Euclidean Distance similarity metric was performed to classify the tumours into groups based on similarity of expression of markers (protein expression profiles). The resulting cluster groups were assessed for their relationship to various clinicopathological features using independent-samples t-tests and Chi square analysis. The association of the cluster groups with disease-specific survival was assessed using Kaplan-Meier survival curves, and both univariate and multivariate Cox regression models. Cox regression models were also used to assess the cluster groups for their associated risk of developing distant metastasis. All *P* values were two sided, and $P < 0.05$ was considered significant.

Unsupervised hierarchical clustering analysis was performed using Cluster version 3.0 [804]. Cluster groups were visualized using Java Treeview version 1.1.5r2 [805]. All other statistical analysis was performed using PASW Statistics for Mac, version 18.0 (SPSS Inc., Chicago, Il, USA). The approach taken to clustering analysis in this manuscript has been utilized for CRC and other cancer types in previous studies [642,802,803].

3.5.5. Results

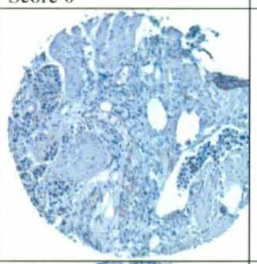
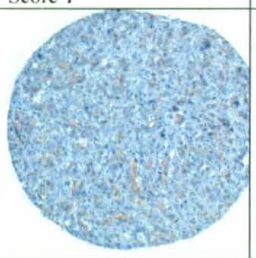
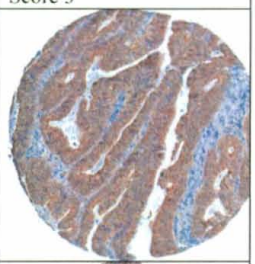
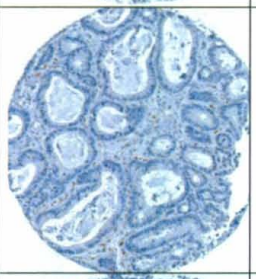
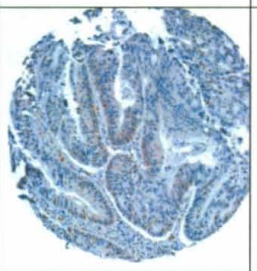
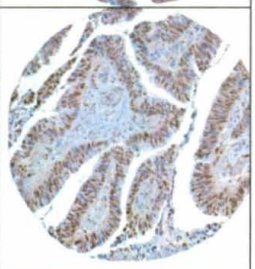
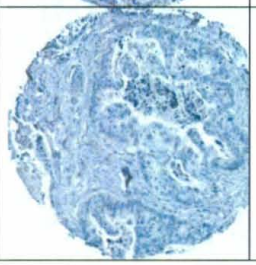
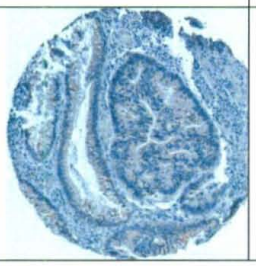
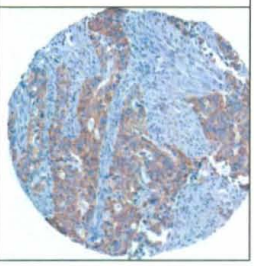
3.5.5.1. Patient Demographics

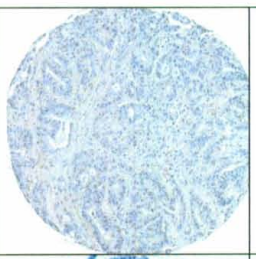
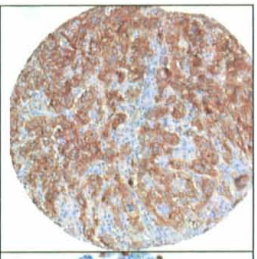
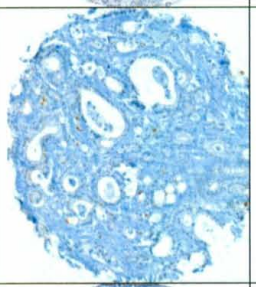
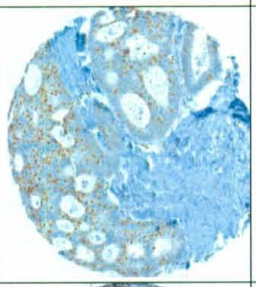
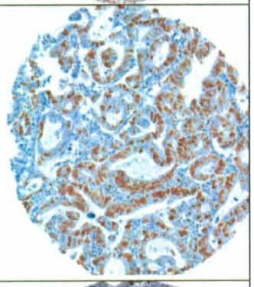
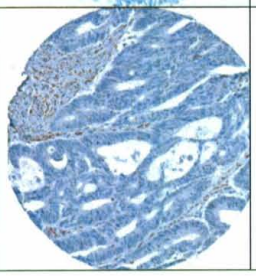
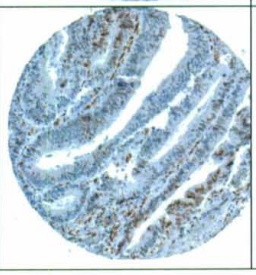
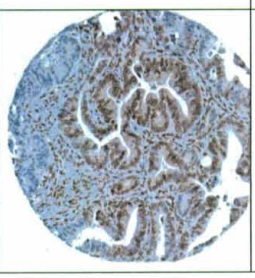
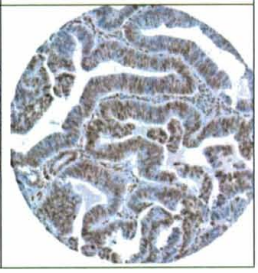
The mean age at diagnosis was 64.5 years (range 21 to 91 years, standard deviation 12.5 years). At diagnosis, 76 tumours (19.4%) were stage I, 127 (32.5%) were stage II, 102 (26.1%) were stage III, and 80 (20.5%) were stage IV. For 6 tumours (1.5%) stage could not be determined. The majority of patients were male (56.5%), most tumours were left-sided (54.0%), and 11.5% were microsatellite instable (MSI-H).

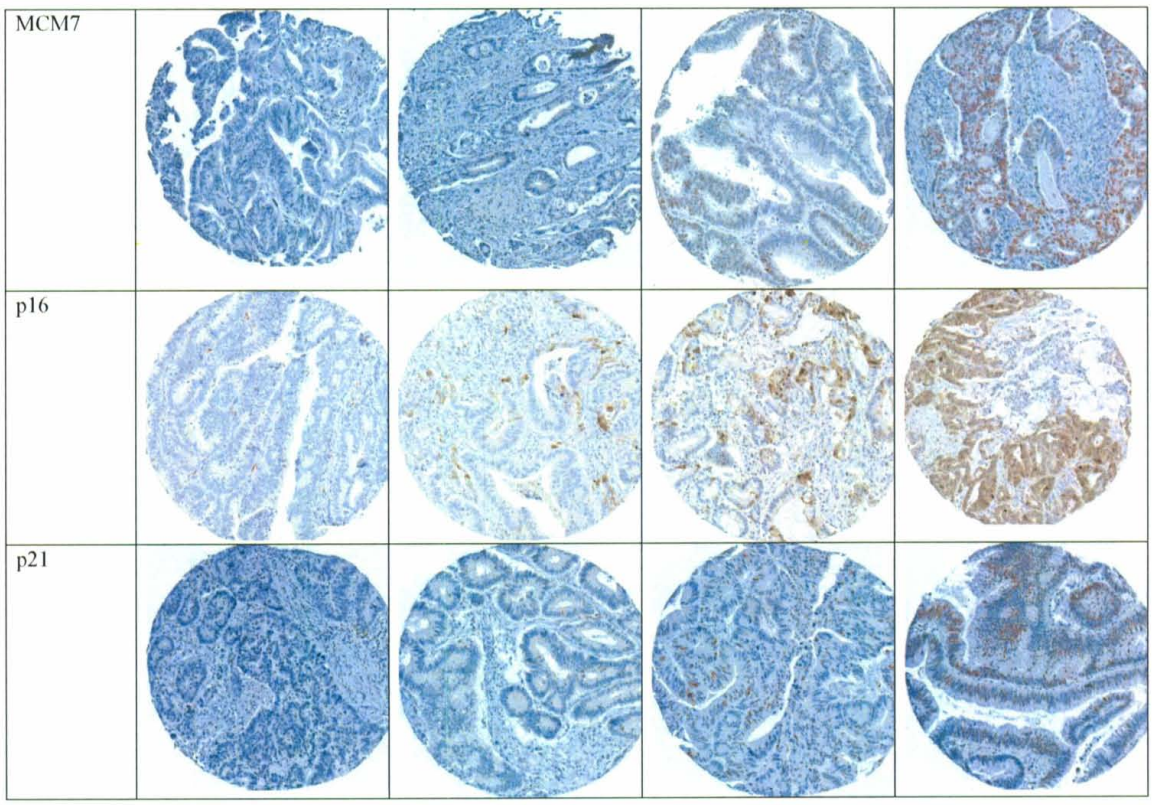
3.5.5.2. Immunostaining

Examples of representative immunostaining results for each of the possible scores for each biomarker are provided in Figure 3.5.

There was a significant proportion of missing data points for each biomarker (5.4% to 22.8%). Common causes of missing data included technical failures such as: tissue cores lost during processing, cores unrepresentative of tumour, and lack of control staining in stromal fibroblast nuclei. In total, 300 of the 391 tumours (76.7%) were considered informative, as they had interpretable data for at least 80% of markers, and were included in further analysis. Comparisons of the informative and non-informative tumours are presented in Table 3.7. These groups differ significantly only for patient age at diagnosis, as those with non-informative tumours were younger, on average, than those with informative tumours.

	Score 0	Score 1	Score 2	Score 3
β-catenin			N/A	
Cyclin D1				
E-cadherin			N/A	

EGFR		N/A	N/A	
Ki67			N/A	
Lamin A/C				


MCM7
p16
p21

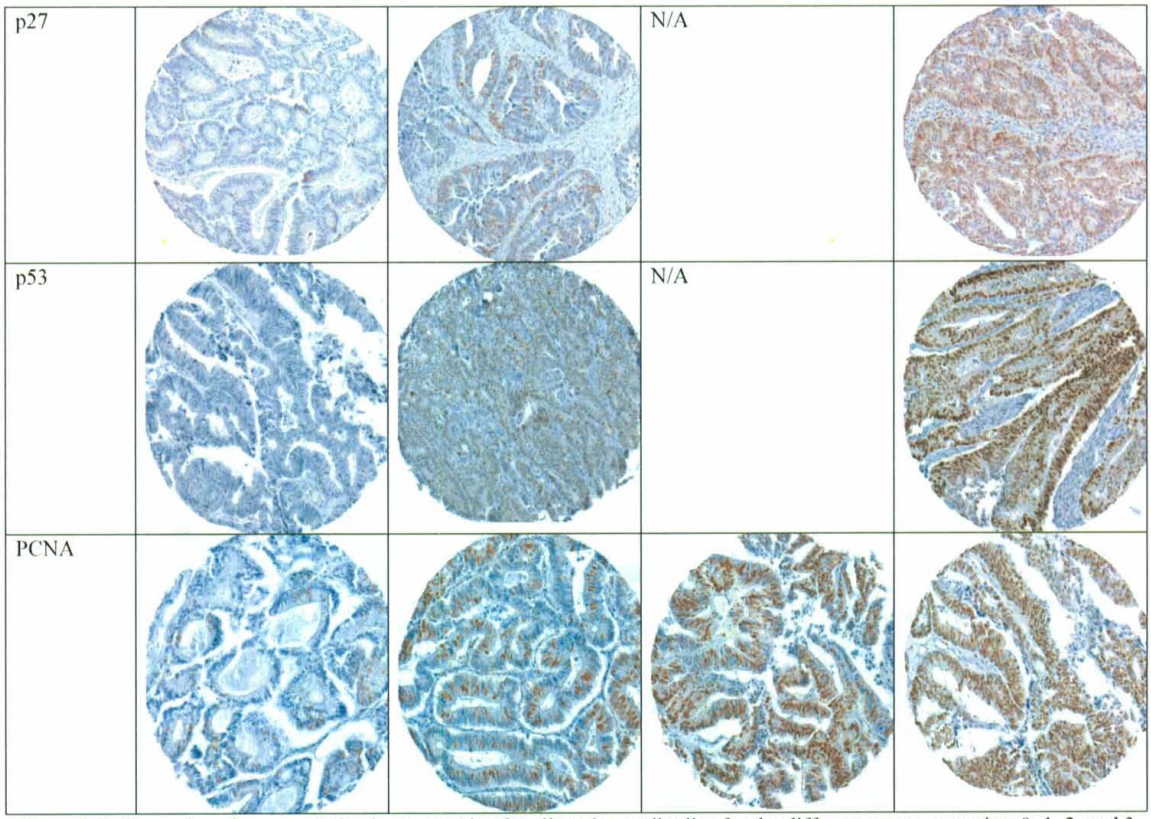


Figure 3.5. Examples of representative immunostains for all twelve antibodies for the different score categories: 0, 1, 2, and 3, as defined in Table 3.6.

Table 3.7. Description and comparison of informative and non-informative patients

Clinicopathological Features	Non-Informative* Tumours n = 91	Informative Tumours n = 300	*P*†
	No. (%)	No. (%)	
Age, mean (SD), years	59.2 (12.0)	66.1 (12.3)	<0.001
Sex			0.73
Female	41 (45.1)	129 (43.0)	
Male	50 (54.9)	171 (57.0)	
Tumour grade			0.06
Well / moderate	72 (79.1)	263 (87.7)	
Poor / undifferentiated	17 (18.7)	35 (11.6)	
Clinical stage	2 (2.2)	2 (0.7)	
I			0.08
II	16 (17.6)	60 (20.0)	
III	27 (29.7)	100 (33.3)	
IV	28 (30.8)	74 (24.7)	
Unknown	16 (17.6)	64 (21.3)	
Tumour location	4 (4.4)	2 (0.7)	
Proximal			0.49
Distal	39 (42.9)	141 (47.0)	
MSI	52 (57.1)	159 (53.0)	
MSS			0.58
MSI-H	77 (84.6)	266 (88.7)	
Unknown	13 (14.3)	32 (10.7)	

* Non-informative patients lacked staining score for >20% of antibodies.

† Comparison between patients; *P* value determined by either Chi-square or Fisher's exact test.

Abbreviations: MSS = micrsoatellite-stable: MSI-H = microsatellite-instability high: SD = standard deviation

3.5.5.3. Prognostic Significance of Individual Markers

Of the 12 markers tested, six were significantly associated with survival ($P<0.05$) as either a positive or negative indicator, and another three markers showed a trend toward significance ($P<0.2$) (Table 3.6). Expression of the remaining three markers, Cyclin D1, EGFR, and p53, were not associated with survival in univariate analysis.

Multivariate Cox regression analyses, including other prognostic indicators (patient age at diagnosis, sex, tumour stage, grade, and MSI status), were performed with each of the markers that were significant in univariate analyses. E-cadherin, MCM7, and PCNA remained significant after adjusting for these indicators (data not shown).

3.5.5.4. Unsupervised Hierarchical Clustering Analysis of Immunostaining Data

Unsupervised hierarchical clustering analysis resulted in the assignment of three well-defined cluster groups, which included all tumours, as seen in Figure 3.6. We used Chi square analysis to assess the contribution to cluster groups made by each marker; this data is given in Table 3.8. Cluster group 2 was characterized by comparatively lower expression of MCM7, p21, and p53 than the other cluster groups. Cluster group 3 had the lowest levels of expression of β-catenin, E-cadherin, and p27.

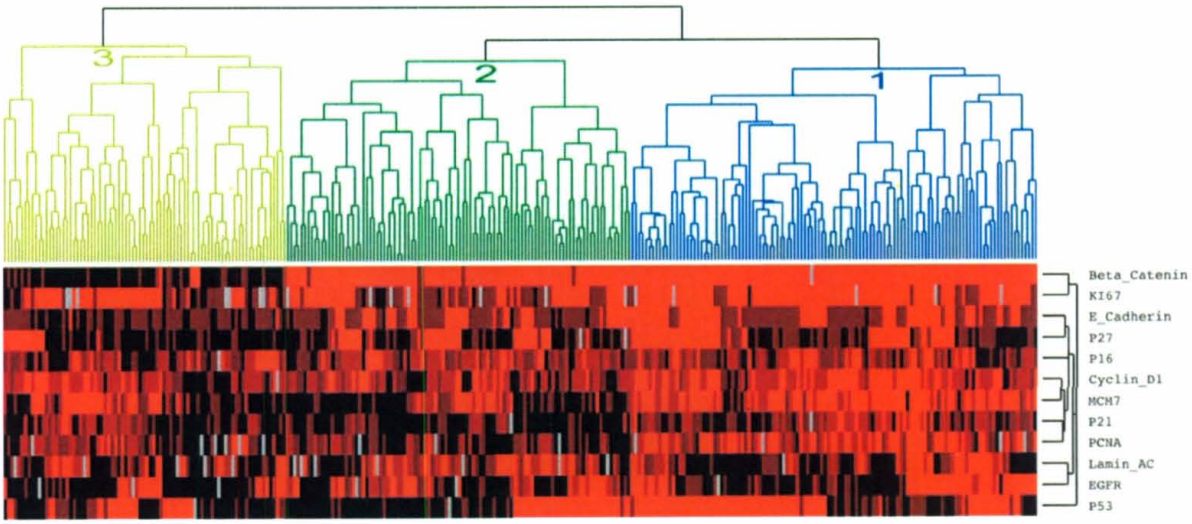


Figure 3.6. Dendrogram and heat map illustrating the cluster groups that result from unsupervised clustering analysis. Each row represents a unique immunomarker, and each column represents a unique tumour. The relatedness of different tumours, based on their immunoprofiles, is illustrated by the length of the horizontal lines separating them at the top of the image.

Table 3.8. Distribution of markers according to cluster assignment

Marker	Cluster 1	Cluster 2	Cluster 3	*P**
	No. (%)	No. (%)	No. (%)	
Beta Catenin				<0.001
Score 0	0 (0)	0 (0)	60 (73.2)	
Score 1	0 (0)	5 (5)	18 (22.0)	
Score 3	117 (100)	95 (95.0)	4 (4.9)	
Cyclin D1				<0.001
Score 0	7 (5.9)	22 (22.0)	16 (19.8)	
Score 1	8 (6.8)	19 (19.0)	11 (13.6)	
Score 2	50 (42.4)	44 (44.0)	39 (48.1)	
Score 3	53 (44.9)	15 (15.0)	15 (18.5)	
E-Cadherin				<0.001
Score 0	1 (0.8)	8 (8.0)	29 (35.4)	
Score 2	59 (50.0)	57 (57.0)	51 (62.2)	
Score 3	58 (49.2)	35 (35.0)	2 (2.4)	
EGFR				<0.001
Score 0	46 (39.3)	37 (37.8)	52 (65.8)	
Score 1	18 (15.4)	15 (15.3)	5 (6.3)	
Score 2	9 (7.7)	21 (21.4)	5 (6.3)	
Score 3	44 (37.6)	25 (25.5)	17 (21.5)	
KI67				<0.001
Score 0	1 (0.9)	7 (7.5)	12 (16.0)	
Score 1	15 (13.3)	41 (44.1)	21 (28.0)	
Score 3	97 (85.8)	45 (48.4)	42 (56.0)	
Lamin A/C				<0.07
Score 0	50 (42.7)	22 (23.7)	35 (45.5)	
Score 1	21 (17.9)	20 (21.5)	12 (15.6)	
Score 2	20 (17.1)	26 (28.0)	14 (18.2)	
Score 3	26 (22.2)	25 (26.9)	16 (20.8)	
MCM7				<0.001
Score 0	6 (5.1)	53 (53.0)	29 (35.4)	
Score 1	6 (5.1)	22 (22.0)	18 (22.0)	
Score 2	30 (25.4)	10 (10.0)	9 (11.0)	
Score 3	76 (64.4)	15 (15.0)	26 (31.7)	
p16				0.15
Score 0	20 (16.9)	13 (13.0)	21 (25.9)	
Score 1	39 (33.1)	29 (29.0)	24 (29.6)	
Score 2	34 (28.8)	42 (42.0)	22 (27.2)	
Score 3	25 (21.2)	16 (16.0)	14 (17.3)	

p21				<0.001
Score 0	31 (26.5)	66 (66.7)	39 (47.6)	
Score 1	36 (30.8)	24 (24.2)	20 (24.4)	
Score 2	37 (31.6)	9 (9.1)	21 (25.60	
Score 3	13 (11.1)	0 (0)	2 (2.4)	
p27				<0.001
Score 0	31 (26.3)	41 (41.8)	53 (64.6)	
Score 1	37 (31.4)	25 (25.5)	16 (19.5)	
Score 2	50 (42.4)	32 (32.7)	13 (15.9)	
p53				<0.001
Score 0	23 (19.5)	55 (55.6)	38 (46.9)	
Score 1	24 (20.3)	11 (11.1)	18 (22.2)	
Score 3	71 (60.2)	33 (33.3)	25 (30.9)	
PCNA				<0.001
Score 0	4 (3.5)	45 (46.4)	26 (34.2)	
Score 1	21 (18.4)	32 (33.0)	25 (32.9)	
Score 2	34 (29.8)	19 (19.6)	16 (21.1)	
Score 3	55 (48.2)	1 (1.0)	9 (11.8)	

* *P* value determined by either Chi-square or Fisher's exact test.

3.5.5.5. Association of Cluster Groups with Clinicopathological Features

There were no differences between cluster groups for patient sex, tumour location, histological differentiation, or MSI status, but there were statistically significant differences among groups for both patient age ($P = 0.003$) and clinical stage at diagnosis ($P = 0.04$)(Table 3.9). Patients were generally youngest at diagnosis in cluster group 1, and oldest in cluster group 3. They were generally of lowest clinical stage in cluster group 1.

Table 3.9. Clinicopathological features of patients according to cluster assignment

Clinicopathological Feature	Cluster 1 n = 118 No.* (%)	Cluster 2 n = 100 No. (%)	Cluster 3 n = 82 No. (%)	*P*†
Age, yrs (SD)	64.0 (11.1)	65.5 (12.6)	69.9 (12.7)	0.003
Sex				0.31
Female	53 (44.9)	37 (37.0)	39 (47.6)	
Male	65 (55.1)	63 (63.0)	43 (52.4)	
Clinical Stage				0.04
Stage I	31 (26.3)	14 (14.0)	15 (18.3)	
Stage II	44 (37.3)	34 (34.0)	22 (26.8)	
Stage III	27 (22.9)	26 (26.0)	21 (25.6)	
Stage IV	16 (13.6)	24 (24.0)	24 (29.3)	
Tumour Location				0.72
Distal	62 (52.5)	56 (56.0)	41 (50.0)	
Proximal	56 (47.5)	44 (44.0)	41 (50.0)	
Differentiation				0.36
Well | Moderate	107 (90.7)	88 (88.9)	68 (84.0)	
Poor | Undiff.	11 (9.3)	11 (11.0)	13 (15.9)	
MSI				0.2
MSS	107 (91.5)	90 (90.9)	69 (84.1)	
MSI-H	10 (8.5)	9 (9.1)	13 (15.9)	

* Patients may not equal the number of patients in clusters because of missing data.
† Comparison of cluster groups and *P* value determined by either independent-samples t-test or Chi-square test.
Abbreviations: MSS = microsatellite-stable; MSI-H = microsatellite-instability high.

3.5.5.6. Association of Cluster Groups with Patient Survival

Disease-free survival was greatest for those assigned to cluster group 1, whereas those in cluster groups 2 and 3 were similar to each other, both having significantly poorer survival rates (Figure 3.7). The increased disease-free survival of cluster group 1 patients was statistically significant ($P = 0.002$ compared to cluster group 2, and $P <$

0.001 compared to cluster group 3). The results were similar when considering overall survival. Again, cluster group 1 had better overall survival, and this was statistically significant when compared to both cluster groups 2 and 3 ($P < 0.001$ for both) (data not shown). The hazard ratio for cluster group 2 versus group 1 was 1.94 (95% confidence interval [CI] 1.23 – 3.01), and for group 3 it was 2.35 (95% CI 1.50 – 3.69) (Table 3.10).

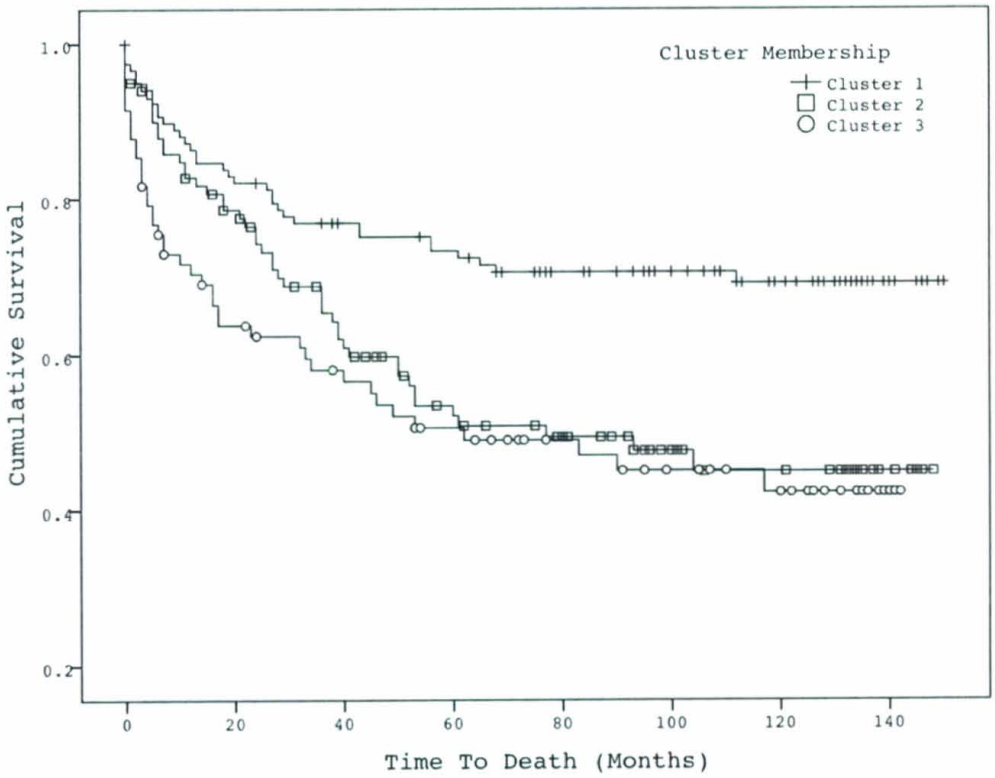


Figure 3.7. Cumulative disease-free survival of colorectal cancer patients based on assignment to cluster groups

Multivariate Cox regression analysis revealed that disease-specific survival in cluster group 3 remained significantly worse than cluster group 1, with hazard ratio (HR)

1.82 (95% CI 1.15 – 2.88), independent of patient age at diagnosis, clinical stage, tumour grade, and MSI status (Table 3.10). However, disease-specific survival in cluster group 2 was not significantly different from group 1 in the multivariate model.

Table 3.10. Disease-specific survival according to cluster assignment

Factor	Univariate Analysis		Multivariate Analysis	
	Hazard Ratio* (95% CI)	*P*	Hazard Ratio* (95% CI)	*P*
Cluster Assignment				
Cluster 1	1.00 (Ref.)		1.00 (Ref.)	
Cluster 2	1.94 (1.23 - 3.01)	0.003	1.13 (0.72 - 1.80)	0.59
Cluster 3	2.35 (1.50 - 3.69)	<0.001	1.82 (1.15 - 2.88)	0.01
Age			1.02 (1.00 - 1.03)	0.02
Clinical Stage			2.90 (2.36 - 3.56)	<0.001
Tumor Grade				
Well or moderately differentiated			1.00 (Ref.)	<0.001
Poorly or undifferentiated			2.38 (1.49 - 3.79)	
MSI				
MSS			1.00 (Ref.)	0.001
MSI-H			0.15 (0.05 - 0.47)	

*Hazard ratio estimated by Cox regression model.

Abbreviations: CI = confidence interval; MSS = microsatellite-stable; MSI-H = microsatellite-instability high; Ref. = reference

3.5.5.7. Risk of Distant Metastasis

Although there was a trend indicating possible increased risk of patients in cluster group 2 developing distant metastasis when compared to cluster group 1, this was not

statistically significant; nor was it when comparing cluster group 3. In a multivariate Cox regression model assessing the hazard of developing distant metastasis in patients diagnosed with CRC at stage I, II, or III, only clinical stage at diagnosis and MSI status were significantly associated (Table 3.11). Increasing clinical stage at diagnosis was associated with an increased risk of developing metastasis (HR 2.32, 95% CI 1.51 – 3.55), while high levels of MSI (MSI-H) was associated with protection against metastasis (HR 0.11, 95% CI 0.02 – 0.83).

3.5.6. Discussion

Colorectal cancer is a common, yet heterogeneous disease. Beyond the TNM classification system there is a paucity of clinically relevant biomarkers in use to effectively predict individual patient outcome and response to therapy. While the majority of work to date has evaluated the utility of single biomarkers for these purposes, the expression profiles of multiple biomarkers may be more appropriate considering the complex, heterogeneous nature of the disease. We used unsupervised hierarchical clustering analysis of 12 biologically relevant immunomarkers on tissue microarrays containing cores of tissue from 391 colorectal tumours to produce prognostically relevant cluster groups. Our results are an important first step in the process of creating such clinically useful profiles.

Table 3.11. Risk of metastasis for patients diagnosed with tumours at stages 1, 2, and 3, according to cluster assignment

Factor	Univariate Analysis		Multivariate Analysis	
	Hazard Ratio* (95% CI)	*P*	Hazard Ratio* (95% CI)	*P*
Cluster Assignment				
Cluster 1	1.00 (Ref.)		1.00 (Ref.)	
Cluster 2	1.83 (0.98 - 3.41)	0.06	1.56 (0.83 - 2.93)	0.16
Cluster 3	1.17 (0.54 - 2.51)	0.7	1.24 (0.56 - 2.71)	0.6
Age			1.01 (0.99 - 1.04)	0.38
Clinical Stage			2.32 (1.51 - 3.55)	<0.001
Tumor Grade				
Well or moderately differentiated			1.00 (Ref.)	0.88
Poorly or undifferentiated			0.92 (0.33 - 2.61)	
MSI				
MSS			1.00 (Ref.)	0.03
MSI-H			0.11 (0.02 - 0.83)	

* Hazard ratio estimated by Cox regression model.

Abbreviations: CI = confidence interval; MSS = microsatellite-stable; MSI-H = microsatellite-instability high; Ref. = reference

A limitation of this study was the multiple missing data points. Despite this, we found that a number of protein markers were associated with the prognosis of patients with colorectal cancer in Newfoundland and Labrador. We found that tumour expression of E-cadherin, MCM7, p16, p27, and PCNA were all significantly associated with improved prognosis, and expression of β-catenin and p21 showed a trend towards significance in univariate analysis. Tumour expression of Ki67 was significantly associated with poorer survival, and expression of Lamin A/C showed this trend as well. When adjusting for other known prognostic indicators, E-cadherin, MCM7, and PCNA maintained their significant association with patient outcome.

Unsupervised hierarchical clustering analysis identified three tumour cluster groups with different patient clinicopathological features and outcomes. One of these cluster groups, group 3, was associated with poorer patient survival independently from other prognostic variables including clinical stage at diagnosis. This cluster group differed from the other two in that it displayed an overall lower level of protein expression, and was associated with the oldest age at onset. Although not statistically significant, cluster group 3 had the highest proportion of female patients, tumours located in the proximal colon, and tumours with a poorly differentiated or undifferentiated histology. All of these features are associated with tumours that develop along the CpG island methylator phenotype (CIMP) pathway of carcinogenesis [263]. Due to the widespread methylation of CpG islands associated with this phenotype, transcription is silenced for many genes, including p16 [273], which is compatible with observations of cluster group 3. CIMP tumours represent approximately three quarters of MSI-H CRCs,

but can also be MSS. CIMP, MSS tumours are associated with a poor prognosis [260], and this is also associated with cluster group 3. Approximately one third of CRCs are CIMP positive, and these are commonly older, female patients. All of this is compatible with cluster group 3 representing CIMP tumours, however, unfortunately methylation and *BRAF* mutation data is not available for us to confirm this theory.

We have shown that unsupervised hierarchical clustering analysis of immunohistochemistry data garnered from tissue microarray analysis can be used to identify prognostically relevant subgroups of colorectal carcinoma. These subgroups may be related to the pathway of tumour carcinogenesis, and are somewhat independent from other prognostic factors, including clinical stage at diagnosis. Future study should be performed using this technology to further explore the potential for developing clinically useful prognostic tests.

3.5.7. Acknowledgements

We would like to thank the Centre for Translational and Applied Genomics at the Provincial Health Services Authority Laboratory of the British Columbia Cancer Agency (Alicja Parker and Julie Lorette) for TMA staining; Susan Ryan and Elizabeth Hunt from the Cancer Care Program, Eastern Health Newfoundland for data collection; Laura Edwards for data entry; the Newfoundland Familial Colorectal Cancer Registry (Project Coordinator Dr. Elizabeth Dicks) for access to the surgical specimens obtained from 1999-2003.

Chapter 4 – Use of Tumour Histology for Prediction of Microsatellite Instability

4.1 Introduction

4.1.1 The Use of Histology in Colorectal Cancer Research

Histology is defined as the study of tissues and cells with the use of a microscope [806]. Similarly, tumour histology is the study of tumour cells and tissues with a microscope. The size, shape, and appearance of tumour cells, as well as the structural morphology in which they appear, can be informative with regards to tumour pathogenesis, prognosis, and drug response. The composition and appearance of the tumour-supporting cells, including the tumour stroma and related immune cells, can also provide information to aid in explaining a particular tumour phenotype. For example, tumour histology and lymphocytic infiltration have been used to predict MSI status in colorectal cancer [61,72,105,106,166,344,474,807-809], to discriminate between sporadic MSI-H tumours and those related to Lynch Syndrome [810-812], and also used as both predictive [69,813] and prognostic markers [63,68,78,96,97,101-103,814-816].

In 1986 the first study [817] was published that compared the histopathologic features of colorectal tumours diagnosed in patients with what was then termed Cancer Family Syndrome (CFS), with those diagnosed in patients without a family history of

CRC. Cancers of the proximal colon and endometrium were the predominant cancers seen in cancer-family syndrome, and families with these types were later redefined as having Lynch Syndrome. In their 1986 paper [817], Mecklin and colleagues compared the histopathology of 100 CRCs from 75 patients with confirmed diagnoses of CFS, with that of control, non-hereditary CRCs. They found that the tumours from CFS patients were both significantly more mucinous, and more poorly differentiated. They also compared the histopathology of the colonic adenomas found in the patients of the same groups, and found that the CFS adenomas were more likely to have moderate or severe dysplasia, and have villous features. From this they made two general conclusions: that the presence of mucinous histology could be used to help identify CFS, and that there might be an accelerated adenoma to carcinoma sequence in CRCs associated with CFS.

Published in 2004, the Revised Bethesda Guidelines (RBG) [72] (Table 1.14.), were the first example of recommendations for the use of histology to aid in the identification of Lynch Syndrome in a population-based setting. These guidelines were discussed previously in section 1.8.1.1, and include both personal and family history of Lynch Syndrome associated cancers, and tumour histology.

As a screening test for MSI status and Lynch Syndrome, the RBG have not been very useful. An article published in 2005 by Piñol and colleagues [818] reported 23.5% of their population-based cohort met RBG. The guidelines identified carriers of germline mutations with a sensitivity of 90.0%, and a specificity of 77.1%. Numbers this high have not been replicated by any other group. Other studies using population-based cohorts found from 25.9% to 42.1% of patients meeting RBG [349,351,352]. One paper [109]

from 2006, reported a sensitivity similar to that of Piñol, with the RBG identifying 94.7% of patients with germline mutations; however, their specificity was much lower, at 37.6%. This group selected CRC patients, and only included those diagnosed younger than age 55 years in their analyses. A number of studies have reported sensitivities in the range of 75% [6,350,351], and one reported a sensitivity of only 50% [349] for identifying germline mutation carriers. Most of these articles did not provide enough information to determine the associated specificities.

A 2008 paper by Hampel and colleagues [350] reported that in their population-based cohort, if tumour analysis had been limited to patients who fulfilled the RBG, 28% of Lynch Syndrome cases would have been missed. Missing more than one in four cases is unacceptable for a screening test.

The RBG have not performed any better when considering the identification of MSI-H tumours. A 2008 paper [819], in which only patients diagnosed younger than age 60 years were included, reported a sensitivity of 75%, but did not report a specificity. A 2010 paper [820], which also studied a cohort diagnosed younger than 60 years, found a similar sensitivity, of 77.8%, and a specificity of 47.2%. Studies which included older patients reported a much lower utility in identifying MSI-H tumours. A 2009 paper [352] reported a sensitivity of 45.4% and a specificity of 64.6%. In this study, of the tumours that did not meet RBG, 11% were MSI-H, whereas 14% of the entire cohort were MSI-H. Finally, a 2010 paper [349] reported a sensitivity of only 31.6%, and a specificity of 75.3% for RBG in identifying MSI-H tumours in their population-based cohort of 336 CRCs.

While the RBG have not been successful in screening for MSI-H tumours, it is possible that parts of it could be. Four of the five criteria are based on personal and family history of Lynch-associated tumours, and age at onset of disease. These criteria were intended to identify the clinical aspects of Lynch Syndrome within a family, whereas the histology-based criterion is based on evidence that the features it includes are associated with MSI-H tumours.

A 2007 paper by Jenkins and colleagues [105] introduced a new screening tool for the identification of MSI-H CRCs. They published an algorithm, MsPath, which included clinical and demographic information about the patient, as well as histological information from the tumour. The histological features included in their algorithm were those from the histology-based criterion of the RBG, however, they weighted them based on their predictive values. They reported a sensitivity of 93% and specificity of 55% for the identification of MSI-H tumours.

The histology features described in the RBG, as well as others not included there, could play an important role in clinical practice as a first step in triaging CRCs to identify cases of Lynch Syndrome. However, identifying cases of Lynch Syndrome is only one of the reasons for identifying MSI-H tumours. As discussed in sections 1.5.1.5 and 1.10, MSI-H CRCs are associated with a better prognosis [61,63-69], and do not appear to benefit from 5-FU, a standard chemotherapeutic agent [62,67,346,453,821].

In the study presented in this chapter, I performed a histology review of the CRCs in our cohort. I assigned scores for all of the histological features included in the RBG, as well as others. Histological features were selected based on observation during

histological review, as well as on their relevance to CRC or other cancers. I used the histological data to assess the utility of the RBG and MsPath at identifying MSI-H tumours in our cohort, and to build a new algorithm for the prediction of these tumours.

4.1.2 Histological Features Included in Analysis

The histological features that I scored can be divided into two categories: those of the tumour and of the host response. The tumour features included: histological heterogeneity, the histological subtype of both predominant and minor components of the tumour, tumour grade, presence of dirty necrosis, and new blood vessel formation. Features considered to be due to the host response were immune-based, and included: Crohn-like lymphocytic response (CLR), peritumoural lymphocytic response, tumour-infiltrating lymphocytes (TILs), and proportion of plasma cells in the stroma. With the exception of increased stromal plasma cells, these features of host immune response were previously described in section 1.5.2.1 as prognostic factors for CRC. CLR and TILs were also included in the Revised Bethesda Guidelines as being predictive of MSI-H tumours [72]. To my knowledge there is no published data on the significance of increased stromal plasma cells in colorectal cancer. We chose to study this feature because of its relevance to other malignancies, such as melanoma and breast cancer, and our observation of variations of stromal plasma cell densities in different colorectal tumours.

We defined histologic heterogeneity as the presence of one or more areas of a tumour that exhibit different morphology from the preponderant morphology of the

tumour. This distinct, second growth pattern could be a minor component, or represent up to half of the tumour. For each histologic subclone identified we scored both histologic subtype and grade. The presence of histologic heterogeneity has been identified by others as being predictive of microsatellite instability in CRC [106,807,808].

While we and others [810] have included mucin as a distinct morphology for this feature, another group did not [106], as they stated that virtually all mucinous tumours would be identified as having histologic heterogeneity. We were particularly interested in the identification of a mucinous component as the minor morphologic feature within a tumour. Tumours with more than 50% mucinous morphology have been identified as more likely than conventional adenocarcinomas to be MSI-H [25,72], but tumours with less than 50% have not been well described. Only one other research group has evaluated focal mucin within a CRC. In both of their studies [106,807] they found focal mucin to be predictive of microsatellite instability.

We observed and recorded components of morphological subtypes not usually described in association with colorectal cancer, including tubular and papillary morphology. A tubular morphology is associated with a rare, genetically distinct group of breast cancers with a favourable outcome [822,823]. In breast cancer it is defined as a very well-differentiated invasive carcinoma, with cells arranged in well defined tubules, surrounded by fibrous stroma [824]. The morphology is similar to that of a non-invasive tubular adenoma of the colon or rectum, but is not recognized as a distinct morphological subtype of invasive carcinomas of the colon and rectum. A papillary morphology is seen in a small number of breast cancers, and these also have good prognosis [825]. A papillary

pattern of growth resembles finger-like projections, papillae, which have fibrovascular cores lined by tumour cells. It is also seen in the majority of thyroid cancers, and again is generally associated with excellent prognosis [826]. A subset of papillary thyroid cancers, however, are associated with a poorer outcome, and these cancers are identified by the activating *BRAF* mutation, p.V300E, the same mutation seen in a subset of colorectal tumours [827].

We observed heterogeneity not only in the histological subtype of some tumours, but also in tumour grade. At the time that the cancers in our cohort were diagnosed, the convention in clinical pathology was to score tumour grade based on the predominant level of differentiation observed within the tumour. However, some evidence suggests that it is more clinically relevant to score tumour grade based on the area of the tumour observed with the least differentiation [25,105,816], so we assigned a score for this as well.

