Characterizing the Prevalence of Chromosomal Instability in Interval Colorectal Cancer

by

Amy L. Cisyk

A Thesis submitted to the Faculty of Graduate Studies of
The University of Manitoba
in partial fulfilment of the requirements of the degree of

MASTER OF SCIENCE

Department of Biochemistry & Medical Genetics
University of Manitoba
Winnipeg

TABLE OF CONTENTS

ABSTRACT

Over 80% of colorectal cancers (CRCs) are sporadic/randomly arising tumors. Interval CRCs represent a subset of sporadic tumors that develop within 6-36 months after a negative colonoscopy. Interval CRCs are suggested to exhibit altered biological properties that contribute to rapid growth and proliferation. We hypothesize that chromosomal instability (CIN), or aberrant chromosome numbers, contributes to the etiology of Interval CRCs.

We have assembled a Manitoban cohort of Interval and sporadic (control) CRC tumor samples, and established a fluorescence *in situ* hybridization approach to characterize CIN by enumerating specific chromosomes.

The results of this study indicate that 75% of Interval CRCs exhibit a CIN phenotype, making CIN the most prevalent contributor to genomic instability in Interval CRCs. Only once we grasp a better understanding of the tumorigenic pathways through which Interval CRCs develop, can we tailor screening strategies and treatment options to specifically identify and combat this subset of sporadic CRC.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest gratitude towards my supervisor and mentor, Dr. Kirk McManus. Through his passion and dedication to research, he has inspired in me the drive and fortitude necessary to accomplish this degree. Without his instrumental guidance this thesis would not have been possible. I could not have asked for a more positive experience and I attribute a great deal of that to Dr. McManus's meticulous and demanding, yet lighthearted and sometimes quirky, supervision.

I acknowledge and thank my committee members Dr. Etienne Leygue, Dr. Sabine Mai and Dr. Harminder Singh, who have dedicated their time and expertise to guide me and critique my work throughout the last 2 ½ years of graduate studies.

In addition, I certainly can't forget Zelda Lichtensztejn, who provided countless hours of training and assistance to me. From day one, she guided me through many of the protocols and techniques that make up the foundation of this thesis (from cell culture, to generating mitotic spreads, to FISH). Furthermore, Signe Penner-Goeke dedicated her summer to study in our lab; her hard work and contributions are greatly appreciated. Zelda, Signe and lab members past and present, including Sajesh Babu, Brent Guppy and Laura Thompson, have been such a pleasure to work with. They have made coming to work exciting and entertaining day-in and day-out.

A special thanks goes out to Andrea Fristensky from the Manitoba Tumor Bank, who cored each of the mini-TMAs and generated the impeccable sections on such short notice. With that said, I need to acknowledge our collaborators Dr. Harminder Singh, Dr. Robert Wightman, and Dr. Zoann Nugent for generating the cohort of colorectal cancer tumor samples.

This research was funded though operating grants from the Canadian Institutes of Health Research, CancerCare Manitoba Foundation, and Colon Cancer Canada, and utilized the Genomics Center for Cancer Research and Diagnosis. This study was supported by the Manitoba Tumor Bank, Winnipeg, Manitoba, funded in part by the CancerCare Manitoba Foundation and the Canadian Institutes of Health Research, and a member of the Canadian Tumour Repository. Additional funding for graduate studies was provided by the Graduate Enhancement of Tri-Council Stipends and the Flying Officer George Finkle Memorial Scholarship.

Last but not least, I must acknowledge the unwavering support from my family. Most days, my mom, dad, and brother had no idea what I was working on (being the only scientist in the family, I guess it comes with the territory), but they never stopped encouraging me to continue with my education. Their love and support has gotten me to where I am today. And I must thank my special Diego, who encouraged me to believe in myself, and to pursue my dreams.

Amy Cisyk

LIST OF TABLES

LIST OF FIGURES

LIST OF ABREVIATIONS

2D	2-dimensional
3D	3-dimensional
5C	5 carbon
µg	microgram (weight)
µL	microliter (volume)
°C	degrees Celsius
ACG	American College of Gastroenterology
aka.	also known as
APC	adenomatous polyposis coli
ASGE	American Society for Gastrointestinal Endoscopy
BRAF	v-raf murine sarcoma viral oncogene homolog B
CEF	complexation enhanced fluorescence
CFP	cyan fluorescent protein
Chr	chromosome
CIMP	CpG island methylator phenotype
CIN	chromosomal instability
CRC	colorectal cancer
Cy3	cyanine 3
DAPI	4',6-diamidino-2-phenylindole
ddH_2O	double distilled water
DMEM	Dulbecco's modified eagle medium
DNA	deoxyribonucleic acid
e.g.	*exempli gratia*, for example
et al.	*et alii*, and others
etc.	*et cetera*, and the rest
FACS	fluorescence activated cell sorting
FAP	familial adenomatous polyposis
FBS	fetal bovine serum
FFPE	formalin fixed paraffin embedded
Fig.	figure
FITC	fluorescein isothiocyanate
FISH	fluorescence *in situ* hybridization
xg	gravitational force
G_1	Gap 1
G_2	Gap 2
GCHP	goblet cell hyperplastic polyp
h	hour(s)
H_2O	water

H&E	hematoxylin and eosin
HCl	hydrochloric acid
HNPCC	hereditary non-polyposis colorectal cancer
hTERT	human telomerase reverse transcriptase
IBD	inflammatory bowel disease
i.e.	*id est*, that is
KCl	Potassium chloride
KRAS	kirsten rat sarcoma viral oncogene homolog
L	litre (volume)
M	molar (concentration)
M	mitosis
mL	millilitre (volume)
mM	millimolar (concentration)
MCR	Manitoba Cancer Registry
M-banding	multicolour banding
$MgCl_2$	Magnesium chloride
Min	minute(s)
MMR	mismatch repair
MPHP	mucin-poor hyperplastic polyp
MSI	microsatellite instability
Mth	month
MVHP	microvesicular hyperplastic polyp
N	normal (concentration)
NaOH	sodium hydroxide
ND	not determined
NI	not indicated
nm	nanometer (wavelength) (distance)
PBS	phosphate buffered saline
PCR	polymerase chain reaction
pg.	page
rpm	revolutions per minute
SSA/P	sessile serrated adenoma/polyp
SSC	saline-sodium citrate
T	time
Thy	thymidine
TMA	tumor micro array
TNM	tumor, node, and metastases
vs.	verses