Dirty necrosis, also known as garland necrosis, is a feature observed in some colorectal and other cancers. It is characterized by inflammation, identified by the presence of polymorphonuclear lymphocytes; as well as devitalised cells, identified by apoptotic bodies and necrosis. In intraductal breast and prostate cancers it is associated with aggressive tumours and a poor prognosis [828-831]. When seen in ovary it raises suspicion that the origin of the cancer may be a metastasis from the colon. In colon and rectal cancers this feature has been associated with microsatellite stable tumours, and inversely with MSI-H [106,807].

Angiogenesis refers to the growth of new blood vessels. It is required by solid tumours to facilitate growth beyond 1 or 2 mm in diameter. Solid tumours require blood

flow to deliver oxygen and nutrients to tumour cells, as well as to remove metabolic waste. New vessel formation also facilitates invasion [832] and metastasis [833]. Furthermore, the endothelial cells of blood vessels release growth factors and cytokines that can act in a paracrine manner to further stimulate tumour growth [834]. A high density of microvessels within the tumour has been associated with poor prognosis [835]. We scored the presence of new vessel formation deep within the tumour, and did not include the superficial vessels associated with ulceration.

4.2 Objectives

The objectives of this study were to assess the utility of the RBG and MsPath, and create an improved algorithm to predict MSI-H CRCs. I used the population-based cohort of CRCs diagnosed before the age of 75 years in Newfoundland and Labrador in 1999 to 2003 (N=719), collected by the NFCCR as part of the CIHR-IHRT project (described in section 1.11) for the assessment of RBG and MsPath, and the development of the algorithm. I used the population-based cohort of CRCs diagnosed at any age on the Avalon Peninsula of Newfoundland in 1997 and 1998 (N=294), collected for the pilot molecular study presented in chapter 2, for validation of the new algorithm.

4.3 Manuscript – "A Histology-Based Model for Predicting Microsatellite Instability in Colorectal Cancers"

The data from this study have been used as the basis of one publication, presented here, and as supporting evidence in a second, which is included as Appendix K. The manuscript included in this chapter describes the development and validation of the PREDICT model for predicting MSI-H CRCs. It was published in the *American Journal of Surgical Pathology*, in 2010 [836].

4.3.1 Authors and Affiliations

Angela Hyde[1], Dan Fontaine[2], Susan Stuckless[3], Roger Green[1], Aaron Pollett[4], Michelle Simms[1], Payal Sipahimalani[1], Patrick Parfrey[3], and H.B. Younghusband[1]

1. Discipline of Genetics, Memorial University of Newfoundland, St. John's, Newfoundland, Canada
2. Department of Pathology, Eastern Health, St. John's, Newfoundland, Canada
3. Clinical Epidemiology Unit, Memorial University of Newfoundland, St. John's, Newfoundland, Canada
4. Department of Pathology, Mount Sinai Hospital, University of Toronto, Toronto, Ontario, Canada

4.3.2 Summary

Background: Identifying colorectal cancers (CRCs) with high levels of microsatellite instability (MSI-H) is clinically important. MSI-H is a positive prognostic marker for CRC, a predictive marker for resistance to standard 5-fluorouracil-based adjuvant chemotherapy, and an important feature for identifying individuals and families with Lynch Syndrome. Our aim was to compare and improve upon the existing predictive pathology models for MSI-H CRCs.

Methods: We tested two existing models used to predict MSI-H tumors, 1) Revised Bethesda Guidelines and 2) MsPath, in our population-based cohort of CRCs diagnosed <75 years from Newfoundland (N=710). We also scored additional histological features not described in the other models.

Results: From this analysis we developed a model for the prediction of MSI-H CRCs; Pathologic RolE in Determination of Instability in Colorectal Tumours (PREDICT). An independent pathologist validated this model in a second cohort of all CRCs (N=276).

Conclusions: Tumour histology was a better predictor of MSI status than was personal and family history of cancer. MsPath identified MSI-H CRCs with a sensitivity of 92.1% and a specificity of 37.8%, while the Revised Bethesda Guidelines had a sensitivity of 81.3% and a specificity of 39.5%. PREDICT included some new histology features.

including peritumoral lymphocytic reaction and increased proportion of plasma cells in the tumour stroma. PREDICT was superior to both existing models in the development cohort with a sensitivity of 97.4% and a specificity of 53.9%. In the validation cohort sensitivity was 96.9% and specificity 76.6%. We conclude that PREDICT is a good predictor of MSI-H colorectal cancer.

4.3.3 Introduction

Colorectal cancer (CRC) is one of the most common forms of malignancy, representing almost 10% of all new cancer diagnoses and 8% of deaths from cancer worldwide [1,794]. Approximately 10 to 15% of colorectal cancers develop due to defects in the DNA mismatch repair (MMR) system [60-62], and have high levels of microsatellite instability (MSI-H). Patients with MMR-deficient CRCs have a better overall prognosis than those with MMR-proficient CRCs [61,63-69], although the tumours respond less well to standard, 5-fluorouracil (5-FU)-based adjuvant chemotherapy [70,71]. Identification of MMR-deficient CRCs therefore has important clinical implications, as MMR-deficiency is both a prognostic and predictive factor.

5-FU is the basis of virtually all first-line adjuvant chemotherapy for colorectal cancer [345,346]. MMR-deficient tumours are resistant to 5-FU [70,71], and patients with MSI-H tumours do not appear to benefit in overall or disease-free survival from 5-FU therapy [62,67,346,453,821]. For this reason it is important to determine the MSI status of a tumour when considering adjuvant treatment.

The majority of MMR-deficient CRCs are sporadic, and develop as a result of silencing of the MMR gene *MLH1* by methylation of its promoter. However, a subgroup of MMR-deficient tumours develops from inherited germline mutations in MMR genes, resulting in an autosomal dominant cancer predisposition called Lynch Syndrome. Lynch Syndrome (also known as Hereditary Non-Polyposis Colorectal Cancer) accounts for up to 3% of all cases of colorectal cancer (CRC) [4-6]. Since CRC is common, with a cumulative lifetime risk of 5% in the general population, it is estimated that as many as 1 in 660 people have germline MMR mutations [4]. It is important to identify cases of Lynch Syndrome since mutation carriers and their relatives are at an increased risk of early onset CRC and other extra-colonic cancers. Identification of this risk allows for increased clinical screening and early detection of disease, which has been shown to decrease morbidity and mortality [418,837-840].

MMR deficiency leads to MSI-H in the tumours. Testing all colorectal tumours for microsatellite status in a community-based setting would be difficult due to cost, although a recent article addresses the importance of doing this [841]. The Revised Bethesda Guidelines (RBG) [72] were developed to identify tumours likely to be MSI-H, that should be tested for MSI and investigated for Lynch Syndrome. Four of the five criteria that make up these guidelines are based on family and personal history of Lynch-associated tumours, while the other criterion is based on tumour histology.

In 2007, Jenkins *et. al.* [105] published the MsPath model for predicting MSI-H colorectal tumours. This model is based on the tumour histology described in the third (histology-based) RBG criterion, and improves upon the sensitivity and specificity of the

RBG. Since these guidelines were developed to identify cases of Lynch Syndrome, they focus on patients diagnosed at a young age: they are only applied to patients diagnosed below age 60 years. The majority of MSI-H CRCs are diagnosed older than age 60 years, and would be missed by these models.

In this study we assessed the utility of both the Revised Bethesda and MsPath models at predicting MSI-H colorectal tumours in our population-based cohort. We assessed the value of the third RBG criterion and MsPath for patients diagnosed older than age 60 years. We also considered whether additional histological features could improve the prediction of MSI-H CRCs, which led us to develop a model to assess the Pathological RolE in the Determination of Instability in Colorectal Tumours (PREDICT). Finally, we validated this model in a second patient cohort, with a different pathologist.

4.3.4 Materials and Methods

4.3.4.1. Subjects

The cohort used to develop the model was from the Newfoundland Familial Colorectal Cancer Registry (NFCCR) [469]. All colorectal cancers confirmed by pathology and diagnosed under the age of 75 years in the province of Newfoundland and Labrador from 1999 to 2003 were identified, and informed consent was obtained from 730 patients or proxies, who provided family history information. Formalin-fixed, paraffin-embedded

tissue samples from pathological specimens were available for 719 tumours from 700 patients. MSI status was not available for nine tumours, so these were not included in the study. An additional 19 tumours had incomplete data and were not analyzed.

The validation cohort consisted of all incident cases of colorectal cancer diagnosed at any age on the Avalon Peninsula of Newfoundland in 1997 or 1998. The Avalon Peninsula accounts for half of the total population of the province. Tumour tissue samples from pathological specimens were available for 276 of 294 patients.

Ethics approval was granted by the Human Investigations Committee of Memorial University of Newfoundland (HIC #01.70).

4.3.4.2. Collection of Data

Patients' demographic and clinical information, including age, sex, and tumour location, was collected from the provincial tumour registry. Information regarding tumour pathology, such as tumour grade and histologic subtype, was collected from the original pathology reports.

4.3.4.3. Microsatellite Instability Analysis

DNA was extracted from matched tumour and normal tissue from formalin-fixed paraffin-embedded pathological tissue blocks using the Qiagen DNeasy Tissue Kit, as per

the manufacturer's instructions (Qiagen Inc., Mississauga, ON, Canada), as previously described [797]. The 99 patients from the validation cohort that were diagnosed age 75 years or older were not tested for all five microsatellite markers, but only for BAT26, as it had a 100% concordance with MSI status in the younger cases from that cohort.

4.3.4.4. Histology Review and Scoring of Histology Features

One representative slide from the original pathology analysis was chosen from each tumour to include the maximum number of viable tumour cells, along with the advancing edge, when possible. Histology review was performed by two of the authors (D.F. and A.H.), with blinding to MSI status and clinical outcomes. Scores were assigned for: Crohn-like lymphocytic reaction, the presence of dirty necrosis, mucinous component, tumour-infiltrating lymphocytes (TILs), new vessel formation, increased stromal plasma cells, peritumoural lymphocytic reaction, and the presence, type and grade of tumour subclones. TILs and Crohn-like reaction were included in the Revised Bethesda and MsPath guidelines, and scored as described previously [105] for the purpose of assessing the sensitivity and specificity of those guidelines. The additional features included in our new model were scored as follows:

Mucinous component was defined as the presence of mucin dissecting into the stroma surrounding a tumour gland associated with a cellular proliferation. Scoring was binary, with tumours scored as having a mucinous component if they either had a mucinous

morphology (>50% of the tumour on the slide examined), or if there was a focal area of dissecting mucin found within a predominantly non-mucinous tumour morphology.

Peritumoural lymphocytic reaction was scored as either present or absent. It was defined as a conspicuous lichenoid banding of lymphocytes found just beyond the advancing edge of the tumour, in the subserosa or mesenteric fat.

Increased plasma cells in the stroma was defined as a greater than normal proportion of plasma cells amongst the immune cells in the tumour stroma. This feature was considered present when plasma cells made up greater than 25% of stromal immune cells.

4.3.4.5. Eligibility

Only primary, invasive colorectal tumours were included in this study. Cases of carcinoma *in-situ*, intramucosal carcinoma, and carcinoid tumours were excluded. While all eligible tumours were used in the development of PREDICT (N=710), only those without any missing scores for relevant clinical variables or histological features were included in analysis of the utility of the model (development cohort N=691, validation cohort N=276). The 967 tumours used for analysis were from 950 patients, as 17 patients had synchronous tumours included in the study.

4.3.4.6. Statistical Analysis

The development and validation cohorts were compared for demographic and clinical variables. Mean age at diagnosis was compared using the independent samples T-test. The proportion of MSI-H tumours, tumours from female patients and right-sided tumours were compared for the two cohorts using Pearson's Chi Squared test. Significance values were 2-tailed, and all variables were considered significant if *P* values were less than 0.05.

Univariate linear logistic regression analyses were performed to determine the value of the clinical and pathological features at predicting MSI-H tumours. Variables found to be significant predictors were used to create a multivariate model; only those features that remained significant were included in the final model.

The utility of the different models at predicting MSI status was compared by assessing the sensitivity, specificity, and positive (PPV) and negative (NPV) predictive values of the different models. These values were calculated using standard definitions [842]. Comparison of the area under the curve (AUC) of the receiver operating characteristics (ROC) curve was also used to compare models where appropriate.

The fit of the PREDICT model was assessed using data from the validation cohort. Multivariate linear regression analyses were performed to assess the validity of the model, as was estimation of the AUC of the ROC curve.

Statistical analyses were performed using SPSS version 15 for Windows (SPSS, Chicago, U.S.A.). ROC curves were generated using MedCalc (MedCalc Software, Mariakerke, Belgium).

4.3.5 Results

4.3.5.1. Description of Population

As outlined in Table 4.1, the development cohort consisted of 710 tumours, of which 78 (11.0%) were MSI-H and 293 (42.1%) were right-sided. The mean age of diagnosis was 60.7 years (range from 20 to 74 years) and 275 (38.7%) were diagnosed in females. The validation cohort consisted of 276 tumours, of which 32 (11.6%) were MSI-H and 122 (44.2%) were right-sided. The mean age at diagnosis was 66.5 years (range from 25 to 91 years) and 129 (46.7%) were from female patients. There was no significant difference between the two cohorts for MSI status or tumour site. However, the validation cohort was older at diagnosis (as expected) and had a significantly higher proportion of females than the development cohort.

Table 4.1. Demographic and clinical variables in development and validation cohorts

	Development cohort (N=710)	Validation cohort (N=276)	P value
Mean age at diagnosis (range)	60.7 years (20 – 74)	66.5 years (25 – 91)	<0.001
Number of MSI-H tumours (%)	78 (11.0)	32 (11.6)	0.785
Number of female patients (%)	275 (38.7)	129 (46.7)	0.022
Number of right-sided† tumours (%)	293 (42.1)*	122 (44.2)	0.550
Geographic region covered	Entire province of Newfoundland and Labrador	Avalon peninsula of Newfoundland (represents approx. half the population of the province)	N/A

*Note: The total for this number is out of 696, not 710, because there were 14 tumours for which site was unknown.
†Right-sided was defined as tumours located at or proximal to the splenic flexure.

4.3.5.2. Utility of Revised Bethesda and MsPath

The utility of the Revised Bethesda Guidelines (RBG) and MsPath to predict MSI-H tumours was assessed in the development cohort. The third RGB criterion (MSI-H histology in tumours diagnosed in patients younger than age 60 years) has not always been included in analysis of the RBG [6,350,352,466]. We assessed the RBG both with and without the third criterion.

The four commonly used RBG criteria, criteria 1, 2, 4, and 5, assess personal and family history of cancer. Fifty one percent of tumours met at least one of these criteria, and so in a clinical setting would be referred for MSI analysis. However, of the 75 MSI-

H tumours in the cohort, only 50 were identified by the criteria, resulting in a sensitivity of 66.7% and a specificity of 50.9%. In practical terms, the RBG, as commonly used, would require MSI testing of just over half of all tumours, and would still miss one third of MSI-H tumours.

The rarely used third criterion, so-called "MSI-H histology" in a tumour diagnosed younger than age 60 years, on its own has a higher sensitivity than the other four criteria combined. This criterion identified 33 of the 40 MSI-H tumours diagnosed younger than age 60 years, giving a sensitivity of 82.5%, however, the specificity was low, at 27.1%. The age limit imposed in the criterion was intended to minimize the number of sporadic MSI-H tumours included, since the RBG were developed to help identify patients with Lynch Syndrome. However, this limit excludes the identification of both older-onset Lynch families, and the non-Lynch related MSI-H tumours. We decided to evaluate the criteria without this age limit. With the age limit removed, the third criterion improves slightly in its utility for identifying MSI-H CRCs. The sensitivity increases to 87.2%, and the specificity increases slightly to 28.2%.

To truly test the RBG, all 5 criteria must be included, as they were intended, with the age limit imposed for the third criterion. These guidelines identified MSI-H tumours with a sensitivity of 81.3% and a specificity of 39.5%. In a clinical setting this would translate to identifying 62.7% of tumours for MSI testing, while missing almost 20% of MSI-H tumours. If we removed the age limit from the third criterion, and again evaluated the full five criteria, only two of the 75 MSI-H tumours would be missed. The sensitivity increases to 97.3%, however, the specificity drops to just 13.1%. Clinically, this would

mean that very few MSI-H tumours would be missed, however, almost 90% of tumours would require MSI testing.

MsPath does not include family history, but rather just the histology features described in the RBG (tumour grade, histological subtype, Crohn-like reaction and TILs), along with tumour location and age at diagnosis. In our population, MsPath, with the recommended cut-off score of 1.0, identified 35 of the 38 MSI-H tumours in the group diagnosed younger than age 60 years. The sensitivity was 92.1%, and specificity 37.8%. When we removed the imposed age limit, the accuracy did not significantly change: 65 of the 71 MSI-H tumours were identified, the sensitivity was 91.5% and the specificity, 38.9%.

4.3.5.3. Development of PREDICT Model

To develop a new model for predicting MSI-H status, we assessed the predictive value of not only the clinical and histological features described in the RBG and MsPath models, but other histological features as well. Univariate analyses demonstrated that tumour-infiltrating lymphocytes (TILs), right-sided tumour location, peritumoural lymphocytic reaction, evidence of dissecting mucin, increased stromal plasma cells, the presence of histologic heterogeneity, lack of dirty necrosis, age of diagnosis less than 50 years, predominant mucinous, signet ring, or medullary histology, Crohn-like lymphocytic reaction, and female sex were each significantly associated with MSI-H

tumour status (Table 4.2). Tumour grade was not a significant predictor of MSI status on univariate analysis.

With multivariate regression analysis of the 11 features found to be significant on univariate analysis, six remained significant. Age below 50 years at diagnosis, right-sided tumour location, the presence of any dissecting mucin, TILs, peritumoural lymphocytic reaction and increased stromal plasma cells (Figure 4.1) were all independent predictors of MSI-H tumours, and were included in PREDICT (Table 4.3). The analysis showed that tumour location was the strongest predictor of MSI-H status, with right-sided tumours having an odds ratio of 5.1.

The rounded β coefficient from the multivariate regression equation for each feature was used as a score in PREDICT (Table 4.4). As with the MsPath model, the scores for each feature of a tumour were added together, resulting in the total PREDICT score for the tumour. PREDICT scores with associated sensitivity and specificity values are given in Table 4.5. To optimize the model for use as a clinical screening test, it was important to increase sensitivity so as to include the maximum number of MSI-H tumours, while keeping the specificity as high as possible to minimize the number of false positives. For this purpose we suggest an arbitrary cut-off of 2.5, and recommend that tumours with a total PREDICT score of at least 2.5 be referred for MSI testing. In our development cohort, using this threshold, the model had a sensitivity of 97.4% and a specificity of 53.9%.

Table 4.2. Univariate analysis of demographic and histologic variables in tumours with and without MSI in the development cohort

Feature	MSI-H tumors N (%)	MSS tumors N (%)	Total with feature N (%)	Odds Ratio (95% CI)	P value	Sensitivity (95% CI)	Specificity (95% CI)
TILs (not incl. subclone-only)	46 (33.6)	91 (66.4)	137 (19.3)	8.530 (5.173 – 14.065)	<0.001	59.0 (49.0 – 68.3)	85.6 (84.3 – 86.7)
Right-sided tumour location	61 (20.8)	232 (79.2)	293 (42.1)	6.360 (3.604 – 11.216)	<0.001	79.2 (69.5 – 86.6)	62.5 (61.3 – 63.4)
TILs (incl. subclone-only)	47 (24.9)	142 (75.1)	189 (26.7)	5.221 (3.207 – 8.500)	<0.001	60.3 (50.0 – 69.7)	77.5 (76.2 – 78.7)
Peritumoral reaction	67 (15.4)	367 (84.6)	434 (61.5)	4.332 (2.268 – 8.268)	<0.001	85.9 (77.0 – 91.9)	41.6 (40.4 – 42.3)
Mucinous component	40 (22.2)	140 (77.8)	180 (25.4)	3.699 (2.291 – 5.974)	<0.001	51.3 (41.2 – 61.2)	77.8 (76.6 – 79.1)
Increased stromal plasma cells	61 (15.5)	332 (84.5)	393 (55.4)	3.242 (1.863 – 5.640)	<0.001	78.2 (68.4 – 85.7)	47.5 (46.3 – 48.4)
Histologic heterogeneity	43 (18.0)	196 (82.0)	239 (33.7)	2.733 (1.701 – 4.390)	<0.001	55.1 (44.8 – 65.1)	69.0 (67.7 – 70.2)
Lack of Dirty necrosis	20 (21.7)	72 (78.3)	92 (13.0)	2.682 (1.533 – 4.695)	<0.001	25.6 (17.8 – 34.9)	88.6 (87.6 – 89.8)
Age at diagnosis <50 yrs	16 (18.4)	71 (81.6)	87 (12.3)	2.039 (1.124 – 3.702)	0.018	20.5 (13.4 – 29.4)	88.8 (87.9 – 89.9)
Mucinous/signet ring/medullary histology	16 (17.6)	75 (82.4)	91 (12.8)	1.917 (1.059 – 3.472)	0.031	20.5 (13.4 – 29.5)	88.1 (87.3 – 89.2)
Crohn-like reaction	54 (13.5)	346 (86.5)	400 (56.7)	1.834 (1.110 – 3.028)	0.018	69.2 (58.9 – 78.0)	44.9 (43.6 – 46.0)
Grade – poorly differentiated	12 (16.2)	62 (83.8)	74 (10.7)	1.800 (0.928 – 3.497)	0.083*	16.7 (10.1 – 25.6)	90.0 (89.2 – 91.0)
Female	40 (14.5)	235 (85.5)	275 (38.7)	1.778 (1.112 – 2.843)	0.016	51.3 (41.0 – 61.5)	62.8 (61.5 – 64.1)

Grade – poorly and well differentiated	23 (13.9)	143 (86.1)	166 (24.0)	1.566 (0.926 – 2.648)	0.095*	31.9 (22.8 – 42.5)	76.9 (75.9 – 78.2)
New vessel formation	39 (9.7)	365 (90.3)	404 (56.9)	0.732 (0.458 – 1.168)	0.192*	50.0 (39.7 – 60.3)	42.2 (41.0 – 43.5)
Tubular component	2 (4.5)	42 (95.5)	44 (6.2)	0.370 (0.097 – 1.413)	0.158*	2.6 (0.7 – 8.2)	93.4 (93.1 – 94.1)

* Not statistically significant.
(MSS: microsatellite stable, CI: confidence interval)
Note that the total number of tumours for some features was less than 710, this is because there were some tumors with missing data.

Table 4.3. A multivariate model for the prediction of MSI-H in the development cohort using the features identified as significant in the univariate analyses

Feature	**β (Standard Error)**	**Odds Ratio (95% CI)**	**P value**
Right-sided	1.625 (0.324)	5.080 (2.692 – 9.585)	<0.001
Mucinous component	1.550 (0.418)	4.711 (2.075– 10.696)	<0.001
TILs (incl. subclone-only)	1.303 (0.314)	3.679 (1.987 – 6.814)	<0.001
Age < 50 years	1.324 (0.384)	3.758 (1.771 – 7.973)	0.001
Peritumoural reaction	1.315 (0.429)	3.723 (1.608 – 8.624)	0.002
Increased stromal plasma cells	0.731 (0.345)	2.078 (1.057 – 4.085)	0.034
Lack of dirty necrosis	0.352 (0.373)	1.422 (0.685 – 2.953)	0.345*
Female sex	0.226 (0.291)	1.254 (0.709 – 2.220)	0.437*
Crohn-like reaction	0.057 (0.332)	1.058 (0.553 – 2.027)	0.864*
Histologic Heterogeneity	-0.103 (0.356)	0.902 (0.449 – 1.813)	0.772*
Mucinous/signet ring/medullary	-0.125 (0.437)	0.882 (0.375 – 2.078)	0.775*

*Not statistically significant

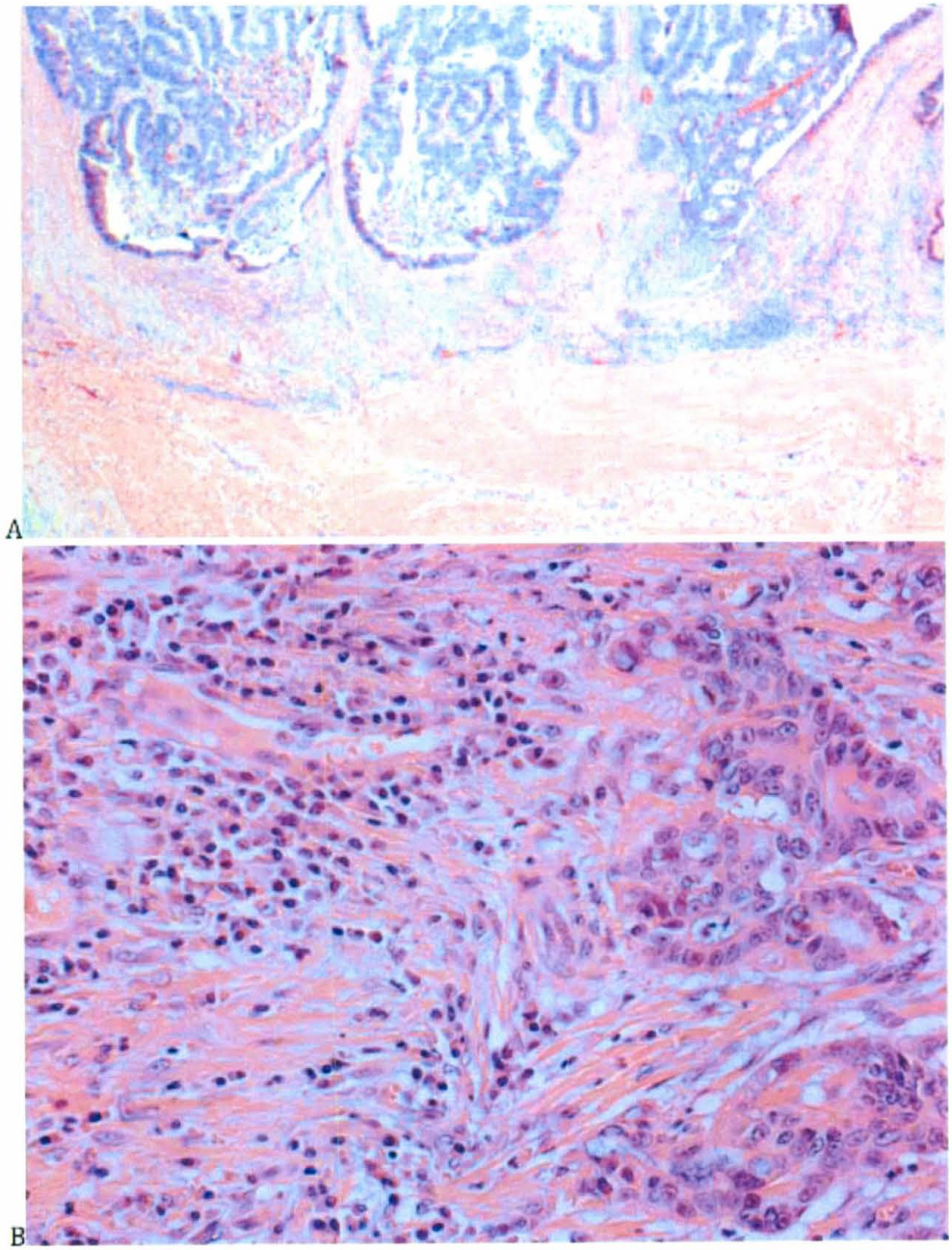


Figure 4.1. Novel features included in PREDICT model
(A), A low-power (4×) view of the advancing tumour edge illustrating the band-like formation of lymphocytes found in a peritumoural lymphocytic reaction. (B), A high-power (40×) view of tumour showing prominent plasma cell infiltrate.

Table 4.4. PREDICT for prediction of MSI-H colorectal cancers

Feature	Score
Right-sided tumour location	1.6
Mucinous component	1.6
Age at diagnosis <50 yrs	1.3
TILs	1.3
Peritumoural reaction	1.3
Increased stromal plasma cells	0.7
Maximum score:	7.8

Table 4.5. The utility of PREDICT at predicting MSI-H status in the development cohort, using different possible cut-off scores

Cut-off (≥)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
1.6	100	23.1	14.0	100
2.0	98.7	32.4	15.5	99.5
2.5	97.4	53.9	20.9	99.4
2.9	94.8	56.2	21.3	98.9
3.2	90.9	65.3	24.7	98.3
3.3	87.0	67.4	25.1	97.6
3.6	81.8	73.9	28.3	97.0
3.9	72.7	82.2	33.9	96.0

A simplified version of this model could be used in which the scores are not calculated, but instead the identification of any two or more of the features included in PREDICT could be sufficient to recommend MSI testing. This method was equivalent to applying a cut-off of 2.0, and had a sensitivity of 98.7% and a specificity of 32.4% in the development cohort.

Another pathologist (A.P.), uninvolved with the development of the model, reported the histology of CRC tumours in the validation cohort (N=276). To minimize

the differences between the development and validation cohorts we first assessed those patients in the validation cohort that were diagnosed younger than age 75 years (n=177). This corrected not only the age discrepancy between the cohorts, but also the sex discrepancy, resulting in a non-significant p value for difference in sex (data not shown). A multivariate logistic regression of the PREDICT features for these tumours is shown in Table 4.6a. The odds ratios increased compared to those from the development cohort for four of the features, and were equivalent for the presence of increased stromal plasma cells, and decreased slightly for peritumoural reaction.

We then assessed the validity of the model in the older portion of the validation cohort, tumours of patients diagnosed age 75 years or older (n=99). The odds ratios remained strong, with all values higher than those from the development cohort (Table 4.6b). This remained the case when the two age groups were combined (N=276), with only peritumoural reaction having an odds ratio slightly lower than that of the development cohort (Table 4.6c).

Using the threshold of 2.5, the sensitivity of PREDICT at predicting MSI-H tumours in the validation cohort was 96.9%, while the specificity was 76.6%. The area under the ROC curve (AUC) was 92.4% (Figure 4.2).

Table 4.7 compares the sensitivity, specificity, and PPV and NPV for all of the models described above, with and without the imposed age limits. The AUC is also given, for the models for which it is appropriate. PREDICT was the best test at predicting MSI-H tumours, with the highest combination of sensitivity and specificity, and the highest PPV, NPV, and AUC.

Table 4.6. Multivariate model for the prediction of MSI-H in the validation cohort

a) Including only tumours of patients diagnosed <75 years old (n=177)

Feature	β (Standard Error)	Odds Ratio (95% CI)	P value
Right-sided	1.868 (0.798)	6.477 (1.354 – 30.977)	0.019
Mucinous component	2.137 (0.755)	8.473 (1.929 – 37.212)	0.005
TILs	3.009 (0.794)	20.271 (4.272 – 96.193)	<0.001
Age < 50 years	1.768 (0.787)	5.861 (1.254 – 27.396)	0.025
Peritumoural reaction	0.570 (0.696)	1.769 (0.452 – 6.915)	0.412
Increased stromal plasma cells	0.738 (0.689)	2.091 (0.542 – 8.069)	0.284

b) Including only tumours of patients diagnosed ≥75 years old (n=99)

Feature	β (Standard Error)	Odds Ratio (95% CI)	P value
Right-sided	1.723 (0.918)	5.600 (0.926 – 33.853)	0.061
Mucinous component	3.500 (1.279)	33.110 (2.698 – 406.392)	0.006
TILs	1.716 (0.827)	5.562 (1.100 – 28.131)	0.038
Peritumoural reaction	2.031 (1.143)	7.621 (0.811 – 71.618)	0.076
Increased stromal plasma cells	1.019 (0.841)	2.769 (0.533 – 14.399)	0.226

Note: Age at diagnosis is not included in this model, because there were no patients in this cohort diagnosed younger than age 50 years.

c) Including all patients in the validation cohort (N=276)

Feature	β (Standard Error)	Odds Ratio (95% CI)	P value
Right-sided	1.673 (0.584)	5.328 (1.697 – 16.725)	0.004
Mucinous component	2.379 (0.583)	10.793 (3.440 – 33.864)	<0.001
TILs	2.409 (0.558)	11.125 (3.725 – 33.224)	<0.001
Age < 50 years	1.576 (0.697)	4.836 (1.233 – 18.968)	0.024
Peritumoural reaction	0.948 (0.533)	2.580 (0.908 – 7.330)	0.075
Increased stromal plasma cells	0.815 (0.510)	2.259 (0.831 – 6.141)	0.110

Table 4.7. The utility of different models at predicting MSI-H colorectal cancer

Model	Population	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	AUC (%)
RBG criteria 1245	<75 yrs	66.7	50.9	14.3	92.5	-
RBG criterion 3	<60 yrs	82.5	27.1	15.5	90.5	-
RBG criterion 3 (without age limit)	<75 yrs	87.2	28.2	13.0	94.7	-
Full RBG	<60 yrs	81.3	39.5	14.2	94.5	-
Full RBG (without age limit)	<75 yrs	97.3	13.1	12.1	97.6	-
MsPath	<60 yrs	92.1	37.8	18.9	96.8	77.6
MsPath (without age limit)	<75 yrs	91.5	38.9	14.9	97.5	80.4
PREDICT (development)	<75 yrs	97.4	53.9	20.9	99.4	86.9
PREDICT (validation)	All ages	96.9	76.6	35.2	99.5	92.4
Simplified PREDICT (development)	All ages	98.7	32.4	15.5	99.5	86.9
Simplified PREDICT (validation)	All ages	96.9	64.8	26.5	99.4	92.4

Note: MsPath was analyzed with the recommended cut-off of 1.0. PREDICT was analyzed with a cut-off of 2.5, with the exception of the "Simplified PREDICT", in which the cut-off was the identification of any two features included in the model.

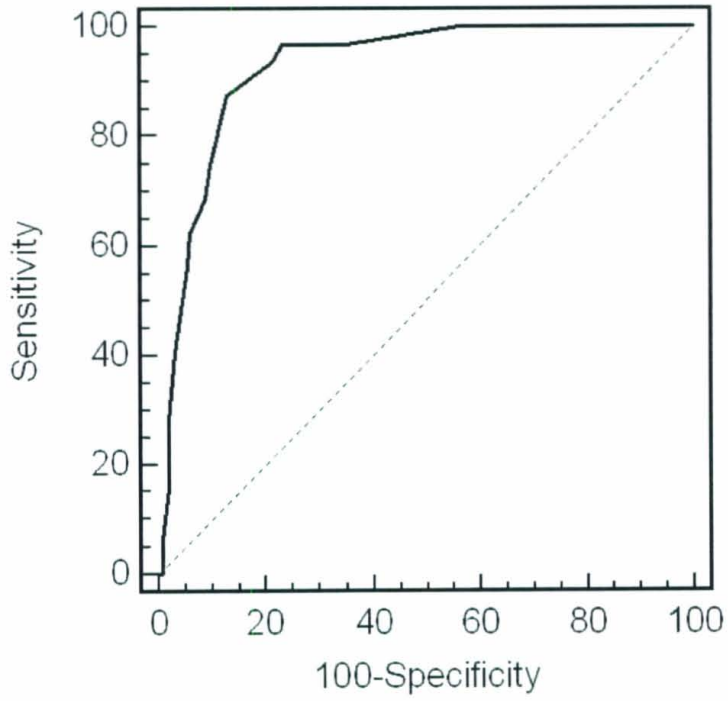


Figure 4.2. Receiver operating characteristics curve for the validation cohort, PREDICT Area under the curve, 92.4%.

4.3.6 Discussion

We assessed and compared existing models for the prediction of MSI-H colorectal tumours in a population-based registry, and developed and validated a new model. The use of the Revised Bethesda Guidelines is a poor predictive test for the identification of MSI-H colorectal tumours. The MsPath model is an improvement, but still has some limitations. Since MsPath was built upon the third RBG criterion, and the intention was to increase the specificity of the RBG at identifying cases of Lynch Syndrome, the only histologic features the authors considered for their model were those

included in the RBG, and they included the age limit imposed by the RBG [105]. This age limit was intended to increase the specificity of the RBG at identifying patients with Lynch Syndrome; however, it was not based on evidence, but rather by vote of the participating members of the panel [72]. We now know that MSI status is critical not only as a first step in identifying Lynch Syndrome, but also as a prognostic and predictive factor for CRC, with important clinical implications. For this reason we suggest that the age limit be removed. Our study has demonstrated that RBG and MsPath are each equally useful at predicting MSI-H tumours in both older and younger populations.

We have developed a model for the prediction of MSI-H CRCs that adds to the tumour histology features included in the RBG, and does not impose an age limit. While PREDICT is similar to MsPath, there are differences that improved the accuracy of the test in our cohorts; histological subtype is assessed quite differently. The MsPath model considered the predominant histological subtype, and categorized tumours into one of two groups: (1) conventional adenocarcinoma and (2) mucinous, signet ring or medullary histology. Using that method of classification, histological subtype was not a significant feature in predicting MSI-H status in our cohort with multivariate regression analysis. However, we found that the identification of mucinous histology, even if seen only focally, was a significant, independent predictive feature in our cohort. Others have reported similar findings[106].

Crohn-like lymphocytic reaction is another feature that was not statistically significant in discriminating MSI and MSS tumours in our cohort. A similar feature, peritumoural lymphocytic reaction, was predictive, and replaced Crohn-like reaction in PREDICT. Both of these features describe the presence of lymphocytes along the

tumour's leading edge. With Crohn-like reaction the lymphocytes form aggregates, whereas with peritumoural reaction they form bands parallel to the tumour's leading edge. In many cases both features were present in the same tumour.

Tumour grade is another feature that was not an independent predictor of MSI-H in our study, while it was in others. MsPath included a poorly-differentiated grade as being more likely associated with MSI-H tumours than a well- or moderately-differentiated grade. Greenson *et. al.* [106] assessed tumour grade differently. They considered well-differentiated tumours to also indicate an increased likelihood of MSI, and so considered well- and poorly-differentiated tumours together. We assessed grade using both of these methods, and found neither to be significant in our cohort (Table 4.2).

We also assessed two different approaches to including tumour-infiltrating lymphocytes (TILs). In the development cohort we identified 51 tumours with histologic heterogeneity in which TILS were only present in the non-preponderant histological type. If these "subclone-only" TILs were included, the specificity of this feature decreased by 8% to 86%, while the sensitivity remained virtually the same (data not shown), as compared to when the "subclone-only" TILs were not included. In our model we included the tumours with "subclone-only" TILS as having the feature, since standard methods for scoring TILs [105] would likely identify these. To our knowledge there have been no other published reports of these "subclone only" TILs. Further studies will be required to determine the significance of these TILs, and if they represent focal microsatellite instability that is not identified by standard methods, and if such foci are clinically relevant.

Lack of dirty necrosis was significantly associated with MSI-H tumours in univariate analysis, but was not significant in multivariate analysis. Greenson *et. al.* [106] found that lack of dirty necrosis remained a strong independent predictor of MSI-H CRCs in their study. They classified tumours as negative even if rare foci of necrosis were seen; however, we used a stricter definition, whereby the presence of any dirty necrosis resulted in a positive score. Evaluation of this feature was also limited by our study of a single representative slide, instead of including all tumour slides.

We assigned an arbitrary cut-off of 2.5 for PREDICT, meaning that any colorectal tumour that had enough features listed in the model to reach a cumulative score of 2.5 or higher be referred for MSI testing. This cut-off was selected to optimize both the sensitivity and specificity of the model for use as a screening test for MSI status. The best cut-off for the model should minimize the number of MSI-H tumours missed (false negatives), while reducing the proportion of tumours that would be identified for further MSI testing (potential false positives). In our population this cut-off resulted in a sensitivity of 97.4% and a specificity of 53.9% for identifying MSI-H colorectal tumours in the development cohort, and a sensitivity of 96.9% and specificity of 76.6% in the validation cohort.

This model could be further simplified or adapted for clinical application as required. For example, rather than assigning and combining scores for the features, note could just be made of which features from the model were present in a given tumour. Any tumour with two or more features would be referred for MSI testing. This simplification slightly increased the sensitivity of the model in our population, but decreased the specificity, as shown in Table 4.7. The implication of this in our

population would be that an extra 5% of all tumours would require further testing for MSI, but one more MSI-H case would be identified. Using the threshold of 2.5 for both the development and validation cohorts, 46.1% of tumours would require further testing for MSI, and three MSI-H tumours (2.8% of MSI-H tumours) would be missed. If the simplified method were used, 51.5% of tumours would be identified for further testing, and only 2 (1.8%) MSI-H tumours would be missed.

As stated, PREDICT with the proposed threshold of 2.5 missed three (2.8%) of 109 MSI-H tumours in our population. The one tumour missed in the validation cohort was from an 85 year old female patient, with an unknown family history. Her tumour attained a PREDICT score of 1.6, on the basis of mucin. The two tumours missed from the development cohort were from patients diagnosed in their late 60s. One female patient was diagnosed at age 69 years with a tumour deficient in both MLH1 and PMS2 by immunohistochemistry, with methylation of the *MLH1* promoter region, and the p.V600E mutation in the *BRAF* gene in DNA from the tumour (data not shown). This patient had a family history that met criteria 2 and 5 of the RBG. Her tumour attained a PREDICT score of 1.6 on the basis of a proximal tumour location.

The final tumour missed by PREDICT was that of a 67 year old male. His tumour attained a score of 2.0 on the basis of both a peritumoural lymphocytic reaction and increased plasma cells in the stroma. Since this tumour had two of the features of our model, it would not be missed by the simplified version given above. The tumour was deficient for PMS2 by immunohistochemistry. This patient had a history that met RBG criterion 2, which is a personal history of other Lynch-related tumours. In fact, this patient had a second, synchronous tumour that was also MSI-H. PREDICT identified this

other tumour, with a score of 2.9. So, although the model missed one of his tumours, this patient would still have been further assessed, as his other tumour would have gone on for MSI testing.

A limitation of this study is that we examined only one slide from each tumour, instead of all slides, representing the entire tumour. Possible tumour heterogeneity could have resulted in our missing certain features that were not present in the slide we examined, but were present in other areas, such as a component of mucin, histologic heterogeneity, or tumour infiltrating lymphocytes. This limitation may influence the threshold of 2.5 and should be considered in future study of this model.

Other limitations include the small size of our validation cohort, and our deliberate use of a different pathologist to score the tumours in this cohort. While the odds ratios were strong, indicating that they were highly likely to be predictive, the features were not all statistically significant in assessment of the validation cohort, probably due to inadequate power. The two new features in our model, peritumoural reaction and increased stromal plasma cells, were not statistically significant in the multivariate models of the validation cohort. While this was partially due to the small size of the cohort, and resulting lack of statistical power, it is possible that there were also differences in how these features were scored between the two cohorts. In the development cohort, 61.5% of tumours had a peritumoural reaction, compared to 31.5% in the validation cohort (data not shown). Similarly, 55.4% of tumours in the development cohort had increased stromal plasma cells, compared to 22.5% of tumours in the validation cohort. While there could be differences in the tumours of the two cohorts, the fact that they were scored by different pathologists may be a factor. The

reproducibility of scoring for these features needs to be explored further. Despite these limitations, PREDICT proved to be a valuable tool in predicting MSI status, with a high sensitivity and specificity, and an area under the ROC curve of 92.4% for the validation cohort.

Finally, we conclude from our comparison of the existing models that histological evaluation is superior to family history for identifying MSI-H CRCs. Identifying histological features in a tumour is an efficient and cost effective method compared to collecting and confirming a detailed family history, and is more sensitive and specific at identifying MSI-H tumours. Pathologic evaluation is also more useful in situations where an informative family history is not available. It would take little additional time for pathologists to provide scores for these features when conducting routine clinical evaluations of CRCs. Such scores would be valuable in identifying cases that require further testing of microsatellite status. This feature is an important prognostic and predictive feature in colorectal cancer, which is increasingly significant in guiding therapeutic options, and an important first step in identifying individuals with Lynch Syndrome.

4.3.7 Acknowledgements

We thank the staff of the Newfoundland Familial Colorectal Cancer Registry (Dr. Elizabeth Dicks, Project Coordinator) for ascertaining the patients, collecting family histories and biospecimens, Cathy Searle for her assistance with pathology specimens, and the colorectal cancer patients who participated for making this study possible.

4.4 Other Work

A second manuscript, included in this thesis as Appendix K, describes the use of the histology data collected in this project to help define and describe a new category of familial colorectal cancer.

Complete family history, pathological and molecular data was available for 552 of the tumours collected as part of the NFCCR project. A mutation of the *BRAF* gene, p.V600E, was identified in 12% of all tumours. This mutation has been associated with the serrated pathway of carcinogenesis [292]. It was found in 44% of all MSI-H tumours, but only 8% of MSS tumours. The *BRAF* mutation was associated with an increased risk of cancer in relatives, and also with a distinct clinical, molecular, and pathologic phenotype. Regardless of MSI status, patients with tumours harbouring the *BRAF* mutation were predominantly female, and were more likely to have proximally located tumours. Family members of these patients had a significantly increased risk of developing colorectal cancer. When compared to the first degree relatives (FDRs) of patients with MSS, *BRAF* wild-type tumours (considered the reference), FDRs of patients with MSI-H, *BRAF* mutant tumours had a hazard ratio (HR) of 2.49, with 95% confidence interval (CI) 1.57 to 3.93, and FDRs of patients with MSS, *BRAF* mutant tumours had a HR of 1.64, with 95% CI 1.01 to 2.66.

The MSI-H tumours with the *BRAF* mutation did not have germline mismatch repair mutations. MSI-H CRCs that do not harbour germline mismatch repair mutations are generally considered to be of a sporadic, and not familial nature. However, we

demonstrated a familial factor involved in approximately one half of these. FDRs of patients with the *BRAF* mutation were also at a significantly higher risk of developing nonmelanoma skin cancer (HR of 2.52, 95% CI 1.31 to 4.86). Furthermore, tumours with the *BRAF* mutation were associated with a poorly differentiated histology, the presence of mucin, increased stromal plasma cells, and tumour-infiltrating lymphocytes.

My role in this project was in providing the data from the histology review, which was used to help define the phenotype of this new category of familial colorectal cancer. This work was published in *Cancer Epidemiology, Biomarkers and Prevention* in 2010 [481].

Chapter 5 – Concluding Remarks

The main goal of this PhD programme was to characterize colorectal cancer in the province of Newfoundland and Labrador, and through this characterization to identify clinically relevant differences in colorectal tumours that could be used to enhance and personalize clinical care. Many different methods were used to gather and then expand upon the knowledge base of the inner workings of CRC in this province. Three different approaches were taken to the characterization of the disease, as described in the previous three chapters, and summarized below. The first two major projects presented here were preliminary studies, setting the groundwork for a larger study to follow. The final project was itself a part of that later, larger study.

The first project took a molecular approach, whereby the function of the DNA mismatch repair pathway in tumours was assessed through microsatellite instability analysis and immunohistochemistry of MMR proteins. This work was correlated with family history information to reveal that while the rate of mismatch repair deficiency is approximately as expected in the province, the rate of familial CRC is much higher, and does not correlate strongly with the molecular data. This work was published in the journal *Clinical Cancer Research* [797]. The results were later replicated and validated by the larger, CIHR-IHRT project that followed [466,469]. The results suggest that there are familiar forms of CRC in this province that contribute greatly to the burden of disease here, yet are distinct from the most common hereditary CRC syndromes. Familial colorectal cancer that meets the Amsterdam Criteria for family risk, but does not have the molecular features of mismatch repair deficiency is currently defined as Familial

Colorectal Cancer Type X. Multiple aetiologies are suspected, but none have yet been elucidated. This work led to a genome-wide scan of multiple CRC families in a search for novel loci associated with the disease. The search was not fruitful, and so was not included in this thesis; however, the search for the cause of familial CRC in these families was continued by others, and is ongoing.

The second project characterized CRC in the province according to protein expression. Expression of a number of proteins biologically relevant to CRC was assessed using tissue microarray technology and immunohistochemistry. Some proteins displayed prognostic utility in our population, including lamin A/C, which showed great potential for use as a prognostic indicator for patients with stage III CRC, and may also have a role in guiding adjuvant therapy in this cohort.

Through unsupervised hierarchical clustering analysis of the protein expression data, cluster groups of tumours with similar expression profiles were identified. These cluster groups of CRC were associated with different patient outcomes. This study has illustrated the clinical potential for this type of work, in the identification of clinically relevant subgroups of cancer. While a larger follow-up study has yet to occur, the preliminary study presented here has provided many interesting avenues along which research can continue.

While the follow-up TMA study has not yet begun, another component of this project has already led to multiple spin-off projects. In order to assess the prognostic and predictive utilities of the various proteins and protein expression profiles, detailed and comprehensive demographic, pathological, clinical, and follow-up information on each patient had to be collected. This resource of patient data has been anonymized and

shared with other researchers, and is currently the basis of six active research projects approved by the local ethics board. These projects include assessment of a number of genotypes with CRC prognosis and drug response, as well as the relationship of vitamin D levels and CRC, and genes related to anxiety and depression and their association with outcome in CRC.

The third project presented in this thesis characterized the histological heterogeneity of CRC, and assessed the potential clinical value of histological differences between tumours. This led to the development and validation of a new algorithm for identifying MSI-H tumours for the purpose of guiding patient prognosis and adjuvant therapy, and as a first step to in the identification of patients with Lynch Syndrome. In our population, this new algorithm, "PREDICT", has a higher sensitivity and specificity for identifying MSI-H tumours than the existing algorithms: the revised Bethesda Guidelines and MsPath. This work was published in the *American Journal of Surgical Pathology* [836]. This new algorithm has already been put into clinical practice at the hospitals affiliated with Grey Bruce Health Services in Ontario, and will soon be implemented in Newfoundland and Labrador as part of an upcoming Community Genetics programme.

Again, the work done as part of this project has also led to numerous other spin-off projects. The histological data supported the sub-classification of CRCs for the identification of a new form of familial CRC associated with a mutation in the *BRAF* gene and the serrated pathway of carcinogenesis. The existence of this new form of familial CRC was later supported by the results of genetic linkage analysis performed by another research group on patients from Australia, and has been named Jass Syndrome

[397]. Our manuscript was published in the journal *Cancer, Epidemiology, Biomarkers and Prevention* [481], and is included as Appendix K. As well, the data is being used as part of a geographical assessment of CRC in the province.

The original CIHR-IHRT grant and interdisciplinary collaboration with Ontario, (of which my programme became a part) was approved for a second phase in 2005, which allowed for continuation and expansion of existing projects, as well as the addition of new ones. A new prospective population-based Community Genetics initiative has also begun, in which all patients diagnosed with CRC in Newfoundland and Labrador are invited to fill out a family history form and other questionnaires. Clinical team members review these forms and results in an effort to identify families at risk of hereditary cancer syndromes, and subsequently organize appropriate screening for family members at risk of developing cancer, and ultimately prevent some of these cancers from occurring. My PREDICT algorithm will be included in a new phase of this Community Genetics initiative, to predict MSI status of tumours, to assist in the identification of families with Lynch Syndrome and help guide patient prognosis and therapy.

The field of colorectal cancer research has evolved greatly over the course of my PhD programme. Research techniques have become less cumbersome and more high-throughput. PCR and other reactions that I initially ran in individual tubes, I later ran on 96 well plates. MSI analysis that I ran manually and electrophoresed on polyacrymamide gels are now run on automated sequencing machines. While I started with immunohistochemistry of whole tissue sections, I later increased throughput with the use of tissue microarray technology. Similarly, this research project has grown during the same time. What started as a small assessment of colorectal cancer on the Avalon

Peninsula of Newfoundland grew into a large, inter-provincial and interdisciplinary collaboration, and later into an association with an even larger, international colorectal cancer consortium. With the increasingly high-throughput capacity of newer technologies, the growing number of available patients and tumours through collaborations, and the growing foundation of knowledge on which to build, the future of colorectal cancer research is bright. We should soon be better able to predict those at risk of CRC and prevent it. We should soon have better markers to predict patient outcome and response to treatment, and be better able to individualize and optimize therapies.

A lot of further work is needed to achieve these goals, but this will be work for other people. It has been my pleasure to be a part of this research, and to have learned from the researchers and clinicians around me. While this thesis is ending, the research does not.

References

1. Parkin DM. Global cancer statistics in the year 2000. Lancet Oncol 2001;2:533-543.

2. Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. CA Cancer J Clin 2005;55:74-108.

3. Lynch HT, de la Chapelle A. Hereditary colorectal cancer. N Engl J Med 2003;348:919-932.

4. de la Chapelle A. The incidence of Lynch syndrome. Fam Cancer 2005;4:233-237.

5. Rustgi AK. The genetics of hereditary colon cancer. Genes Dev 2007;21:2525-2538.

6. Hampel H, Frankel WL, Martin E, et al. Screening for the Lynch syndrome (hereditary nonpolyposis colorectal cancer). N Engl J Med 2005;352:1851-1860.

7. Warthin A. Heredity with Reference to Carcinoma: As Shown by the Study of the Cases Examined in the Pathological Laboratory of the University of Michigan, 1895-1913. Arch Intern Med 1913:XII:546-555.

8. Levin I. The influence of heredity on cancer. J Cancer Res Clin Oncol 1912;11:547-558.

9. Lynch HT. Aldred Scott Warthin, M.D., Ph.D. (1866-1931). CA Cancer J Clin 1985;35:345-347.

10. Warthin A. The Further Study of a Cancer Family. J Cancer Res 1925;9:279-286.

11. Lynch HT, Lynch JF. Lynch syndrome: History and current status. Dis Markers 2004;20:181-198.

12. Lynch HT, Shaw MW, Magnuson CW, Larsen AL, Krush AJ. Hereditary Factors in Cancer: Study of Two Large Midwestern Kindreds. Arch Intern Med 1966;117:206-212.

13. Lynch HT, Krush AJ, Larsen AL. Heredity and multiple primary malignant neoplasms: six cancer families. Am J Med Sci 1967;254:322-9.

14. Lynch HT, Krush AJ. The cancer family syndrome and cancer control. Surg Gynecol Obstet 1971;132:247-250.

15. Lynch HT, Krush AJ. Cancer family "G" revisited: 1895-1970. Cancer 1971;27:1505-1511.

16. Yan H, Papadopoulos N, Marra G, et al. Conversion of diploidy to haploidy. Nature 2000;403:723-724.

17. Douglas JA, Gruber SB, Meister KA, et al. History and Molecular Genetics of Lynch Syndrome in Family G: A Century Later. JAMA 2005;294:2195-2202.

18. Peltomaki P, Aaltonen LA, Sistonen P, et al. Genetic mapping of a locus predisposing to human colorectal cancer. Science 1993;260:810-812.

19. Cotran RS, Kumar V, Collins T. Robbins Pathologic Basis of Disease. U.S.A.: W.B. Saunders Company, 1999.

20. Ross M, Romrell LJ, Kaye GI. Histology: A Text and Atlas. Maryland, USA: Williams & Wilkins, 1995.

21. Canadian Cancer Society's Steering Committee on Cancer Statistics. Canadian Cancer Statistics 2011. 2011;.

22. American Joint Committee on Cancer (AJCC). AJCC Cancer Staging Manual. Springer, 2002.

23. Zeki SS, Graham TA, Wright NA. Stem cells and their implications for colorectal cancer. Nat Rev Gastroenterol Hepatol 2011;8:90-100.

24. Stead L, Stead M, Kaufman M, Sotsky-Kent T. First Aid for the Surgery Clerkship: A Student to Student Guide. McGraw-Hill Companies, Inc, 2003.

25. Hamilton SR, Aaltonen LA, eds. Pathology & Genetics: Tumours of the Digestive Tract. Lyon, France: International Agency for Research on Cancer, World Health Organization, 2000.

26. Chang AE, Karnell LH, Menck HR. The National Cancer Data Base report on cutaneous and noncutaneous melanoma. Cancer 1998;83:1664-1678.

27. Brady MS, Kavolius JP, Quan SH. Anorectal melanoma. A 64-year experience at Memorial Sloan-Kettering Cancer Center. Dis Colon Rectum 1995;38:146-151.

28. Théraux J, Bretagnol F, Guedj N, Cazals-Hatem D, Panis Y. Colorectal breast carcinoma metastasis diagnosed as an obstructive colonic primary tumor. A case report and review of the literature. Gastroentérologie clinique et biologique 2009;33:1114-7.

29. Lockhart-Mummery JP. Two hundred cases of cancer of the rectum treated by perineal excision. Br J Surg 1926;14:110-124.

30. Gordon-Watson C, Dukes C. The Treatment of Carcinoma of the Rectum with Radium. Br J Surg 1930;17:643-669.

31. Dukes CE. The Classification of Cancer of the Rectum. J Pathol Bacteriol 1932;35:323-332.

32. Kirklin JW, Dockerty MB, Waugh JM. The role of the peritoneal reflection in the prognosis of carcinoma of the rectum and sigmoid colon. Surg Gynecol Obstet 1949;88:326-331.

33. Astler VB, Coller FA. The prognostic significance of direct extension of carcinoma of the colon and rectum. Ann Surg 1954;139:846-852.

34. Dukes CE, Bussey HJ. The spread of rectal cancer and its effect on prognosis. Br J Cancer 1958;12:309-20.

35. Turnbull RB,Jr, Kyle K, Watson FR, Spratt J. Cancer of the colon: the influence of the no-touch isolation technic on survival rates. Ann Surg 1967;166:420-427.

36. Jass JR, Morson BC. Reporting colorectal cancer. J Clin Pathol 1987;40:1016-1023.

37. International Union Against Cancer (UICC). TNM Classification of Malignant Tumours. John Wiley & Sons, 2010.

38. Doll R. The Pierre Denoix Memorial Lecture: Nature and nurture in the control of cancer. Eur J Cancer 1999;35:16-23.

39. Denoix P. Nomenclature Classification des Cancers. Bull Inst Nat Hyg 1950;5:81-84.

40. Denoix P. Nomenclature Classification des Cancers. Bull Inst Nat Hyg 1952;7:743-748.

41. American Joint Committee for Cancer Staging and End Results Reporting. Manual for Staging Cancer. 1977.

42. American Joint Committee on Cancer (AJCC). AJCC Cancer Staging Manual. Springer, 2010.

43. American Joint Committee on Cancer (AJCC). AJCC Cancer Staging Handbook. Springer, 2010.

44. Washington K, Berlin J, Branton P, et al. Protocol for the Examination of Specimens from Patients with Primary Carcinoma of the Colon and Rectum. College of American Pathologists; 2009. Available at: www.cap.org.

45. American Joint Committee on Cancer (AJCC). AJCC Cancer Staging Atlas. Springer, 2006.

46. Ratto C, Sofo L, Ippoliti M, et al. Accurate lymph-node detection in colorectal specimens resected for cancer is of prognostic significance. Dis Colon Rectum 1999;42:143-54; discussion 154-8.

47. Pocard M, Panis Y, Malassagne B, Nemeth J, Hautefeuille P, Valleur P. Assessing the effectiveness of mesorectal excision in rectal cancer: prognostic value of the number of lymph nodes found in resected specimens. Dis Colon Rectum 1998;41:839-845.

48. Tepper JE, O'Connell MJ, Niedzwiecki D, et al. Impact of number of nodes retrieved on outcome in patients with rectal cancer. J Clin Oncol 2001;19:157-163.

49. George S, Primrose J, Talbot R, et al. Will Rogers revisited: prospective observational study of survival of 3592 patients with colorectal cancer according to number of nodes examined by pathologists. Br J Cancer 2006;95:841-847.

50. Scott KW, Grace RH. Detection of lymph node metastases in colorectal carcinoma before and after fat clearance. Br J Surg 1989;76:1165-1167.

51. Gold P, Freedman SO. Specific carcinoembryonic antigens of the human digestive system. J Exp Med 1965;122:467-481.

52. Thomas P, Toth CA, Saini KS, Jessup JM, Steele G. The structure, metabolism and function of the carcinoembryonic antigen gene family. Biochim Biophys Acta 1990;1032:177-89.

53. Thompson JA. Molecular cloning and expression of carcinoembryonic antigen gene family members. Tumour Biol 1995;16:10-16.

54. Paschos KA, Canovas D, Bird NC. The role of cell adhesion molecules in the progression of colorectal cancer and the development of liver metastasis. Cell Signal 2009;21:665-674.

55. Khoo SK, Mackay IR. Carcinoembryonic antigen in serum in diseases of the liver and pancreas. J Clin Pathol 1973;26:470-475.

56. Lurie BB, M SL, Zamcheck N. Elevated carcinoembryonic antigen levels and biliary tract obstruction. JAMA : the journal of the American Medical Association 1975;233:326-30.

57. Sajid KM, Parveen R, Durr-e-Sabih, Chaouachi K, Naeem A, Mahmood R. Carcinoembryonic antigen (CEA) levels in hookah smokers, cigarette smokers and non-smokers. JPMA The Journal of the Pakistan Medical Association 2007;57:595-9.

58. Vider M, Kashmiri R, Hunter L, Moses B, Meeker WR, Utley JF. Carcinoembryonic antigen (CEA) monitoring in the management of radiotherapeutic patients. Oncology 1974;30:257-72.

59. Goldstein NS, Turner JR. Pericolonic tumor deposits in patients with T3N+MO colon adenocarcinomas: markers of reduced disease free survival and intra-abdominal metastases and their implications for TNM classification. Cancer 2000;88:2228-2238.

60. Aaltonen LA, Peltomaki P, Leach FS, et al. Clues to the pathogenesis of familial colorectal cancer. Science 1993;260:812-816.

61. Boland CR, Thibodeau SN, Hamilton SR, 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;58:5248-5257.

62. Jover R, Zapater P, Castells A, et al. The efficacy of adjuvant chemotherapy with 5-fluorouracil in colorectal cancer depends on the mismatch repair status. Eur J Cancer 2009;45:365-373.

63. Guidoboni M, Gafa R, Viel A, et al. Microsatellite instability and high content of activated cytotoxic lymphocytes identify colon cancer patients with a favorable prognosis. Am J Pathol 2001;159:297-304.

64. Lukish JR, Muro K, DeNobile J, et al. Prognostic significance of DNA replication errors in young patients with colorectal cancer. Ann Surg 1998;227:51-56.

65. Gryfe R, Kim H, Hsieh ET, et al. Tumor microsatellite instability and clinical outcome in young patients with colorectal cancer. N Engl J Med 2000;342:69-77.

66. Samowitz WS, Curtin K, Ma KN, et al. Microsatellite instability in sporadic colon cancer is associated with an improved prognosis at the population level. Cancer Epidemiol Biomarkers Prev 2001;10:917-923.

67. Popat S, Hubner R, Houlston RS. Systematic review of microsatellite instability and colorectal cancer prognosis. J Clin Oncol 2005;23:609-618.

68. Chang EY, Dorsey PB, Frankhouse J, et al. Combination of microsatellite instability and lymphocytic infiltrate as a prognostic indicator in colon cancer. Arch Surg 2009;144:511-515.

69. Kumar S, Chang EY, Frankhouse J, Dorsey PB, Lee RG, Johnson N. Combination of Microsatellite Instability and Lymphocytic Infiltrate as a Prognostic Indicator for Adjuvant Therapy in Colon Cancer. Arch Surg 2009;144:835-840.

70. Carethers JM, Chauhan DP, Fink D, et al. Mismatch repair proficiency and in vitro response to 5-fluorouracil. Gastroenterology 1999;117:123-131.

71. Tajima A, Hess MT, Cabrera BL, Kolodner RD, Carethers JM. The mismatch repair complex hMutS alpha recognizes 5-fluorouracil-modified DNA: implications for chemosensitivity and resistance. Gastroenterology 2004;127:1678-1684.

72. Umar A, Boland CR, Terdiman JP, et al. Revised Bethesda Guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. J Natl Cancer Inst 2004;96:261-268.

73. Minsky BD, Mies C, Rich TA, Recht A. Lymphatic vessel invasion is an independent prognostic factor for survival in colorectal cancer. Int J Radiat Oncol Biol Phys 1989;17:311-318.

74. Talbot IC, Ritchie S, Leighton MH, Hughes AO, Bussey HJ, Morson BC. The clinical significance of invasion of veins by rectal cancer. Br J Surg 1980;67:439-442.

75. Talbot IC, Ritchie S, Leighton MH, Hughes AO, Bussey HJ, Morson BC. Spread of rectal cancer within veins. Histologic features and clinical significance. Am J Surg 1981;141:15-17.

76. Amado RG, Wolf M, Peeters M, et al. Wild-type KRAS is required for panitumumab efficacy in patients with metastatic colorectal cancer. J Clin Oncol 2008;26:1626-1634.

77. De Roock W, Piessevaux H, De Schutter J, et al. KRAS wild-type state predicts survival and is associated to early radiological response in metastatic colorectal cancer treated with cetuximab. Ann Oncol 2008;19:508-515.

78. MacCarty WC. Principles of Prognosis in Cancer
. J Am Med Assoc 1931;96:30-33.

79. Compton CC, Fielding LP, Burgart LJ, et al. Prognostic factors in colorectal cancer. College of American Pathologists Consensus Statement 1999. Arch Pathol Lab Med 2000;124:979-994.

80. Ueno H, Murphy J, Jass JR, Mochizuki H, Talbot IC. Tumour 'budding' as an index to estimate the potential of aggressiveness in rectal cancer. Histopathology 2002;40:127-132.

81. Wassermann S, Scheel SK, Hiendlmeyer E, et al. p16INK4a is a beta-catenin target gene and indicates low survival in human colorectal tumors. Gastroenterology 2009;136:196-205.e2.

82. Prall F, Ostwald C. High-degree tumor budding and podia-formation in sporadic colorectal carcinomas with K-ras gene mutations. Hum Pathol 2007;38:1696-1702.

83. Zlobec I, Lugli A, Baker K, et al. Role of APAF-1, E-cadherin and peritumoral lymphocytic infiltration in tumour budding in colorectal cancer. J Pathol 2007;212:260-268.

84. Zlobec I, Lugli A. Prognostic and predictive factors in colorectal cancer. J Clin Pathol 2008;61:561-569.

85. Benjamini E, Coico R, Sunshine G. Immunology A Short Course. UNITED STATES: John Wiley & Sons, Inc, 2000.

86. Linnebacher M, Gebert J, Rudy W, et al. Frameshift peptide-derived T-cell epitopes: a source of novel tumor-specific antigens. Int J Cancer 2001;93:6-11.

87. Saeterdal I, Gjertsen MK, Straten P, Eriksen JA, Gaudernack G. A TGF betaRII frameshift-mutation-derived CTL epitope recognised by HLA-A2-restricted CD8+ T cells. Cancer Immunol Immunother 2001;50:469-476.

88. Saeterdal I, Bjorheim J, Lislerud K, et al. Frameshift-mutation-derived peptides as tumor-specific antigens in inherited and spontaneous colorectal cancer. Proc Natl Acad Sci U S A 2001;98:13255-13260.

89. Ishikawa T, Fujita T, Suzuki Y, et al. Tumor-specific immunological recognition of frameshift-mutated peptides in colon cancer with microsatellite instability. Cancer Res 2003;63:5564-5572.

90. Ripberger E, Linnebacher M, Schwitalle Y, Gebert J, von Knebel Doeberitz M. Identification of an HLA-A0201-restricted CTL epitope generated by a tumor-specific frameshift mutation in a coding microsatellite of the OGT gene. J Clin Immunol 2003;23:415-423.

91. Schwitalle Y, Linnebacher M, Ripberger E, Gebert J, von Knebel Doeberitz M. Immunogenic peptides generated by frameshift mutations in DNA mismatch repair-deficient cancer cells. Cancer Immun 2004;4:14.

92. Schwitalle Y, Kloor M, Eiermann S, et al. Immune response against frameshift-induced neopeptides in HNPCC patients and healthy HNPCC mutation carriers. Gastroenterology 2008;134:988-997.

93. Tougeron D, Fauquembergue E, Rouquette A, et al. Tumor-infiltrating lymphocytes in colorectal cancers with microsatellite instability are correlated with the number and spectrum of frameshift mutations. Mod Pathol 2009;22:1186-1195.

94. Linnebacher M, Wienck A, Boeck I, Klar E. Identification of an MSI-H tumor-specific cytotoxic T cell epitope generated by the (-1) frame of U79260(FTO). Journal of biomedicine 2010;2010:.

95. Banerjea A, Bustin SA, Dorudi S. The immunogenicity of colorectal cancers with high-degree microsatellite instability. World J Surg Oncol 2005;3:26.

96. Svennevig JL, Lunde OC, Holter J, Bjorgsvik D. Lymphoid infiltration and prognosis in colorectal carcinoma. Br J Cancer 1984;49:375-377.

97. Jass JR. Lymphocytic infiltration and survival in rectal cancer. J Clin Pathol 1986;39:585-589.

98. Ropponen KM, Eskelinen MJ, Lipponen PK, Alhava E, Kosma VM. Prognostic value of tumour-infiltrating lymphocytes (TILs) in colorectal cancer. J Pathol 1997;182:318-24.

99. Galon J, Costes A, Sanchez-Cabo F, et al. Type, Density, and Location of Immune Cells Within Human Colorectal Tumors Predict Clinical Outcome. Science 2006;313:1960-1964.

100. Ohtani H. Focus on TILs: prognostic significance of tumor infiltrating lymphocytes in human colorectal cancer. Cancer Immun 2007;7:4.

101. Ogino S, Nosho K, Irahara N, et al. Lymphocytic Reaction to Colorectal Cancer Is Associated with Longer Survival, Independent of Lymph Node Count, Microsatellite Instability, and CpG Island Methylator Phenotype. Clinical Cancer Research 2009;15:6412-6420.

102. Graham DM, Appelman HD. Crohn's-like lymphoid reaction and colorectal carcinoma: a potential histologic prognosticator. Mod Pathol 1990;3:332-335.

103. Harrison JC, Dean PJ, el-Zeky F, Vander Zwaag R. Impact of the Crohn's-like lymphoid reaction on staging of right-sided colon cancer: results of multivariate analysis. Hum Pathol 1995;26:31-38.

104. Baker K, Zlobec I, Tornillo L, Terracciano L, Jass JR, Lugli A. Differential significance of tumour infiltrating lymphocytes in sporadic mismatch repair deficient versus proficient colorectal cancers: a potential role for dysregulation of the transforming growth factor-beta pathway. Eur J Cancer 2007;43:624-631.

105. Jenkins MA, Hayashi S, O'Shea AM, et al. Pathology Features in Bethesda Guidelines Predict Colorectal Cancer Microsatellite Instability: A Population-Based Study. Gastroenterology 2007;133:48-56.

106. Greenson JK, Huang SC, Herron C, et al. Pathologic predictors of microsatellite instability in colorectal cancer. Am J Surg Pathol 2009;33:126-133.

107. Nielsen HJ, Hansen U, Christensen IJ, Reimert CM, BrÃ¼nner N, Moesgaard F. Independent prognostic value of eosinophil and mast cell infiltration in colorectal cancer tissue. J Pathol 1999;189:487-95.

108. Wijnen JT, Vasen HF, Khan PM, et al. Clinical findings with implications for genetic testing in families with clustering of colorectal cancer. N Engl J Med 1998;339:511-518.

109. Barnetson RA, Tenesa A, Farrington SM, et al. Identification and survival of carriers of mutations in DNA mismatch-repair genes in colon cancer. N Engl J Med 2006;354:2751-2763.

110. Balmana J, Stockwell DH, Steyerberg EW, et al. Prediction of MLH1 and MSH2 mutations in Lynch syndrome. JAMA 2006;296:1469-1478.

111. Alon U, Barkai N, Notterman DA, et al. Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by oligonucleotide arrays. Proc Natl Acad Sci U S A 1999;96:6745-6750.

112. Lascorz J, Chen B, Hemminki K, Asta Försti. Consensus pathways implicated in prognosis of colorectal cancer identified through systematic enrichment analysis of gene expression profiling studies. PloS one 2011;6:.

113. Cardoso J, Boer J, Morreau H, Fodde R. Expression and genomic profiling of colorectal cancer. Biochim Biophys Acta 2007;1775:103-137.

114. Wolin KY, Tuchman H. Physical activity and gastrointestinal cancer prevention. Recent Results Cancer Res 2011;186:73-100.

115. Campbell PT, Cotterchio M, Dicks E, Parfrey P, Gallinger S, McLaughlin JR. Excess Body Weight and Colorectal Cancer Risk in Canada: Associations in Subgroups of Clinically Defined Familial Risk of Cancer. Cancer Epidemiology Biomarkers & Prevention 2007;16:1735-1744.

116. Campbell PT, Elizabeth TJ, Cornelia MU, Jane CF, Jenny NP, John RM. Case-control study of overweight, obesity, and colorectal cancer risk, overall and by tumor microsatellite instability status. J Natl Cancer Inst 2010;102:391-400.

117. Burkitt DP. Related disease--related cause? Lancet 1969;2:1229-1231.

118. Kritchevsky D. Epidemiology of fibre, resistant starch and colorectal cancer. Eur J Cancer Prev 1995;4:345-352.

119. Trock B, Lanza E, Greenwald P. Dietary fiber, vegetables, and colon cancer: critical review and meta-analyses of the epidemiologic evidence. J Natl Cancer Inst 1990;82:650-661.

120. Howe GR, Benito E, Castelleto R, et al. Dietary intake of fiber and decreased risk of cancers of the colon and rectum: evidence from the combined analysis of 13 case-control studies. J Natl Cancer Inst 1992;84:1887-1896.

121. Levi F, Pasche C, Lucchini F, La Vecchia C. Dietary fibre and the risk of colorectal cancer. Eur J Cancer 2001;37:2091-2096.

122. Bingham SA, Day NE, Luben R, et al. Dietary fibre in food and protection against colorectal cancer in the European Prospective Investigation into Cancer and Nutrition (EPIC): an observational study. Lancet 2003;361:1496-1501.

123. Park Y, Hunter DJ, Spiegelman D, et al. Dietary fiber intake and risk of colorectal cancer: a pooled analysis of prospective cohort studies. JAMA 2005;294:2849-2857.

124. Koushik A, Hunter DJ, Spiegelman D, et al. Fruits, vegetables, and colon cancer risk in a pooled analysis of 14 cohort studies. J Natl Cancer Inst 2007;99:1471-1483.

125. Chan AT, Giovannucci EL. Primary Prevention of Colorectal Cancer. Gastroenterology 2010;138:2029-2043.e10.

126. Giovannucci E. Insulin and colon cancer. Cancer Causes Control 1995;6:164-179.

127. Larsson SC, Orsini N, Wolk A. Diabetes mellitus and risk of colorectal cancer: a meta-analysis. J Natl Cancer Inst 2005;97:1679-1687.

128. Vinikoor LC, Long MD, Keku TO, Martin CF, Galanko JA, Sandler RS. The association between diabetes, insulin use, and colorectal cancer among Whites and African Americans. Cancer Epidemiol Biomarkers Prev 2009;18:1239-1242.

129. He J, Stram DO, Kolonel LN, Henderson BE, Le Marchand L, Haiman CA. The association of diabetes with colorectal cancer risk: the Multiethnic Cohort. Br J Cancer 2010;103:120-126.

130. Wilkinson JE, Culpepper L. Associations between colorectal cancer screening and glycemic control in people with diabetes, Boston, Massachusetts, 2005-2010. Preventing Chronic Disease 2011;8:.

131. Giovannucci E, Rimm EB, Ascherio A, Stampfer MJ, Colditz GA, Willett WC. Alcohol, low-methionine--low-folate diets, and risk of colon cancer in men. J Natl Cancer Inst 1995;87:265-273.

132. Giovannucci E, Stampfer MJ, Colditz GA, et al. Folate, methionine, and alcohol intake and risk of colorectal adenoma. J Natl Cancer Inst 1993;85:875-884.

133. Choi SW, Stickel F, Baik HW, Kim YI, Seitz HK, Mason JB. Chronic alcohol consumption induces genomic but not p53-specific DNA hypomethylation in rat colon. J Nutr 1999;129:1945-1950.

134. Poynter JN, Haile RW, Siegmund KD, et al. Associations between Smoking, Alcohol Consumption, and Colorectal Cancer, Overall and by Tumor Microsatellite Instability Status. Cancer Epidemiology Biomarkers & Prevention 2009;18:2745-2750.

135. Samowitz WS, Albertsen H, Sweeney C, et al. Association of smoking, CpG island methylator phenotype, and V600E BRAF mutations in colon cancer. J Natl Cancer Inst 2006;98:1731-1738.

136. Curtin K, Samowitz WS, Wolff RK, et al. Assessing tumor mutations to gain insight into base excision repair sequence polymorphisms and smoking in colon cancer. Cancer Epidemiol Biomarkers Prev 2009;18:3384-3388.

137. Liang PS, Chen TY, Giovannucci E. Cigarette smoking and colorectal cancer incidence and mortality: systematic review and meta-analysis. Int J Cancer 2009;124:2406-2415.

138. Cole BF, Logan RF, Halabi S, et al. Aspirin for the chemoprevention of colorectal adenomas: meta-analysis of the randomized trials. J Natl Cancer Inst 2009;101:256-266.

139. Hayden M, Pignone M, Phillips C, Mulrow C. Aspirin for the Primary Prevention of Cardiovascular Events: A Summary of the Evidence for the U.S. Preventive Services Task Force. Annals of Internal Medicine 2002;136:161-172.

140. Dube C, Rostom A, Lewin G, et al. The use of aspirin for primary prevention of colorectal cancer: a systematic review prepared for the U.S. Preventive Services Task Force. Ann Intern Med 2007;146:365-375.

141. Cuzick J, Otto F, Baron JA, et al. Aspirin and non-steroidal anti-inflammatory drugs for cancer prevention: an international consensus statement. Lancet Oncol 2009;10:501-507.

142. Bertagnolli MM, Eagle CJ, Zauber AG, et al. Celecoxib for the prevention of sporadic colorectal adenomas. N Engl J Med 2006;355:873-884.

143. Zell JA, Pelot D, Chen WP, McLaren CE, Gerner EW, Meyskens FL. Risk of cardiovascular events in a randomized placebo-controlled, double-blind trial of difluoromethylornithine plus sulindac for the prevention of sporadic colorectal adenomas. Cancer Prev Res (Phila) 2009;2:209-212.

144. Eberhart CE, Coffey RJ, Radhika A, Giardiello FM, Ferrenbach S, DuBois RN. Up-regulation of cyclooxygenase 2 gene expression in human colorectal adenomas and adenocarcinomas. Gastroenterology 1994;107:1183-1188.

145. Luciani MG, Campregher C, Fortune JM, Kunkel TA, Gasche C. 5-ASA affects cell cycle progression in colorectal cells by reversibly activating a replication checkpoint. Gastroenterology 2007;132:221-235.

146. Palmer RH. Bile acid heterogeneity and the gastrointestinal epithelium: from diarrhea to colon cancer. J Lab Clin Med 1979;94:655-660.

147. Newmark HL, Wargovich MJ, Bruce WR. Colon cancer and dietary fat, phosphate, and calcium: a hypothesis. J Natl Cancer Inst 1984;72:1323-1325.

148. Fedirko V, Bostick RM, Flanders WD, et al. Effects of vitamin d and calcium on proliferation and differentiation in normal colon mucosa: a randomized clinical trial. Cancer Epidemiol Biomarkers Prev 2009;18:2933-2941.

149. Fedirko V, Bostick RM, Flanders WD, et al. Effects of vitamin D and calcium supplementation on markers of apoptosis in normal colon mucosa: a randomized, double-blind, placebo-controlled clinical trial. Cancer Prev Res (Phila) 2009;2:213-223.

150. Cho E, Smith-Warner SA, Spiegelman D, et al. Dairy foods, calcium, and colorectal cancer: a pooled analysis of 10 cohort studies. J Natl Cancer Inst 2004;96:1015-1022.

151. Wu K, Willett WC, Fuchs CS, Colditz GA, Giovannucci EL. Calcium intake and risk of colon cancer in women and men. J Natl Cancer Inst 2002;94:437-446.

152. Grau MV, Baron JA, Sandler RS, et al. Vitamin D, calcium supplementation, and colorectal adenomas: results of a randomized trial. J Natl Cancer Inst 2003;95:1765-1771.

153. Ahmed FE,. Effect of diet, life style, and other environmental/chemopreventive factors on colorectal cancer development, and assessment of the risks. Journal of environmental science and health Part C, Environmental carcinogenesis 2004;22:91-147.

154. Posaci C, Altunyurt S, Islekel H, Onvural A. Effects of HRT on serum levels of IGF-I in postmenopausal women. Maturitas 2001;40:69-74.

155. Grodstein F, Martinez ME, Platz EA, et al. Postmenopausal hormone use and risk for colorectal cancer and adenoma. Ann Intern Med 1998;128:705-712.

156. Rossouw JE, Anderson GL, Prentice RL, et al. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results From the Women's Health Initiative randomized controlled trial. JAMA 2002;288:321-333.

157. Wei EK, Giovannucci E, Wu K, et al. Comparison of risk factors for colon and rectal cancer. Int J Cancer 2004;108:433-442.

158. Platz EA, Willett WC, Colditz GA, Rimm EB, Spiegelman D, Giovannucci E. Proportion of colon cancer risk that might be preventable in a cohort of middle-aged US men. Cancer Causes Control 2000;11:579-588.

159. Riddell RH, Goldman H, Ransohoff DF, et al. Dysplasia in inflammatory bowel disease: standardized classification with provisional clinical applications. Hum Pathol 1983;14:931-968.

160. Yardley JH, Ransohoff DF, Riddell RH, Goldman H. Cancer in inflammatory bowel disease: how serious is the problem and what should be done about it? Gastroenterology 1983;85:197-200.

161. Munkholm P. Review article: the incidence and prevalence of colorectal cancer in inflammatory bowel disease. Aliment Pharmacol Ther 2003;18 Suppl 2:1-5.

162. Potack J, Itzkowitz SH. Colorectal cancer in inflammatory bowel disease. Gut Liver 2008;2:61-73.

163. Lakatos L, Mester G, Erdelyi Z, et al. Risk factors for ulcerative colitis-associated colorectal cancer in a Hungarian cohort of patients with ulcerative colitis: results of a population-based study. Inflamm Bowel Dis 2006;12:205-211.

164. Eaden JA, Abrams KR, Mayberry JF. The risk of colorectal cancer in ulcerative colitis: a meta-analysis. Gut 2001;48:526-535.

165. Bernstein CN, Blanchard JF, Kliewer E, Wajda A. Cancer risk in patients with inflammatory bowel disease: a population-based study. Cancer 2001;91:854-862.

166. Alexander J, Watanabe T, Wu TT, Rashid A, Li S, Hamilton SR. Histopathological identification of colon cancer with microsatellite instability. Am J Pathol 2001;158:527-535.

167. Lakatos PL, Lakatos L. Risk for colorectal cancer in ulcerative colitis: changes, causes and management strategies. World J Gastroenterol 2008;14:3937-3947.

168. Terzić J, Grivennikov S, Karin E, Karin M. Inflammation and Colon Cancer. Gastroenterology 2010;138:2101-2114.e5.

169. Winther KV, Jess T, Langholz E, Munkholm P, Binder V. Long-term risk of cancer in ulcerative colitis: a population-based cohort study from Copenhagen County. Clin Gastroenterol Hepatol 2004;2:1088-1095.

170. Gellad ZF, Provenzale D. Colorectal cancer: national and international perspective on the burden of disease and public health impact. Gastroenterology 2010;138:2177-2190.

171. Schulmann K, Mori Y, Croog V, et al. Molecular phenotype of inflammatory bowel disease-associated neoplasms with microsatellite instability. Gastroenterology 2005;129:74-85.

172. Ozaki K, Nagasaka T, Notohara K, et al. Heterogeneous microsatellite instability observed within epithelium of ulcerative colitis. Int J Cancer 2006;119:2513-2519.