CHAPTER 1: INTRODUCTION

1.1 INTRODUCTION

Colorectal Cancer (CRC) is the second leading cause of cancer related deaths, after lung cancer, among both men and women in North America[1]. According to the American Cancer Society, it is estimated that over 140,000 Americans will be diagnosed with CRC in 2013, and an additional 50,000 will succumb to the disease. The lifetime risk of developing CRC in the United States is just over 1 in 20[2]. The majority of CRC cases (up to 85%) are sporadic, or randomly arising[3], as defined as occurring within individuals without familial conditions or other high risk factors (such as Inflammatory Bowel Disease [IBD]).

Screening for CRC and its precursor lesions has been shown to reduce the incidence and mortality from CRC[4, 5]. Not only are CRCs that are detected at early stages more responsive to treatment, but screening also diminishes the incidence of CRC by detecting precursor lesions before they progress into cancerous neoplasms[2]. As further insight into the disease is gained, screening programs can be adapted and enhanced to increase the reduction in CRC incidence and mortality.

Colonoscopy has become the preferred initial test for CRC screening in the United States[6] and is the essential follow-up test after other initial CRC screening tests[7, 8]. Other commonly used screening strategies include the fecal occult blood test and sigmoidoscopy. Unfortunately, a large proportion of CRCs and CRC related deaths are not prevented by any of these techniques. Even with colonoscopies there remains a portion of CRCs, termed Interval CRCs, that are either not easily detected and/or develop rapidly. Interval CRCs are a subset of sporadic CRCs that are diagnosed within a

relatively short time period after a negative (i.e. complete clearing) colonoscopy. There is evidence to suggest a proportion of Interval CRCs may result from sporadic CRCs missed during the colonoscopy, often referred to as an incomplete or false-negative colonoscopy[9-14]. Recently however, studies have suggested that a large proportion of Interval CRCs may arise through rapid tumor development, as a result of a *de novo* mutation or altered tumor biology[9, 13, 15-17]. It is likely that Interval CRCs result from a combination of missed sporadic CRCs and rapid tumor progression.

To date, there is no universally accepted definition for Interval CRCs[18], and this subset of CRC has been reported to be diagnosed anywhere from 6 months to 10 years after a negative colonoscopy[19, 20] (Table 1). The inconsistency in the parameters surrounding the definition of Interval CRCs make it challenging to compare and contrast studies and build a fundamental understanding of the unique biology that defines and contributes to the development of Interval CRCs. There are an expanding number of studies on Interval CRCs, from which a description of the biological differences between sporadic and Interval CRC biology has begun to emerge. It remains necessary to further establish and expand our understanding of Interval CRC pathogenesis, in order to develop/modify appropriate screening programs and specific treatment options to combat this unique form of CRC. This chapter describes the tumor biology that has been associated with Interval CRCs and describes how both false negative colonoscopies and altered tumorigenic pathways likely contribute to this subset of sporadic CRC. Furthermore, the altered molecular biology of Interval CRCs relative to other sporadic CRCs as well as emerging evidence of precursor lesions will be discussed.

Table 1. Definitions and Rates of Interval Colorectal Cancer.

Location	Time Frame (years)	Exclusion (Inclusion) Criteria[A]	Rate of Interval CRC	N	Reference
Dundee, UK	2[B]	NI	31.2-58.9%	635	21
Christchurch, New Zealand	2.3	Patients with anastomotic recurrent carcinoma	5.9%	17	10
Toronto, Canada	0.5 – 3	IBD, patients <20 years	2.1-5.9%[D]	430	19
Winnipeg, Canada	0.5 – 3	Resective colorectal surgery, IBD (50-70 years)	7.9%	388	22
Indiana, USA	3	NI	5%	47	23, 24
Fukui, Japan	3	IBD, prior cancer or polyps	6.0%	15	12
Arizona, USA	0.5 – 4	Prior CRC (only patients under postpolypectomy surveillance)	11.8%	1082	25
Pennsylvania, USA	1, 4	FAP, IBD, prior CRC, surgical resection of adenoma, body weight >150% of recommended weight	ND	13	9
Vermont, USA	4	NI	0.9%	26	26
Vermont, USA	4[C]	NI	ND	53	13

Table 1 Continued. Definitions and Rates of Interval Colorectal Cancer.

Location	Time Frame (years)	Exclusion (Inclusion) Criteria[A]	Rate of Interval CRC	N	Reference
Warsaw, Poland	5	Inadequate bowel preparation, prior cancer (patients with/without family history of CRC)	ND	42	18
Minnesota, USA	5	FAP, IBD (patients with/without previous [completely removed] polyps)	ND	51	16
Minnesota, USA	5	FAP, IBD (patients with/without previous [completely removed] polyps)	ND	63	15, 17, 27
Pennsylvania, USA	5	IBD, prior/recurrent cancer, residual polyps	5.2%	29	11
Minnesota, USA	5	FAP, HNPCC, IBD	5.4%	45	28
Oregon, USA	5.5	Prior history of colon disease (50-75 years)	2.4%	7	29
Heidelberg, Germany	1 – 10	IBD	ND	78	20

N = Number of Interval CRCs analysed

[A]Abbreviations; NI (Not Identified), IBD (inflammatory bowel disease), ND (Not Determined), FAP (familial adenomatous polyposis), HNPCC (hereditary non-polyposis CRC)

[B]After a negative fecal occult blood test instead of after a negative colonoscopy