173. Tahara T, Inoue N, Hisamatsu T, et al. Clinical significance of microsatellite instability in the inflamed mucosa for the prediction of colonic neoplasms in patients with ulcerative colitis. J Gastroenterol Hepatol 2005;20:710-715.

174. Issa JP, Ahuja N, Toyota M, Bronner MP, Brentnall TA. Accelerated age-related CpG island methylation in ulcerative colitis. Cancer Res 2001;61:3573-3577.

175. Nowell PC. The clonal evolution of tumor cell populations. Science 1976;194:23-28.

176. Tomlinson I, Bodmer W. Selection, the mutation rate and cancer: ensuring that the tail does not wag the dog. Nat Med 1999;5:11-12.

177. Nucci MR, Robinson CR, Longo P, Campbell P, Hamilton SR. Phenotypic and genotypic characteristics of aberrant crypt foci in human colorectal mucosa. Hum Pathol 1997;28:1396-1407.

178. Shpitz B, Hay K, Medline A, et al. Natural history of aberrant crypt foci. A surgical approach. Dis Colon Rectum 1996;39:763-767.

179. Lindgren G, Liljegren A, Jaramillo E, Rubio C, Lindblom A. Adenoma prevalence and cancer risk in familial non-polyposis colorectal cancer. Gut 2002;50:228-234.

180. Aust DE, Baretton GB, Members of the Working Group GI-Pathology of the German Society of Pathology. Serrated polyps of the colon and rectum (hyperplastic polyps, sessile serrated adenomas, traditional serrated adenomas, and mixed polyps)-proposal for diagnostic criteria. Virchows Arch 2010;457:291-297.

181. Iino H, Jass JR, Simms LA, et al. DNA microsatellite instability in hyperplastic polyps, serrated adenomas, and mixed polyps: a mild mutator pathway for colorectal cancer? J Clin Pathol 1999;52:5-9.

182. Jass JR, Biden KG, Cummings MC, et al. Characterisation of a subtype of colorectal cancer combining features of the suppressor and mild mutator pathways. J Clin Pathol 1999;52:455-460.

183. Spring KJ, Zhao ZZ, Karamatic R, et al. High prevalence of sessile serrated adenomas with BRAF mutations: a prospective study of patients undergoing colonoscopy. Gastroenterology 2006;131:1400-1407.

184. Longacre TA, Fenoglio-Preiser CM. Mixed hyperplastic adenomatous polyps/serrated adenomas. A distinct form of colorectal neoplasia. Am J Surg Pathol 1990;14:524-537.

185. Jass JR. Classification of colorectal cancer based on correlation of clinical, morphological and molecular features. Histopathology 2007;50:113-130.

186. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. Cell 1990;61:759-767.

187. Lengauer C, Kinzler KW, Vogelstein B. Genetic instability in colorectal cancers. Nature 1997;386:623-7.

188. Pino MS, Chung DC. The chromosomal instability pathway in colon cancer. Gastroenterology 2010;138:2059-2072.

189. Hanson CA, Miller JR. Non-traditional roles for the Adenomatous Polyposis Coli (APC) tumor suppressor protein. Gene 2005;361:1-12.

190. Fodde R, Kuipers J, Rosenberg C, et al. Mutations in the APC tumour suppressor gene cause chromosomal instability. Nat Cell Biol 2001;3:433-438.

191. Powell SM, Zilz N, Beazer-Barclay Y, Bryan TM, Hamilton SR, Thibodeau SN. APC mutations occur early during colorectal tumorigenesis. Nature 1992;359:235-7.

192. Smith KJ, Johnson KA, Bryan TM, et al. The APC gene product in normal and tumor cells. Proceedings of the National Academy of Sciences of the United States of America 1993;90:2846-2850.

193. Rowan AJ, Lamlum H, Ilyas M, et al. APC mutations in sporadic colorectal tumors: A mutational "hotspot" and interdependence of the "two hits". Proceedings of the National Academy of Sciences of the United States of America 2000;97:3352-3357.

194. Vlad A, Rohrs S, Klein-Hitpass L, Muller O. The first five years of the Wnt targetome. Cell Signal 2008;20:795-802.

195. He TC, Sparks AB, Rago C, et al. Identification of c-MYC as a target of the APC pathway. Science 1998;281:1509-1512.

196. Tetsu O, McCormick F. Beta-catenin regulates expression of cyclin D1 in colon carcinoma cells. Nature 1999;398:422-426.

197. Morin PJ, Sparks AB, Korinek V, et al. Activation of beta-catenin-Tcf signaling in colon cancer by mutations in beta-catenin or APC. Science 1997;275:1787-1790.

198. Markowitz SD, Bertagnolli MM. Molecular origins of cancer: Molecular basis of colorectal cancer. N Engl J Med 2009;361:2449-2460.

199. Saif MW, Chu E. Biology of colorectal cancer. Cancer J 2010;16:196-201.

200. Santini D, Loupakis F, Vincenzi B, et al. High concordance of KRAS status between primary colorectal tumors and related metastatic sites: implications for clinical practice. Oncologist 2008;13:1270-1275.

201. Vogelstein B, Fearon ER, Hamilton SR, et al. Genetic alterations during colorectal-tumor development. N Engl J Med 1988;319:525-532.

202. Fearon ER, Cho KR, Nigro JM, et al. Identification of a chromosome 18q gene that is altered in colorectal cancers. Science 1990;247:49-56.

203. Keino-Masu K, Masu M, Hinck L, et al. Deleted in Colorectal Cancer (DCC) encodes a netrin receptor. Cell 1996;87:175-185.

204. Fazeli A, Dickinson SL, Hermiston ML, et al. Phenotype of mice lacking functional Deleted in colorectal cancer (Dcc) gene. Nature 1997;386:796-804.

205. Xu Y, Pasche B. TGF-beta signaling alterations and susceptibility to colorectal cancer. Hum Mol Genet 2007;16 Spec No 1:R14-20.

206. Wu CL, Kirley SD, Xiao H, Chuang Y, Chung DC, Zukerberg LR. Cables enhances cdk2 tyrosine 15 phosphorylation by Wee1, inhibits cell growth, and is lost in many human colon and squamous cancers. Cancer Res 2001;61:7325-7332.

207. Park do Y, Sakamoto H, Kirley SD, et al. The Cables gene on chromosome 18q is silenced by promoter hypermethylation and allelic loss in human colorectal cancer. Am J Pathol 2007;171:1509-1519.

208. Kargozaran H, Kahlenberg M, Khatri VP. The implications of colorectal cancer molecular biology in clinical practice. Surg Oncol Clin N Am 2008;17:341-55, viii-ix.

209. Sjoblom T, Jones S, Wood LD, et al. The consensus coding sequences of human breast and colorectal cancers. Science 2006;314:268-274.

210. Wood LD, Parsons DW, Jones S, et al. The genomic landscapes of human breast and colorectal cancers. Science 2007;318:1108-1113.

211. Pritchard CC, Grady WM. Colorectal cancer molecular biology moves into clinical practice. Gut 2011;60:116-129.

212. Garber JE, Goldstein AM, Kantor AF, Dreyfus MG, Fraumeni JF,Jr, Li FP. Follow-up study of twenty-four families with Li-Fraumeni syndrome. Cancer Res 1991;51:6094-6097.

213. Kinzler KW, Vogelstein B. Lessons from Hereditary Colorectal Cancer. Cell 1996;87:159-170.

214. Modrich P. Mechanisms in eukaryotic mismatch repair. J Biol Chem 2006;281:30305-30309.

215. Potocnik U, Glavac D, Golouh R, Ravnik-Glavac M. Causes of microsatellite instability in colorectal tumors: implications for hereditary non-polyposis colorectal cancer screening. Cancer Genet Cytogenet 2001;126:85-96.

216. Kane MF, Loda M, Gaida GM, et al. Methylation of the hMLH1 promoter correlates with lack of expression of hMLH1 in sporadic colon tumors and mismatch repair-defective human tumor cell lines. Cancer Res 1997;57:808-811.

217. Iacopetta B, Grieu F, Amanuel B. Microsatellite instability in colorectal cancer. Asia Pac J Clin Oncol 2010;6:260-269.

218. Arana ME, Kunkel TA. Mutator phenotypes due to DNA replication infidelity. Semin Cancer Biol 2010;20:304-311.

219. Arzimanoglou II, Gilbert F, Barber HR. Microsatellite instability in human solid tumors. Cancer 1998;82:1808-1820.

220. Preston BD, Albertson TM, Herr AJ. DNA replication fidelity and cancer. Semin Cancer Biol 2010;20:281-293.

221. Peinado MA, Malkhosyan S, Velazquez A, Perucho M. Isolation and characterization of allelic losses and gains in colorectal tumors by arbitrarily primed polymerase chain reaction. Proc Natl Acad Sci U S A 1992;89:10065-10069.

222. Ionov Y, Peinado MA, Malkhosyan S, Shibata D, Perucho M. Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. Nature 1993;363:558-561.

223. Thibodeau SN, Bren G, Schaid D. Microsatellite instability in cancer of the proximal colon. Science 1993;260:816-819.

224. Levinson G, Gutman GA. High frequencies of short frameshifts in poly-CA/TG tandem repeats borne by bacteriophage M13 in Escherichia coli K-12. Nucleic Acids Res 1987;15:5323-5338.

225. Strand M, Prolla TA, Liskay RM, Petes TD. Destabilization of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair. Nature 1993;365:274-276.

226. Grilley M, Holmes J, Yashar B, Modrich P. Mechanisms of DNA-mismatch correction. Mutat Res 1990;236:253-267.

227. Parsons R, Li GM, Longley MJ, Fang WH, Papadopoulos N, Jen J. Hypermutability and mismatch repair deficiency in RER tumor cells. Cell 1993;75:1227-1236.

228. Leach FS, Nicolaides NC, Papadopoulos N, et al. Mutations of a mutS homolog in hereditary nonpolyposis colorectal cancer. Cell 1993;75:1215-1225.

229. Fishel R, Lescoe MK, Rao MR, et al. The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer. Cell 1993;75:1027-1038.

230. Su SS, Modrich P. Escherichia coli mutS-encoded protein binds to mismatched DNA base pairs. Proc Natl Acad Sci U S A 1986;83:5057-5061.

231. Miyaki M, Konishi M, Tanaka K, et al. Germline mutation of MSH6 as the cause of hereditary nonpolyposis colorectal cancer. Nat Genet 1997;17:271-272.

232. Papadopoulos N, Nicolaides NC, Wei YF, et al. Mutation of a mutL homolog in hereditary colon cancer. Science 1994;263:1625-1629.

233. Bronner CE, Baker SM, Morrison PT, et al. Mutation in the DNA mismatch repair gene homologue hMLH1 is associated with hereditary non-polyposis colon cancer. Nature 1994;368:258-261.

234. Nicolaides NC, Papadopoulos N, Liu B, et al. Mutations of two PMS homologues in hereditary nonpolyposis colon cancer. Nature 1994;371:75-80.

235. Duval A, Hamelin R. Mutations at coding repeat sequences in mismatch repair-deficient human cancers: toward a new concept of target genes for instability. Cancer Res 2002;62:2447-2454.

236. Acharya S, Wilson T, Gradia S, et al. hMSH2 forms specific mispair-binding complexes with hMSH3 and hMSH6. Proc Natl Acad Sci U S A 1996;93:13629-13634.

237. McCulloch SD, Gu L, Li GM. Bi-directional processing of DNA loops by mismatch repair-dependent and -independent pathways in human cells. J Biol Chem 2003;278:3891-3896.

238. Jiricny J. The multifaceted mismatch-repair system. Nat Rev Mol Cell Biol 2006;7:335-346.

239. Zhou XP, Hoang JM, Li YJ, et al. Determination of the replication error phenotype in human tumors without the requirement for matching normal DNA by analysis of mononucleotide repeat microsatellites. Genes Chromosomes Cancer 1998;21:101-107.

240. Suraweera N, Duval A, Reperant M, et al. Evaluation of tumor microsatellite instability using five quasimonomorphic mononucleotide repeats and pentaplex PCR. Gastroenterology 2002;123:1804-1811.

241. Hoang JM, Cottu PH, Thuille B, Salmon RJ, Thomas G, Hamelin R. BAT-26, an indicator of the replication error phenotype in colorectal cancers and cell lines. Cancer Res 1997;57:300-303.

242. Xicola RM, Llor X, Pons E, et al. Performance of different microsatellite marker panels for detection of mismatch repair-deficient colorectal tumors. J Natl Cancer Inst 2007;99:244-252.

243. Ku JL, Yoon KA, Kim DY, Park JG. Mutations in hMSH6 alone are not sufficient to cause the microsatellite instability in colorectal cancer cell lines. Eur J Cancer 1999;35:1724-1729.

244. Berends MJ, Wu Y, Sijmons RH, et al. Molecular and clinical characteristics of MSH6 variants: an analysis of 25 index carriers of a germline variant. Am J Hum Genet 2002;70:26-37.

245. Wu Y, Berends MJ, Mensink RG, et al. Association of hereditary nonpolyposis colorectal cancer-related tumors displaying low microsatellite instability with MSH6 germline mutations. Am J Hum Genet 1999;65:1291-1298.

246. Overbeek LI, Ligtenberg MJ, Willems RW, et al. Interpretation of immunohistochemistry for mismatch repair proteins is only reliable in a specialized setting. Am J Surg Pathol 2008;32:1246-1251.

247. Drescher KM, Sharma P, Watson P, Gatalica Z, Thibodeau SN, Lynch HT. Lymphocyte recruitment into the tumor site is altered in patients with MSI-H colon cancer. Fam Cancer 2009;8:231-239.

248. Kouri M, Laasonen A, Mecklin JP, Jarvinen H, Franssila K, Pyrhonen S. Diploid predominance in hereditary nonpolyposis colorectal carcinoma evaluated by flow cytometry. Cancer 1990;65:1825-1829.

249. Frei JV. Hereditary nonpolyposis colorectal cancer (Lynch syndrome II). Diploid malignancies with prolonged survival. Cancer 1992;69:1108-1111.

250. Sankila R, Aaltonen LA, Jarvinen HJ, Mecklin JP. Better survival rates in patients with MLH1-associated hereditary colorectal cancer. Gastroenterology 1996;110:682-687.

251. Drescher KM, Sharma P, Lynch HT. Current hypotheses on how microsatellite instability leads to enhanced survival of Lynch Syndrome patients. Clin Dev Immunol 2010;2010:170432.

252. Dolcetti R, Viel A, Doglioni C, 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;154:1805-1813.

253. Samowitz WS, Curtin K, Wolff RK, Tripp SR, Caan BJ, Slattery ML. Microsatellite instability and survival in rectal cancer. Cancer Causes Control 2009;20:1763-1768.

254. Boland CR. Clinical uses of microsatellite instability testing in colorectal cancer: an ongoing challenge. J Clin Oncol 2007;25:754-756.

255. Chen H, Taylor NP, Sotamaa KM, et al. Evidence for heritable predisposition to epigenetic silencing of MLH1. Int J Cancer 2007;120:1684-1688.

256. Chan TL, Yuen ST, Kong CK, et al. Heritable germline epimutation of MSH2 in a family with hereditary nonpolyposis colorectal cancer. Nat Genet 2006;38:1178-1183.

257. Ligtenberg MJ, Kuiper RP, Chan TL, et al. Heritable somatic methylation and inactivation of MSH2 in families with Lynch syndrome due to deletion of the 3' exons of TACSTD1. Nat Genet 2009;41:112-117.

258. Kovacs ME, Papp J, Szentirmay Z, Otto S, Olah E. Deletions removing the last exon of TACSTD1 constitute a distinct class of mutations predisposing to Lynch syndrome. Hum Mutat 2009;30:197-203.

259. Goecke T, Schulmann K, Engel C, et al. Genotype-phenotype comparison of German MLH1 and MSH2 mutation carriers clinically affected with Lynch syndrome: a report by the German HNPCC Consortium. J Clin Oncol 2006;24:4285-4292.

260. Samowitz WS, Sweeney C, Herrick J, et al. Poor survival associated with the BRAF V600E mutation in microsatellite-stable colon cancers. Cancer Res 2005;65:6063-6069.

261. Li WQ, Kawakami K, Ruszkiewicz A, Bennett G, Moore J, Iacopetta B. BRAF mutations are associated with distinctive clinical, pathological and molecular features of colorectal cancer independently of microsatellite instability status. Mol Cancer 2006;5:2.

262. Ogino S, Nosho K, Kirkner GJ, et al. CpG island methylator phenotype, microsatellite instability, BRAF mutation and clinical outcome in colon cancer. Gut 2009;58:90-96.

263. Hawkins N, Norrie M, Cheong K, et al. CpG island methylation in sporadic colorectal cancers and its relationship to microsatellite instability. Gastroenterology 2002;122:1376-1387.

264. Bird AP. CpG-rich islands and the function of DNA methylation. Nature 1986;321:209-213.

265. Bonasio R, Tu S, Reinberg D. Molecular signals of epigenetic states. Science 2010;330:612-616.

266. Latham KE. X chromosome imprinting and inactivation in the early mammalian embryo. Trends Genet 1996;12:134-138.

267. Li E, Beard C, Jaenisch R. Role for DNA methylation in genomic imprinting. Nature 1993;366:362-365.

268. Issa JP, Ottaviano YL, Celano P, Hamilton SR, Davidson NE, Baylin SB. Methylation of the oestrogen receptor CpG island links ageing and neoplasia in human colon. Nat Genet 1994;7:536-540.

269. Oberle I, Rousseau F, Heitz D, et al. Instability of a 550-base pair DNA segment and abnormal methylation in fragile X syndrome. Science 1991;252:1097-1102.

270. Hansen RS, Gartler SM, Scott CR, Chen SH, Laird CD. Methylation analysis of CGG sites in the CpG island of the human FMR1 gene. Hum Mol Genet 1992;1:571-578.

271. Sakai T, Toguchida J, Ohtani N, Yandell DW, Rapaport JM, Dryja TP. Allele-specific hypermethylation of the retinoblastoma tumor-suppressor gene. Am J Hum Genet 1991;48:880-888.

272. Baylin SB, Herman JG, Graff JR, Vertino PM, Issa JP. Alterations in DNA methylation: a fundamental aspect of neoplasia. Adv Cancer Res 1998;72:141-196.

273. Merlo A, Herman JG, Mao L, et al. 5' CpG island methylation is associated with transcriptional silencing of the tumour suppressor p16/CDKN2/MTS1 in human cancers. Nat Med 1995;1:686-692.

274. Toyota M, Ahuja N, Ohe-Toyota M, Herman JG, Baylin SB, Issa JP. CpG island methylator phenotype in colorectal cancer. Proc Natl Acad Sci U S A 1999;96:8681-8686.

275. Ahuja N, Mohan AL, Li Q, et al. Association between CpG island methylation and microsatellite instability in colorectal cancer. Cancer Res 1997;57:3370-3374.

276. Esteller M, Hamilton SR, Burger PC, Baylin SB, Herman JG. Inactivation of the DNA repair gene O6-methylguanine-DNA methyltransferase by promoter hypermethylation is a common event in primary human neoplasia. Cancer Res 1999;59:793-797.

277. Whitehall VL, Walsh MD, Young J, Leggett BA, Jass JR. Methylation of O-6-methylguanine DNA methyltransferase characterizes a subset of colorectal cancer with low-level DNA microsatellite instability. Cancer Res 2001;61:827-830.

278. Issa JJ, Baylin SB, Belinsky SA. Methylation of the Estrogen Receptor CpG Island in Lung Tumors Is Related to the Specific Type of Carcinogen Exposure. Cancer Research 1996;56:3655-3658.

279. Frigola J, Song J, Stirzaker C, Hinshelwood RA, Peinado MA, Clark SJ. Epigenetic remodeling in colorectal cancer results in coordinate gene suppression across an entire chromosome band. Nat Genet 2006;38:540-549.

280. Eguchi K, Kanai Y, Kobayashi K, Hirohashi S. DNA hypermethylation at the D17S5 locus in non-small cell lung cancers: its association with smoking history. Cancer Res 1997;57:4913-4915.

281. Belinsky SA, Palmisano WA, Gilliland FD, et al. Aberrant promoter methylation in bronchial epithelium and sputum from current and former smokers. Cancer Res 2002;62:2370-2377.

282. Kim DH, Nelson HH, Wiencke JK, et al. p16(INK4a) and histology-specific methylation of CpG islands by exposure to tobacco smoke in non-small cell lung cancer. Cancer Res 2001;61:3419-3424.

283. Soria JC, Rodriguez M, Liu DD, Lee JJ, Hong WK, Mao L. Aberrant promoter methylation of multiple genes in bronchial brush samples from former cigarette smokers. Cancer Res 2002;62:351-355.

284. Jarmalaite S, Kannio A, Anttila S, Lazutka JR, Husgafvel-Pursiainen K. Aberrant p16 promoter methylation in smokers and former smokers with nonsmall cell lung cancer. Int J Cancer 2003;106:913-918.

285. Feinberg AP, Ohlsson R, Henikoff S. The epigenetic progenitor origin of human cancer. Nat Rev Genet 2006;7:21-33.

286. Irizarry RA, Ladd-Acosta C, Wen B, et al. The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. Nat Genet 2009;41:178-186.

287. Makinen MJ. Colorectal serrated adenocarcinoma. Histopathology 2007;50:131-150.

288. Yamashita K, Dai T, Dai Y, Yamamoto F, Perucho M. Genetics supersedes epigenetics in colon cancer phenotype. Cancer Cell 2003;4:121-131.

289. Anacleto C, Leopoldino AM, Rossi B, et al. Colorectal cancer "methylator phenotype": fact or artifact? Neoplasia 2005;7:331-335.

290. Weisenberger DJ, Siegmund KD, Campan M, et al. CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. Nat Genet 2006;38:787-793.

291. Deng G, Bell I, Crawley S, et al. BRAF Mutation Is Frequently Present in Sporadic Colorectal Cancer with Methylated hMLH1, But Not in Hereditary Nonpolyposis Colorectal Cancer. Clinical Cancer Research 2004;10:191-195.

292. Kambara T, Simms LA, Whitehall VLJ, et al. BRAF mutation is associated with DNA methylation in serrated polyps and cancers of the colorectum. Gut 2004;53:1137-1144.

293. Samowitz WS, Albertsen H, Herrick J. Evaluation of a large, population-based sample supports a CpG island methylator phenotype in colon cancer. Gastroenterology 2005;129:837-845.

294. Chan AO, Issa JP, Morris JS, Hamilton SR, Rashid A. Concordant CpG island methylation in hyperplastic polyposis. Am J Pathol 2002;160:529-536.

295. Wynter CV, Walsh MD, Higuchi T, Leggett BA, Young J, Jass JR. Methylation patterns define two types of hyperplastic polyp associated with colorectal cancer. Gut 2004;53:573-580.

296. Vaughn CP, Wilson AR, Samowitz WS. Quantitative evaluation of CpG island methylation in hyperplastic polyps. Mod Pathol 2009;23:151-156.

297. Higuchi T, Sugihara K, Jass JR. Demographic and pathological characteristics of serrated polyps of colorectum. Histopathology 2005;47:32-40.

298. Tateyama H, Li W, Takahashi E, Miura Y, Sugiura H, Eimoto T. Apoptosis index and apoptosis-related antigen expression in serrated adenoma of the colorectum: the saw-toothed structure may be related to inhibition of apoptosis. Am J Surg Pathol 2002;26:249-256.

299. Torlakovic E, Snover DC. Serrated adenomatous polyposis in humans. Gastroenterology 1996;110:748-755.

300. Jass JR, Whitehall VL, Young J, Leggett BA. Emerging concepts in colorectal neoplasia. Gastroenterology 2002;123:862-876.

301. Noffsinger AE. Serrated polyps and colorectal cancer: new pathway to malignancy. Annu Rev Pathol 2009;4:343-364.

302. Jass JR, Baker K, Zlobec I, et al. Advanced colorectal polyps with the molecular and morphological features of serrated polyps and adenomas: concept of a 'fusion' pathway to colorectal cancer. Histopathology 2006;49:121-131.

303. Jass JR, Iino H, Ruszkiewicz A, et al. Neoplastic progression occurs through mutator pathways in hyperplastic polyposis of the colorectum. Gut 2000;47:43-49.

304. Rajagopalan H, Bardelli A, Lengauer C, Kinzler KW, Vogelstein B, Velculescu VE. Tumorigenesis: RAF/RAS oncogenes and mismatch-repair status. Nature 2002;418:934.

305. Nagasaka T, Sasamoto H, Notohara K, et al. Colorectal cancer with mutation in BRAF, KRAS, and wild-type with respect to both oncogenes showing different patterns of DNA methylation. J Clin Oncol 2004;22:4584-4594.

306. Zlobec I, Bihl MP, Schwarb H, Terracciano L, Lugli A. Clinicopathological and protein characterization of BRAF- and K-RAS-mutated colorectal cancer and implications for prognosis. Int J Cancer 2010;127:367-380.

307. Davies H, Bignell GR, Cox C, et al. Mutations of the BRAF gene in human cancer. Nature 2002;417:949-954.

308. Tie J, Gibbs P, Lipton L, et al. Optimizing targeted therapeutic development: Analysis of a colorectal cancer patient population with the BRAF(V600E) mutation. International journal of cancer Journal international du cancer 2010;.

309. Barault L, Charon-Barra C, Jooste V, et al. Hypermethylator phenotype in sporadic colon cancer: study on a population-based series of 582 cases. Cancer Res 2008;68:8541-8546.

310. Nagasaka T, Koi M, Kloor M, et al. Mutations in Both KRAS and BRAF May Contribute to the Methylator Phenotype in Colon Cancer. Gastroenterology 2008;.

311. Ogino S, Goel A. Molecular classification and correlates in colorectal cancer. J Mol Diagn 2008;10:13-27.

312. Leggett B, Whitehall V. Role of the serrated pathway in colorectal cancer pathogenesis. Gastroenterology 2010;138:2088-2100.

313. Arain MA, Sawhney M, Sheikh S, et al. CIMP status of interval colon cancers: another piece to the puzzle. Am J Gastroenterol 2010;105:1189-1195.

314. Snover DC. Update on the serrated pathway to colorectal carcinoma. Hum Pathol 2011;42:1-10.

315. Sawhney MS, Farrar WD, Gudiseva S, et al. Microsatellite instability in interval colon cancers. Gastroenterology 2006;131:1700-1705.

316. Jass JR. Molecular heterogeneity of colorectal cancer: Implications for cancer control. Surg Oncol 2007;16 Suppl 1:S7-9.

317. Cappuzzo F, Varella-Garcia M, Finocchiaro G, et al. Primary resistance to cetuximab therapy in EGFR FISH-positive colorectal cancer patients. Br J Cancer 2008;99:83-89.

318. Di Nicolantonio F, Martini M, Molinari F, et al. Wild-type BRAF is required for response to panitumumab or cetuximab in metastatic colorectal cancer. J Clin Oncol 2008;26:5705-5712.

319. Issa JP, Kantarjian HM. Targeting DNA methylation. Clin Cancer Res 2009;15:3938-3946.

320. Curtin K, Slattery ML, Samowitz WS. CpG island methylation in colorectal cancer: past, present and future. Patholog Res Int 2011;2011:902674.

321. McCubrey JA, Steelman LS, Abrams SL, et al. Emerging Raf inhibitors. Expert Opin Emerg Drugs 2009;14:633-648.

322. Wilhelm SM, Carter C, Tang L, et al. BAY 43-9006 exhibits broad spectrum oral antitumor activity and targets the RAF/MEK/ERK pathway and receptor tyrosine kinases involved in tumor progression and angiogenesis. Cancer Res 2004;64:7099-7109.

323. Reya T, Morrison SJ, Clarke MF, Weissman IL. Stem cells, cancer, and cancer stem cells. Nature 2001;414:105-111.

324. O'Brien CA, Pollett A, Gallinger S, Dick JE. A human colon cancer cell capable of initiating tumour growth in immunodeficient mice. Nature 2007;445:106-110.

325. Luebeck EG, Moolgavkar SH. Multistage carcinogenesis and the incidence of colorectal cancer. Proc Natl Acad Sci U S A 2002;99:15095-15100.

326. Potten CS, Booth C, Pritchard DM. The intestinal epithelial stem cell: the mucosal governor. Int J Exp Pathol 1997;78:219-43.

327. Kosinski C, Li VS, Chan AS, et al. Gene expression patterns of human colon tops and basal crypts and BMP antagonists as intestinal stem cell niche factors. Proc Natl Acad Sci U S A 2007;104:15418-15423.

328. Shackleton M, Quintana E, Fearon ER, Morrison SJ. Heterogeneity in cancer: cancer stem cells versus clonal evolution. Cell 2009;138:822-829.

329. Lapidot T, Sirard C, Vormoor J, et al. A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. Nature 1994;367:645-648.

330. Bonnet D, Dick JE. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. Nat Med 1997;3:730-737.

331. Singh SK, Hawkins C, Clarke ID, et al. Identification of human brain tumour initiating cells. Nature 2004;432:396-401.

332. Al-Hajj M, Wicha MS, Benito-Hernandez A, Morrison SJ, Clarke MF. Prospective identification of tumorigenic breast cancer cells. Proc Natl Acad Sci U S A 2003;100:3983-3988.

333. Ricci-Vitiani L, Lombardi DG, Pilozzi E, et al. Identification and expansion of human colon-cancer-initiating cells. Nature 2007;445:111-115.

334. Todaro M, Francipane MG, Medema JP, Stassi G. Colon cancer stem cells: promise of targeted therapy. Gastroenterology 2010;138:2151-2162.

335. Schatton T, Murphy GF, Frank NY, et al. Identification of cells initiating human melanomas. Nature 2008;451:345-349.

336. Quintana E, Shackleton M, Sabel MS, Fullen DR, Johnson TM, Morrison SJ. Efficient tumour formation by single human melanoma cells. Nature 2008;456:593-598.

337. Willis ND, Cox TR, Rahman-Casans SF, et al. Lamin A/C is a risk biomarker in colorectal cancer. PLoS One 2008;3:e2988.

338. Brabletz T, Jung A, Spaderna S, Hlubek F, Kirchner T. Opinion: migrating cancer stem cells - an integrated concept of malignant tumour progression. Nat Rev Cancer 2005;5:744-749.

339. Brabletz T, Hlubek F, Spaderna S, et al. Invasion and metastasis in colorectal cancer: epithelial-mesenchymal transition, mesenchymal-epithelial transition, stem cells and beta-catenin. Cells Tissues Organs 2005;179:56-65.

340. Jass JR, Barker M, Fraser L, et al. APC mutation and tumour budding in colorectal cancer. J Clin Pathol 2003;56:69-73.

341. Vasen HF, Watson P, Mecklin JP, Lynch HT. New clinical criteria for hereditary nonpolyposis colorectal cancer (HNPCC, Lynch syndrome) proposed by the International Collaborative group on HNPCC. Gastroenterology 1999;116:1453-1456.

342. Lynch HT, Boland CR, Gong G, et al. Phenotypic and genotypic heterogeneity in the Lynch syndrome: diagnostic, surveillance and management implications. Eur J Hum Genet 2006;14:390-402.

343. Vasen HF, Mecklin JP, Khan PM, Lynch HT. The International Collaborative Group on Hereditary Non-Polyposis Colorectal Cancer (ICG-HNPCC). Dis Colon Rectum 1991;34:424-425.

344. Rodriguez-Bigas MA, Boland CR, Hamilton SR, et al. A National Cancer Institute Workshop on Hereditary Nonpolyposis Colorectal Cancer Syndrome: meeting highlights and Bethesda guidelines. J Natl Cancer Inst 1997;89:1758-1762.

345. Andre T, Boni C, Mounedji-Boudiaf L, et al. Oxaliplatin, fluorouracil, and leucovorin as adjuvant treatment for colon cancer. N Engl J Med 2004;350:2343-2351.

346. Des Guetz G, Schischmanoff O, Nicolas P, Perret GY, Morere JF, Uzzan B. Does microsatellite instability predict the efficacy of adjuvant chemotherapy in colorectal cancer? A systematic review with meta-analysis. Eur J Cancer 2009;45:1890-1896.

347. Parry S, Win AK, Parry B, et al. Metachronous colorectal cancer risk for mismatch repair gene mutation carriers: the advantage of more extensive colon surgery. Gut 2011;60:950-957.

348. Terdiman JP, Gum JR,Jr, Conrad PG, et al. Efficient detection of hereditary nonpolyposis colorectal cancer gene carriers by screening for tumor microsatellite instability before germline genetic testing. Gastroenterology 2001;120:21-30.

349. Trano G, Sjursen W, Wasmuth HH, Hofsli E, Vatten LJ. Performance of clinical guidelines compared with molecular tumour screening methods in identifying possible Lynch syndrome among colorectal cancer patients: a Norwegian population-based study. Br J Cancer 2010;.

350. Hampel H, Frankel WL, Martin E, et al. Feasibility of Screening for Lynch Syndrome Among Patients With Colorectal Cancer. J Clin Oncol 2008;26:5783-5788.

351. Julie C, Tresallet C, Brouquet A, et al. Identification in daily practice of patients with Lynch syndrome (hereditary nonpolyposis colorectal cancer): revised Bethesda guidelines-based approach versus molecular screening. Am J Gastroenterol 2008;103:2825-35; quiz 2836.

352. Bapat B, Lindor NM, Baron J, et al. The association of tumor microsatellite instability phenotype with family history of colorectal cancer. Cancer Epidemiol Biomarkers Prev 2009;18:967-75.

353. Lindor NM, Rabe K, Petersen GM, et al. Lower Cancer Incidence in Amsterdam-I Criteria Families Without Mismatch Repair Deficiency: Familial Colorectal Cancer Type X. JAMA 2005;293:1979-1985.

354. Valle L, Perea J, Carbonell P, et al. Clinicopathologic and pedigree differences in amsterdam I-positive hereditary nonpolyposis colorectal cancer families according to tumor microsatellite instability status. J Clin Oncol 2007;25:781-786.

355. Lindor NM. Familial colorectal cancer type X: the other half of hereditary nonpolyposis colon cancer syndrome. Surg Oncol Clin N Am 2009;18:637-645.

356. Mueller-Koch Y, Vogelsang H, Kopp R, et al. Hereditary non-polyposis colorectal cancer: clinical and molecular evidence for a new entity of hereditary colorectal cancer. Gut 2005;54:1733-1740.

357. Llor X, Pons E, Xicola RM, et al. Differential features of colorectal cancers fulfilling Amsterdam criteria without involvement of the mutator pathway. Clin Cancer Res 2005;11:7304-7310.

358. Abdel-Rahman WM, Ollikainen M, Kariola R, et al. Comprehensive characterization of HNPCC-related colorectal cancers reveals striking molecular features in families with no germline mismatch repair gene mutations. Oncogene 2005;24:1542-1551.

359. Sanchez-de-Abajo A, de la Hoya M, van Puijenbroek M, et al. Molecular analysis of colorectal cancer tumors from patients with mismatch repair proficient hereditary nonpolyposis colorectal cancer suggests novel carcinogenic pathways. Clin Cancer Res 2007;13:5729-5735.

360. Aaltonen L, Johns L, Jarvinen H, Mecklin JP, Houlston R. Explaining the familial colorectal cancer risk associated with mismatch repair (MMR)-deficient and MMR-stable tumors. Clin Cancer Res 2007;13:356-361.

361. Jasperson KW, Tuohy TM, Neklason DW, Burt RW. Hereditary and familial colon cancer. Gastroenterology 2010;138:2044-2058.

362. Bulow S, Bjork J, Christensen IJ, et al. Duodenal adenomatosis in familial adenomatous polyposis. Gut 2004;53:381-386.

363. Spigelman AD, Williams CB, Talbot IC, Domizio P, Phillips RK. Upper gastrointestinal cancer in patients with familial adenomatous polyposis. Lancet 1989;2:783-785.

364. Lynch HT, Smyrk T, McGinn T, et al. Attenuated familial adenomatous polyposis (AFAP). A phenotypically and genotypically distinctive variant of FAP. Cancer 1995;76:2427-2433.

365. Gardner RJ, Kool D, Edkins E, et al. The clinical correlates of a 3' truncating mutation (codons 1982-1983) in the adenomatous polyposis coli gene. Gastroenterology 1997;113:326-331.

366. Giardiello FM, Brensinger JD, Luce MC, et al. Phenotypic expression of disease in families that have mutations in the 5' region of the adenomatous polyposis coli gene. Ann Intern Med 1997;126:514-519.

367. Soravia C, Berk T, Madlensky L, et al. Genotype-phenotype correlations in attenuated adenomatous polyposis coli. Am J Hum Genet 1998;62:1290-1301.

368. Scott RJ, van der Luijt R, Spycher M, et al. Novel germline APC gene mutation in a large familial adenomatous polyposis kindred displaying variable phenotypes. Gut 1995;36:731-736.

369. Heinen CD. Genotype to phenotype: Analyzing the effects of inherited mutations in colorectal cancer families. Mutat Res 2010;693:32-45.

370. Krush AJ, Traboulsi EI, Offerhaus JA, Maumenee IH, Yardley JH, Levin LS. Hepatoblastoma, pigmented ocular fundus lesions and jaw lesions in Gardner syndrome. Am J Med Genet 1988;29:323-32.

371. Hamilton SR, Liu B, Parsons RE, et al. The molecular basis of Turcot's syndrome. N Engl J Med 1995;332:839-847.

372. Paraf F, Jothy S, Van Meir EG. Brain tumor-polyposis syndrome: two genetic diseases? Journal of Clinical Oncology 1997;15:2744-58.

373. Herrera L, Kakati S, Gibas L, Pietrzak E, Sandberg AA. Brief Clinical Report: Gardner Syndrome in a Man With an Interstitial Deletion of 5q. Am J Med Genet 1986;25:473.

374. Kinzler KW, Nilbert MC, Su LK, et al. Identification of FAP locus genes from chromosome 5q21. Science 1991;253:661-665.

375. Nishisho I, Nakamura Y, Miyoshi Y, et al. Mutations of chromosome 5q21 genes in FAP and colorectal cancer patients. Science 1991;253:665-9.

376. Groden J, Thliveris A, Samowitz W, et al. Identification and characterization of the familial adenomatous polyposis coli gene. Cell 1991;66:589-600.

377. Laurent-Puig P, Beroud C, Soussi T. APC gene: database of germline and somatic mutations in human tumors and cell lines. Nucleic Acids Res 1998;26:269-270.

378. Olschwang S, Tiret A, Laurent-Puig P, Muleris M, Parc R, Thomas G. Restriction of ocular fundus lesions to a specific subgroup of APC mutations in adenomatous polyposis coli patients. Cell 1993;75:959-968.

379. Davies DR, Armstrong JG, Thakker N, et al. Severe Gardner syndrome in families with mutations restricted to a specific region of the APC gene. Am J Hum Genet 1995;57:1151-1158.

380. Nieuwenhuis MH, Vasen HF. Correlations between mutation site in APC and phenotype of familial adenomatous polyposis (FAP): a review of the literature. Crit Rev Oncol Hematol 2007;61:153-161.

381. Al-Tassan N, Chmiel NH, Maynard J, et al. Inherited variants of MYH associated with somatic G:C-->T:A mutations in colorectal tumors. Nat Genet 2002;30:227-232.

382. Jones S, Emmerson P, Maynard J, et al. Biallelic germline mutations in MYH predispose to multiple colorectal adenoma and somatic G:C-->T:A mutations. Hum Mol Genet 2002;11:2961-2967.

383. Sieber OM, Lipton L, Crabtree M, et al. Multiple colorectal adenomas, classic adenomatous polyposis, and germ-line mutations in MYH. N Engl J Med 2003;348:791-799.

384. Sampson JR, Dolwani S, Jones S, et al. Autosomal recessive colorectal adenomatous polyposis due to inherited mutations of MYH. Lancet 2003;362:39-41.

385. Lipton L, Halford SE, Johnson V, et al. Carcinogenesis in MYH-associated polyposis follows a distinct genetic pathway. Cancer Res 2003;63:7595-7599.

386. Chow E, Lipton L, Lynch E, et al. Hyperplastic polyposis syndrome: phenotypic presentations and the role of MBD4 and MYH. Gastroenterology 2006;131:30-39.

387. Jenkins MA, Croitoru ME, Monga N, et al. Risk of colorectal cancer in monoallelic and biallelic carriers of MYH mutations: a population-based case-family study. Cancer Epidemiol Biomarkers Prev 2006;15:312-314.

388. Theodoratou E, Campbell H, Tenesa A, et al. A large-scale meta-analysis to refine colorectal cancer risk estimates associated with MUTYH variants. Br J Cancer 2010;103:1875-1884.

389. Williams GT, Arthur JF, Bussey HJ, Morson BC. Metaplastic polyps and polyposis of the colorectum. Histopathology 1980;4:155-170.

390. Rashid A, Houlihan PS, Booker S, Petersen GM, Giardiello FM, Hamilton SR. Phenotypic and molecular characteristics of hyperplastic polyposis. Gastroenterology 2000;119:323-332.

391. Leggett BA, Devereaux B, Biden K, Searle J, Young J, Jass J. Hyperplastic polyposis: association with colorectal cancer. Am J Surg Pathol 2001;25:177-184.

392. Yeoman A, Young J, Arnold J, Jass J, Parry S. Hyperplastic polyposis in the New Zealand population: a condition associated with increased colorectal cancer risk and European ancestry. N Z Med J 2007;120:U2827.

393. Boparai KS, Dekker E, van Eeden S, et al. Hyperplastic Polyps and Sessile Serrated Adenomas as a Phenotypic Expression of MYH-Associated Polyposis. Gastroenterology 2008;135:2014-2018.

394. Buchanan D, Young J. A perspective on bi-allelic MUTYH mutations in patients with hyperplastic polyposis syndrome. Gastroenterology 2009;136:2407-2408.

395. Frazier ML, Xi L, Zong J, et al. Association of the CpG island methylator phenotype with family history of cancer in patients with colorectal cancer. Cancer Res 2003;63:4805-4808.

396. Young J, Jenkins M, Parry S, et al. Serrated pathway colorectal cancer in the population: genetic consideration. Gut 2007;56:1453-1459.