[C]4 year median

[D]Depending on colon region; 2.1% (distal), 2.3% (rectosigmoid), 5.5% (transverse), 5.9% (proximal)

1.2 BIOLOGICAL PROPERTIES OF INTERVAL COLORECTAL CANCER

Interval CRCs are an aggressive subset of sporadic CRCs that have only recently acquired due recognition for their unique biological attributes. While there is limited research on the pathogenesis of Interval CRCs, emerging evidence suggests that Interval CRCs display an altered biology relative to the non-interval sporadic CRCs. Due to the relatively lower incidence rates of Interval CRCs (0.9-11.8% of sporadic CRCs, see Table 1) the remaining non-interval sporadic CRCs are often collectively referred to as sporadic CRCs (Fig. 1). Many population-based studies from Canada[19, 22, 30], Germany[20], Japan[12], New Zealand[10], Poland[18] and the United Kingdom[21] have compared characteristics of Interval tumors to those of sporadic CRCs. Summarizing the physical characteristics of Interval tumors (e.g. gender, tumor location, tumor grade, *etc*.) from these studies begins to establish a unique profile for Interval CRCs.

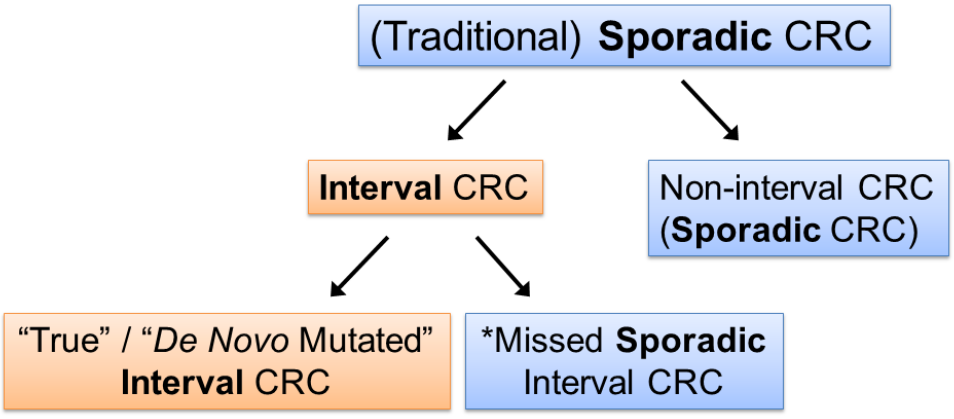

Figure 1. Outline of Sporadic Colorectal Cancer Sub-categories.
Recent evidence suggests that traditional sporadic CRCs can be further subcategorized into Interval CRCs and non-interval CRCs. Blue and red colored boxes identify categories of CRC that exhibit (*or are presumed to possess) two different biological and molecular profiles.

One of the most striking differences between Interval and sporadic CRC is their predominant location in the colon. To identify the location of tumor presentation in the colon, the large bowel is classically divided into four major sections; 1) proximal (cecum and ascending/right colon), 2) transverse, 3) distal (descending/left colon), and 4) rectosigmoid (sigmoid colon and rectum). According to the World Health Organization[31] the majority of sporadic CRCs occur in the rectosigmoid colon. Interval CRCs on the other hand, are unique in that they present most frequently within the proximal colon[10, 19, 20, 30]. For example, data from multiple Canadian colonoscopy cohorts indicate that more than 50% of Interval CRCs arise in the proximal colon, compared to less than 30% for sporadic CRCs[20, 30].

The biological environment varies throughout the length of the colon, which influences the pathways in which polyps and CRCs develop. During fetal development, the proximal colon originates from the embryonic midgut, whereas the distal colon is derived from the hindgut[32]. Blood supply[32], mucin pH[33], and average crypt length[34] are examples of biological features that differ along the colonic tract. Consequently the environmental and physical properties of the proximal colon may contribute to the etiology of Interval CRCs. Knowing these tumors occur disproportionally in the proximal colon supports the hypothesis that these tumors possess an altered biology, and this knowledge of presentation is beneficial in identifying the molecular origins of Interval CRCs.

There is also a gender difference in the occurrence of Interval CRCs. A population-based study from Germany identified a greater than 2-fold higher risk of Interval CRC in women compared to men[20]. This higher occurrence of Interval CRC in

women[18, 19, 22, 30], is opposite to what is observed in sporadic CRC, where there is a 20% increased incidence of colon cancer and 50% increased incidence of rectal cancer in men[31]. Analyzing gender and location of CRCs in the United Kingdom, Steele *et al*[21] identified the proportion of proximal tumors to be 10% higher in women compared to men for both the sporadic and Interval CRCs. This report also reconfirmed that Interval CRCs present most often in the proximal colon, and suggested that this may partially explain the increased occurrence of Interval CRCs in women. Indeed in another study from Manitoba, women were more likely to have Interval CRC when the analysis was not adjusted for the site of CRC, and had similar risk as men when the analysis was adjusted for the site of CRC, thus supporting that increased occurrence of CRC in the proximal colon among women is responsible for at least some of the increased rates of Interval CRC among women[22].

Patient age and tumor grade have also been associated with Interval CRCs, but with conflicting results. The proportion of sporadic CRCs in the proximal colon has been shown to increase with age[35]. Therefore it is not surprising that some studies have identified older age to be a risk factor for Interval CRCs[18, 19]. In Japan on the other hand, the risk of Interval CRC is highest for people between the ages of 60-69, compared to patients aged 59 and under, or 70 and older[12]. In either case, Interval CRCs, similar to the sporadic CRCs, generally appear to be a disease of the elderly. It has also been reported that Interval CRCs are often identified at a more advanced grade (e.g. grade 3 or 4) compared to the sporadic CRCs (e.g. most commonly grade 1), and this fraction of advanced tumors is even greater in the Interval CRCs presenting in the proximal colon[20, 21]. However, there still remains debate over the theory that sporadic proximal tumors

present at a more advanced stage[32]. Therefore, it is inconclusive, if poorly differentiated tumors (i.e. worse grade) result from a unique pathogenesis of Interval CRC, or are a result of tumor location.

Finally, additional risk factors have been identified in connection with particular disease states and/or prior surgeries. For example, it has been reported that patients who have a history of diverticular disease[19, 22] or pelvic/abdominal surgery[19] have a higher risk of being diagnosed with Interval CRC. These traits can hinder adequate examination of the entire colon on colonoscopy and lead to higher rates of incomplete colonoscopy, and potentially false negative colonoscopies. Thus these conditions contribute more significantly to the missed proportion of Interval CRCs than to the *de novo* mutated Interval CRCs. Patients with these or similar conditions have been excluded in some studies evaluating the biological properties of Interval CRCs. Other exclusion criteria often comprise patients with IBD, resective colorectal surgery (i.e. for previous polyps), and/or a family history of CRCs, including familial adenomatous polyposis (FAP) (Table 1).

In summary, there are several clinical and biological characteristics that differ between Interval and sporadic CRCs, which further support that Interval CRCs represent a unique subset of sporadic CRCs that may arise as a result of an aberrant biological pathway that drives their rapid development. False-negative colonoscopies also contribute to the occurrence of Interval CRCs (see below).

1.3 FALSE-NEGATIVE COLONOSCOPIES CONTRIBUTE TO THE DEVELOPMENT OF INTERVAL COLORECTAL CANCERS

Although the above section highlights the divergent biological properties associated with Interval CRCs, there is a proportion of Interval CRCs which could be attributed to false negative colonoscopies[9-14]. A sporadic CRC missed during an initial colonoscopy but subsequently detected within a relatively short time period (3-10 years, depending on the definition, see Table 1) has classically been categorized as an Interval CRC. Consequently, while some of Interval CRCs may possess a unique biology, a portion resulting merely from misdiagnosis would possess sporadic CRC biology, and this confounds efforts to identify the tumorigenic pathways and molecular biology of Interval CRCs (Fig. 1). An understanding of the many factors that contribute to false-negative colonoscopies is necessary to comprehend the effect of inaccurate or inadequate colonoscopy on the occurrence of Interval CRCs.

In general, five major factors have been identified to contribute to false-negative colonoscopies resulting in missed CRCs, and include; 1) incomplete colonoscopy, 2) poor bowel preparation, 3) insufficient withdrawal time, 4) colonoscopy examiner characteristics and procedure location of the endoscopy unit, and 5) incomplete resection of previous neoplasms. To prevent inadequate colonoscopies a number of key quality indicators have been proposed by many societies, most prominently by the American Society for Gastrointestinal Endoscopy/American College of Gastroenterology (ASGE/ACG). The quality indicators are categorised as pre-procedural (e.g. bowel preparation), intra-procedural (e.g. cecal intubation rates) and post-procedural (e.g.

recommendation for follow-up/surveillance colonoscopy)[36], and are designed to address the contributors to false-negative colonoscopies.

1.3.1 Incomplete Colonoscopy

A potential contributor to missed adenomas and thus Interval CRCs, is an incomplete colonoscopy[10, 12]. A complete colonoscopy is defined as one that reaches the cecal landmarks, including the ileocecal valve, appendiceal orifice, or terminal ileum[37, 38]. According to the ASGE/ACG, photo documentation of the landmarks of the endpoint of the colonoscopy examination should be recorded with every procedure[36]. The ileocecal valve is recognized as most dependable of the three cecal landmarks, as it can be identified in 98% of complete colonoscopies (i.e. colonoscopies that reached the cecum) confirmed by fluoroscopy[38]. If an endoscopist fails to intubate the cecum or to identify the cecal landmarks, surveillance of the most proximal regions of the colon may be compromised, and any lesion within unexamined regions would ultimately be missed. Many studies have shown that colonoscopies are less effective at detecting proximal verses distal CRCs[23, 39, 40]. Incomplete colonoscopy may partially explain the higher proportion of Interval CRCs occurring within the proximal colon. In one large widely quoted study from Poland, cecal intubation rates (i.e. rates of complete colonoscopy) were not associated with the risk of Interval CRCs[18], which suggested that incomplete colonoscopies may contribute less to the proportion of missed sporadic CRCs than the other risk factors. However, in another study from Ontario, Canada, colonoscopies performed by endoscopies with lower colonoscopy completion rates had a much higher rate of subsequent Interval CRC than those performed by endoscopists with higher colonoscopy completion rates[41]. Colonoscopies can be performed without sedation,

under conscious sedation, or under deep sedation, and incomplete colonoscopies due to patient intolerance are most common in colonoscopies performed without sedation. A large proportion (~40%) of incomplete colonoscopies has been attributed to patient intolerance[37]. Increasing the number of colonoscopies performed under sedation may enhance surveillance, raise the success rates, reduce the number of incomplete colonoscopies, and consequently potentially reduce the number of Interval CRCs attributed to an incomplete colonoscopy.

1.3.2 Inadequate Bowel Preparation

Inadequate bowel preparation (i.e. cleansing of the bowel prior to colonoscopy) not only impacts detection of both small[42] and large[42, 43] polyps, but is also an underlying factor of incomplete colonoscopies and therefore by both of these mechanisms potentially impacts the occurrence of Interval CRCs. It has been reported that 13% of incomplete colonoscopies can be accredited to poor bowel preparation, as the intraluminal contents hinder visibility and can mask the cecal landmarks[37]. Bowel cleansing is performed through the administration of a laxative regimen (e.g. aqueous sodium phosphate, polyethylene glycol electrolyte solution, *etc*.) between 12-24 hours prior to the colonoscopy. ASGE/ACG recommends that the quality of bowel preparation, whether "excellent", "good", "fair", or "poor" be documented in the procedural documentation notes of every colonoscopy. These terms refer to the intraluminal contents that remain after the laxative regimens[36]. A validated bowel preparation quality scale should be used and will allow standardization among the different providers. Even with excellent colonic cleansing there remain regions of the colon that are technically difficult to survey, including under folds, in the cecum, and at the hepatic and splenic flexures. Gorski *et*

al[11] determined that 14% of Interval CRCs presented in the hepatic and splenic flexures, and 31% of Interval CRCs occurred in the cecum, suggesting that hindered visibility in these regions may contribute to the missed sporadic proportion of Interval CRCs. Generally, suboptimal bowel cleansing (those categorized as "fair" or "poor") decreases the diagnostic accuracy of the colonoscopy and these patients often require repeat examinations at a short interval. Bowel preparation is an important factor that impacts CRC screening efficacy, and thus, poor bowel preparation may contribute to the missed sporadic portion of Interval CRCs.