397. Roberts A, Nancarrow D, Clendenning M, et al. Linkage to chromosome 2q32.2-q33.3 in familial serrated neoplasia (Jass syndrome). Fam Cancer 2011;10:245-254.

398. Manfredi M,. Hereditary hamartomatous polyposis syndromes: understanding the disease risks as children reach adulthood. Gastroenterology 2010;6:185-96.

399. Coburn MC, Pricolo VE, DeLuca FG, Bland KI. Malignant potential in intestinal juvenile polyposis syndromes. Ann Surg Oncol 1995;2:386-391.

400. Jass JR, Williams CB, Bussey HJ, Morson BC. Juvenile polyposis--a precancerous condition. Histopathology 1988;13:619-630.

401. Howe JR, Mitros FA, Summers RW. The risk of gastrointestinal carcinoma in familial juvenile polyposis. Ann Surg Oncol 1998;5:751-756.

402. Houlston R, Bevan S, Williams A, et al. Mutations in DPC4 (SMAD4) cause juvenile polyposis syndrome, but only account for a minority of cases. Hum Mol Genet 1998;7:1907-1912.

403. Howe JR, Roth S, Ringold JC, et al. Mutations in the SMAD4/DPC4 gene in juvenile polyposis. Science 1998;280:1086-1088.

404. Hahn SA, Schutte M, Hoque AT, et al. DPC4, a candidate tumor suppressor gene at human chromosome 18q21.1. Science 1996;271:350-353.

405. Mendelsohn J, Howley PM, Israel MA, Gray JW, Thompson CB, eds. The Molecular Basis of Cancer. U.S.A.: Saunders, Elsevier Science, 2008.

406. Marsh DJ, Kum JB, Lunetta KL, et al. PTEN mutation spectrum and genotype-phenotype correlations in Bannayan-Riley-Ruvalcaba syndrome suggest a single entity with Cowden syndrome. Hum Mol Genet 1999;8:1461-1472.

407. Gorlin RJ, Cohen MM,Jr, Condon LM, Burke BA. Bannayan-Riley-Ruvalcaba syndrome. Am J Med Genet 1992;44:307-314.

408. Volikos E, Robinson J, Aittomaki K, et al. LKB1 exonic and whole gene deletions are a common cause of Peutz-Jeghers syndrome. J Med Genet 2006;43:e18.

409. Risch HA, McLaughlin JR, Cole DEC, et al. Prevalence and Penetrance of Germline BRCA1 and BRCA2 Mutations in a Population Series of 649 Women with Ovarian Cancer. The American Journal of Human Genetics 2001;68:700-710.

410. Xiang HP, Geng XP, Ge WW, Li H. Meta-analysis of CHEK2 1100delC variant and colorectal cancer susceptibility. European journal of cancer 2011;.

411. Marvin ML, Mazzoni SM, Herron CM, Edwards S, Gruber SB, Petty EM. AXIN2-associated autosomal dominant ectodermal dysplasia and neoplastic syndrome. American Journal of Medical Genetics 2011;155A:898-902.

412. Bapat B, Xia L, Madlensky L, et al. The genetic basis of Muir-Torre syndrome includes the hMLH1 locus. Am J Hum Genet 1996;59:736-9.

413. Barana D, van der Klift H, Wijnen J, et al. Spectrum of genetic alterations in Muir-Torre syndrome is the same as in HNPCC. American Journal of Medical Genetics 2004;125A:318-9.

414. Zanke BW, Greenwood CM, Rangrej J, et al. Genome-wide association scan identifies a colorectal cancer susceptibility locus on chromosome 8q24. Nat Genet 2007;39:989-994.

415. Poynter JN, Figueiredo JC, Conti DV, et al. Variants on 9p24 and 8q24 are associated with risk of colorectal cancer: results from the Colon Cancer Family Registry. Cancer Res 2007;67:11128-11132.

416. Gray-McGuire C, Guda K, Adrianto I, et al. Confirmation of linkage to and localization of familial colon cancer risk haplotype on chromosome 9q22. Cancer Res 2010;70:5409-5418.

417. Guda K, Moinova H, He J, et al. Inactivating germ-line and somatic mutations in polypeptide N-acetylgalactosaminyltransferase 12 in human colon cancers. Proc Natl Acad Sci U S A 2009;106:12921-12925.

418. Järvinen HJ, Aarnio M, Mustonen H, et al. Controlled 15-year trial on screening for colorectal cancer in families with hereditary nonpolyposis colorectal cancer. Gastroenterology 2000;118:829-834.

419. Burch JA, Soares-Weiser K, St John DJ, et al. Diagnostic accuracy of faecal occult blood tests used in screening for colorectal cancer: a systematic review. J Med Screen 2007;14:132-137.

420. Ouyang DL, Chen JJ, Getzenberg RH, Schoen RE. Noninvasive testing for colorectal cancer: a review. Am J Gastroenterol 2005;100:1393-1403.

421. Colorectal Cancer Association of Canada - CCAC. Colorectal Cancer Screening. Available at: http://www.colorectal-cancer.ca/en/screening/screening-tests/. Accessed 3/8/2011, 2011.

422. Canadian Task Force on Preventive Health Care. Colorectal cancer screening. Recommendation statement from the Canadian Task Force on Preventive Health Care. CMAJ 2001;165:206-208.

423. Leddin D, Hunt R, Champion M, Cockeram A, Flook N, Gould M. Canadian Association of Gastroenterology and the Canadian Digestive Health Foundation: Guidelines on colon cancer screening. Can J Gastroenterol 2004;18:93-9.

424. Lansdorp-Vogelaar I, Kuntz KM, Knudsen AB, Wilschut JA, Zauber AG, van Ballegooijen M. Stool DNA Testing to Screen for Colorectal Cancer in the Medicare Population. Annals of Internal Medicine 2010;153:368-377.

425. Imperiale TF, Ransohoff DF, Itzkowitz SH, Turnbull BA, Ross ME, Colorectal Cancer Study Group. Fecal DNA versus fecal occult blood for colorectal-cancer screening in an average-risk population. N Engl J Med 2004;351:2704-2714.

426. Kalimutho M, Del Vecchio Blanco G, Cretella M, et al. A simplified, non-invasive fecal-based DNA integrity assay and iFOBT for colorectal cancer detection. Int J Colorectal Dis 2011;.

427. Segnan N, Senore C, Andreoni B, et al. Randomized trial of different screening strategies for colorectal cancer: patient response and detection rates. J Natl Cancer Inst 2005;97:347-357.

428. Cotterchio M, Manno M, Klar N, McLaughlin J, Gallinger S. Colorectal screening is associated with reduced colorectal cancer risk: a case-control study within the population-based Ontario Familial Colorectal Cancer Registry. Cancer Causes Control 2005;16:865-875.

429. Schoenfeld P, Cash B, Flood A, Dobhan R, Eastone J, Coyle W. Colonoscopic screening of average-risk women for colorectal neoplasia. N Engl J Med 2005;352:2061-8.

430. Winawer SJ, Stewart ET, Zauber AG, et al. A comparison of colonoscopy and double-contrast barium enema for surveillance after polypectomy. National Polyp Study Work Group. N Engl J Med 2000;342:1766-1772.

431. Philip AK, Meghan GL, Harms B. Computed tomographic colonography. Surg Clin North Am 2011;91:127-39.

432. Zarychanski R, Chen Y, Bernstein CN, Hebert PC. Frequency of colorectal cancer screening and the impact of family physicians on screening behaviour. CMAJ 2007;177:593-597.

433. NIH consensus conference. Adjuvant therapy for patients with colon and rectal cancer. JAMA 1990;264:1444-1450.

434. Benson AB,3rd, Schrag D, Somerfield MR, et al. American Society of Clinical Oncology recommendations on adjuvant chemotherapy for stage II colon cancer. J Clin Oncol 2004;22:3408-3419.

435. Figueredo A, Charette ML, Maroun J, Brouwers MC, Zuraw L. Adjuvant therapy for stage II colon cancer: a systematic review from the Cancer Care Ontario Program in evidence-based care's gastrointestinal cancer disease site group. J Clin Oncol 2004;22:3395-3407.

436. O'Connor ES, Greenblatt DY, LoConte NK, et al. Adjuvant chemotherapy for stage II colon cancer with poor prognostic features. J Clin Oncol 2011;29:3381-3388.

437. Andre T, Boni C, Navarro M, et al. Improved overall survival with oxaliplatin, fluorouracil, and leucovorin as adjuvant treatment in stage II or III colon cancer in the MOSAIC trial. J Clin Oncol 2009;27:3109-3116.

438. Wirtzfeld DA, Mikula L, Gryfe R, et al. Concordance with clinical practice guidelines for adjuvant chemotherapy in patients with stage I-III colon cancer: experience in 2 Canadian provinces. Can J Surg 2009;52:92-97.

439. Jo W, John MC. Chemotherapeutic implications in microsatellite unstable colorectal cancer. Cancer biomarkers : section A of Disease markers 2006;2:51-60.

440. Carethers JM, Hawn MT, Chauhan DP, et al. Competency in mismatch repair prohibits clonal expansion of cancer cells treated with N-methyl-N'-nitro-N-nitrosoguanidine. J Clin Invest 1996;98:199-206.

441. Hawn MT, Umar A, Carethers JM, et al. Evidence for a connection between the mismatch repair system and the G2 cell cycle checkpoint. Cancer Res 1995;55:3721-3725.

442. Casorelli I, Russo MT, Bignami M. Role of mismatch repair and MGMT in response to anticancer therapies. Anticancer Agents Med Chem 2008;8:368-380.

443. Fink D, Nebel S, Aebi S, et al. The role of DNA mismatch repair in platinum drug resistance. Cancer Res 1996;56:4881-4886.

444. de Gramont A, Figer A, Seymour M, et al. Leucovorin and Fluorouracil With or Without Oxaliplatin as First-Line Treatment in Advanced Colorectal Cancer. J Clin Oncol 2000:18:2938-2947.

445. Parker WB, Cheng YC. Metabolism and mechanism of action of 5-fluorouracil. Pharmacol Ther 1990;48:381-395.

446. Major PP, Egan E, Herrick D, Kufe DW. 5-Fluorouracil incorporation in DNA of human breast carcinoma cells. Cancer Res 1982;42:3005-3009.

447. Arnold CN, Goel A, Boland CR. Role of hMLH1 promoter hypermethylation in drug resistance to 5-fluorouracil in colorectal cancer cell lines. Int J Cancer 2003;106:66-73.

448. Plumb JA, Strathdee G, Sludden J, Kaye SB, Brown R. Reversal of drug resistance in human tumor xenografts by 2'-deoxy-5-azacytidine-induced demethylation of the hMLH1 gene promoter. Cancer Res 2000;60:6039-6044.

449. Elsaleh H, Joseph D, Grieu F, Zeps N, Spry N, Iacopetta B. Association of tumour site and sex with survival benefit from adjuvant chemotherapy in colorectal cancer. Lancet 2000;355:1745-1750.

450. Elsaleh H, Powell B, Soontrapornchai P, et al. p53 gene mutation, microsatellite instability and adjuvant chemotherapy: impact on survival of 388 patients with Dukes' C colon carcinoma. Oncology 2000;58:52-59.

451. Bras-Goncalves RA, Rosty C, Laurent-Puig P, Soulie P, Dutrillaux B, Poupon MF. Sensitivity to CPT-11 of xenografted human colorectal cancers as a function of microsatellite instability and p53 status. Br J Cancer 2000;82:913-923.

452. Hemminki A, Mecklin JP, Jarvinen H, Aaltonen LA, Joensuu H. Microsatellite instability is a favorable prognostic indicator in patients with colorectal cancer receiving chemotherapy. Gastroenterology 2000;119:921-928.

453. Ribic CM, Sargent DJ, Moore MJ, et al. Tumor microsatellite-instability status as a predictor of benefit from fluorouracil-based adjuvant chemotherapy for colon cancer. N Engl J Med 2003;349:247-257.

454. Valentini AM, Armentano R, Pirrelli M, Caruso ML. Chemotherapeutic agents for colorectal cancer with a defective mismatch repair system: The state of the art. Cancer Treat Rev 2006;32:607-618.

455. Fedier A, Schwarz VA, Walt H, Carpini RD, Haller U, Fink D. Resistance to topoisomerase poisons due to loss of DNA mismatch repair. Int J Cancer 2001;93:571-576.

456. Rahman P, Jones A, Curtis J, et al. The Newfoundland population: a unique resource for genetic investigation of complex diseases. Hum Mol Genet 2003;12 Spec No 2:R167-72.

457. Mannion JJ, ed. The Peopling of Newfoundland; Essays in Historical Geography. St. John's, Newfoundland: Institute of Social and Economic Research, Memorial University of Newfoundland, 1977.

458. Bear JC, Nemec TF, Kennedy JC, et al. Persistent genetic isolation in outport Newfoundland. Am J Med Genet 1987;27:807-830.

459. Young TL, Woods MO, Parfrey PS, Green JS, Hefferton D, Davidson WS. A Founder Effect in the Newfoundland Population Reduces the Bardet-Biedl Syndrome I (BBS1) Interval to 1 cM. The American Journal of Human Genetics 1999;65:1680-1687.

460. Young TL, Penney L, Woods MO, et al. A Fifth Locus for Bardet-Biedl Syndrome Maps to Chromosome 2q31. The American Journal of Human Genetics 1999;64:900-904.

461. Xie YG, Zheng H, Leggo J, Scully MF, Lillicrap D. A founder factor VIII mutation, valine 2016 to alanine, in a population with an extraordinarily high prevalence of mild hemophilia A. Thromb Haemost 2002;87:178-9.

462. Olufemi SE, Green JS, Manickam P, Guru SC, Agarwal SK, Kester MB. Common ancestral mutation in the MEN1 gene is likely responsible for the prolactinoma variant of MEN1 (MEN1Burin) in four kindreds from Newfoundland. Hum Mutat 1998;11:264-9.

463. Grewal KK, Stefanelli MG, Meijer IA, Hand CK, Rouleau GA, Ives EJ. A founder effect in three large Newfoundland families with a novel clinically variable spastic ataxia and supranuclear gaze palsy. American Journal of Medical Genetics 2004;131:249-54.

464. Kaurah P, MacMillan A, Boyd N, et al. Founder and recurrent CDH1 mutations in families with hereditary diffuse gastric cancer. JAMA 2007;297:2360-2372.

465. Service S, DeYoung J, Karayiorgou M, et al. Magnitude and distribution of linkage disequilibrium in population isolates and implications for genome-wide association studies. Nat Genet 2006;38:556-560.

466. Woods MO, Younghusband HB, Parfrey PS, et al. The genetic basis of colorectal cancer in a population-based incident cohort with a high rate of familial disease. Gut 2010;59:1369-1377.

467. Stuckless S, Parfrey PS, Woods MO, et al. The phenotypic expression of three MSH2 mutations in large Newfoundland families with Lynch syndrome. Fam Cancer 2007;6:1-12.

468. Spirio L, Green J, Robertson J, et al. The identical 5' splice-site acceptor mutation in five attenuated APC families from Newfoundland demonstrates a founder effect. Hum Genet 1999;105:388-398.

469. Green RC, Green JS, Buehler SK, Robb JD, Daftary D, Gallinger S. Very high incidence of familial colorectal cancer in Newfoundland: a comparison with Ontario and 13 other population-based studies. Fam Cancer 2007;6:53-62.

470. Cotterchio M, McKeown-Eyssen G, Sutherland H, et al. Ontario familial colon cancer registry: methods and first-year response rates. Chronic Dis Can 2000;21:81-86.

471. Newcomb PA, Baron J, Cotterchio M, et al. Colon Cancer Family Registry: an international resource for studies of the genetic epidemiology of colon cancer. Cancer Epidemiol Biomarkers Prev 2007;16:2331-2343.

472. McLaughlin JR, Dryer D, Fields ALA, Mao Y, Villeneuve G, Whylie B. Canadian Cancer Statistics 2000. Canadian Cancer Society, National Cancer Institute of Canada, Statistics Canada, Provincial/Territorial Cancer Registries, and Health Canada, 2000.

473. Aaltonen LA, Salovaara R, Kristo P, et al. Incidence of hereditary nonpolyposis colorectal cancer and the feasibility of molecular screening for the disease. N Engl J Med 1998;338:1481-1487.

474. Kim H, Jen J, Vogelstein B, Hamilton SR. Clinical and pathological characteristics of sporadic colorectal carcinomas with DNA replication errors in microsatellite sequences. Am J Pathol 1994;145:148-156.

475. de la Chapelle A. Testing tumors for microsatellite instability. Eur J Hum Genet 1999;7:407-408.

476. Mangold E, Pagenstecher C, Friedl W, et al. Tumours from MSH2 mutation carriers show loss of MSH2 expression but many tumours from MLH1 mutation carriers exhibit weak positive MLH1 staining. J Pathol 2005;207:385-395.

477. Raevaara TE, Vaccaro C, Abdel-Rahman WM, et al. Pathogenicity of the hereditary colorectal cancer mutation hMLH1 del616 linked to shortage of the functional protein. Gastroenterology 2003;125:501-509.

478. Zighelboim I, Powell MA, Babb SA, Whelan AJ, Schmidt AP, Clendenning M. Epitope-positive truncating MLH1 mutation and loss of PMS2: implications for IHC-directed genetic testing for Lynch syndrome. Familial cancer 2009;8:501-4.

479. Barrow E, Alduaij W, Robinson L, et al. Colorectal cancer in HNPCC: cumulative lifetime incidence, survival and tumour distribution. A report of 121 families with proven mutations. Clin Genet 2008;74:233-42.

480. Iacopetta B, Li WQ, Grieu F, Ruszkiewicz A, Kawakami K. BRAF mutation and gene methylation frequencies of colorectal tumours with microsatellite instability increase markedly with patient age. Gut 2006;55:1213-1214.

481. Wish TA, Hyde AJ, Parfrey PS, et al. Increased cancer predisposition in family members of colorectal cancer patients harboring the p.V600E BRAF mutation: a population-based study. Cancer Epidemiol Biomarkers Prev 2010;19:1831-1839.

482. Wijnen J, van der Klift H, Vasen H, et al. MSH2 genomic deletions are a frequent cause of HNPCC. Nat Genet 1998;20:326-328.

483. Nakagawa H, Hampel H, de la Chapelle A. Identification and characterization of genomic rearrangements of MSH2 and MLH1 in Lynch syndrome (HNPCC) by novel techniques. Hum Mutat 2003;22:258.

484. Taylor CF, Charlton RS, Burn J, Sheridan E, Taylor GR. Genomic deletions in MSH2 or MLH1 are a frequent cause of hereditary non-polyposis colorectal cancer: identification of novel and recurrent deletions by MLPA. Hum Mutat 2003;22:428-433.

485. Jaskowski L, Young J, Jackson L, et al. Stability of BAT26 in Lynch syndrome colorectal tumours. Eur J Hum Genet 2007;15:139-41; author reply 141-2.

486. Pastrello C, Baglioni S, Tibiletti MG, et al. Stability of BAT26 in tumours of hereditary nonpolyposis colorectal cancer patients with MSH2 intragenic deletion. Eur J Hum Genet 2006;14:63-68.

487. Cannon-Albright LA, Skolnick MH, Bishop DT, Lee RG, Burt RW. Common inheritance of susceptibility to colonic adenomatous polyps and associated colorectal cancers. N Engl J Med 1988;319:533-537.

488. Houlston RS, Collins A, Slack J, Morton NE. Dominant genes for colorectal cancer are not rare. Ann Hum Genet 1992;56:99-103.

489. Lichtenstein P, Holm NV, Verkasalo PK, et al. Environmental and heritable factors in the causation of cancer--analyses of cohorts of twins from Sweden, Denmark, and Finland. N Engl J Med 2000;343:78-85.

490. Giardiello FM, Brensinger JD, Petersen GM. AGA technical review on hereditary colorectal cancer and genetic testing. Gastroenterology 2001;121:198-213.

491. Park JG, Vasen HF, Park KJ, et al. Suspected hereditary nonpolyposis colorectal cancer: International Collaborative Group on Hereditary Non-Polyposis Colorectal Cancer (ICG-HNPCC) criteria and results of genetic diagnosis. Dis Colon Rectum 1999;42:710-5; discussion 715-6.

492. Heinimann K, Scott RJ, Buerstedde JM, et al. Influence of selection criteria on mutation detection in patients with hereditary nonpolyposis colorectal cancer. Cancer 1999;85:2512-2518.

493. Wei SC, Yu CY, Tsai-Wu JJ, et al. Low mutation rate of hMSH2 and hMLH1 in Taiwanese hereditary non-polyposis colorectal cancer. Clin Genet 2003;64:243-251.

494. Lynch HT, Smyrk TC, Watson P, et al. Genetics, natural history, tumor spectrum, and pathology of hereditary nonpolyposis colorectal cancer: an updated review. Gastroenterology 1993;104:1535-1549.

495. Salovaara R, Loukola A, Kristo P, et al. Population-based molecular detection of hereditary nonpolyposis colorectal cancer. J Clin Oncol 2000;18:2193-2200.

496. Samowitz WS, Curtin K, Lin HH, et al. The colon cancer burden of genetically defined hereditary nonpolyposis colon cancer. Gastroenterology 2001;121:830-838.

497. Akiyama Y, Sato H, Yamada T, et al. Germ-line mutation of the hMSH6/GTBP gene in an atypical hereditary nonpolyposis colorectal cancer kindred. Cancer Res 1997;57:3920-3923.

498. Liu T, Yan H, Kuismanen S, et al. The role of hPMS1 and hPMS2 in predisposing to colorectal cancer. Cancer Res 2001;61:7798-7802.

499. De Rosa M, Fasano C, Panariello L, et al. Evidence for a recessive inheritance of Turcot's syndrome caused by compound heterozygous mutations within the PMS2 gene. Oncogene 2000;19:1719-1723.

500. Nakagawa H, Lockman JC, Frankel WL, et al. Mismatch repair gene PMS2: disease-causing germline mutations are frequent in patients whose tumors stain negative for PMS2 protein, but paralogous genes obscure mutation detection and interpretation. Cancer Res 2004;64:4721-4727.

501. Liu HX, Zhou XL, Liu T, et al. The role of hMLH3 in familial colorectal cancer. Cancer Res 2003;63:1894-1899.

502. Hienonen T, Laiho P, Salovaara R, et al. Little evidence for involvement of MLH3 in colorectal cancer predisposition. Int J Cancer 2003;106:292-296.

503. de Jong MM, Hofstra RM, Kooi KA, et al. No association between two MLH3 variants (S845G and P844L)and colorectal cancer risk. Cancer Genet Cytogenet 2004;152:70-71.

504. Martin LJ, Crawford MH, Koertvelyessy T, Keeping D, Collins M, Huntsman R. The population structure of ten Newfoundland outports. Hum Biol 2000;72:997-1016.

505. Parfrey PS, Davidson WS, Green JS. Clinical and genetic epidemiology of inherited renal disease in Newfoundland. Kidney Int 2002;61:1925-1934.

506. Green J, O'Driscoll M, Barnes A, et al. Impact of gender and parent of origin on the phenotypic expression of hereditary nonpolyposis colorectal cancer in a large Newfoundland kindred with a common MSH2 mutation. Dis Colon Rectum 2002;45:1223-1232.

507. Thibodeau SN, French AJ, Cunningham JM, et al. Microsatellite instability in colorectal cancer: different mutator phenotypes and the principal involvement of hMLH1. Cancer Res 1998;58:1713-1718.

508. Bapat BV, Madlensky L, Temple LK, et al. Family history characteristics, tumor microsatellite instability and germline MSH2 and MLH1 mutations in hereditary colorectal cancer. Hum Genet 1999;104:167-176.

509. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 1988;16:1215.

510. Schouten JP, McElgunn CJ, Waaijer R, Zwijnenburg D, Diepvens F, Pals G. Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. Nucleic Acids Res 2002;30:e57.

511. Deng G, Chen A, Hong J, Chae HS, Kim YS. Methylation of CpG in a small region of the hMLH1 promoter invariably correlates with the absence of gene expression. Cancer Res 1999;59:2029-2033.

512. Herman JG, Umar A, Polyak K, et al. Incidence and functional consequences of hMLH1 promoter hypermethylation in colorectal carcinoma. Proc Natl Acad Sci U S A 1998;95:6870-6875.

513. Veigl ML, Kasturi L, Olechnowicz J, et al. Biallelic inactivation of hMLH1 by epigenetic gene silencing, a novel mechanism causing human MSI cancers. Proc Natl Acad Sci U S A 1998;95:8698-8702.

514. Huang J, Kuismanen SA, Liu T, et al. MSH6 and MSH3 are rarely involved in genetic predisposition to nonpolypotic colon cancer. Cancer Res 2001;61:1619-1623.

515. Thiffault I, Hamel N, Pal T, et al. Germline truncating mutations in both MSH2 and BRCA2 in a single kindred. Br J Cancer 2004;90:483-491.

516. Wang Y, Friedl W, Lamberti C, et al. Hereditary nonpolyposis colorectal cancer: frequent occurrence of large genomic deletions in MSH2 and MLH1 genes. Int J Cancer 2003;103:636-641.

517. Ng PC, Henikoff S. Accounting for human polymorphisms predicted to affect protein function. Genome Res 2002;12:436-446.

518. Moslein G, Tester DJ, Lindor NM, et al. Microsatellite instability and mutation analysis of hMSH2 and hMLH1 in patients with sporadic, familial and hereditary colorectal cancer. Hum Mol Genet 1996;5:1245-1252.

519. Liu B, Parsons R, Papadopoulos N, et al. Analysis of mismatch repair genes in hereditary non-polyposis colorectal cancer patients. Nat Med 1996;2:169-174.

520. Borresen AL, Lothe RA, Meling GI, et al. Somatic mutations in the hMSH2 gene in microsatellite unstable colorectal carcinomas. Hum Mol Genet 1995;4:2065-2072.

521. Viel A, Genuardi M, Capozzi E, et al. Characterization of MSH2 and MLH1 mutations in Italian families with hereditary nonpolyposis colorectal cancer. Genes Chromosomes Cancer 1997;18:8-18.

522. Studamire B, Quach T, Alani E. Saccharomyces cerevisiae Msh2p and Msh6p ATPase activities are both required during mismatch repair. Mol Cell Biol 1998;18:7590-7601.

523. Kolodner RD, Tytell JD, Schmeits JL, et al. Germ-line msh6 mutations in colorectal cancer families. Cancer Res 1999;59:5068-5074.

524. Ellison AR, Lofing J, Bitter GA. Human MutL homolog (MLH1) function in DNA mismatch repair: a prospective screen for missense mutations in the ATPase domain. Nucleic Acids Res 2004;32:5321-5338.

525. Nystrom-Lahti M, Wu Y, Moisio AL, et al. DNA mismatch repair gene mutations in 55 kindreds with verified or putative hereditary non-polyposis colorectal cancer. Hum Mol Genet 1996;5:763-769.

526. Weber TK, Conlon W, Petrelli NJ, et al. Genomic DNA-based hMSH2 and hMLH1 mutation screening in 32 Eastern United States hereditary nonpolyposis colorectal cancer pedigrees. Cancer Res 1997;57:3798-3803.

527. Croitoru ME, Cleary SP, Di Nicola N, et al. Association between biallelic and monoallelic germline MYH gene mutations and colorectal cancer risk. J Natl Cancer Inst 2004;96:1631-1634.

528. McGarrity TJ, Ruggiero FM, Chey WY, Bajaj R, Kelly JE, Kauffman GL,Jr. Giant fundic polyp complicating attenuated familial adenomatous polyposis. Am J Gastroenterol 2000;95:1824-1828.

529. Scott RJ, McPhillips M, Meldrum CJ, et al. Hereditary nonpolyposis colorectal cancer in 95 families: differences and similarities between mutation-positive and mutation-negative kindreds. Am J Hum Genet 2001;68:118-127.

530. Chan TL, Curtis LC, Leung SY, et al. Early-onset colorectal cancer with stable microsatellite DNA and near-diploid chromosomes. Oncogene 2001;20:4871-4876.

531. Wan WH, Fortuna MB, Furmanski P. A rapid and efficient method for testing immunohistochemical reactivity of monoclonal antibodies against multiple tissue samples simultaneously. J Immunol Methods 1987;103:121-129.

532. Kononen J, Bubendorf L, Kallioniemi A, et al. Tissue microarrays for high-throughput molecular profiling of tumor specimens. Nat Med 1998;4:844-847.

533. Camp RL, Neumeister V, David LR. A decade of tissue microarrays: progress in the discovery and validation of cancer biomarkers. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2008;26:5630-7.

534. Conway C, Dobson L, O'Grady A, Kay E, Costello S, O'Shea D. Virtual microscopy as an enabler of automated/quantitative assessment of protein expression in TMAs. Histochem Cell Biol 2008;130:447-463.

535. Shergill IS, Shergill NK, Arya M, Patel HR. Tissue microarrays: a current medical research tool. Curr Med Res Opin 2004;20:707-712.

536. Camp RL, Charette LA, Rimm DL. Validation of tissue microarray technology in breast carcinoma. Lab Invest 2000;80:1943-1949.

537. Torhorst J, Bucher C, Kononen J, et al. Tissue microarrays for rapid linking of molecular changes to clinical endpoints. Am J Pathol 2001;159:2249-2256.

538. Jensen KC, Mariappan MR, Putcha GV, et al. Microsatellite instability and mismatch repair protein defects in ovarian epithelial neoplasms in patients 50 years of age and younger. Am J Surg Pathol 2008;32:1029-1037.

539. Liu CL, Prapong W, Natkunam Y, et al. Software tools for high-throughput analysis and archiving of immunohistochemistry staining data obtained with tissue microarrays. Am J Pathol 2002;161:1557-1565.

540. Atkinson AJ, Colburn WA, DeGruttola VG, et al. Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework*. Clin Pharmacol Ther 2001;69:89-95.

541. Liotta LA, Ferrari M, Petricoin E. Clinical proteomics: Written in blood. Nature 2003;425:905-905.

542. Kulasingam V, Diamandis EP. Strategies for discovering novel cancer biomarkers through utilization of emerging technologies. Nat Clin Prac Oncol 2008;5:588-599.

543. van de Rijn M, Gilks CB. Applications of microarrays to histopathology. Histopathology 2004;44:97-108.

544. Hartwell LH, Culotti J, Pringle JR, Reid BJ. Genetic control of the cell division cycle in yeast. Science 1974;183:46-51.

545. Hartwell LH, Weinert TA. Checkpoints: controls that ensure the order of cell cycle events. Science 1989;246:629-634.

546. Pardee AB. G1 events and regulation of cell proliferation. Science 1989;246:603-608.

547. Sherr CJ. G1 phase progression: cycling on cue. Cell 1994;79:551-555.

548. Hartwell L. Defects in a cell cycle checkpoint may be responsible for the genomic instability of cancer cells. Cell 1992;71:543-546.

549. Collins I, Garrett MD. Targeting the cell division cycle in cancer: CDK and cell cycle checkpoint kinase inhibitors. Curr Opin Pharmacol 2005;5:366-373.

550. Tumor Protein p53; TP53 - OMIM Result. Available at: http://www.ncbi.nlm.nih.gov/omim/191170. Accessed 1/24/2011, 2011.

551. Rodrigues NR, Rowan A, Smith ME, et al. P53 Mutations in Colorectal Cancer. Proc Natl Acad Sci U S A 1990;87:7555-7559.

552. Levine AJ, Momand J, Finlay CA. The p53 tumour suppressor gene. Nature 1991;351:453-456.

553. Vogelstein B, Lane D, Levine AJ. Surfing the p53 network. Nature 2000;408:307-310.

554. Lee JM, Bernstein A. p53 mutations increase resistance to ionizing radiation. Proceedings of the National Academy of Sciences of the United States of America 1993;90:5742-5746.

555. Seoane J, Le HV, Massague J. Myc suppression of the p21(Cip1) Cdk inhibitor influences the outcome of the p53 response to DNA damage. Nature 2002;419:729-734.

556. El-Deiry WS, Tokino T, Velculescu VE, et al. WAF1, a potential mediator of p53 tumor suppression. Cell 1993;75:817-825.

557. Wade Harper J, Adami GR, Wei N, Keyomarsi K, Elledge SJ. The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases. Cell 1993;75:805-816.

558. Maddika S, Ande SR, Panigrahi S, et al. Cell survival, cell death and cell cycle pathways are interconnected: implications for cancer therapy. Drug Resist Updat 2007;10:13-29.

559. Xiong Y, Zhang H, Beach D. D type cyclins associate with multiple protein kinases and the DNA replication and repair factor PCNA. Cell 1992;71:505-514.

560. Li R, Waga S, Hannon GJ, Beach D, Stillman B. Differential effects by the p21 CDK inhibitor on PCNA-dependent DNA replication and repair. Nature 1994;371:534-537.

561. Bunz F, Dutriaux A, Lengauer C, et al. Requirement for p53 and p21 to Sustain G2 Arrest After DNA Damage. Science 1998;282:1497-1501.

562. Allegra CJ, Paik S, Colangelo LH, et al. Prognostic value of thymidylate synthase, Ki-67, and p53 in patients with Dukes' B and C colon cancer: a National Cancer Institute-National Surgical Adjuvant Breast and Bowel Project collaborative study. J Clin Oncol 2003;21:241-250.

563. Zavrides H, Zizi-Sermpetzoglou A, Elemenoglou I, et al. Immunohistochemical expression of bcl-2 in Dukes' stage B and C colorectal carcinoma patients: correlation with p53 and ki-67 in evaluating prognostic significance. Pol J Pathol 2005;56:179-185.

564. Ismail HM, El-Baradie M, Moneer M, Khorshid O, Touny A. Clinico-Pathological and Prognostic Significance of p53, Bcl-2 and Her-2/neu Protein Markers in Colorectal Cancer Using Tissue Microarray. J Egypt Natl Canc Inst 2007;19:3-14.

565. Ahnen DJ, Feigl P, Quan G, et al. Ki-ras mutation and p53 overexpression predict the clinical behavior of colorectal cancer: a Southwest Oncology Group study. Cancer Res 1998;58:1149-1158.

566. Garrity MM, Burgart LJ, Mahoney MR, et al. Prognostic value of proliferation, apoptosis, defective DNA mismatch repair, and p53 overexpression in patients with resected Dukes' B2 or C colon cancer: a North Central Cancer Treatment Group Study. J Clin Oncol 2004;22:1572-1582.

567. Allegra CJ, Parr AL, Wold LE, et al. Investigation of the prognostic and predictive value of thymidylate synthase, p53, and Ki-67 in patients with locally advanced colon cancer. J Clin Oncol 2002;20:1735-1743.

568. Zlobec I, Baker K, Terracciano L, et al. Two-marker protein profile predicts poor prognosis in patients with early rectal cancer. Br J Cancer 2008;99:1712-1717.

569. Cohen T, Prus D, Shia J, et al. Expression of P53, P27 and KI-67 in colorectal cancer patients of various ethnic origins: clinical and tissue microarray based analysis. J Surg Oncol 2008;97:416-422.

570. Stott FJ, Bates S, James MC, et al. The alternative product from the human CDKN2A locus, p14(ARF), participates in a regulatory feedback loop with p53 and MDM2. EMBO J 1998;17:5001-5014.

571. Collado M, Blasco MA, Serrano M. Cellular senescence in cancer and aging. Cell 2007;130:223-233.

572. Hayflick L, Moorhead PS. The serial cultivation of human diploid cell strains. Exp Cell Res 1961;25:585-621.

573. Hayflick L. The limited in vitro lifetime of human diploid cell strains, Exp Cell Res 1965;37:614-636.

574. Serrano M, Lin AW, McCurrach ME, Beach D, Lowe SW. Oncogenic ras provokes premature cell senescence associated with accumulation of p53 and p16INK4a. Cell 1997;88:593-602.

575. Lin AW, Barradas M, Stone JC, van Aelst L, Serrano M, Lowe SW. Premature senescence involving p53 and p16 is activated in response to constitutive MEK/MAPK mitogenic signaling. Genes Dev 1998;12:3008-3019.

576. Zhu J, Woods D, McMahon M, Bishop JM. Senescence of human fibroblasts induced by oncogenic Raf. Genes Dev 1998;12:2997-3007.

577. Ohtani N, Zebedee Z, Huot TJ, et al. Opposing effects of Ets and Id proteins on p16INK4a expression during cellular senescence. Nature 2001;409:1067-1070.

578. Michaloglou C, Vredeveld LC, Soengas MS, et al. BRAFE600-associated senescence-like cell cycle arrest of human naevi. Nature 2005;436:720-724.

579. Collado M, Gil J, Efeyan A, et al. Tumour biology: senescence in premalignant tumours. Nature 2005;436:642.

580. Herman JG, Merlo A, Mao L, et al. Inactivation of the CDKN2/p16/MTS1 gene is frequently associated with aberrant DNA methylation in all common human cancers. Cancer Res 1995;55:4525-4530.

581. Shima K, Nosho K, Baba Y, et al. Prognostic significance of CDKN2A (p16) promoter methylation and loss of expression in 902 colorectal cancers: Cohort study and literature review. Int J Cancer 2011;128:1080-1094.

582. Mitomi H, Fukui N, Tanaka N, et al. Aberrant p16((INK4a)) methylation is a frequent event in colorectal cancers: prognostic value and relation to mRNA expression and immunoreactivity. J Cancer Res Clin Oncol 2010;136:323-331.

583. Paya A, Alenda C, Perez-Carbonell L, et al. Utility of p16 immunohistochemistry for the identification of Lynch syndrome. Clin Cancer Res 2009;15:3156-3162.

584. Xie D, Sham JS, Zeng WF, et al. Heterogeneous expression and association of beta-catenin, p16 and c-myc in multistage colorectal tumorigenesis and progression detected by tissue microarray. Int J Cancer 2003;107:896-902.

585. Maeda K, Kawakami K, Ishida Y, Ishiguro K, Omura K, Watanabe G. Hypermethylation of the CDKN2A gene in colorectal cancer is associated with shorter survival. Oncol Rep 2003;10:935-938.

586. Sanz-Casla MT, Maestro ML, Vidaurreta M, Maestro C, Arroyo M, Cerdan J. p16 Gene methylation in colorectal tumors: correlation with clinicopathological features and prognostic value. Dig Dis 2005;23:151-155.

587. Shen L, Catalano PJ, Benson AB,3rd, O'Dwyer P, Hamilton SR, Issa JP. Association between DNA methylation and shortened survival in patients with advanced colorectal cancer treated with 5-fluorouracil based chemotherapy. Clin Cancer Res 2007;13:6093-6098.

588. Grau AM, Zhang L, Wang W, et al. Induction of p21waf1 expression and growth inhibition by transforming growth factor beta involve the tumor suppressor gene DPC4 in human pancreatic adenocarcinoma cells. Cancer Res 1997;57:3929-3934.

589. Zhang H, Hannon GJ, Beach D. P21-Containing Cyclin Kinases Exist in both Active and Inactive States. Genes Dev 1994;8:1750-1758.

590. Cheng M, Olivier P, Diehl JA, et al. The p21(Cip1) and p27(Kip1) CDK 'inhibitors' are essential activators of cyclin D-dependent kinases in murine fibroblasts. EMBO J 1999;18:1571-1583.

591. LaBaer J, Garrett MD, Stevenson LF, et al. New functional activities for the p21 family of CDK inhibitors. Genes Dev 1997;11:847-862.

592. Jung Y, Qian Y, Chen X. Examination of the expanding pathways for the regulation of p21 expression and activity. Cell Signal 2010;22:1003-1012.

593. Shiohara M, el-Deiry WS, Wada M, et al. Absence of WAF1 mutations in a variety of human malignancies. Blood 1994;84:3781-3784.

594. Davis T, Singhrao SK, Wyllie FS, et al. Telomere-based proliferative lifespan barriers in Werner-syndrome fibroblasts involve both p53-dependent and p53-independent mechanisms. J Cell Sci 2003;116:1349-1357.

595. Ropponen KM, Kellokoski JK, Lipponen PK, et al. p21/WAF1 expression in human colorectal carcinoma: association with p53, transcription factor AP-2 and prognosis. Br J Cancer 1999;81:133-140.

596. Holland TA, Elder J, McCloud JM, et al. Subcellular localisation of cyclin D1 protein in colorectal tumours is associated with p21(WAF1/CIP1) expression and correlates with patient survival. Int J Cancer 2001;95:302-306.

597. Tornillo L, Lugli A, Zlobec I, et al. Prognostic value of cell cycle and apoptosis regulatory proteins in mismatch repair-proficient colorectal cancer: a tissue microarray-based approach. Am J Clin Pathol 2007;127:114-123.

598. Ioachim E. Expression patterns of cyclins D1, E and cyclin-dependent kinase inhibitors p21waf1/cip1, p27kip1 in colorectal carcinoma: correlation with other cell cycle regulators (pRb, p53 and Ki-67 and PCNA) and clinicopathological features. Int J Clin Pract 2008;62:1736-1743.

599. Zlobec I, Baker K, Terracciano LM, Lugli A. RHAMM, p21 Combined Phenotype Identifies Microsatellite Instability-High Colorectal Cancers with a Highly Adverse Prognosis. Clin Cancer Res 2008;14:3798-3806.

600. Cacina C, Ilhan Yaylim-Eraltan, Arikan S, Esra KS, Zeybek U, Isbir T. Association between CDKN1A Ser31Arg and C20T gene polymorphisms and colorectal cancer risk and prognosis. In Vivo 2010;24:179-83.

601. Polyak K, Kato JY, Solomon MJ, et al. p27Kip1, a cyclin-Cdk inhibitor, links transforming growth factor-beta and contact inhibition to cell cycle arrest. Genes Dev 1994;8:9-22.

602. Polyak K, Lee MH, Erdjument-Bromage H, et al. Cloning of p27Kip1, a cyclin-dependent kinase inhibitor and a potential mediator of extracellular antimitogenic signals. Cell 1994;78:59-66.

603. Toyoshima H, Hunter T. p27, a novel inhibitor of G1 cyclin-Cdk protein kinase activity, is related to p21. Cell 1994;78:67-74.