1.3.3 Insufficient Withdrawal Time

The colonoscopy procedure can be divided into two segments, insertion and withdrawal, and the time spent for both portions of the procedure can be measured[36]. The colonoscopic insertion time is the time elapsed from the insertion of the colonoscope into the rectum to the identification of the cecum. The colonoscopic withdrawal time is the recommended regulated component of the procedure, and is the time elapsed from the identification of the cecum to when the colonoscope is withdrawn across the anus. It is during this stage of the procedure where the colon is critically assessed for neoplasms/polyps. Withdrawal time is a surrogate marker for adenomatous polyp detection rate, which in turn has been associated with risk of Interval CRC. Colonoscopy withdrawal time has been positively associated with polyp (including adenomatous polyp) detection rates[44, 45], and therefore insufficient colonoscopic intubation rates (i.e. rapid assessments) may contribute to missed sporadic CRCs, particularly in the areas of hindered/decreased visibility. The minimum recommended withdrawal time for a standard colonoscopy in which no polyps are removed is 6 minutes[36]. Barclay *et al*[46]

compared the adenomatous polyp detection rate of gastroenterologists with a mean withdrawal time of <6 minutes (11.8% of patients) to those with a mean withdrawal time of ≥6 minutes (28.3% of patients). They noted a positive association that was further compared in a subsequent study that evaluated withdrawal times <8 minutes to withdrawal time ≥8 minutes and found adenomatous polyp detection rates of 23.3% and 37.8%, respectively[47]. Again, the longer withdrawal times were associated with a higher adenomatous polyp detection rate[47], suggesting that insufficient intubation rates may contribute to missed CRCs. Therefore, insufficient withdrawal times, like that of incomplete colonoscopies and inadequate bowel preparation, may be a contributing factor to both false-negative colonoscopies as well as a portion of Interval CRCs.

1.3.4 Colonoscopy Examiner Characteristics and Procedure Location of the Endoscopy Unit

Many studies have suggested colonoscopy providers' medical specialty and their adenomatous polyp detection rate as well as the examination settings of the initial colonoscopy may influence the rates of missed neoplasms[19, 22-24, 48]. Bressler *et al*[19] from Toronto, Canada looked into the relationship between physician specialty (e.g. gastroenterology, general surgery, and internal medicine/family medicine) of the endoscopist and the rate of Interval CRCs. They identified performance of colonoscopy by internists and family physicians as being independently associated with an increased occurrence of Interval CRCs. Specifically, 5.5% of patients of internists and family physicians that were diagnosed with CRC had a negative colonoscopy performed by them 6-36 months prior to diagnosis, compared to 3.3% for general surgeons and 3.0% for gastroenterologists. Variation among endoscopists in the screening procedure and/or

techniques to identify lesions, likely contributes to an increased rate of Interval CRCs for the procedures performed by specific medical specialties. The study by Bressler *et al*[19] also looked into the colonoscopy setting (i.e. academic hospital, non-academic hospital, and office), and found that colonoscopies performed in an office were associated with an increased risk for subsequent identification of Interval CRCs. Specifically, 5.9% of patients whose procedures were performed in an office and were diagnosed with CRC, had a negative colonoscopy performed in an office 6-36 months prior to diagnosis, compared to 2.5% and 3.4% of procedures in academic and non-academic hospitals respectively. While they indicated that physician specialty and procedural location may be modifiable risks of Interval CRCs, these findings were not consistent with a comparable study from Warsaw, Poland[18]. This report found no significant association between endoscopist specialty and the risk of Interval CRC, whether it be gastroenterology, internal medicine or surgery. However, they did determine that an endoscopist's adenoma detection rate was an independent risk factor for Interval CRC. The adenoma detection rate is the percentage of screened patients that have at least one adenoma detected during the colonoscopy. The risk of Interval CRC was considerably higher for those examined by an endoscopist with an adenoma detection rate of <20%, compared to those with a detection rate of ≥20%[18]. It is possible that the increased Interval CRC risk factor for family physicians identified by Bressler *et al*[19] may be a consequence of a decreased adenoma detection rate in this particular subset of endoscopists.

1.3.5 Incomplete Resection of Previous Neoplasms

Finally, invasive adenomatous polyps that are incompletely excised have been reported to result in Interval CRCs[9, 28]. However, depending on the operational definition for Interval CRC, individuals with incomplete resections and subsequent diagnosis of CRC are sometimes excluded from Interval CRC studies. This is due to the fact that Interval CRCs may be defined as those detected after a negative colonoscopy, and thus the previous colonoscopy would not have identified a neoplasm. However in some studies, if a polyp is detected and successfully removed during the colonoscopy, the procedure is still categorised as a negative colonoscopy even though a polyp was detected, because no CRC was seen and the polyp was removed. Therefore, labelling a carcinoma that developed in a region where a previous polyp was removed as an Interval CRC is somewhat controversial, and this debate results from the numerous definitions for Interval CRCs.

While local recurrence of CRCs after endoscopic resection does occur, the probability of recurrence has been linked to the resection method. The colonic lesions can be removed by *en-bloc* resection (i.e. tumor/polyp removed all in one piece) or by piecemeal resection (i.e. tumor removed in pieces). The risk of recurrence of CRC tumors is approximately 4 times greater after piecemeal resection compared to en-bloc resection[49]. This suggests inadequate resection of the precursor lesions may be a contributing factor to the occurrence of Interval CRCs. One research team from Vermont reported 21% of Interval CRCs were due to an incomplete resection of a precursor lesion[13], and another study identified 5 Interval tumors that were diagnosed in a segment of the colon where a large adenoma was previously removed[26]. Interestingly, Lieberman

et al[29] compared the rate of Interval CRC after a previous negative colonoscopy to previous colonoscopies with polypectomy. They identified an ~10-fold increase in the diagnosis of Interval CRC when the previous colonoscopy was performed with polypectomy. The cancers occurring due to the transformation of incompletely resected precursor lesions are neither *de novo*, nor missed sporadic, Interval CRCs. Ideally, a more specific definition for Interval CRCs should be established, excluding patients with previous history of CRC and/or colonoscopies performed with polypectomy. Alternatively, these two subgroups (i.e. those with and those without prior polypectomy) should be evaluated separately, as the etiopathogenesis of Interval CRC is likely different in these two subgroups.

1.4 ALTERED TUMOR BIOLOGY AND ITS ASSOCIATION TO INTERVAL COLORECTAL CANCERS

While the above sections highlight the contributions that inaccurate or inadequate colonoscopy may have on the occurrence of Interval CRCs, there is a growing body of evidence to suggest that Interval CRCs possess an aberrant biology that could lead to their rapid growth and proliferation. While the number of studies on Interval CRCs is limited, the majority of the published research to date has focused on the physical characteristics of these tumors. As detailed above, it is globally recognised that Interval CRCs differ from sporadic CRC by presenting most frequently in the proximal colon and are more often among women. These characteristics among others, not only help identify variables in colonoscopies that may contribute to missed neoplasms, but also suggest that Interval CRCs may harbour an altered molecular biology. Brenner *et al*[50] analysed the effectiveness of colonoscopies and the risk of developing CRC after a negative

colonoscopy. They identified that exclusion of incomplete colonoscopies only lowered the risk of Interval CRC slightly, and suggested that additional factors, such as aberrant biology, may have a more substantial impact on the development of these tumors. Below is an introduction to the tumorigenic pathways observed in sporadic CRC, from which aberrations and variations in Interval CRCs have recently been identified.

1.5 MOLECULAR TUMORIGENIC PATHWAYS IN COLORECTAL CANCER

Genomic instability is a hallmark of virtually all tumor types and is arguably best characterized in sporadic CRC. In general, genomic instability is a phenotype that enables the acquisition of tumor-associated genetic alterations[3] and arises though one of three different tumorigenic pathways; chromosomal instability (CIN), microsatellite instability (MSI), or CpG island methylator phenotype (CIMP) (Fig. 2). While these pathways have distinct molecular properties and were classically believed to be mutually exclusive[51, 52], it has recently been determined that a small proportion of CRCs possess more than one genomic instability phenotype[53-55]. The relationships between the tumorigenic pathways in sporadic CRCs are discussed following introduction to the mechanisms and key players in each of the three pathways. An understanding of the tumorigenic pathways in sporadic CRCs, including the molecular players, frequencies and detection methods, is necessary to appreciate the molecular properties contributing to Interval CRCs.