604. Lee J, Sung SK. The function of p27 KIP1 during tumor development. Exp Mol Med 2009;41:765-71.

605. Tsihlias J, Kapusta L, Slingerland J. The prognostic significance of altered cyclin-dependent kinase inhibitors in human cancer. Annu Rev Med 1999;50:401-423.

606. Porter PL, Malone KE, Heagerty PJ, et al. Expression of cell-cycle regulators p27Kip1 and cyclin E, alone and in combination, correlate with survival in young breast cancer patients. Nat Med 1997;3:222-225.

607. Tan P, Cady B, Wanner M, et al. The cell cycle inhibitor p27 is an independent prognostic marker in small (T1a,b) invasive breast carcinomas. Cancer Res 1997;57:1259-1263.

608. Catzavelos C, Bhattacharya N, Ung YC, et al. Decreased levels of the cell-cycle inhibitor p27Kip1 protein: prognostic implications in primary breast cancer. Nat Med 1997;3:227-230.

609. Tsihlias J, Kapusta LR, DeBoer G, et al. Loss of cyclin-dependent kinase inhibitor p27Kip1 is a novel prognostic factor in localized human prostate adenocarcinoma. Cancer Res 1998;58:542-548.

610. Yang RM, Naitoh J, Murphy M, et al. Low p27 expression predicts poor disease-free survival in patients with prostate cancer. J Urol 1998;159:941-945.

611. Mori M, Mimori K, Shiraishi T, et al. P27 Expression and Gastric Carcinoma. Nat Med 1997;3:593.

612. Esposito V, Baldi A, De Luca A, et al. Prognostic role of the cyclin-dependent kinase inhibitor p27 in non-small cell lung cancer. Cancer Res 1997;57:3381-3385.

613. Singh SP, Lipman J, Goldman H, et al. Loss or altered subcellular localization of p27 in Barrett's associated adenocarcinoma. Cancer Res 1998;58:1730-1735.

614. Pellegata NS, Quintanilla-Martinez L, Siggelkow H, Samson E, Bink K, Hofler H. Germ-line mutations in p27Kip1 cause a multiple endocrine neoplasia syndrome in rats and humans. Proc Natl Acad Sci U S A 2006;103:15558-63.

615. Georgitsi M. MEN-4 and other multiple endocrine neoplasias due to cyclin-dependent kinase inhibitors (p27Kip1 and p18INK4C) mutations. Best Practice & Research Clinical Endocrinology & Metabolism 2010;24:425-437.

616. Loda M, Cukor B, Tam SW, et al. Increased proteasome-dependent degradation of the cyclin-dependent kinase inhibitor p27 in aggressive colorectal carcinomas. Nat Med 1997;3:231-234.

617. Bottini C, Platini F, Rinaldi M, et al. p27Kip1 is inactivated in human colorectal cancer by cytoplasmic localization associated with activation of Akt/PKB. Int J Oncol 2009;34:69-77.

618. Sarli L, Bottarelli L, Azzoni C, et al. Loss of p27 expression and microsatellite instability in sporadic colorectal cancer. Surg Oncol 2006;15:97-106.

619. Weinberg RA. The retinoblastoma protein and cell cycle control. Cell 1995;81:323-330.

620. Harbour JW, Luo RX, Dei Santi A, Postigo AA, Dean DC. Cdk phosphorylation triggers sequential intramolecular interactions that progressively block Rb functions as cells move through G1. Cell 1999;98:859-869.

621. Alt JR, Gladden AB, Diehl JA. p21(Cip1) Promotes cyclin D1 nuclear accumulation via direct inhibition of nuclear export. J Biol Chem 2002;277:8517-8523.

622. Diehl JA. Cycling to cancer with cyclin D1. Cancer Biol Ther 2002;1:226-231.

623. Motokura T, Bloom T, Kim HG, et al. A novel cyclin encoded by a bcl1-linked candidate oncogene. Nature 1991;350:512-515.

624. Williams ME, Swerdlow SH, Rosenberg CL, Arnold A. Chromosome 11 translocation breakpoints at the PRAD1/cyclin D1 gene locus in centrocytic lymphoma. Leukemia 1993;7:241-245.

625. Bergsagel PL, Kuehl WM. Chromosome translocations in multiple myeloma. Oncogene 2001;20:5611-5622.

626. Yamanouchi H, Furihata M, Fujita J, et al. Expression of cyclin E and cyclin D1 in non-small cell lung cancers. Lung Cancer 2001;31:3-8.

627. Jin M, Inoue S, Umemura T, et al. Cyclin D1, p16 and retinoblastoma gene product expression as a predictor for prognosis in non-small cell lung cancer at stages I and II. Lung Cancer 2001;34:207-218.

628. Zhu J, Yu L, Zhan P, Song Y, Wang Q. The relationships between cyclin D1 expression and prognosis of non-small cell lung cancer. Chin J Lung Cancer 2010;13:803-808.

629. Namazie A, Sassan A, Olopade OI, et al. Cyclin D1 amplification and p16(MTS1/CDK4I) deletion correlate with poor prognosis in head and neck tumors. Laryngoscope 2002;112:472-481.

630. Kaminagakura E, Werneck da Cunha I, Soares FA, Nishimoto IN, Kowalski LP. CCND1 amplification and protein overexpression in oral squamous cell carcinoma of young patients. Head Neck 2011;.

631. Gillett C, Fantl V, Smith R, et al. Amplification and overexpression of cyclin D1 in breast cancer detected by immunohistochemical staining. Cancer Res 1994;54:1812-1817.

632. Gansauge S, Gansauge F, Ramadani M, et al. Overexpression of cyclin D1 in human pancreatic carcinoma is associated with poor prognosis. Cancer Res 1997;57:1634-1637.

633. Reis-Filho JS, Savage K, Lambros MBK, et al. Cyclin D1 protein overexpression and CCND1 amplification in breast carcinomas: an immunohistochemical and chromogenic in situ hybridisation analysis. Mod Pathol 2006;19:999-1009.

634. Pankaj RG, Pratt N, Purdie CA, et al. High CCND1 amplification identifies a group of poor prognosis women with estrogen receptor positive breast cancer. Int J Cancer 2010;127:355-360.

635. Arber N, Hibshoosh H, Moss SF, et al. Increased expression of cyclin D1 is an early event in multistage colorectal carcinogenesis. Gastroenterology 1996;110:669-674.

636. Ogino S, Nosho K, Irahara N, et al. A cohort study of cyclin D1 expression and prognosis in 602 colon cancer cases. Clin Cancer Res 2009;15:4431-4438.

637. Maeda K, Chung YS, Kang SM, et al. Overexpression of cyclin D1 and p53 associated with disease recurrence in colorectal adenocarcinoma. Int J Cancer 1997;74:310-315.

638. Bahnassy AA, Zekri AR, El-Houssini S, et al. Cyclin A and cyclin D1 as significant prognostic markers in colorectal cancer patients. BMC Gastroenterol 2004;4:22.

639. Bukholm IK, Nesland JM. Protein expression of p53, p21 (WAF1/CIP1), bcl-2, Bax, cyclin D1 and pRb in human colon carcinomas. Virchows Arch 2000;436:224-228.

640. Bhatavdekar JM, Patel DD, Chikhlikar PR, et al. Molecular markers are predictors of recurrence and survival in patients with Dukes B and Dukes C colorectal adenocarcinoma. Dis Colon Rectum 2001;44:523-533.

641. McKay JA, Douglas JJ, Ross VG, et al. Analysis of key cell-cycle checkpoint proteins in colorectal tumours. J Pathol 2002;196:386-393.

642. Knosel T, Emde A, Schluns K, et al. Immunoprofiles of 11 biomarkers using tissue microarrays identify prognostic subgroups in colorectal cancer. Neoplasia 2005;7:741-747.

643. Hilska M, Collan YU, O Laine VJ, et al. The significance of tumor markers for proliferation and apoptosis in predicting survival in colorectal cancer. Dis Colon Rectum 2005;48:2197-2208.

644. Derenzini M, Trere D, Pession A, Govoni M, Sirri V, Chieco P. Nucleolar size indicates the rapidity of cell proliferation in cancer tissues. J Pathol 2000;191:181-186.

645. Gerdes J, Lemke H, Baisch H, Wacker H, Schwab U, Stein H. Cell cycle analysis of a cell proliferation-associated human nuclear antigen defined by the monoclonal antibody Ki-67. J Immunol 1984;133:1710-1715.

646. Hofmann K, Bucher P. The FHA domain: a putative nuclear signalling domain found in protein kinases and transcription factors. Trends Biochem Sci 1995;20:347-349.

647. Schluter C, Duchrow M, Wohlenberg C, et al. The cell proliferation-associated antigen of antibody Ki-67: a very large, ubiquitous nuclear protein with numerous repeated elements, representing a new kind of cell cycle-maintaining proteins. J Cell Biol 1993;123:513-522.

648. Scholzen T, Gerdes J. The Ki-67 protein: From the known and the unknown. J Cell Physiol 2000;182:311-322.

649. Hall PA, McKee PH, Menage HD, Dover R, Lane DP. High levels of p53 protein in UV-irradiated normal human skin. Oncogene 1993;8:203-207.

650. Georgescu CV, Saftoiu A, Georgescu CC, Ciurea R, Ciurea T. Correlations of proliferation markers, p53 expression and histological findings in colorectal carcinoma. J Gastrointestin Liver Dis 2007;16:133-139.

651. Ihmann T, Liu J, Schwabe W, et al. High-level mRNA quantification of proliferation marker pKi-67 is correlated with favorable prognosis in colorectal carcinoma. J Cancer Res Clin Oncol 2004;130:749-756.

652. Duchrow M, Ziemann T, Windhovel U, Bruch HP, Broll R. Colorectal carcinomas with high MIB-1 labelling indices but low pKi67 mRNA levels correlate with better prognostic outcome. Histopathology 2003;42:566-574.

653. Brown DC, Gatter KC. Ki67 protein: the immaculate deception? Histopathology 2002;40:2-11.

654. Funaioli C, Pinto C, Mutri V, Di Fabio F, Ceccarelli C, Martoni AA. Does biomolecular characterization of stage II/III colorectal cancer have any prognostic value? Clin Colorectal Cancer 2006;6:38-45.

655. Rosati G, Chiacchio R, Reggiardo G, De Sanctis D, Manzione L. Thymidylate synthase expression, p53, bcl-2, Ki-67 and p27 in colorectal cancer: relationships with tumor recurrence and survival. Tumour Biol 2004;25:258-263.

656. Palmqvist R, Sellberg P, Oberg A, Tavelin B, Rutegard JN, Stenling R. Low tumour cell proliferation at the invasive margin is associated with a poor prognosis in Dukes' stage B colorectal cancers. Br J Cancer 1999;79:577-581.

657. Kimura T, Tanaka S, Haruma K, et al. Clinical significance of MUC1 and E-cadherin expression, cellular proliferation, and angiogenesis at the deepest invasive portion of colorectal cancer. Int J Oncol 2000;16:55-64.

658. Canna K, Hilmy M, McMillan DC, et al. The relationship between tumour proliferative activity, the systemic inflammatory response and survival in patients undergoing curative resection for colorectal cancer. Colorectal Disease 2008;10:663-667.

659. Anjomshoaa A, Lin YH, Black MA, et al. Reduced expression of a gene proliferation signature is associated with enhanced malignancy in colon cancer. Br J Cancer 2008;99:966-973.

660. Hall PA, Levison DA, Woods AL, et al. Proliferating cell nuclear antigen (PCNA) immunolocalization in paraffin sections: an index of cell proliferation with evidence of deregulated expression in some neoplasms. J Pathol 1990;162:285-294.

661. Celis JE, Celis A. Cell cycle-dependent variations in the distribution of the nuclear protein cyclin proliferating cell nuclear antigen in cultured cells: subdivision of S phase. Proc Natl Acad Sci U S A 1985;82:3262-3266.

662. Moldovan GL, Pfander B, Jentsch S. PCNA, the maestro of the replication fork. Cell 2007;129:665-679.

663. Guzinska-Ustymowicz K, Stepien E, Kemona A. MCM-2, Ki-67 and PCNA protein expressions in pT3G2 colorectal cancer indicated lymph node involvement. Anticancer Res 2008;28:451-457.

664. Yang HB, Hsu PI, Chan SH, Lee JC, Shin JS, Chow NH. Growth kinetics of colorectal adenoma-carcinoma sequence: an immunohistochemical study of proliferating cell nuclear antigen expression. Hum Pathol 1996;27:1071-1076.

665. Umar A, Buermeyer AB, Simon JA, et al. Requirement for PCNA in DNA mismatch repair at a step preceding DNA resynthesis. Cell 1996;87:65-73.

666. Hoege C, Pfander B, Moldovan GL, Pyrowolakis G, Jentsch S. RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and SUMO. Nature 2002;419:135-141.

667. Arakawa H, Moldovan GL, Saribasak H, Saribasak NN, Jentsch S, Buerstedde JM. A role for PCNA ubiquitination in immunoglobulin hypermutation. PLoS Biol 2006;4:e366.

668. He X, Wei C, Song T, et al. Proliferating cell nuclear antigen destabilizes c-Abl tyrosine kinase and regulates cell apoptosis in response to DNA damage. Apoptosis 2009;14:268-275.

669. al-Sheneber IF, Shibata HR, Sampalis J, Jothy S. Prognostic significance of proliferating cell nuclear antigen expression in colorectal cancer. Cancer 1993;71:1954-1959.

670. Sun XF, Carstensen JM, Stal O, Zhang H, Nordenskjold B. Proliferating cell nuclear antigen (PCNA) in relation to ras, c-erbB-2,p53, clinico-pathological variables and prognosis in colorectal adenocarcinoma. Int J Cancer 1996;69:5-8.

671. Guerra A, Borda F, Javier Jimenez F, Martinez-Penuela JM, Larrinaga B. Multivariate analysis of prognostic factors in resected colorectal cancer: a new prognostic index. Eur J Gastroenterol Hepatol 1998;10:51-58.

672. Mayer A, Takimoto M, Fritz E, Schellander G, Kofler K, Ludwig H. The prognostic significance of proliferating cell nuclear antigen, epidermal growth factor receptor, and mdr gene expression in colorectal cancer. Cancer 1993;71:2454-2460.

673. Onodera H, Maetani S, Kawamoto K, Kan S, Kondo S, Imamura M. Pathologic significance of tumor progression in locally recurrent rectal cancer: different nature from primary cancer. Dis Colon Rectum 2000;43:775-781.

674. Kawamoto K, Onodera H, Kondo S, et al. Expression of insulin-like growth factor-2 can predict the prognosis of human colorectal cancer patients: correlation with tumor progression, proliferative activity and survival. Oncology 1998;55:242-248.

675. Kanazawa Y, Onda M, Tanaka N, Seya T. Proliferating cell nuclear antigen and p53 protein expression in submucosal invasive colorectal carcinoma. J Nippon Med Sch 2000;67:242-249.

676. Leonardi E, Girlando S, Serio G, et al. PCNA and Ki67 expression in breast carcinoma: correlations with clinical and biological variables. J Clin Pathol 1992;45:416-419.

677. Vestweber D, Kemler R. Some structural and functional aspects of the cell adhesion molecule uvomorulin. Cell Differ 1984;15:269-273.

678. Nagafuchi A, Takeichi M, Tsukita S. The 102 kd cadherin-associated protein: similarity to vinculin and posttranscriptional regulation of expression. Cell 1991;65:849-857.

679. Takeichi M. The cadherins: cell-cell adhesion molecules controlling animal morphogenesis. Development 1988;102:639-655.

680. Ozawa M, Baribault H, Kemler R. The cytoplasmic domain of the cell adhesion molecule uvomorulin associates with three independent proteins structurally related in different species. EMBO J 1989;8:1711-1717.

681. Rimm DL, Koslov ER, Kebriaei P, Cianci CD, Morrow JS. Alpha 1(E)-catenin is an actin-binding and -bundling protein mediating the attachment of F-actin to the membrane adhesion complex. Proc Natl Acad Sci U S A 1995;92:8813-8817.

682. Heuberger J, Birchmeier W. Interplay of cadherin-mediated cell adhesion and canonical Wnt signaling. Cold Spring Harb Perspect Biol 2010;2:a002915.

683. Ben-Ze'ev A, Shtutman M, Zhurinsky J. The integration of cell adhesion with gene expression: the role of beta-catenin. Exp Cell Res 2000;261:75-82.

684. MacDonald BT, Tamai K, He X. Wnt/beta-catenin signaling: components, mechanisms, and diseases. Dev Cell 2009;17:9-26.

685. Lucero OM, Dawson DW, Moon RT, Chien AJ. A re-evaluation of the "oncogenic" nature of Wnt/beta-catenin signaling in melanoma and other cancers. Curr Oncol Rep 2010;12:314-318.

686. Zhang X, Gaspard JP, Chung DC. Regulation of vascular endothelial growth factor by the Wnt and K-ras pathways in colonic neoplasia. Cancer Res 2001;61:6050-6054.

687. Katoh Y, Katoh M. Comparative genomics on PROM1 gene encoding stem cell marker CD133. Int J Mol Med 2007;19:967-970.

688. Horst D, Kriegl L, Engel J, Jung A, Kirchner T. CD133 and nuclear beta-catenin: the marker combination to detect high risk cases of low stage colorectal cancer. Eur J Cancer 2009;45:2034-2040.

689. Vermeulen L, Todaro M, de Sousa Mello F, et al. Single-cell cloning of colon cancer stem cells reveals a multi-lineage differentiation capacity. Proc Natl Acad Sci U S A 2008;105:13427-13432.

690. Lugli A, Zlobec I, Minoo P, et al. Prognostic significance of the wnt signalling pathway molecules APC, beta-catenin and E-cadherin in colorectal cancer: a tissue microarray-based analysis. Histopathology 2007;50:453-464.

691. Horst D, Reu S, Kriegl L, Engel J, Kirchner T, Jung A. The intratumoral distribution of nuclear beta-catenin is a prognostic marker in colon cancer. Cancer 2009;115:2063-2070.

692. Ilyas M, Tomlinson IP. The interactions of APC, E-cadherin and beta-catenin in tumour development and progression. J Pathol 1997;182:128-137.

693. Chetty R, Serra S, Sylvia LA. Loss of membrane localization and aberrant nuclear E-cadherin expression correlates with invasion in pancreatic endocrine tumors. Am J Surg Pathol 2008;32:413-9.

694. Dorudi S, Hanby AM, Poulsom R, Northover J, Hart IR. Level of expression of E-cadherin mRNA in colorectal cancer correlates with clinical outcome. Br J Cancer 1995;71:614-616.

695. Mohri Y. Prognostic significance of E-cadherin expression in human colorectal cancer tissue. Surg Today 1997;27:606-612.

696. Dorudi S, Sheffield JP, Poulsom R, Northover JM, Hart IR. E-cadherin expression in colorectal cancer. An immunocytochemical and in situ hybridization study. Am J Pathol 1993;142:981-986.

697. Gagliardi G, Kandemir O, Liu D, et al. Changes in E-cadherin immunoreactivity in the adenoma-carcinoma sequence of the large bowel. Virchows Arch 1995;426:149-154.

698. Kinsella AR, Green B, Lepts GC, Hill CL, Bowie G, Taylor BA. The role of the cell-cell adhesion molecule E-cadherin in large bowel tumour cell invasion and metastasis. Br J Cancer 1993;67:904-909.

699. Blechschmidt K, Sassen S, Schmalfeldt B, Schuster T, Höfler H, Becker KF. The E-cadherin repressor Snail is associated with lower overall survival of ovarian cancer patients. Br J Cancer 2008;98:489-495.

700. Oda T, Kanai Y, Oyama T, et al. E-cadherin gene mutations in human gastric carcinoma cell lines. Proc Natl Acad Sci U S A 1994;91:1858-1862.

701. Becker KF, Atkinson MJ, Reich U, Becker I, Nekarda H, Siewert JR. E-cadherin gene mutations provide clues to diffuse type gastric carcinomas. Cancer Res 1994;54:3845-52.

702. Guilford P, Hopkins J, Harraway J, et al. E-cadherin germline mutations in familial gastric cancer. Nature 1998;392:402-405.

703. Huntsman DG, Carneiro F, Lewis FR, et al. Early Gastric Cancer in Young, Asymptomatic Carriers of Germ-Line E-Cadherin Mutations. N Engl J Med 2001;344:1904-1909.

704. Brooks-Wilson AR, Kaurah P, Suriano G, et al. Germline E-cadherin mutations in hereditary diffuse gastric cancer: assessment of 42 new families and review of genetic screening criteria. Journal of Medical Genetics 2004;41:508-517.

705. Gamallo C, Palacios J, Suarez A, et al. Correlation of E-cadherin expression with differentiation grade and histological type in breast carcinoma. Am J Pathol 1993;142:987-993.

706. Berx G, Cleton-Jansen AM, Nollet F, et al. E-cadherin is a tumour/invasion suppressor gene mutated in human lobular breast cancers. EMBO J 1995;14:6107-6115.

707. Masciari S, Larsson N, Senz J, et al. Germline E-cadherin mutations in familial lobular breast cancer. Journal of Medical Genetics 2007;44:726-731.

708. Aebi U, Cohn J, Buhle L, Gerace L. The nuclear lamina is a meshwork of intermediate-type filaments. Nature 1986;323:560-564.

709. Liu J, Rolef Ben-Shahar T, Riemer D, et al. Essential roles for Caenorhabditis elegans lamin gene in nuclear organization, cell cycle progression, and spatial organization of nuclear pore complexes. Mol Biol Cell 2000;11:3937-3947.

710. Gerace L, Blum A, Blobel G. Immunocytochemical localization of the major polypeptides of the nuclear pore complex-lamina fraction. Interphase and mitotic distribution. The Journal of Cell Biology 1978;79:546-566.

711. Lebkowski JS, Laemmli UK. Non-histone proteins and long-range organization of HeLa interphase DNA. J Mol Biol 1982;156:325-344.

712. Hutchison CJ,. Lamins: building blocks or regulators of gene expression? Nature reviews Molecular cell biology 2002;3:848-58.

713. Goldman RD, Gruenbaum Y, Moir RD, Shumaker DK, Spann TP. Nuclear lamins: building blocks of nuclear architecture. Genes Dev 2002;16:533-547.

714. Rober RA, Weber K, Osborn M. Differential timing of nuclear lamin A/C expression in the various organs of the mouse embryo and the young animal: a developmental study. Development 1989;105:365-378.

715. Vergnes L, Peterfy M, Bergo MO, Young SG, Reue K. Lamin B1 is required for mouse development and nuclear integrity. Proc Natl Acad Sci U S A 2004;101:10428-10433.

716. Osouda S, Nakamura Y, de Saint Phalle B, et al. Null mutants of Drosophila B-type lamin Dm(0) show aberrant tissue differentiation rather than obvious nuclear shape distortion or specific defects during cell proliferation. Dev Biol 2005;284:219-232.

717. Furukawa K, Hotta Y. cDNA cloning of a germ cell specific lamin B3 from mouse spermatocytes and analysis of its function by ectopic expression in somatic cells. EMBO J 1993;12:97-106.

718. Furukawa K, Inagaki H, Hotta Y. Identification and cloning of an mRNA coding for a germ cell-specific A-type lamin in mice. Exp Cell Res 1994;212:426-430.

719. Bridger J, Kill I, O'Farrell M, Hutchison C. Internal lamin structures within G1 nuclei of human dermal fibroblasts. J Cell Sci 1993;104:297-306.

720. Goldman AE, Moir RD, Montag-Lowy M, Stewart M, Goldman RD. Pathway of incorporation of microinjected lamin A into the nuclear envelope. The Journal of Cell Biology 1992;119:725-735.

721. Sasseville A, Raymond Y. Lamin A precursor is localized to intranuclear foci. J Cell Sci 1995;108:273-285.

722. Jagatheesan G, Thanumalayan S, Muralikrishna B, Rangaraj N, A AK, V KP. Colocalization of intranuclear lamin foci with RNA splicing factors. J Cell Sci 1999;112:4651-61.

723. Dorner D, Gotzmann J, Foisner R. Nucleoplasmic lamins and their interaction partners, LAP2alpha, Rb, and BAF, in transcriptional regulation. FEBS J 2007;274:1362-1373.

724. Bridger JM, Foeger N, Kill IR, Herrmann H. The nuclear lamina. Both a structural framework and a platform for genome organization. FEBS J 2007;274:1354-1361.

725. Hozak P, Hassan AB, Jackson DA, Cook PR. Visualization of replication factories attached to nucleoskeleton. Cell 1993;73:361-373.

726. Hozak P, Jackson DA, Cook PR. Replication factories and nuclear bodies: the ultrastructural characterization of replication sites during the cell cycle. J Cell Sci 1994;107 (Pt 8):2191-2202.

727. Hozak P, Sasseville AM, Raymond Y, Cook PR. Lamin proteins form an internal nucleoskeleton as well as a peripheral lamina in human cells. J Cell Sci 1995;108 (Pt 2):635-644.

728. Holubec H, Payne CM, Bernstein H, et al. Assessment of apoptosis by immunohistochemical markers compared to cellular morphology in ex vivo-stressed colonic mucosa. J Histochem Cytochem 2005;53:229-235.

729. Rao L, Perez D, White E. Lamin proteolysis facilitates nuclear events during apoptosis. J Cell Biol 1996;135:1441-1455.

730. Broers JL, Ramaekers FC, Bonne G, Yaou RB, Hutchison CJ. Nuclear lamins: laminopathies and their role in premature ageing. Physiol Rev 2006;86:967-1008.

731. Cross T, Griffiths G, Deacon E, et al. PKC-delta is an apoptotic lamin kinase. Oncogene 2000;19:2331-2337.

732. Slee EA, Adrain C, Martin SJ. Executioner caspase-3, -6, and -7 perform distinct, non-redundant roles during the demolition phase of apoptosis. J Biol Chem 2001;276:7320-7326.

733. Lee SC, Chan J, Clement MV, Pervaiz S. Functional proteomics of resveratrol-induced colon cancer cell apoptosis: caspase-6-mediated cleavage of lamin A is a major signaling loop. Proteomics 2006;6:2386-2394.

734. Markiewicz E, Dechat T, Foisner R, Quinlan RA, Hutchison CJ. Lamin A/C binding protein LAP2alpha is required for nuclear anchorage of retinoblastoma protein. Mol Biol Cell 2002;13:4401-4413.

735. Markiewicz E, Tilgner K, Barker N, et al. The inner nuclear membrane protein emerin regulates beta-catenin activity by restricting its accumulation in the nucleus. EMBO J 2006;25:3275-3285.

736. Hutchison CJ, Worman HJ. A-type lamins: guardians of the soma? Nat Cell Biol 2004;6:1062-1067.

737. Moss SF, Krivosheyev V, de Souza A, et al. Decreased and aberrant nuclear lamin expression in gastrointestinal tract neoplasms. Gut 1999;45:723-729.

738. Foster CR, Przyborski SA, Wilson RG, Hutchison CJ. Lamins as cancer biomarkers. Biochem Soc Trans 2010;38:297-300.

739. Oguchi M, Sagara J, Matsumoto K, Saida T, Taniguchi S. Expression of lamins depends on epidermal differentiation and transformation. Br J Dermatol 2002;147:853-858.

740. Wu Z, Wu L, Weng D, Xu D, Geng J, Zhao F. Reduced expression of lamin A/C correlates with poor histological differentiation and prognosis in primary gastric carcinoma. J Exp Clin Cancer Res 2009;28:8.

741. Broers JL, Machiels BM, Kuijpers HJ, et al. A- and B-type lamins are differentially expressed in normal human tissues. Histochem Cell Biol 1997;107:505-517.

742. Ivorra C, Kubicek M, Gonzalez JM, et al. A mechanism of AP-1 suppression through interaction of c-Fos with lamin A/C. Genes Dev 2006;20:307-320.

743. Scaffidi P, Misteli T. Lamin A-dependent misregulation of adult stem cells associated with accelerated ageing. Nat Cell Biol 2008;10:452-459.

744. Willis ND, Wilson RG, Hutchison CJ. Lamin A: a putative colonic epithelial stem cell biomarker which identifies colorectal tumours with a more aggressive phenotype. Biochem Soc Trans 2008;36:1350-1353.

745. Gibson SI, Surosky RT, Tye BK. The phenotype of the minichromosome maintenance mutant mcm3 is characteristic of mutants defective in DNA replication. Mol Cell Biol 1990;10:5707-5720.

746. Hennessy KM, Lee A, Chen E, Botstein D. A group of interacting yeast DNA replication genes. Genes Dev 1991;5:958-969.

747. Yan H, Gibson S, Tye BK. Mcm2 and Mcm3, two proteins important for ARS activity, are related in structure and function. Genes Dev 1991;5:944-957.

748. Chen Y, Hennessy KM, Botstein D, Tye BK. CDC46/MCM5, a yeast protein whose subcellular localization is cell cycle-regulated, is involved in DNA replication at autonomously replicating sequences. Proc Natl Acad Sci U S A 1992;89:10459-10463.

749. Labib K, Tercero JA, Diffley JFX. Uninterrupted MCM2-7 Function Required for DNA Replication Fork Progression. Science 2000;288:1643-1647.

750. Blow JJ, Laskey RA. A role for the nuclear envelope in controlling DNA replication within the cell cycle. Nature 1988;332:546-548.

751. Koonin EV. A common set of conserved motifs in a vast variety of putative nucleic acid-dependent ATPases including MCM proteins involved in the initiation of eukaryotic DNA replication. Nucleic Acids Res 1993;21:2541-2547.

752. Brewster AS, Chen XS. Insights into the MCM functional mechanism: lessons learned from the archaeal MCM complex. Crit Rev Biochem Mol Biol 2010;45:243-256.

753. Bochman ML, Schwacha A. The Mcm2-7 complex has in vitro helicase activity. Mol Cell 2008;31:287-293.

754. Moyer SE, Lewis PW, Botchan MR. Isolation of the Cdc45/Mcm2-7/GINS (CMG) complex, a candidate for the eukaryotic DNA replication fork helicase. Proc Natl Acad Sci U S A 2006;103:10236-10241.

755. Kearsey SE, Labib K. MCM proteins: evolution, properties, and role in DNA replication. Biochim Biophys Acta 1998;1398:113-136.

756. Tye BK. MCM proteins in DNA replication. Annu Rev Biochem 1999;68:649-686.

757. Bochman ML, Schwacha A. The Mcm complex: unwinding the mechanism of a replicative helicase. Microbiol Mol Biol Rev 2009;73:652-683.

758. Sterner JM, Dew-Knight S, Musahl C, Kornbluth S, Horowitz JM. Negative regulation of DNA replication by the retinoblastoma protein is mediated by its association with MCM7. Mol Cell Biol 1998;18:2748-2757.

759. Kukimoto I, Aihara S, Yoshiike K, Kanda T. Human papillomavirus oncoprotein E6 binds to the C-terminal region of human minichromosome maintenance 7 protein. Biochem Biophys Res Commun 1998;249:258-262.

760. Gladden AB, Diehl JA. The cyclin D1-dependent kinase associates with the pre-replication complex and modulates RB.MCM7 binding. J Biol Chem 2003;278:9754-9760.

761. Facoetti A, Ranza E, Benericetti E, Ceroni M, Tedeschi F, Nano R. Minichromosome maintenance protein 7: a reliable tool for glioblastoma proliferation index. Anticancer Res 2006;26:1071-5.

762. Giaginis C, Vgenopoulou S, Vielh P, Theocharis S. MCM proteins as diagnostic and prognostic tumor markers in the clinical setting. Histol Histopathol 2010;25:351-370.

763. Freeman A, Morris LS, Mills AD, et al. Minichromosome maintenance proteins as biological markers of dysplasia and malignancy. Clin Cancer Res 1999;5:2121-2132.

764. Brake T, Connor JP, Petereit DG, Lambert PF. Comparative analysis of cervical cancer in women and in a human papillomavirus-transgenic mouse model: identification of minichromosome maintenance protein 7 as an informative biomarker for human cervical cancer. Cancer Res 2003;63:8173-8180.

765. Padmanabhan V, Callas P, Philips G, Trainer TD, Beatty BG. DNA replication regulation protein Mcm7 as a marker of proliferation in prostate cancer. J Clin Pathol 2004;57:1057-1062.

766. Cromer A, Carles A, Millon R, et al. Identification of genes associated with tumorigenesis and metastatic potential of hypopharyngeal cancer by microarray analysis. Oncogene 2004;23:2484-2498.

767. Fujioka S, Shomori K, Nishihara K, Yamaga K, Nosaka K, Araki K. Expression of minichromosome maintenance 7 (MCM7) in small lung adenocarcinomas (pT1): Prognostic implication. Lung Cancer 2009;65:223-9.

768. Honeycutt KA, Chen Z, Koster MI, et al. Deregulated minichromosomal maintenance protein MCM7 contributes to oncogene driven tumorigenesis. Oncogene 2006;25:4027-4032.

769. Nishihara K, Shomori K, Fujioka S, et al. Minichromosome maintenance protein 7 in colorectal cancer: implication of prognostic significance. Int J Oncol 2008;33:245-251.

770. Nishihara K, Shomori K, Tamura T, Fujioka S, Ogawa T, Ito H. Immunohistochemical expression of geminin in colorectal cancer: Implication of prognostic significance. Oncol Rep 2009;21:1189-1195.

771. Pillaire MJ, Selves J, Gordien K, et al. A 'DNA replication' signature of progression and negative outcome in colorectal cancer. Oncogene 2010;29:876-887.

772. Yarden Y. The EGFR family and its ligands in human cancer. signalling mechanisms and therapeutic opportunities. Eur J Cancer 2001;37 Suppl 4:S3-8.

773. Rizzo S, Bronte G, Fanale D, et al. Prognostic vs predictive molecular biomarkers in colorectal cancer: is KRAS and BRAF wild type status required for anti-EGFR therapy? Cancer Treat Rev 2010;36:S56-S61.

774. Tedesco KL, Lockhart AC, Berlin JD. The epidermal growth factor receptor as a target for gastrointestinal cancer therapy. Curr Treat Options Oncol 2004;5:393-403.

775. Threadgill DW, Dlugosz AA, Hansen LA, et al. Targeted disruption of mouse EGF receptor: effect of genetic background on mutant phenotype. Science 1995;269:230-234.

776. Sibilia M, Wagner EF. Strain-dependent epithelial defects in mice lacking the EGF receptor. Science 1995;269:234-238.

777. Miettinen PJ, Berger JE, Meneses J, et al. Epithelial immaturity and multiorgan failure in mice lacking epidermal growth factor receptor. Nature 1995;376:337-341.

778. Nicholson RI, Gee JM, Harper ME. EGFR and cancer prognosis. Eur J Cancer 2001;37 Suppl 4:S9-15.

779. Koretz K, Schlag P, Moller P. Expression of epidermal growth factor receptor in normal colorectal mucosa, adenoma, and carcinoma. Virchows Arch A Pathol Anat Histopathol 1990;416:343-349.

780. McKay JA, Murray LJ, Curran S, et al. Evaluation of the epidermal growth factor receptor (EGFR) in colorectal tumours and lymph node metastases. Eur J Cancer 2002;38:2258-2264.

781. De Jong KP, Stellema R, Karrenbeld A, et al. Clinical relevance of transforming growth factor alpha, epidermal growth factor receptor, p53, and Ki67 in colorectal liver metastases and corresponding primary tumors. Hepatology 1998;28:971-979.

782. Zlobec I, Vuong T, Hayashi S, et al. A simple and reproducible scoring system for EGFR in colorectal cancer: application to prognosis and prediction of response to preoperative brachytherapy. Br J Cancer 2007;96:793-800.

783. Steele RJ, Kelly P, Ellul B, Eremin O. Epidermal growth factor receptor expression in colorectal cancer. Br J Surg 1990;77:1352-1354.

784. Kluftinger AM, Robinson BW, Quenville NF, Finley RJ, Davis NL. Correlation of epidermal growth factor receptor and c-erbB2 oncogene product to known prognostic indicators of colorectal cancer. Surg Oncol 1992;1:97-105.

785. Harari PM. Epidermal growth factor receptor inhibition strategies in oncology. Endocr Relat Cancer 2004;11:689-708.

786. Chung KY, Shia J, Kemeny NE, et al. Cetuximab shows activity in colorectal cancer patients with tumors that do not express the epidermal growth factor receptor by immunohistochemistry. J Clin Oncol 2005;23:1803-1810.

787. Lievre A, Bachet JB, Le Corre D, et al. KRAS mutation status is predictive of response to cetuximab therapy in colorectal cancer. Cancer Res 2006;66:3992-3995.

788. Karapetis CS, Khambata-Ford S, Jonker DJ, et al. K-ras mutations and benefit from cetuximab in advanced colorectal cancer. N Engl J Med 2008;359:1757-1765.

789. Van Cutsem E, Kohne CH, Hitre E, et al. Cetuximab and chemotherapy as initial treatment for metastatic colorectal cancer. N Engl J Med 2009;360:1408-1417.

790. Wellcome Trust Sanger institute. Catalogue of Somatic Mutations in Cancer. Available at: www.sanger.ac.uk/perl/genetics/CGP/cosmic.

791. Samuels Y, Wang Z, Bardelli A, et al. High frequency of mutations of the PIK3CA gene in human cancers. Science 2004;304:554.

792. Barault L, Veyrie N, Jooste V, et al. Mutations in the RAS-MAPK, PI(3)K (phosphatidylinositol-3-OH kinase) signaling network correlate with poor survival in a population-based series of colon cancers. Int J Cancer 2008;122:2255-2259.

793. Goel A, Arnold CN, Niedzwiecki D, et al. Frequent inactivation of PTEN by promoter hypermethylation in microsatellite instability-high sporadic colorectal cancers. Cancer Res 2004;64:3014-3021.

794. International Agency for Research on Cancer. World Cancer Report 2008. Lyon, France: World Health Organization, 2008.

795. Foster CR, Robson JL, Simon WJ, et al. The role of Lamin A in cytoskeleton organization in colorectal cancer cells: a proteomic investigation. Nucleus 2011;2:434-443.

796. Belt EJ, Fijneman RJ, van den Berg EG, et al. Loss of lamin A/C expression in stage II and III colon cancer is associated with disease recurrence. Eur J Cancer 2011;47:1837-1845.

797. Woods MO, Hyde AJ, Curtis FK, et al. High frequency of hereditary colorectal cancer in Newfoundland likely involves novel susceptibility genes. Clin Cancer Res 2005;11:6853-6861.

798. Camp RL, Dolled-Filhart M, Rimm DL. X-tile: a new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization. Clinical cancer research : an official journal of the American Association for Cancer Research 2004;10:7252-9.

799. Clark-Langone KM, Wu JY, Sangli C, et al. Biomarker discovery for colon cancer using a 761 gene RT-PCR assay. BMC Genomics 2007;8:279.

800. O'Connell MJ, Lavery I, Yothers G, et al. Relationship between tumor gene expression and recurrence in four independent studies of patients with stage II/III colon cancer treated with surgery alone or surgery plus adjuvant fluorouracil plus leucovorin. J Clin Oncol 2010;28:3937-3944.

801. Watanabe T, Kobunai T, Yamamoto Y, et al. Gene expression signature and response to the use of leucovorin, fluorouracil and oxaliplatin in colorectal cancer patients. Clin Transl Oncol 2011;13:419-425.

802. Makretsov NA, Huntsman DG, Nielsen TO, et al. Hierarchical clustering analysis of tissue microarray immunostaining data identifies prognostically significant groups of breast carcinoma. Clin Cancer Res 2004;10:6143-6151.

803. Grossi F, Spizzo R, Bordo D, et al. Prognostic stratification of stage IIIA pN2 non-small cell lung cancer by hierarchical clustering analysis of tissue microarray immunostaining data: an Alpe Adria Thoracic Oncology Multidisciplinary Group study (ATOM 014). J Thorac Oncol 2010;5:1354-1360.

804. de Hoon MJ, Imoto S, Nolan J, Miyano S. Open source clustering software. Bioinformatics 2004;20:1453-1454.

805. Saldanha AJ. Java Treeview--extensible visualization of microarray data. Bioinformatics 2004;20:3246-3248.

806. NCI Dictionary of Cancer Terms. Available at: http://www.cancer.gov/dictionary/. Accessed 5/5/2010, 2010.

807. Greenson JK, Bonner JD, Ben-Yzhak O, 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;27:563-570.

808. Shia J, Ellis NA, Paty PB, et al. Value of histopathology in predicting microsatellite instability in hereditary nonpolyposis colorectal cancer and sporadic colorectal cancer. Am J Surg Pathol 2003;27:1407-1417.

809. Roman R, Verdu M, Calvo M, et al. Microsatellite instability of the colorectal carcinoma can be predicted in the conventional pathologic examination. A prospective

multicentric study and the statistical analysis of 615 cases consolidate our previously proposed logistic regression model. Virchows Arch 2010;.

810. Young J, Simms LA, Biden KG, et al. Features of colorectal cancers with high-level microsatellite instability occurring in familial and sporadic settings: parallel pathways of tumorigenesis. Am J Pathol 2001;159:2107-2116.

811. Jass JR, Walsh MD, Barker M, Simms LA, Young J, Leggett BA. Distinction between familial and sporadic forms of colorectal cancer showing DNA microsatellite instability. Eur J Cancer 2002;38:858-866.

812. Jass JR,. HNPCC and sporadic MSI-H colorectal cancer: a review of the morphological similarities and differences. Familial cancer 2004;3:93-100.

813. Halama N, Michel S, Kloor M, et al. The localization and density of immune cells in primary tumors of human metastatic colorectal cancer shows an association with response to chemotherapy. Cancer Immun 2009;9:1.

814. Roxburgh CS, Salmond JM, Horgan PG, Oien KA, McMillan DC. Comparison of the prognostic value of inflammation-based pathologic and biochemical criteria in patients undergoing potentially curative resection for colorectal cancer. Ann Surg 2009;249:788-793.

815. Zlobec I, Minoo P, Baumhoer D, et al. Multimarker phenotype predicts adverse survival in patients with lymph node-negative colorectal cancer. Cancer 2008;112:495-502.

816. Purdie CA, Piris J. Histopathological grade, mucinous differentiation and DNA ploidy in relation to prognosis in colorectal carcinoma. Histopathology 2000;36:121-126.

817. Mecklin JP, Sipponen P, Jarvinen HJ. Histopathology of colorectal carcinomas and adenomas in cancer family syndrome. Dis Colon Rectum 1986;29:849-853.

818. Pinol V, Castells A, Andreu M, et al. Accuracy of revised Bethesda guidelines, microsatellite instability, and immunohistochemistry for the identification of patients with hereditary nonpolyposis colorectal cancer. JAMA 2005;293:1986-1994.

819. Truta B, Chen YY, Blanco AM, et al. Tumor histology helps to identify Lynch syndrome among colorectal cancer patients. Fam Cancer 2008;.

820. Chou CL, Lin JK, Wang HS, Yang SH, Li AF, Chang SC. Microsatellite instability screening should be done for right-sided colon cancer patients less than 60 years of age. Int J Colorectal Dis 2010;25:47-52.

821. Carethers JM, Smith EJ, Behling CA, et al. Use of 5-fluorouracil and survival in patients with microsatellite-unstable colorectal cancer. Gastroenterology 2004;126:394-401.

822. Carstens PHB, Gfeenberg RA, Francis D, Lyon H. Tubular carcinoma of the breast. A long term follow-up. Histopathology 1985;9:271-280.

823. Oberman HA, Fidler WJ,Jr. Tubular carcinoma of the breast. Am J Surg Pathol 1979;3:387-395.

824. Waldman FM, Hwang ES, Etzell J, et al. Genomic alterations in tubular breast carcinomas. Hum Pathol 2001;32:222-226.

825. Ueng S, Mezzetti T, Tavassoli FA. Papillary Neoplasms of the Breast: A Review. Arch Pathol Lab Med 2009;133:893-907.

826. Randolph GW, Thompson GB, Branovan DI, Tuttle RM. Treatment of Thyroid Cancer: 2007—A Basic Review. International Journal of Radiation Oncology*Biology*Physics 2007;69:S92-S97.

827. Xing M. BRAF mutation in papillary thyroid cancer: pathogenic role, molecular bases, and clinical implications. Endocr Rev 2007;28:742-62.

828. Yagata H, Harigaya K, Suzuki M, et al. Comedonecrosis is an unfavorable marker in node-negative invasive breast carcinoma. Pathol Int 2003;53:501-506.

829. Silverstein MJ, Poller DN, Waisman JR, et al. Prognostic classification of breast ductal carcinoma-in-situ. Lancet 1995;345:1154-1157.

830. Guo CC, Epstein JI. Intraductal carcinoma of the prostate on needle biopsy: Histologic features and clinical significance. Mod Pathol 2006;19:1528-1535.

831. Wilcox G, Soh S, Chakraborty S, Scardino PT, Wheeler TM. Patterns of high-grade prostatic intraepithelial neoplasia associated with clinically aggressive prostate cancer. Hum Pathol 1998;29:1119-1123.

832. Skobe M, Rockwell P, Goldstein N, Vosseler S, Fusenig NE. Halting angiogenesis suppresses carcinoma cell invasion. Nat Med 1997;3:1222-1227.

833. Liotta LA, Kleinerman J, Saidel GM. Quantitative relationships of intravascular tumor cells, tumor vessels, and pulmonary metastases following tumor implantation. Cancer Res 1974;34:997-1004.

834. Rak J, Filmus J, Kerbel RS. Reciprocal paracrine interactions between tumour cells and endothelial cells: the 'angiogenesis progression' hypothesis. Eur J Cancer 1996;32A:2438-2450.

835. Vermeulen PB, Van den Eynden GG, Huget P, et al. Prospective study of intratumoral microvessel density, p53 expression and survival in colorectal cancer. Br J Cancer 1999;79:316-322.

836. Hyde A, Fontaine D, Stuckless S, et al. A histology-based model for predicting microsatellite instability in colorectal cancers. Am J Surg Pathol 2010;34:1820-1829.

837. Green SE, Chapman PD, Burn J, Bishop DT, Varma JS. Clinical impact of colonoscopic screening in first-degree relatives of patients with hereditary non-polyposis colorectal cancer. Br J Surg 1995;82:1338-1340.

838. Renkonen-Sinisalo L, Aarnio M, Mecklin JP, Jarvinen HJ. Surveillance improves survival of colorectal cancer in patients with hereditary nonpolyposis colorectal cancer. Cancer Detect Prev 2000;24:137-142.

839. Jarvinen HJ, Mecklin JP, Sistonen P. Screening reduces colorectal cancer rate in families with hereditary nonpolyposis colorectal cancer. Gastroenterology 1995;108:1405-1411.

840. Vasen HF, Abdirahman M, Brohet R, et al. One to 2-Year Surveillance Intervals Reduce Risk of Colorectal Cancer in Families With Lynch Syndrome. Gastroenterology 2010;138:2300-2306.

841. de la Chapelle A, Hampel H. Clinical Relevance of Microsatellite Instability in Colorectal Cancer. J Clin Oncol 2010;28:3380-3387.

842. Knottnerus JA, van Weel C, Muris JWM. Evidence base of clinical diagnosis: Evaluation of diagnostic procedures. BMJ 2002;324:477-480.

Appendices

APPENDIX A: Microsatellite Primer Sequences used for MSI Analysis

BAT26

Forward: 5' – TGACTACTTTTGACTTCAGCC – 3'

Reverse: 5' – AACCATTCAACATTTTTAACCC – 3'

BAT25

Forward: 5' – TCGCCTCCAAGAATGTAAGT – 3'

Reverse: 5' – TCTGCATTTTAACTATGGCTC – 3'

D2S123

Forward: 5' – AAACAGGATGCCTGCCTTTA – 3'

Reverse: 5' – GGACTTTCCACCTATGGGAC – 3'

D5S346

Forward: 5' – ACTCACTCTAGTGATAAATCGGG – 3'

Reverse: 5' – AGCAGATAAGACAGTATTACTAGTT – 3'

D17S250

Forward: 5' – GGAAGAATCAAATAGACAAT – 3'

Reverse: 5' – GCTGGCCATATATATATTTAAACC – 3'

BAT40

Forward: 5' – ATTAACTTCCTACACCACAAC – 3'

Reverse: 5' – GTAGAGCAAGACCACCTTG – 3'

D7S519

Forward: 5' – ACAGCCAAGCATTTCTGCTG – 3'

Reverse: 5' – ACAGACCAGGACTCAGCCAG – 3'

D17S787

Forward: 5' – TGGGCTCAACTATATGAACC – 3'

Reverse: 5' – TTGATACCTTTTTGAAGGGG – 3'

D18S58

Forward: 5' – GCTCCCGGCTGGTTTT – 3'

Reverse: 5' – GCAGGAAATCGCAGGAACTT – 3'

D20S100

Forward: 5' – ATTGGGTTTACTTGTGCCTT – 3'

Reverse: 5' – CGTGATTTCATTTCTTGCTG – 3'

APPENDIX B: Recipes for Solutions and Buffers

MSI Analysis

6% Acrylamide, 7M Urea Gel Solution

189 g urea

45 ml 5x TBE

67.5 ml 40% acrylamide:bis 19:1

I stirred the above until the urea dissolved, then filtered the solution under vacuum suction, and stored at 4°C in a dark bottle.

40 cm x 60 cm Polyacrylamide Gel

For the plug:

30 ml	6% acrylamide, 7M urea gel solution
175 µl	10% ammonium persulfate
175 µl	TEMED

For the gel:

120 ml	6% acrylamide, 7M urea gel solution
1 ml	10% ammonium persulfate
100 µl	TEMED

Silver Staining – Fix/Stop Solution

1800 ml	deionized water
200 ml	glacial acetic acid

I stirred the water and acid to mix, and then stored the solution at room temperature.

Silver Staining – Staining Solution

2 L	deionized water
3 ml	formaldehyde
2 g	silver nitrate

I stirred the ingredients until the silver nitrate had dissolved, then stored the solution at room temperature in the dark.

Silver Staining – Developing Solution

2 L	deionized water
3 ml	formaldehyde
400 µl	10mg/ml sodium thiosulfate
60 g	sodium carbonate

I stirred the ingredients until all solids were dissolved, then stored the solution at 4°C.

Immunohistochemistry Analysis

Phosphate Buffered Saline (PBS) – 1 L

8.0 g	sodium chloride
1.3 g	disodium phosphate
4.0 g	monosodium phosphate

I dissolved the salts in deionized water to just less than one litre total volume. I then adjusted the pH to 7.4 using 2M NaOH, and added enough deionized water to bring the volume up to one litre. I stored the solution at room temperature.

10mM Citrate Buffer – 1 L

2.1 g	citric acid

I dissolved the citric acid in deionized water to just less than one litre volume, and then used 2M NaOH to adjust pH to 6.0. I added enough deionized water to bring the volume up to one litre, and then stored the solution at room temperature.

Scotts Tap Water – 2 L

I made up two solutions and stored them separately. The bicarbonate solution contained 7 g of sodium bicarbonate dissolved in one litre of deionozed water. The magnesium sulfate solution contained 40 g of magnesium sulfate dissolved in one litre of water. I stored these solutions separately until I was ready to use them, at which time I mixed them in a one to one ratio.

Mayer's Haematoxylin – 2 L

I dissolved 180 g of aluminum ammonium sulfate in 1200 ml of deionized water, and 12 g of haematoxylin in 75 ml of 95% ethanol, then combined the two solutions. I left this mixture exposed to light and air for one month, then filtered it, and added 300 ml of glycerin and 300 ml of 95% ethanol. I left the solution exposed to light until the colour was sufficiently dark, then filtered it again, and stored it in a tightly sealed dark bottle at room temperature.

APPENDIX C: Physician Contact Letter

<<Date>>

<<Title>> <<First Name>> <<Last Name>>
<<Address1>>
<<Address2>>
<<City>>, <<Province>>
<<Postal Code>>

Dear Dr._____:

RE: THE GENETICS OF COLO-RECTAL CANCER IN NEWFOUNDLAND

I am writing on behalf of a research team of Memorial University to request your participation in this study of the genetics of colo-rectal cancer in Newfoundland. The investigators include Dr. P. Parfrey, Dr. J. Green, Dr. R. Green, Dr. B. Younghusband and myself. Much of the work will be done by graduate students – Fiona Curtis and Angela Hyde, who will have direct contact with patients and their families.

This portion of the study wishes to determine the contribution of genetic factors to the incidence of col-rectal cancer in Newfoundland. It will involve direct contact with patients and their families by the graduate students who will take detailed family histories. Patients and family members may be asked to contribute blood samples. This has been approved by the Human Investigation Committee of Memorial University of Newfoundland and the Health Care Cooperation of St. John's. Part of the protocol requires that the initial contact with the patient should be initiated by their attending physician. I am writing to enquire whether you would be willing to sign a standardized latter (sample enclosed) which will be sent to your patients or their families asking for their participation in this research.

If you agree to participate, addressed postage paid letters will be provided for you to sign which will then be sent to patients or their families. Patients will be asked to respond either directly to the research team or your office.

.../2

Dr. _____
<<Date>>
RE: The Genetics of Colo-rectal Cancer in Newfoundland

We believe that this research may have important benefits in elucidating the epidemiology of colo-rectal cancer in Newfoundland and may result in significant clinical benefits for patients and their families.

Thank you for considering this request. Please let me know if you are willing to participate at your earliest convenience. If you have any questions or concerns please contact me directly.

Your sincerely,

W.G. Pollett, MD, FRCSC
Professor & Chair
Discipline of Surgery
Memorial University of Newfoundland
Health Sciences Centre

Phone: (709) 737-6612
Fax: (709) 737-5050

Email address:
wpollett@morgan.ucs.mun.ca

WGP/sw

Enclosure

APPENDIX D: Physician Letters to Proband and Proxy

HealthCare
Corporation of St. John's

Genetics of Colorectal Cancer in Newfoundland-Research Project

DATE:_____

DEAR_____

Researchers at Memorial University are interested in learning more about the hereditary nature of colon and rectal cancer. Most cancers happen by chance but some kinds of cancer are inherited. Most colon and rectum cancers are diagnosed in people over 65. About 10% of colon and rectum cancers are inherited, and they often occur before the age of 50.

They need to find out how many people with a diagnosis of colon and rectum cancer in a single year have the inherited kind. If a family is found to have an inherited kind then research can be used to identify the kind of gene that may cause the cancer in that family. This could lead to an earlier diagnosis and treatment for families that have the gene.

I would like to ask you to take part in the study on behalf of the research team (led by Dr Patrick Parfrey, Dr. William Pollett, and Dr. Jane Green). I know that your _____, _____ was diagnosed with colon and rectum cancer. Family members can give us valuable information for this research. Your participation is entirely voluntary. You may decide not to take part or you may withdraw from the study at any time.

If you wish to take part in the study you may call the research team, collect at (709) 777-7622, or you can contact our office and we will pass your name on to the research team and they will contact you by telephone. They will ask you for a convenient time to meet with you. During this meeting the researcher will ask you to help trace your family history. If there seems to be a family history of colon and rectum cancer, you and your family members will be asked to have a simple blood test. This test may or may not be able to tell if you have a gene that put you at an increased risk for colon and rectum cancer.

If you do not wish to take part in the research project please call my office between the hours of _____, and give your name to my secretary.

Your participation would be greatly appreciated. With your help the research team hopes to learn more about colon and rectum cancer, the role genes play, and how often colon and rectum cancer occurs in Newfoundland.

Yours truly,

_____

APPENDIX E: Consent Form for Proband

FACULTY OF MEDICINE – MEMORIAL UNIVERSITY OF NEWFOUNDLAND
AND
HEALTH CARE CORPORATION OF ST. JOHN'S

Consent to Participate in Bio-medical Research

TITLE: The genetics of colorectal cancer in Newfoundland

INVESTIGATORS: Fiona Curtis, Angela Hyde, Dr. P. Parfrey, Dr. J. Green, Dr. W. Pollett, Dr. R. Green, Dr. B. Younghusband

You have been asked to take part in a research study. Participation in this study is entirely voluntary. You may decide not to take part or may withdraw from the study at any time without affecting your normal treatment.

Information obtained from you or about you during this study, which could identify you, will be kept confidential by the investigators. The investigator will be available during the study at all times should you have any problems or questions about the study.

1. Purpose of the study:

   Most cancers affect individuals, but some cancers occur in several family members. These cancers may be inherited, that is they "run in the family" as a result of changes in some genes. The purpose of this study is to find out how many people with a diagnosis of colorectal cancer have an inherited form of the disease and to develop laboratory tests to identify these cases.

2. Description of procedures and tests:

   You will meet with the investigator at your convenience in person or by telephone if necessary. You will help trace your family medical history. If there is an indication of inherited colorectal cancer in your family then you will receive a referral to the Medical Genetics Program for yourself and your family members for further consultation. We will ask your permission to review your medical records, obtain some of your cancer tissue which was removed during surgery, and if you agree, a blood sample will be taken for further genetic studies.

3. Duration of participant's involvement:

   The overall duration of the project is 2 years. Your direct involvement will be to meet with the researcher, in person or by telephone, to discuss your family history and to give a blood sample.

4. Possible risks, discomforts, or inconveniences:

   - The interview may be an inconvenience
   - A blood sample taken may lead to some bruising

5. Benefits which the participant may receive:

   If there is evidence of a hereditary cancer syndrome in your family, you will be given a referral to a geneticist which could then lead to screening and early detection and treatment of cancer in your family members.

6. Alternative procedures or treatment for those not entering the study:

   If you decide not to enter this study, you will receive your normal care from your physician.

7. Liability statement:

   Your signature indicates your consent and that you have understood the information regarding the research study. In no way does this waive your legal rights nor release the investigators or involved agencies from their legal and professional responsibilities.

8. Statement of genetic studies:

In order to interpret the results of the genetic research, we need to have correct information about parents. Sometimes the research shows new information about birth parents. This could happen in the case of an adoption or a mistake in the identity of the mother or father. This information will not be given to anyone, including yourself or family members.

9. Future use of tissue/DNA samples:

In order to preserve a valuable resource, your tissue/DNA samples may be stored at the end of this research project. It is possible that these samples may be useful in a future research project which may or may not be related to the current research project. Any future research would have to be approved by the Research Ethics Board (REB).

Please tick **one** of the following options:

1. ☐ I agree that my tissue/DNA samples with my name can be used for any approved research project without asking me again.
2. ☐ I agree that my tissue/DNA samples with my name can be used for any approved research project but only if I am contacted again to give consent for the new project.
3. ☐ I agree that my tissue/DNA samples with my name can be used for any approved research project but only if my name* cannot be linked to the sample.
4. ☐ Under no circumstances may my tissue/DNA samples be used for future research. My samples must be destroyed at the end of the present project.

*Includes name, MCP number or any other identifying information.

The tissue/DNA samples from this study will be stored in St. John's, NL, for an indefinite period of time.

Signature:_____ Date:_____

Witness:_____

Signature Page

Title of Project: The Genetics of Colorectal Cancer in Newfoundland

Name of Principal Investigators: Fiona Curtis and Dr. Jane Green (supervisor)

To be signed by participant:

I, , the undersigned, agree to my participation or to the participation of (my child, ward, relative) in the research study described above. Any questions have been answered and I understand what is involved in the study. I realise that participation is voluntary and that there is no guarantee that I will benefit from my involvement.

I acknowledge that a copy of this form has been given to me.

(Signature of Participant) (Date)

(Signature of Witness) (Date)

To be signed by investigator:

To the best of my ability I have fully explained the nature of this research study. I have invited questions and provided answers. I believe that the participant fully understands the implications and voluntary nature of the study.

(Signed by Investigator) (Date)

Phone number:

Assent of minor participant (if appropriate)

(Signature of Minor Participant) (Age_____)

Relationship to Participant Named Above

APPENDIX F: Consent Form for Proxy

FACULTY OF MEDICINE – MEMORIAL UNIVERSITY OF NEWFOUNDLAND
AND
HEALTH CARE CORPORATION OF ST. JOHN'S

Consent to Participate in Bio-medical Research

TITLE: The genetics of colorectal cancer in Newfoundland

INVESTIGATORS: Fiona Curtis, Angela Hyde, Dr. P. Parfrey, Dr. J. Green, Dr. W. Pollett, Dr. R. Green, Dr. B. Younghusband

You have been asked to take part in a research study on behalf of your family member, _____. Participation in this study is entirely voluntary. You may decide not to take part or may withdraw from the study at any time.

Information obtained from you or about you and family members during this study, which could identify you and other family members, will be kept confidential by the investigators. The investigator will be available during the study at all times should you have any problems or questions about the study.

1. Purpose of the study:

   Most cancers affect individuals, but some cancers occur in several family members. These cancers may be inherited, that is they "run in the family" as a result of changes in some genes. The purpose of this study is to find out how many people with a diagnosis of colorectal cancer have an inherited form of the disease and to develop laboratory tests to identify these cases.

2. Description of procedures and tests:

   You will meet with the investigator at your convenience in person or by telephone if necessary. You will help trace your family medical history. If there is an indication of inherited colorectal cancer in your family then you will receive a referral to the Medical Genetics Program for yourself and your family members for further consultation. We will ask your permission to review your _____ medical records, obtain some of your _____ cancer tissue which was removed during surgery.

3. Duration of participant's involvement:

   The overall duration of the project is 2 years. Your direct involvement will be to meet with the researcher, in person or by telephone, to discuss your family history. If you are willing a blood sample may also be taken.

4. Possible risks, discomforts, or inconveniences:

   - The interview may be an inconvenience
   - A blood sample taken may lead to some bruising

5. Benefits which the participant may receive:

   If there is evidence of a hereditary cancer syndrome in your family, you will be given a referral to a geneticist which could then lead to screening and early detection and treatment of cancer in you and your family members.

6. Alternative procedures or treatment for those not entering the study:

   NA

7. Liability statement:

   Your signature indicates your consent and that you have understood the information regarding the research study. In no way does this waive your legal rights nor release the investigators or involved agencies from their legal and professional responsibilities.

8. Statement of genetic studies:

In order to interpret the results of the genetic research, we need to have correct information about parents. Sometimes the research shows new information about birth parents. This could happen in the case of an adoption or a mistake in the identity of the mother or father. This information will not be given to anyone, including yourself or family members.

9. Future use of tissue/DNA samples:

In order to preserve a valuable resource, your _____ tissue/DNA samples may be stored at the end of this research project. It is possible that these samples may be useful in a future research project which may or may not be related to the current research project. Any future research would have to be approved by the Research Ethics Board (REB).

Please tick **one** of the following options:

1. ☐ I agree that my _____ tissue/DNA samples with my _____ name can be used for any approved research project without asking me again.
2. ☐ I agree that my _____ tissue/DNA samples with my _____ name can be used for any approved research project but only if I am contacted again to give consent for the new project.
3. ☐ I agree that my _____ tissue/DNA samples with my _____ name can be used for any approved research project but only if my _____ name* cannot be linked to the sample.
4. ☐ Under no circumstances may my _____ tissue/DNA samples be used for future research. My _____ samples must be destroyed at the end of the present project.

*Includes name, MCP number or any other identifying information.

The tissue/DNA samples from this study will be stored in St. John's, NL, for an indefinite period of time.

Signature:_____ Date:_____

Witness:_____

Signature Page

Title of Project: The Genetics of Colorectal Cancer in Newfoundland

Name of Principal Investigators: Fiona Curtis and Dr. Jane Green (supervisor)

To be signed by participant:

I, , the undersigned, agree to my participation or to the participation of (my child, ward, relative) in the research study described above. Any questions have been answered and I understand what is involved in the study. I realise that participation is voluntary and that there is no guarantee that I will benefit from my involvement.

I acknowledge that a copy of this form has been given to me.

(Signature of Participant) (Date)

(Signature of Witness) (Date)

To be signed by investigator:

To the best of my ability I have fully explained the nature of this research study. I have invited questions and provided answers. I believe that the participant fully understands the implications and voluntary nature of the study.

(Signed by Investigator) (Date)

Phone number:

Assent of minor participant (if appropriate)

(Signature of Minor Participant) (Age_____)

Relationship to Participant Named Above

APPENDIX G: Newfoundland and Labrador Medical Genetics Program DNA Consent Form

Newfoundland and Labrador Medical Genetics Program

Information and Appointments
Janeway Genetics Clinic (709) 778-4363
Health Science Centre Genetics Clinic (709) 737-6807
Outreach Program Genetics Clinic:
Central Region Phone (709) 651-3308
Western Region Phone (709) 637-5253

Janeway Child Health Centre
Janeway Place
St. John's, NF
A1A 1R8

DNA Consent Form

I agree that a sample of blood may be drawn on _____
for genetic testing and/or banking and that the DNA obtained from the sample will be stored in the NLMGP Molecular Laboratory.

I understand that:

1. The sample will involve giving 10-15 mls of blood for preparation of the DNA.
2. This process is ***voluntary*** and I may withdraw my consent at any time without penalty.
3. ***If there is no test currently available*** for the condition of concern in my family, the sample will be stored as DNA. This does not guarantee a test will be available in future. Banking this sample may be of no direct benefit to my family or me.
4. ***If a test becomes available in the future***, the NLMGP may contact me to discuss whether further DNA testing is appropriate.
5. ***If a test is currently available*** for the condition of concern in my family, the possible results will be discussed with me. This includes the possibility that the testing may be inconclusive and a result will not be available.
6. The ***test results*** will be strictly confidential. They may, however, be used anonymously to help interpret test results for other members of my family.
7. These tests can sometimes point out ***discrepancies in parentage***. Such information will be kept confidential and will not be released even to family members directly involved.
8. The ***accuracy*** of any result is based on the current knowledge of the disease gene in question. Due to limitations in our understanding, human error, or technical difficulties, there is a small chance that a result could prove to be incorrect.
9. It is my responsibility to keep the NLMGP informed of any ***change of address*** so that they will be able to contact me in future if the need arises.

My signature on this form indicates that I have read the information on the previous page and understand that: (tick one)

☐ Direct DNA analysis will be done on this sample to test for _____

☐ Linkage analysis comparing this sample to others in my family will be done to test for _____

☐ No test is available at this time and the sample will be stored as DNA in the NLMGP Molecular Diagnostic Laboratory. but the DNA will be studied to identify any changes in colon cancer genes.

Future use of this DNA sample: (tick all that apply)

☐ I would like the DNA sample destroyed following completion of the above test.

☐ I would like the DNA to remain banked for the future use of other family members.

☐ I agree to the anonymous use of this DNA for any future research approved by a Research Ethics Board

☐ This DNA may be used at any time without restriction.

_____ _____

Patient's signature Date

_____ _____

Signature of legal guardian / Relationship to patient Date

_____ _____

Signature of Witness Date

APPENDIX H: Newfoundland and Labrador Medical Genetics Program Authorization for Release of Medical Information

nd Appointments
enetics Clinic (709) 778-4363
ience Centre Genetics Clinic (709) 737-6807
atellite Genetics Clinic (709) 651-3306
Satellite Genetics Clinic (709) 637-2260

Janeway Child Health Centre
Janeway Place
St. John's, NF
A1A 1R8

Authorization for Release of Medical Information

** Please return a copy of this form with the medical record**

Medical Records Requested:
- ✓ Clinical Diagnostic Summaries
- ✓ Birth Record
- ✓ Consultants Reports
- ✓ Laboratory reports
- ✓ Diagnostic Imaging reports
- ✓ Operative and/or Pathology reports
- ___ Other: _____

I, _____ hereby consent to the release of information to the Medical Genetics Program c/o: Medical Genetics Program, Janeway Child Health Centre, Janeway Place, St. John's, Nfld., A1A 1R8, from the medical records of:

NAME: _____ MAIDEN NAME: _____

ADDRESS: _____ PARENTS/SPOUSE _____

DATE OF BIRTH: _____ DATE OF DEATH: _____

MCP#: _____ CHART# _____

DATE OF HOSPITALIZATION: _____

NAME / ADDRESS OF HOSPITAL: _____

_____	_____	_____
SIGNATURE	RELATIONSHIP	DATE

Medical Genetics program Use: NLMGP# _____

Patient Name: _____

Geneticist / Genetic Counsellor: _____

FAX (709) 778-4190
cience Centre FAX (709) 737-3374

Outreach FAXES - Central (709) 651-3341
Western (709) 637-2616

APPENDIX I: Medical Chart Data Extraction Form

Colorectal Chart Extraction Form

Date:_____

ID #_____ MCP #_____

DOB_____

Address_____

_____

Telephone #_____

Sex_____

Physician_____

Attending Physician_____

Centre Physician_____

Diagnosis_____

Date of Diagnosis_____ Age at Diagnosis_____

Diagnosing Hospital_____

Means of Diagnosis_____

_____

Presenting Symptoms_____

_____

Site Code_____ Behaviour Code_____

Topology Code_____ Histology Code_____

Grade_____

Staging Method_____ Stage Revision_____

Clinical Stage_____ Pathological Stage_____

Presence of Polyps Y/N if yes, #_____ Location_____

_____

Metastases Y/N if yes, Location_____

_____

Previous Cancer Y/N if yes, Type_____

_____

Age and year at diagnosis_____

Treatment: Surgery Y/N Type_____

Date_____

Hospital_____

Chemo Y/N Type and Duration_____

_____

Radiation Y/N Type and Duration_____

_____

Patient Status_____

Death Date_____

Family History_____

_____

_____

_____

_____

Comments_____

_____

_____

APPENDIX J: Supplemental Data for Woods *et. al.*, 2005: "High Frequency of Hereditary Colorectal Cancer in Newfoundland likely Involves Novel Susceptibility Genes"

Table S2.1 Cumulative percentage of family members at 50% risk who developed colorectal cancer, by age and by HNPCC risk classification*.

Age (yrs)	Amsterdam I+II (N=79)	Age/Cancer Modified Am II (N=111)	Revised Bethesda (N=211)	Low Risk (N=370)
40	6.5	0.0	0.0	0.0
50	17.8	0.0	0.0	0.0
60	22.3	4.7	1.4	0.3
70	47.2	17.5	9.3	2.1
80	53.0	24.1	10.6	2.8
Events	24	16	11	7
Hazard Ratio†, C.I.	31.0 (13.2, 73.1) p-value $\leq$ 0.000	9.9 (4.1, 24.2) p-value $\leq$ 0.000	3.8 (1.5, 9.9) p-value $\leq$ 0.006	1.0

*Proband not included

†Reference category is the Low Risk family members

Table S2.2 Cumulative percentage of family members at 50% risk who developed an HNPCC cancer, by age and by HNPCC risk classification*.

Age (yrs)	Amsterdam I+II (N=79)	Age/Cancer Modified Am II (N=111)	Revised Bethesda (N=211)	Low Risk (N=370)
40	6.5	0.0	0.0	0.0
50	19.1	0.0	0.6	0.3
60	27.4	7.8	2.7	1.3
70	51.1	26.3	14.3	3.5
80	56.6	37.1	21.2	5.8
Events	27	27	21	14
Hazard Ratio†, C.I.	18.0 (9.3, 35.0) p-value $\leq$ 0.000	8.6 (4.5, 16.4) p-value $\leq$ 0.000	3.9 (2.0, 7.6) p-value $\leq$ 0.000	1.0

*Proband not included

†Reference category is the Low Risk family members

Table S2.3 Genomic variants identified in *MLH1*, *MSH2* and *MSH6* from both population-based and referred CRC probands

Gene	Nucleotide change (c.)	Amino acid Involved (p.)	Associated exon	Previously published
MLH1	65G>C	Gly22Ala	1	-
MLH1	306+40G>A	-	3	-
MLH1	453+9G>A	-	5	-
MLH1	545+72T>A	-	6	-
MLH1	545+80T>A	-	6	-
MLH1	655A>G	Ile219Val	8	(1, 2)
MLH1	790+62G>A	-	9	-
MLH1	1558+14G>A	-	13	(3)
MLH1	1668-19A>G	-	15	(1, 3)
MLH1	2104-11G>A	-	19	-
MLH1	2104-7T>G	-	19	-
MSH2	211+8C>G	-	1	(4, 5)
MSH2	366+43G>A	-	2	-
MSH2	366+53A>C	-	2	-
MSH2	399C>T	Asp133Asp	3	-
MSH2	793-29A>T	-	5	-
MSH2	965G>A	Gly322Asp	6	(6, 7)
MSH2	1077-10T>C	-	7	(5, 8)
MSH2	1386+73G>A	-	8	-
MSH2	1387-9T>A	-	10	(5, 9)
MSH2	1661+12A>G	-	10	(10, 11)
MSH2	1787A>G	Asn596Ser	12	(12)
MSH2	2005+25T>G	-	12	-
MSH2	2006-6T>C	-	13	(11, 13)
MSH2	2210+77T>A	-	13	-
MSH2	2634+30delT	-	15	-
MSH6	116G>A	Gly39Glu	1	(14, 15)
MSH6	186C>A	Arg62Arg	1	(14, 15)
MSH6	260+22C>G	-	1	(16)
MSH6	260+25A>C	-	1	-
MSH6	260+43G>A	-	1	-
MSH6	260+96A>C	-	1	-
MSH6	260+101C>A	-	1	-
MSH6	260+131C>A	-	1	-
MSH6	276A>G	Pro92Pro	2	(14, 15)
MSH6	457+13A>G	-	2	(14)
MSH6	540T>C	Asp180Asp	3	(15, 17)
MSH6	628-56C>T	-	4	(14, 18)
MSH6	642C>T	Tyr214Tyr	4	(15, 17)
MSH6	660A>C	Glu220Asp	4	(15)

MSH6	1257G>A	Gln419Gln	4	-
MSH6	2239C>T	Leu747Leu	4	-
MSH6	2633T>C	Val878Ala	4	(19)
MSH6	3246G>T	Pro1082Pro	5	(15,19)
MSH6	3650G>A	Arg1217Lys	8	-

References for Table S2.3 (Appendix J):

1. Liu, B, Farrington,SM, Petersen, GM, et. al. Genetic instability occurs in the majority of young patients with colorectal cancer. Nat Med, 1995; 1(4):348-352

2. Moslein, G, Tester, DJ, Lindor, NM, et. al. Microsatellite instability and mutation analysis of hMSH2 and hMLH1 in p[atients with sporadic, familial and hereditary colorectal cancer. Hum Mol Genet, 1996; 5(9): 1245-1252

3. Tannergrad, P, Lipford, JR, Kolodner, R, et. al. A mutation screening in the hMLH1 gene in Swedish hereditary nonpolyposis colon cancer families. Cancer Res, 1995; 55(24): 6092-6096

4. Bubb, VJ, Curtis, LJ, Cunningham, C, et al. Microsatellite instability and the role of hMSH2 in sporadic colorectal cancer. Oncogene, 1996; 12(12): 2641-2649

5. Holinski-Feder, E, Muller-Koch, Y, Friedl, W, et al. DHPLC mutation analysis of the hereditary nonpolyposis colon cancer (HNPCC) genes hMLH1 and hMSH2. J Biochem Biophys Methods, 2001; 47(1-2): 21-32

6. Maliaka, YK, Chudina, AP, Belev, NF, et al. CpG dinucleotides in the hMSH2 and hMLH1 genes are hotspots for HNPCC mutations. Hum Genet. 1996; 97(2): 251-255

7. Nystrom-Lahti, M, Wu, Y, Moisio, AL, et al. DNA mismatch repair gene mutations in 55 kindreds with verified or putative hereditary non-polyposis colorectal cancer. Hum Mol Genet, 1996; 5(6): 763-769

8. Swensen, J, Lewis, CM, and Cannon-Albright, LA. Identification of a one-base germline deletion (codon 888delC) and an intron splice acceptor site polymorphism in hMSH2. Hum Mutat, 1997; 10(1): 80-81

9. Borresen, AL, Lothe, RA, Meling, GI, et al. Somatic mutations in the hMSH2 gene in microsatellite unstable colorectal carcinomas. Hum Mol Genet, 1995; 4(11): 2065-2072

10. Wijnen, J, Fodde, R, and Khan, PM. DGGE polymorphism in intron 10 of MSH2, the HNPCC gene. Hum Mol Genet, 1994; 3(12): 2268

11. Wahlberg, SS, Nystrom-Lahti, M, Kane, MF, et al. A low frequency of hMHS2 mutations in Swedish HNPCC families. Int J Cancer, 1997; 74(1): 134-137

12. Viel, A, Genuardi, M, Capozzi, E, et al. Characterization of MSH2 and MLH1 mutations in Italian families with hereditary nonpolyposis colorectal cancer. Genes Chromosomes Cancer, 1997; 18(1): 8-18.

13. Hall, NR, Taylor, GR, Finan, PJ, et al. Intron splice acceptor site sequence variation in the hereditary non-polyposis colorectal cancer gene hMSH2. Eur J Cancer, 1994; 30A(10): 1550-1552.

14. Nicolaides, NC, Palombo, F, Kinzler, KW, Vogelstein, B, and Jiricny, J Molecular cloning of the N-terminus of GTBP. Genomics, 1996; 31(3): 395-397.

15. Kolodner, RD, Tytell, JD, Schmeits, JL, et al. Germ-line msh6 mutations in colorectal cancer families. Cancer Res, 1999; 59(20): 5068-5074.

16. Peterlongo, P, Nafa, K, Lerman, GS, et al. MSH6 germline mutations are rare in colorectal cancer families. Int J Cancer, 2003; 107(4): 571-579.

17. Plaschke, J, Kruppa, C, Tischler, R, et al. Sequence analysis of the mismatch repair gene hMSH6 in the germline of patients with familial and sporadic colorectal cancer. Int J Cancer, 2000; 85(5): 606-613.

18. Parc, YR, Halling, KC, Burgart, LJ, et al. Microsatellite instability and hMLH1/hMSH2 expression in young endometrial carcinoma patients: associations with family history and histopathology. Int J Cancer, 2000; 86(1): 60-66.

19. Wijnen, J, de Leeuw, W, Vasen, H, et al. Familial endometrial cancer in female carriers of MSH6 germline mutations. Nat Genet, 1999; 23(2): 142-144.

Appendix K: Wish *et. al.*, 2010: "Increased Cancer Predisposition in Family Members of Colorectal Cancer Patients Harboring the p.V600E *BRAF* Mutation: A Population-based Study"

Wish T[1], Hyde A[2], Parfrey PS[1], Green J[2], Younghusband HB[2], Michelle Simms[2], Fontaine D[3], Dicks EL[1], Stuckless S[1], Gallinger S[4, 5], McLaughlin J[4, 5], Woods MO[2, 6], Green RC[2, 6]

1. Clinical Epidemiology Unit, Memorial University of Newfoundland. St. John's, Newfoundland, Canada.
2. Discipline of Genetics, Memorial University of Newfoundland. St. John's, Newfoundland, Canada.
3. Eastern Health Department of Pathology. St. John's, Newfoundland, Canada.
4. Samuel Lunenfeld Research Institute, Mount Sinai Hospital, University of Toronto. Toronto, Ontario, Canada.
5. Cancer Care Ontario. Toronto, Ontario, Canada.
6. These authors contributed equally to this work

Abstract

Background: The serrated pathway represents a distinct molecular pathway of colorectal carcinogenesis and is associated with the p.V600E *BRAF* mutation. The objective of this study is to characterize the cancer family history and clinicopathologic features of colorectal cancer (CRC) patients according to the microsatellite-instability (MSI) and *BRAF* mutation status of their tumors.

Methods: The tumors from 558 population-based CRC patients underwent pathological examination and molecular analysis for MSI, *BRAF*, and germline mutations in mismatch repair genes, *MUTYH* and *APC*. The cancer history in first-degree relatives (FDRs) of index patients was ascertained.

Results: The risk of CRC in FDRs of index patients with MSI-H *BRAF* mutation (HR = 2.49; 95% CI, 1.57 - 3.93) and microsatellite-stable (MSS) *BRAF* mutation tumors (HR = 1.64; 95% CI, 1.01 - 2.66) was significantly elevated compared to FDRs of index patients with MSS *BRAF* wild type tumors. The incidence of non-melanoma skin cancer was also significantly elevated in FDRs of patients with *BRAF* mutation CRC (HR = 2.52; 95% CI, 1.31 - 4.86). Furthermore, *BRAF* mutation CRC was associated with a distinct clinical, molecular, and pathological phenotype.

Conclusions: The increased incidence of cancer in FDRs of index CRC patients with the p.V600E *BRAF* mutation may be explained by a genetic predisposition to develop cancer via the serrated pathway of colorectal carcinogenesis.

Impact: Family members of *BRAF* CRC patients have an increased predisposition to develop cancer. Future work should aim to identify the causative genetic factors.

Introduction

Colorectal cancer (CRC) is a heterogenous disease that can arise via several different molecular pathways. The two most common are the chromosomal-instability (CIN) and microsatellite-instability (MSI) pathways, both of which promote tumorigenesis by causing genomic instability. CIN is observed in approximately 85% of colorectal tumors, and is characterized by aneuploidy and mutations of *APC*, *KRAS*, and *p53*. The MSI pathway occurs in tumors with DNA mismatch repair (MMR) deficiency, caused either by germline MMR gene mutations (Lynch Syndrome) or by inactivation of *MLH1* by aberrant promoter methylation (1).

More recently, the serrated pathway of CRC carcinogenesis has emerged as a novel molecular pathway, which is characterized by a distinct pathological, clinical, and molecular phenotype for which there is evidence of a genetic predisposition (2). The serrated pathway is tightly associated with the p.V600E *BRAF* mutation and with aberrant DNA methylation (3, 4), which is commonly referred to as the CpG island methylator phenotype (CIMP) (5). The p.V600E *BRAF* mutation is a T-to-A transversion at nucleotide 1796 that causes constitutive activation, and which promotes proliferation and inhibits apoptosis via the *Ras/Raf/MEK/MAPK* signal transduction pathway (6). The mutation is reported in approximately 10 – 18% of all colorectal tumors and in 30 - 50% that are microsatellite-instable (MSI-H) (7-9). This mutation is acknowledged as an early and primary genetic event in the serrated pathway of colorectal carcinogenesis (6, 10).

The precursor lesions of the serrated pathway include sessile serrated adenomas (SSAs), traditional serrated adenomas (TSAs), as well as other lesions (11). SSAs are

relatively common colonic neoplasms (9% of all colonic neoplasms) that are strongly associated with *BRAF*, CIMP, female sex, and proximal tumor location (12-14). Colorectal tumors expressing the *BRAF* mutation are associated with MSI-H, CIMP, *MLH1* promoter methylation, poor differentiation, mucinous morphology, and infiltrating lymphocytes (10, 15). However, it appears that inactivation of *MLH1* is not critical for carcinogenesis in the serrated pathway, since 5 – 10% of microsatellite-stable (MSS) tumors also display *BRAF* (7-9). Furthermore, CRC tumors with the *BRAF* mutation rarely have *KRAS* mutations and do not occur in the context of Lynch syndrome, which supports the notion that the serrated pathway represents a distinct molecular pathway of carcinogenesis (3). In addition to the serrated pathway, the *BRAF* mutation has been recognized as a marker for poor outcomes (9, 16) and resistance to anti-EGFR therapies (17). Inhibition of the mutated *BRAF* gene represents a novel and attractive therapy for patients with *BRAF* ^{Mut} CRC and several clinical trials are underway to evaluate this approach (18).

There is evidence for a genetic predisposition to tumorigenesis via the serrated pathway (2, 12). In addition, there is evidence to suggest that colorectal cancer patients with the somatic *BRAF* mutation have a stronger family history of cancer compared to patients with *BRAF* wild type tumors (9, 19, 20). However, the evidence is limited and patients who have MSI-H *BRAF* mutation tumors are generally regarded as sporadic cases (11).

As only one population-based study (9) has examined the familial basis of *BRAF* mutation CRC we examined the clinicopathologic features and cancer family histories of

CRC patients according to the MSI and *BRAF* status of their tumors, using a series of population-based colorectal cancer patients.