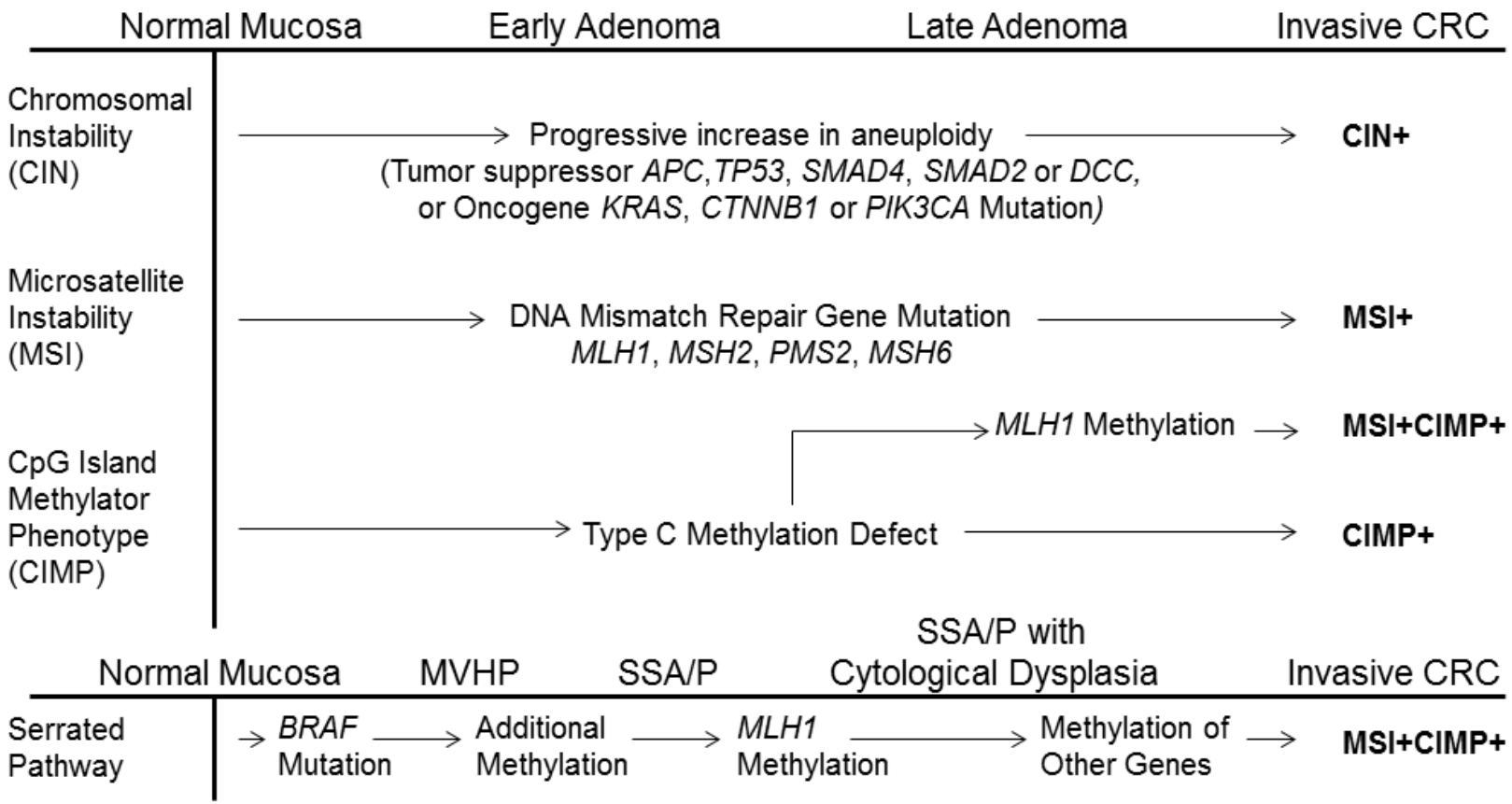

Figure 2. The Molecular Pathways to Colorectal Cancer Tumorigenesis.

A schematic depiction of the molecular alterations that characterize the transition of normal mucosa into CRC for each of the genomic instability pathways (CIN, MSI and CIMP), and the serrated pathway. Each pathway (or a combination of pathways) results in an invasive CRC positive (+) for a particular phenotype.

Abbreviations; MVHP (microvesicular hyperplastic polyp), SSA/P (sessile serrated adenoma/polyp)

1.5.1 Chromosomal Instability

CIN refers to an increase in the rate of which whole chromosomes or large chromosomal fragments are gained or lost. CIN results in a progressive increase in aneusomy (i.e. abnormal chromosome numbers) and can occur numerically as a result of defects in chromosomal segregation, or structurally as a result of chromosomal rearrangements including duplications, deletions and translocations[56]. The CIN pathway is the most prevalent form of genomic instability in CRC and is associated with up to 85% of sporadic CRCs and most heritable forms of the disease[51]. A typical sporadic CIN tumor may have anywhere from 60 to 90 chromosomes per cell. Genes that encode proteins responsible for the DNA (deoxyribonucleic acid) damage response, telomere regulation, cell cycle checkpoints or chromosomal segregation have been identified as CIN genes[57, 58]. It now appears that no single gene accounts for the CIN phenotype, but rather there is a spectrum of genes that can independently initiate this form of genomic instability. In fact, a few genes have been identified as chromosomal stability genes, which when mutated, result in CIN (e.g. *MRE11A*, *SMC1A*, *CDC4*, *NIPBL*)[59, 60]. The relatively low mutation rates of each of these genes in sporadic CRC (i.e. 11% for *CDC4* and 4% for *SMC1A* and *NIPBL*) suggest that there may be ~30-50 aberrantly expressed genes that drive the CIN phenotype.

A specific CIN pathway was first characterized by Volgelstein *et al*[61], and was thought to progress through a multi-step acquisition of specific molecular alterations. While the proposed pathway is still frequently referenced, it is certainly not the only mechanism through which CRCs acquire CIN. The earliest mutation in Volgelstein's proposed pathway was that of the *adenomatous polyposis coli* (*APC*) tumor suppressor

gene, followed by mutations acquired in the proto-oncogene *KRAS* and tumor suppressor gene *P53,* and these genes continue to be prevalently mutated in CIN positive CRCs[57]. However, some CIN positive tumors have been shown to possess wild-type and functional APC, while *APC* mutations have also been identified in CRCs that do not exhibit CIN[62]. Many genes have been identified as mutated in sporadic CRCs exhibiting CIN; the most prevalent mutations observed include the tumor-suppressor genes *APC, TP53*, *SMAD4*, *SMAD2*, and *DCC*, and the oncogenes *KRAS*, *CTNNB1* and *PIK3CA*[57]. It has been suggested that the biological properties of CRCs are determined by the total accumulation of genetic mutations, more so than the order in which the mutations accrue[63].

In addition to a critical role in the development of tumors, the presence of CIN can also be employed as a molecular prognostic biomarker[64-66], as high levels of CIN correlate with poor prognosis in sporadic CRC patients[66-68]. Additionally, CRC cell lines with CIN frequently harbour intrinsic multidrug resistance unlike comparable chromosomal stable lineages[69]. Loss of heterozygosity is a characteristic associated with high metastatic potential in sporadic CRC and along with loss of heterozygosity, aneuploid karyotype and gross chromosomal abnormalities (large deletions/duplications) are the most predominant phenotypic characteristics that define CIN positive CRCs[70]. Aneusomy progressively increases from proximal to distal tumors, and adenocarcinomas that develop CIN arise most predominately in the distal colon[71]. Approximately 20% of proximal tumors are aneuploid, compared to 60% of distal tumors[72]. While this trend is opposite to the predominant location of presentation in Interval CRCs, there is currently no research published on the prevalence of CIN in Interval CRCs.

In order to study the prevalence of CIN, several methods of detection exist. However, the technical definition of CIN as an increase in the *rate* of numerical and structural chromosomal changes makes identifying CIN tumors particularly challenging as most assays used to evaluate CIN are end point analyses. Nevertheless, aneusomy or abnormal chromosome numbers is often used as a metric for CIN, and therefore many of the CIN detection methods analyse chromosomal copy numbers[73, 74]. For instance, utilizing fluorescence *in situ* hybridization (FISH) and chromosome-specific probes, numerical aneusomy can be quantitatively analysed. Karyotyping, such as spectral karyotyping or M-banding, can identify structural aneusomy, however generating mitotic spreads from solid tumors is extremely challenging and often impossible[75]. Other detection methods assess total DNA content such as flow cytometry, and copy number variants such as comparative genome hybridization[73]. Polymerase chain reaction (PCR) based approaches can be used to analyse gains or losses in particular regions of DNA[75]. While each of the tests have individual strengths, collectively they share a specific weakness. To accurately identify the presence and extent of CIN, the detection method would ideally measure the rate in which chromosomal abnormalities are acquired through multiple cellular divisions.