Methods:

Study Population. We identified 1,173 incident CRC patients (ICD-9 code; colon: 153.0 - 153.9, excluding 153.5; rectum: 154.0 – 154.1) from the Newfoundland Cancer Registry that were diagnosed before the age of 75 years, in the 5-year period between January 1, 1999 and December 31, 2003. Of these, 750 (64%) patients (or their proxies) consented to take part in the study. Index Patients were invited to complete a family history questionnaire, which enabled pedigrees to be constructed. From the family history questionnaire, we recorded the type of cancer, age at diagnosis, and age at death or age at last follow-up for all first-degree relatives (FDRs). Cancer diagnoses in consenting family members were confirmed by medical records whenever possible. 552 patients (47% of all eligible patients) completed a family history questionnaire and all of the molecular analyses. Ethics approval was obtained from the Human Investigation Committee of Memorial University.

Family History Criteria. A positive family history of CRC was defined as having at least one FDR affected by CRC at any age. We identified patients with a strong CRC family history if their family history fulfilled the Amsterdam 1 criteria (21). Additionally, we defined patients with a strong family history that had a later-age of onset if they met the Amsterdam 1 criteria, except did not have a CRC 50 < years. We refer to these families as CRC-Triad.

Molecular Analysis. Tumors from index patients underwent pathological review, and molecular analysis for microsatellite-instability (MSI), immunohistochemistry (IHC), *MLH1* promoter methylation, and the p.V600E *BRAF* mutation. DNA from patients whose tumors had mismatch repair (MMR) deficiency, as indicated by MSI or IHC, underwent MMR gene sequencing to identify mutations in *MLH1*, *MSH2*, *MSH6*, and *PMS2*. Patients with a personal or family history of polyposis were tested for *APC* mutations. DNA from most patients (97.6% of those which provided a blood sample) was screened for *MUTYH* mutations. Six patients with a germline mutation in either *APC* or *MUTYH* were excluded from this study.

The protocol for MSI analysis and IHC staining of the MMR proteins (*MLH1, MSH2, MSH6, and PMS2)* in colorectal tumours has previously been described in detail (22). Briefly, formalin-fixed, paraffin-embedded tissues were sectioned, deparaffinized and rehydrated using xylene and alcohol. The slides underwent microwave antigen retrieval followed by incubation with the appropriate monoclonal antibodies.

For MSI analyses, DNA was extracted from thick sections cut from paraffin-embedded tissue blocks that were determined by histology to contain predominantly tumor tissue. Five microsatellite markers were used for the analysis: *BAT25*, *BAT26*, *D2S123*, *D5S346* and *D17S250*. MSI was detected by the presence of additional bands in the PCR-amplified product from the tumor tissue, compared with those from DNA from matched normal colon tissue. For some samples MSI status was determined using the five mononucleotide markers - *BAT-25*, *BAT-26*, *NR-21*, *NR-24* and *MONO-27* - supplied in the MSI Analysis System kit Version 1.1 (Promega Corp, Madison, WI, USA), following

the manufacturers instructions. MSI status was assigned as MSI-high (MSI-H, >30% of markers tested unstable), MSI-low (MSI-L, >0% and <30% of markers unstable), or microsatellite stable (MSS, <30% markers unstable); however since only 25 tumors were identified as MSI-Low we combined them with the MSS tumors and thus assigned tumors as either MSI-H (>30% of markers unstable) or MSS (<30% of markers unstable).

Methylation was detected using the MS-MLPA kit ME001B (MRC-Holland, Amsterdam, Netherlands). This kit can generate two methylation-dependent signals from the *MLH1* promoter: a 166 bp fragment is produced if the *Hha1I* site at position -7 (relative to the ATG start codon) is methylated and a 463 bp fragment is generated if the *Hha1I* sites at -378 and -401 are both methylated. We scored the tumor DNA sample as methylated if either of these fragments was present at a normalized ratio of 0.15 of the peak area of the sample before digestion with *Hha1I* (23). To identify the c.1799T>A (p.Val600Glu) variant in *BRAF*, we used a protocol previously described (24). This is an allele-specific PCR assay that includes a set of primers for *GAPDH* as an internal positive control.

DNA sequencing of *MLH1, MSH2, MSH6* and *PMS2*. Alterations of *MLH1*, *MSH2* and *MSH6* were determined by sequencing all 45 exons and intron/exon boundaries. Genomic DNA from probands who had tumors deficient in an MMR protein was sequenced, except those cases with tumors that were deficient due to *MLH1* promoter methylation. Automated sequencing was performed on an ABI 3700 DNA Analyzer (Applied Biosystems, Foster City, CA). Sequence information of the coding region was

derived from RefSeq NM_000249.2 (*MLH1*), NM_000251.1 (*MSH2*) and NM_000179.1 (*MSH6*). Primer sequences and intronic nucleotide information were derived from genomic sequences from NCBI - AC011816.17 (MLH1), AC079775.6 (MSH2), AC006509.15 (MSH6). Primer sequences are available from the authors upon request. The *PMS2* (RefSeq NM_ 000535.4) variants were detected as described previously (25), with the modifications explained in Clendenning *et al.* (26).

Rearrangements within *MLH1* and *MSH2*. Exon deletions and duplications in *MSH2* and *MLH1* were detected by MLPA (27) in DNA from patients whose tumors were deficient in either MLH1 (unrelated to *MLH1* methylation) or MSH2 using kit SALSA P003 according to the protocol provided by MRC-Holland (Amsterdam, The Netherlands). All rearrangements identified by MLPA were confirmed in other affected relatives by MLPA and, when possible, cDNA analyses.

Pathology Review. One representative tumor slide from each patient's tumor was reviewed and scored for several histological features, including Crohn's-like lymphocytic reaction and tumor-infiltrating lymphocytes (TILS), as described previously (28). Tumor grade and histology were determined from the original pathology reports. Tumor location was obtained from the Newfoundland Cancer Registry. Proximal location was defined as proximal to the splenic flexure. Mucinous component was defined as the presence of any mucin dissecting into stroma surrounding a tumor gland. This definition includes tumors with a mucinous histology, but also those with histologic heterogeneity, in which any area of the tumor displays dissecting mucin. We defined an increase in stromal plasma cells to

be positive when plasma cells made up more than 25% of stromal immune cells. Pathology was reviewed for all available tumors, which included all MSI-H tumors and 98% of MSS tumors.

Statistical Analysis. Comparison of continuous variables was analyzed with independent-samples t test or one-way ANOVA. Categorical data were analyzed with the Pearson's Chi-square test. Multinomial logistic regression analysis was used to compare the clinicopathological features of patients with *BRAF* mutation tumors to patients with *BRAF* wild-type tumors. The analysis includes a univariate model and a multivariate model, which includes all variables as well as age at diagnosis of CRC as a covariate. A Cox regression model was used to estimate hazard ratios (HR) for developing cancer in FDRs according to the molecular classification of the index patient. Index patients were excluded from the estimates of cancer risk. Age at diagnosis or age at last follow-up was used as the time variable. Hazard ratios are adjusted for the age and sex of the index patient, and the sex of the FDR was entered as a strata variable. A univariate Cox regression analysis was used to estimate the risk of developing colorectal cancer in FDRs according to the sex and type of FDR. The cumulative lifetime risk (< 80 years of age) of developing cancer in FDRs was estimated with Kaplan-Meier survival analysis. All *P* values were two-sided and $P \leq 0.05$ was considered significant. All analyses were performed with the SPSS software package, version 17.0 (Chicago, IL).

Results

In 552 population-based colorectal cancer patients we identified the p.V600E *BRAF* mutation in 12% (n = 65) of all patients' tumors and microsatellite-instability in 11% (n = 61) of all patients' tumors. Nearly half of all MSI-H tumors harbored the p.V600E *BRAF* mutation (n = 27, 44%) and the remaining MSI-H *BRAF* ^{Wt} tumors either harbored a pathogenic MMR gene mutation (n = 17) or were of unknown etiology (n = 17).

The majority of patients' tumors were MSS *BRAF* ^{Wt} (n = 453, 82%). These patients were predominantly male (36% female) with distally located tumors (33% proximal) (Table 1). The p.V600E *BRAF* mutation was identified in 8% (n = 38) of 491 index patients with a MSS tumor. In contrast to patients with MSS *BRAF* ^{Wt} tumors, MSS *BRAF* ^{Mut} patients were predominantly female (63%), with proximally located tumors (81%). Similarly, patients with a MSI-H *BRAF* ^{Mut} tumor were predominantly female (67%), with proximally located tumors (96%), however in addition they had a high incidence of multiple tumors (25%).

Table 1. The clinical, molecular, and pathological features of index patients classified according to the molecular characteristics of their tumors

	MSS		**MSI-H**		
	BRAF ^{Wt}	***BRAF*** ^{Mut}	***BRAF*** ^{Mut}	**MMR** ^{Mut}	
	n (%)	*n (%)*	*n (%)*	*n (%)*	*p**
Index Patients	453	38	27	17	
Mean age at CRC diagnosis (SD), yrs	61.3 (9.0)	61.0 (7.3)	66.2 (6.4)	51.1 (10.3)	<0.001
Sex (Female)	165 (36.4)	24 (63.2)	18 (66.7)	5 (29.4)	<0.001
Multiple tumors[†]	34 (7.5)	4 (10.5)	7 (25.9)	7 (41.2)	<0.001
Family History of CRC	121 (26.7)	15 (39.5)	14 (51.9)	13 (76.5)	<0.001
Amsterdam 1 criteria	13 (2.9)	2 (5.3)	0 (0)	12 (70.6)	<0.001
CRC - Triad	16 (3.5)	0 (0)	6 (22.2)	0 (0)	<0.001
Molecular (Index Patients)					

MLH1 deficient	0	1 (2.6)	24 (88.9)	3 (17.6)	<0.001
MLH1 methylation	1 (0.2)	2 (5.3)	26 (96.3)	-	<0.001
Pathology (Index Patients)					
Crohn's-like reaction	248 (56.6)	22 (59.5)	19 (70.4)	8 (53.3)	0.55
Increased stromal plasma cells	249 56.3)	14 (37.8)	22 (81.5)	12 (80.0)	0.001
Presence of TILS	100 (22.6)	11 (29.7)	17 (63.0)	9 (60.0)	<0.001
Mucinous component	95 (21.5)	17 (45.9)	17 (63.0)	5 (33.3)	<0.001
Differentiation (Poor)	40 (9.2)	7 (19.4)	7 (29.2)	2 (14.3)	0.007
Histology (Adenocarcinoma NOS)	392 (88.7)	29 (78.4)	20 (74.1)	11 (73.3)	0.02
Tumor location (Proximal)	145 (32.8)	30 (81.1)	26 (96.3)	8 (53.3)	<0.001
First-Degree Relatives (FDRs)	4337	370	290	154	
FDRs diagnosed with CRC	153 (3.5)	19 (5.1)	22 (7.6)	33 (21.4)	<0.001
Mean age at CRC diagnosis (SD), yrs	63.9 (12.5)	59.0 (10.8)	63.7 (11.6)	48.0 (14.1)	<0.001

* Comparison of all four groups using One-way ANOVA or Pearson's Chi-square test as appropriate.

† Multiple tumors: synchronous or metachronous tumors.

Seventeen patients with a MSI-H tumor were found to have a germline mutation in *MLH1*, *MSH2*, *MSH6*, or *PMS2* (3.1% of the entire cohort). These patients are characterized by having multiple tumors (41%), early age onset colorectal cancer (mean age = 51.1 yrs), and a family history of CRC (77%).

A multinomial logistical regression analysis was conducted in order to investigate the clinicopathological features of patients according to the MSI and *BRAF* mutation status of their tumors' (Table 2). Patients with MSS *BRAF* ^{Wt} tumors were assigned as the reference category. Patients with MSS *BRAF* ^{Mut} tumors are significantly more likely to be female (OR = 3.1; 95% CI, 1.4 – 6.8) and have proximally located tumors (OR = 6.3; 95% CI, 2.6 – 15.1), which have a mucinous component (OR = 2.9; 95% CI, 1.1 – 7.4). Similarly, patients with MSI-H *BRAF* ^{Mut} tumors are significantly more likely to have proximally located (OR = 27.5; 95% CI, 3.5 - 213), mucinous tumors (OR = 5.8; 95% CI, 1.8 – 18.4), which presented with TILS (OR = 5.2; 95% CI, 1.8 – 15.5). However, the association with female sex did not reach statistical significance (OR = 2.7;

95% CI, 0.9 – 2.1). The tumors of patients with Lynch syndrome are characterized by having TILS (OR = 5.0; 95% CI, 0.9 – 28.6) and a mucinous component (OR = 3.9; 95% CI, 1.0 – 14.5), however the association with TILS is not statistically significant.

Table 2. Comparison of the clinicopathologic features of index patients according to the molecular characteristics of the index patients' tumor

	MSS *BRAF* ^{Wt}	**MSS *BRAF* ^{Mut}**			
	OR (95% CI)	*OR* (95% CI)*	*p*	*OR† (95% CI)*	*p*
Sex (Female)	1.0 (ref.)	3.0 (1.5 - 5.9)	0.002	3.1 (1.4 - 6.8)	0.004
Crohn's-like reaction	1.0 (ref.)	1.1 (0.6 - 2.2)	0.74	1.1 (0.5 - 2.3)	0.81
Increased stromal plasma cells	1.0 (ref.)	0.5 (0.2 - 0.9)	0.03	0.5 (0.2 - 1.0)	0.06
Presence of TILS	1.0 (ref.)	1.5 (0.7 - 3.0)	0.33	1.2 (0.5 - 3.0)	0.64
Mucinous component	1.0 (ref.)	3.1 (1.6 - 6.2)	0.001	2.9 (1.1 - 7.4)	0.03
Differentiation (Poor)	1.0 (ref.)	2.4 (1.0 - 5.8)	0.06	2.3 (0.9 - 6.1)	0.10
Histology (Adenocarcinoma NOS)	1.0 (ref.)	0.5 (0.2 - 1.1)	0.07	1.9 (0.6 - 6.5)	0.29
Tumor location (Proximal)	1.0 (ref.)	8.8 (3.8 - 20.5)	<0.001	6.3 (2.6 - 15.1)	<0.001

* Univariate model

† Multivariate model

Table 2 (cntd)

MSI-H *BRAF* ^{Mut}				**MMR Mutation**			
OR (95% CI)*	*p*	*OR† (95% CI)*	*p*	*OR* (95% CI)*	*p*	*OR† (95% CI)*	*p*
3.5 (1.5 - 8.0)	0.003	2.7 (0.9 - 7.7)	0.07	0.7 (0.3 - 2.1)	0.60	0.6 (0.2 - 2.0)	0.39
1.8 (0.8 - 4.2)	0.17	2.2 (0.7 - 7.7)	0.17	0.9 (0.3 - 2.5)	0.80	0.7 (0.2 - 2.5)	0.63
3.4 (1.3 - 9.2)	0.02	3.3 (0.9 - 12.0)	0.08	3.1 (0.9 - 11.1)	0.08	5.0 (0.9 - 28.6)	0.07
5.8 (2.6 - 13.1)	<0.001	5.2 (1.8 - 15.5)	0.003	5.1 (1.8 - 14.8)	0.002	3.9 (1.0 - 14.5)	0.05
6.2 (2.8 - 14.0)	<0.001	5.8 (1.8 - 18.4)	0.003	1.8 (0.6 - 5.5)	0.28	0.9 (0.1 - 8.1)	0.93
4.1 (1.6 - 10.3)	0.004	2.8 (0.8 - 9.6)	0.09	1.6 (0.4 - 7.6)	0.53	0.6 (0.1 - 3.4)	0.53
0.4 (0.2 - 0.9)	0.03	1.2 (0.3 - 5.2)	0.85	0.4 (0.1 - 1.1)	0.08	0.2 (0.02 - 2.3)	0.19
53.3 (7.2 - 396)	<0.001	27.5 (3.5 - 213)	0.002	2.3 (0.8 - 6.6)	0.11	2.2 (0.6 - 7.7)	0.21

Approximately 5% (n = 2) of patients with a MSS *BRAF* ^{Mut} tumor reported a strong family history by meeting the Amsterdam 1 criteria, however none fulfilled the definition of a CRC-Triad. In patients with a MSI-H *BRAF* ^{Mut} tumor a family history

that satisfied the CRC-Triad was reported by 22% (n = 6) of patients, and no patients satisfied the Amsterdam I criteria.

We estimated the risk of developing cancer in first-degree relatives (FDRs) according to the molecular classification of the index patients' tumor with a Cox regression model. The hazard of developing CRC in FDRs of patients with a *BRAF* ^{Mut} tumor was significantly greater compared to the FDRs of index patients with a MSS *BRAF* ^{Wt} tumor (HR = 1.93; 95% CI 1.35 – 2.75) (Table 3; Fig. 1). Furthermore, the lifetime risk of developing CRC was significantly greater in the FDRs of patients with a *BRAF* ^{Mut} tumor compared to the FDRs of patients with MSS *BRAF* ^{Wt} tumors (LR% = 19% vs. 11%; Log Rank, $P < 0.001$). The risk of CRC in FDRs of patients with Lynch syndrome was significantly greater compared to all other molecular classifications.

Table 3. The risk of developing colorectal cancer in FDRs according to the molecular characteristics of the index patients' tumor

	FDRs at-risk	**CRC Events**	**Risk of CRC**	
	n	*n*	*Hazard Ratio (95% CI)*	*Lifetine Risk % (95% CI)*
MSS *BRAF* ^{wt}	4337	153	1.0 (ref.)	11 (9 - 13)
All *BRAF* ^{Mut}	660	41	1.93 (1.35 - 2.75)	19 (13 - 25)
MSS *BRAF* ^{Mut}	370	19	1.64 (1.01 - 2.66)	17 (9 - 25)
MSI-H *BRAF* ^{Mut}	290	22	2.49 (1.57 - 3.93)	21 (12 - 30)
MMR Mutation	154	33	9.57 (6.22 - 14.73)	46 (32 - 61)

Wish T. Figure 1.

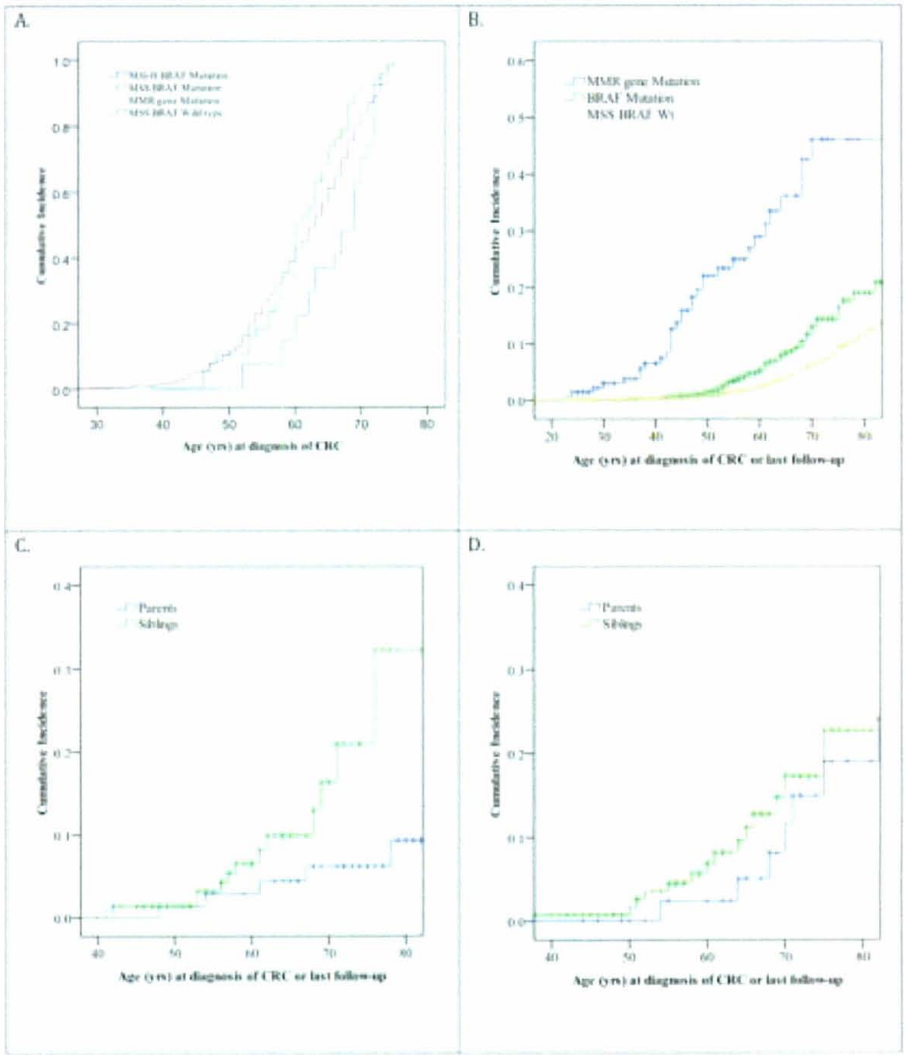


Figure 1. The cumulative age of onset of CRC in index patients (A) and FDRs according to the *BRAF* mutation status of the index patients' tumor (B). The cumulative incidence

of CRC in siblings and parents of index patients with MSS *BRAF* ^{Mut} tumors (C) and MSI-H *BRAF* ^{Mut} tumors (D).

We investigated the risk of developing other common malignancies in FDRs according to the molecular classification of the index patients' tumor (Table 4). The risk of non-melanoma skin cancer was significantly elevated in the FDRs of patients with a MSS *BRAF* ^{Mut} tumor (HR = 2.61; 95% CI, 1.15 – 5.92) and MSI-H *BRAF* ^{Mut} tumor (HR = 2.81; 95% CI, 1.07 – 7.41) compared to the FDRs of patients with MSS *BRAF* ^{Wt} tumors. The incidence of other extracolonic tumors did not differ between FDRs of patients with or without the p.V600E *BRAF* mutation.

In order to evaluate the inheritance pattern of CRC in FDRs we compared the incidence of CRC between parents and siblings according to the molecular classification of the index patients' tumor (Table 5, Fig 1). In the FDRs of index patients with a MSS *BRAF* ^{Mut} tumor, the risk of CRC was significantly greater in siblings compared to parents (HR = 3.28; 95% CI, 1.09 – 9.89). This was not observed in the FDRs of patients with MSI-H *BRAF* ^{Mut} tumors (HR = 1.23; 95% CI, 0.65 – 3.06).

Table 4. The risk of developing extracolonic cancers in FDRs according to the molecular characteristics of the index patients' tumor

	FDRs at-risk	Risk of Breast Cancer		Risk of Lung Cancer		Risk of Skin* Cancer		Risk of Prostate Cancer	
	n	*n*	*HR (95% CI)*	*n*	*HR (95% CI)*	*n*	*HR (95% CI)*	*n*	*HR (95% CI)*
MSS *BRAF*^{wt}	4337	65	1.0 (ref.)	67	1.0 (ref.)	37	1.0 (ref.)	38	1.0 (ref.)
All *BRAF*^{Mut}	660	7	0.90 (0.40 - 1.99)	8	0.90 (0.43 - 1.90)	12	2.52 (1.31 - 4.86)	5	1.14 (0.45 - 2.97)
MSS *BRAF*^{Mut}	370	3	0.71 (0.22 - 2.26)	6	1.15 (0.49 - 2.66)	7	2.61 (1.15 - 5.92)	4	1.41 (0.50 - 3.98)
MSI-H *BRAF*^{Mut}	290	4	1.14 (0.41 - 3.20)	2	0.52 (0.13 - 2.16)	5	2.81 (1.07 - 7.41)	1	0.65 (0.09 - 4.84)
MMR Mutation	154	1	0.35 (0.05 - 2.74)	2	1.24 (0.29 - 5.25)	0	-	0	-

* non-melanom skin cancer

† HR = Hazard Ratio

Table 5. The risk of developing colorectal cancer in FDRs according to type and sex of FDR

	Sex of FDR	FDRs at-risk	CRC	Risk of CRC	Type of FDR	FDRs at-risk	CRC	Risk of CRC
		n	*n*	*HR (95% CI)*		*n*	*n*	*HR (95% CI)*
MSS *BRAF*^{wt}	Female	2149	60	1.0 (ref.)	Parent	866	67	1.0 (ref.)
	Male	2188	93	1.84 (1.33 – 2.55)	Sibling	2047	83	1.61 (1.14 – 2.28)
All *BRAF*^{Mut}	Female	329	16	1.0 (ref.)	Parent	119	13	1.0 (ref.)
	Male	331	25	1.71 (0.91 – 3.21)	Sibling	286	27	2.07 (1.03 – 4.16)
MSS *BRAF*^{Mut}	Female	176	6	1.0 (ref.)	Parent	71	5	1.0 (ref.)
	Male	194	13	2.01 (0.76 – 5.31)	Sibling	151	13	3.28 (1.09 – 9.89)
MSI-H *BRAF*^{Mut}	Female	153	10	1.0 (ref.)	Parent	48	8	1.0 (ref.)
	Male	137	12	1.62 (0.70 – 3.77)	Sibling	135	14	1.23 (0.65 – 3.06)
MMR Mutation	Female	82	16	1.0 (ref.)	Parent	29	11	1.0 (ref.)
	Male	72	17	1.52 (0.76 - 3.04)	Sibling	84	18	1.13 (0.51 - 2.53)

* HR = Hazard Ratio

Discussion

In this population-based study we found evidence to suggest that FDRs of colorectal cancer patients with *BRAF* ^{Mut} tumors have a significantly elevated predisposition to develop colorectal cancer and non-melanoma skin cancer compared to patients with *BRAF* ^{Wt} tumors. The risk of developing colorectal cancer was significantly greater in FDRs of index patients with *BRAF* ^{Mut} tumors compared to index patients with MSS *BRAF* ^{Wt} tumors (HR = 1.93; 95% CI, 1.35 – 2.75). Furthermore, the risk was significantly elevated in the family members of index patients with either a MSI-H *BRAF* ^{Mut} or MSS *BRAF* ^{Mut} tumor. We also observed a significant increase in the risk of non-melanoma skin cancer in the FDRs of patients with *BRAF* ^{Mut} tumors, compared to the FDRs of patients with *BRAF* ^{Wt} tumors. In addition, this analysis provides further evidence that CRC patients with the somatic p.V600E *BRAF* mutation are characterized by a distinct clinical, molecular, and pathological phenotype.

The p.V600E *BRAF* mutation is recognized as a marker of the serrated pathway (29), and previous studies (7-9) have reported that approximately 10 – 18% of all CRCs and up to 50% of MSI-H tumors harbor the *BRAF* mutation. In our study, we observed the *BRAF* mutation in 12% of all tumors and in 47% of MSI-H tumors, which is comparable to previous reports (7-9). The increased incidence of cancer in FDRs of patients harboring the p.V600E *BRAF* mutation may be explained by an inherited predisposition to develop cancer via the serrated pathway of carcinogenesis. Evidence for an inherited predisposition to develop colorectal cancer via the serrated pathway is linked

to two colorectal cancer predisposition syndromes, namely hyperplastic polyposis (HPS) and serrated pathway syndrome (SPS) (30). The latter was described by Young *et. al.* (19), who provided evidence for an autosomal dominant predisposition in a number of families, which develop advanced serrated lesions and MSI-variable colorectal cancer that were associated with the *BRAF* mutation, female sex, and right-sided tumors. Interestingly, in our study some patients with *BRAF* ^{Mut} tumors exhibited a family history that is consistent with autosomal dominant disease. Six index patients with a MSI-H *BRAF* ^{Mut} tumor fulfilled the definition of a CRC-Triad and two patients with MSS *BRAF* ^{Mut} tumors reported a family history that fulfills the familial colorectal type X (FCCTX) criteria (31). Further evidence for a genetic predisposition to the serrated pathway comes from a recent study (12), which reported that patients with sessile serrated adenomas, compared to patients with other colonic lesions, were more likely to have a family history of colorectal cancer (42% vs. 25%; $P > 0.05$) and a greater polyp burden ($P < 0.001$).

A small number of studies (9, 19, 20) have investigated the association between colorectal cancer patients with *BRAF* ^{Mut} tumors and a family history of cancer, only one of which used a population-based approach (9). A correlation between a family history of colorectal cancer and patients with MSI-H *BRAF* ^{Mut} tumors has been previously reported (20). However, that study (20) used a selected cohort of familial colorectal patients, and only eight MSI-H tumors were evaluated. Samowitz *et. al.* (9), utilized a large number of unselected population-based colorectal cancer patients and reported a significant association between patients with MSS *BRAF* ^{Mut} tumors and a family history of colorectal cancer (OR = 4.23; 95% CI, 1.65 – 10.84). However, no association was found amongst patients with MSI-H *BRAF* ^{Mut} tumors (OR = 0.64; 95% CI, 0.18 – 2.19). To

our knowledge, our study is the first to report an association between patients with MSI-H *BRAF* ^{Mut} tumors and a family history of CRC, using an unselected series of population-based CRC patients. In our study, the risk of developing CRC was significantly elevated in the FDRs of index patients with either MSS *BRAF* ^{Mut} or MSI-H *BRAF* ^{Mut} tumors.

The incidence of Lynch syndrome in this population (3.1%) is consistent with estimates that attribute 3-5% of the total CRC burden to Lynch syndrome (32). Lynch syndrome is a highly penetrant autosomal dominant condition that is characterized by early-age, right-sided CRC, as well as various other extracolonic malignancies (33). Our findings are consistent with studies reporting that MMR mutations are highly penetrant (34), since the FDRs of Lynch syndrome patients in this study were affected at an early age, and nearly 50% were diagnosed with CRC by age 80 years.

In FDRs of patients with MSS *BRAF* ^{Mut} tumors the risk of CRC was found to be significantly greater in siblings compared to parents, where as in FDRs of patients with MSI-H *BRAF* ^{Mut} tumors, the risk was similar for parents and siblings. These results suggest that the genetic factors that cause *BRAF* ^{Mut} CRC predisposition may be recessively inherited in families with MSS *BRAF* ^{Mut} CRC. In contrast, the incidence of CRC in families with MSI-H *BRAF* ^{Mut} CRC is more suggestive of dominant or multi-factorial disease. It has been postulated that the burden of colorectal cancer arising from the serrated pathway could be explained by the presence of common co-dominant or recessive alleles causing defective epigenetic regulation (30). In a co-dominant model of inheritance, an inherited predisposition to develop sessile serrated adenomas and colorectal cancer may be attributable to carriers of one allele, and the more severe phenotype associated with HPS may be the result of homozygous carriers or a carrier

paired with a recessive allele. Studies (35, 36) have demonstrated that CRC risk is significantly greater amongst siblings than for parent – offspring, which is suggestive of recessive inheritance. Interestingly, in a large population-based study (35) the greatest CRC risk was associated with an affected sibling who had a right-sided colon cancer. Our results would suggest that MSS *BRAF* ^{Mut} CRC and the serrated pathway could account for some of the excess CRC risk observed in the siblings of patients with right-sided CRC.

In addition to CRC, we observed an increase in the incidence of non-melanoma skin cancer in the FDRs of patients with *BRAF* ^{Mut} CRC. An association between CRC and skin cancer has recently been observed in a prospective cohort study (37), which reported a two-fold increase in the risk of CRC following a diagnosis of non-melanoma skin cancer. Risk factors for non-melanoma skin cancer and *BRAF* ^{Mut} CRC include environmental exposures such as UV radiation and smoking (38). Furthermore, susceptibility to both non-melanoma skin cancer and *BRAF* ^{Mut} CRC has been shown (39, 40) to be modified by the interaction between environmental exposures and polymorphisms in the base excision repair (BER) genes, *XRCC1* and *OGG1*, respectively. A recent case-control study (40) reported that ever-smokers homozygous for the *OGG1* (*S326C)* polymorphism were twice as likely to have the *BRAF* mutation CRC. Polymorphisms of various DNA repair genes have been linked (41) to cancer susceptibility and for some cancer types the risk is exacerbated in individuals who are exposed to environmental carcinogens such as smoking. The association observed in this study might be explained by a germline variant that would increase susceptibility to both cancer types by attenuating DNA repair capability.

Relative to patients with *BRAF* ^{Wt} tumors, those harboring the p.V600E *BRAF* mutation were characterized by females with poorly differentiated, proximally located tumors, which had a mucin component. Furthermore, patients with MSI-H *BRAF* ^{Mut} tumors were older at diagnosis, more likely to have synchronous or metachronous tumors, and to have tumors that presented with increased stromal plasma cells and tumor-infiltrating lymphocytes, compared to patients with MSS *BRAF* ^{Mut} tumors. These findings support the notion that *BRAF* ^{Mut} colorectal cancer represents a distinct clinical, molecular, and pathological phenotype (9, 10, 15).

There are a number of limitations to this study that should be acknowledged. Although we ascertained population-based patients, only those diagnosed before 75 years of age were included, which may have enriched the study cohort with patients having a greater genetic predisposition. The statistical power to detect a difference in clinicopathological features and cancer risk may have been limited due to the small number of patients with MSI-H and *BRAF* ^{Mut} tumors, and by the small number of cancer events in FDRs. We have made numerous comparisons and it is possible that some of the findings may be due to chance. Although we made every effort to verify all cancers reported in FDRs, family history was self-reported and therefore susceptible to recall bias. However, family history information was obtained prior to knowledge of the molecular characteristics of the index patients' tumors, and therefore should not have systematically biased the results.

This analysis provides further evidence that patients with *BRAF* ^{Mut} CRC have a distinct clinical, molecular, and pathological phenotype. However, in contrast to the notion that these patients represent sporadic cases, we observed a significant association

with a family history of cancer. Approximately 12% of all index patients were found to harbor the *BRAF* mutation and the risk of CRC in FDRs of these patients was significantly elevated. In addition, we observed that the CRC risk to siblings and parents of patients with *BRAF* ^{Mut} CRC was dependent on the MSI status of the patients' tumor. Furthermore, we observed an association with non-melanoma skin cancer in FDRs of patients with *BRAF* ^{Mut} CRC. Additional large population-based analyses are needed to confirm these findings, which could have clinically relevant implications for cancer screening.

References for manuscript (Appendix K):

1. Aguilera A, Gomez-Gonzalez B. Genome instability: a mechanistic view of its causes and consequences. Nature reviews 2008; 9:204-17.

2. Young J, Jass JR. The case for a genetic predisposition to serrated neoplasia in the colorectum: hypothesis and review of the literature. Cancer Epidemiol Biomarkers Prev 2006; 15:1778-84.

3. Noffsinger AE. Serrated polyps and colorectal cancer: new pathway to malignancy. Annu Rev Pathol 2009; 4:343-64.

4. Makinen MJ. Colorectal serrated adenocarcinoma. Histopathology 2007; 50:131-50.

5. Weisenberger DJ, Siegmund KD, Campan M, et al. CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. Nature genetics 2006; 38:787-93.

6. Davies H, Bignell GR, Cox C, et al. Mutations of the BRAF gene in human cancer. Nature 2002; 417:949-54.

7. Domingo E, Espin E, Armengol M, et al. Activated BRAF targets proximal colon tumors with mismatch repair deficiency and MLH1 inactivation. Genes Chromosomes Cancer 2004; 39:138-42.

8. Lubomierski N, Plotz G, Wormek M, et al. BRAF mutations in colorectal carcinoma suggest two entities of microsatellite-unstable tumors. Cancer 2005; 104:952-61.

9. Samowitz WS, Sweeney C, Herrick J, et al. Poor survival associated with the BRAF V600E mutation in microsatellite-stable colon cancers. Cancer research 2005; 65:6063-9.

10. Velho S, Moutinho C, Cirnes L, et al. BRAF, KRAS and PIK3CA mutations in colorectal serrated polyps and cancer: primary or secondary genetic events in colorectal carcinogenesis? BMC Cancer 2008; 8:255.

11. Jass JR. Classification of colorectal cancer based on correlation of clinical, morphological and molecular features. Histopathology 2007; 50:113-30.

12. Spring KJ, Zhao ZZ, Karamatic R, et al. High prevalence of sessile serrated adenomas with BRAF mutations: a prospective study of patients undergoing colonoscopy. Gastroenterology 2006; 131:1400-7.

13. Goldstein NS. Clinical significance of (sessile) serrated adenomas: Another piece of the puzzle. Am J Clin Pathol 2005; 123:329-30.

14. Carr NJ, Mahajan H, Tan KL, Hawkins NJ, Ward RL. Serrated and non-serrated polyps of the colorectum: their prevalence in an unselected case series and correlation of BRAF mutation analysis with the diagnosis of sessile serrated adenoma. J Clin Pathol 2009; 62:516-8.

15. Li WQ, Kawakami K, Ruszkiewicz A, Bennett G, Moore J, Iacopetta B. BRAF mutations are associated with distinctive clinical, pathological and molecular features of colorectal cancer independently of microsatellite instability status. Mol Cancer 2006; 5:2.

16. Ogino S, Nosho K, Kirkner GJ, et al. CpG island methylator phenotype, microsatellite instability, BRAF mutation and clinical outcome in colon cancer. Gut 2009; 58:90-6.

17. Bardelli A, Siena S. Molecular mechanisms of resistance to cetuximab and panitumumab in colorectal cancer. J Clin Oncol 2010; 28:1254-61.

18. Garber K. Trial offers early test case for personalized medicine. Journal of the National Cancer Institute 2009; 101:136-8.

19. Young J, Barker MA, Simms LA, et al. Evidence for BRAF mutation and variable levels of microsatellite instability in a syndrome of familial colorectal cancer. Clin Gastroenterol Hepatol 2005; 3:254-63.

20. Vandrovcova J, Lagerstedt-Robinsson K, Pahlman L, Lindblom A. Somatic BRAF-V600E mutations in familial colorectal cancer. Cancer Epidemiol Biomarkers Prev 2006; 15:2270-3.

21. Vasen HF, Mecklin JP, Khan PM, Lynch HT. The International Collaborative Group on Hereditary Non-Polyposis Colorectal Cancer (ICG-HNPCC). Diseases of the colon and rectum 1991; 34:424-5.

22. Woods MO, Hyde AJ, Curtis FK, et al. High frequency of hereditary colorectal cancer in Newfoundland likely involves novel susceptibility genes. Clin Cancer Res 2005; 11:6853-61.

23. Gylling A, Abdel-Rahman WM, Juhola M, et al. Is gastric cancer part of the tumour spectrum of hereditary non-polyposis colorectal cancer? A molecular genetic study. Gut 2007; 56:926-33.

24. Loughrey MB, Waring PM, Tan A, et al. Incorporation of somatic BRAF mutation testing into an algorithm for the investigation of hereditary non-polyposis colorectal cancer. Fam Cancer 2007; 6:301-10.

25. Clendenning M, Hampel H, LaJeunesse J, et al. Long-range PCR facilitates the identification of PMS2-specific mutations. Human mutation 2006; 27:490-5.

26. Clendenning M, Senter L, Hampel H, et al. A frame-shift mutation of PMS2 is a widespread cause of Lynch syndrome. J Med Genet 2008; 45:340-5.

27. Schouten JP, McElgunn CJ, Waaijer R, Zwijnenburg D, Diepvens F, Pals G. Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. Nucleic acids research 2002; 30:e57.

28. Jenkins MA, Hayashi S, O'Shea AM, et al. Pathology features in Bethesda guidelines predict colorectal cancer microsatellite instability: a population-based study. Gastroenterology 2007; 133:48-56.

29. Kim YH, Kakar S, Cun L, Deng G, Kim YS. Distinct CpG island methylation profiles and BRAF mutation status in serrated and adenomatous colorectal polyps. International journal of cancer 2008; 123:2587-93.

30. Young J, Jenkins M, Parry S, et al. Serrated pathway colorectal cancer in the population: genetic consideration. Gut 2007; 56:1453-9.

31. Lindor NM, Rabe K, Petersen GM, et al. Lower cancer incidence in Amsterdam-I criteria families without mismatch repair deficiency: familial colorectal cancer type X. JAMA 2005; 293:1979-85.

32. Lynch HT, de la Chapelle A. Hereditary colorectal cancer. The New England journal of medicine 2003; 348:919-32.

33. Vasen HF, Watson P, Mecklin JP, Lynch HT. New clinical criteria for hereditary nonpolyposis colorectal cancer (HNPCC, Lynch syndrome) proposed by the International Collaborative group on HNPCC. Gastroenterology 1999; 116:1453-6.

34. Stoffel E, Mukherjee B, Raymond VM, et al. Calculation of risk of colorectal and endometrial cancer among patients with lynch syndrome. Gastroenterology 2009; 137:1621-7.

35. Hemminki K, Chen B. Familial risks for colorectal cancer show evidence on recessive inheritance. International journal of cancer 2005; 115:835-8.

36. Boardman LA, Morlan BW, Rabe KG, et al. Colorectal cancer risks in relatives of young-onset cases: is risk the same across all first-degree relatives? Clin Gastroenterol Hepatol 2007; 5:1195-8.

37. Chen J, Ruczinski I, Jorgensen TJ, et al. Nonmelanoma skin cancer and risk for subsequent malignancy. Journal of the National Cancer Institute 2008; 100:1215-22.

38. Samowitz WS, Albertsen H, Sweeney C, et al. Association of smoking, CpG island methylator phenotype, and V600E BRAF mutations in colon cancer. Journal of the National Cancer Institute 2006; 98:1731-8.

39. Nelson HH, Kelsey KT, Mott LA, Karagas MR. The XRCC1 Arg399Gln polymorphism, sunburn, and non-melanoma skin cancer: evidence of gene-environment interaction. Cancer research 2002; 62:152-5.

40. Curtin K, Samowitz WS, Wolff, RK, et al. Assessing tumor mutations to gain insight into base excision repair sequence polymorphisms and smoking in colon cancer. Cancer Epidemiol Biomarkers Prev 2009; 18:3384-8.

41. Paz-Elizur T, Sevilya Z, Leitner-Dagan Y, Elinger D, Roisman LC, Livneh Z. DNA repair of oxidative DNA damage in human carcinogenesis: potential application for cancer risk assessment and prevention. Cancer Lett 2008; 266:60-72.