1.5.2 Microsatellite Instability

Microsatellites are regions in the genome that contain repetitive sequences of DNA. The repeated sequences typically range from 1-6 base pairs, and are prone to errors in DNA replication such as single base-base mismatch, insertions and deletions. While the number of repeats is variable from person to person, individuals have microsatellites of a set length referred to as one's DNA fingerprint[76].

MSI results from dysfunctional DNA mismatch repair (MMR) proteins, causing the accumulation of genomic mutations in the highly repetitive microsatellite DNA[76]. Properly functioning DNA MMR proteins work to identify and correct microsatellite replication errors. In the presence of aberrantly expressed MMR proteins these errors go unresolved.

With the accumulation of expanded/contracted microsatellites, MSI elicits subtle genomic alterations throughout the genome, and hence this condition is also referred to as the mutator pathway[77]. MSI is associated with around 15% of sporadic CRCs as well as with hereditary non-polyposis CRC (HNPCC) also known as Lynch Syndrome[3, 78]. In CRC, the most common defects in DNA MMR result from epigenetic silencing or genetic mutation of the known MMR genes, which were first identified as germline mutations in HNPCC patients. The four most commonly mutated MMR genes in HNPCC include, *MLH1*, *MSH2*, *PMS2*, and *MSH6*[79]. With that said, this form of hereditary CRC constitutes only 1-2% of CRCs, and the majority of MSI tumors are sporadic[80, 81]. The predominant cause in sporadic, MSI positive, CRCs is epigenetic silencing of *MLH1*[82-85].

The MSI pathway results in subtle genomic alterations, and therefore these tumors typically remain diploid or near diploid, rarely showing gains or losses in large

chromosomal segments[86]. The MSI phenotype is a positive predictive marker for CRC patients, as high levels of MSI is associated with a better prognosis, than low MSI and/or high CIN levels[87-92]. MSI predominates in the proximal colon[87, 89, 93, 94], and the typical mutations observed with CIN, such as loss of heterozygosity, *APC* loss and *KRAS* mutations are less frequent in these tumors, but can still occur[90, 94, 95]. The MSI and CIN pathways are often thought to be mutually exclusive[52], but recent evidence suggest this may not be the case, as ~10-15% of CIN positive CRCs have been shown to possess high levels of MSI[53, 54]. MSI in sporadic CRC is associated with older age, female sex, and the acquisition of the oncogenic *BRAF* mutation[93]. Interval CRCs appear to have an altered prevalence of MSI compared to sporadic CRCs[16], and this relationship is discussed below (Section 1.7).

MSI can be used as a molecular prognostic marker[64, 65], and consequently two different MSI detection methods have been established: a PCR centered analysis and an immunohistochemical based approach. The PCR based analysis of microsatellite markers compares the extent of microsatellite amplification in a tumor to that of normal tissue. It should be noted that this technique specifically detects mutations at the nucleotide level that result from MSI (i.e. expansion/contraction in microsatellite DNA as a result of mutations in DNA MMR proteins), and is not designed to detect epigenetic silencing of the MMR genes. The National Cancer Institute[96] recommends a panel of 5 microsatellite markers to be analysed, and for further clarification they have suggested an additional 19 supplemental markers. Upon analysis, MSI status can be categorized as either, MSI high (≥2/5 or ≥30% of the 24 markers show instability); MSI low (1/5 or 10-30% of markers show instability); or microsatellite stable (0/5 or <10% of markers show

instability). The alternative MSI detection method utilizes immunohistochemistry techniques against the four most commonly mutated or epigenetically silenced MMR proteins, namely MLH1, MSH2, PMS2, and MSH6[97]. This is a rapid, cost-effective system that analyses MMR protein expression and localization, and is extremely specific and sensitive towards DNA MMR mutations that confer decreased protein expression. With that said, immunohistochemistry is not able to detect rare missense mutations in MMR genes that affect protein function, unless it also impacts protein localization. Often, PCR-based MSI detection is a useful supplemental test to ensure positive immunohistochemistry results are accurately identifying functional MMR proteins.

1.5.3 CpG Island Methylator Phenotype

The CIMP pathway is an epigenetic phenomenon characterized by hypermethylation of CpG islands within the promoter regions of tumor suppressor genes, resulting in their transcriptional silencing. DNA sequences (~500-2000 base pairs) rich in CpG nucleotide repeats (i.e. CpG islands) have been found in the 5' regions of a significant proportion of human genes[98]. CpG methylation is the product of DNA methyltransferases, which work by adding a methyl group to the 5C position of cytosine nucleotides. CpG hypermethylation is thought to recruit histone deacetylases and chromatin remodelling proteins that result in DNA compaction and the generation of heterochromatin, and thus gene silencing[99]. Methylation of cytosine in CpG islands is a mechanism of transcription regulation, and is observed in normal physiological processes such as genetic imprinting and X chromosome inactivation[100, 101]. CpG island methylation is also a natural progression of colonic epithelial cell aging, and this global methylation increases with age and is referred to as type-A methylation[102].

More extensive CpG island methylation, denoted hypermethylation, has been attributed to gene inactivation in numerous tumor types including CRC. Type-C (i.e. cancer-related) hypermethylation, is frequently observed silencing hundreds of genes per cancer cell, including many tumor-suppressor genes[102]. While the patterns of hypermethylation are tumor specific, it remains unclear why certain regions have a higher tendency to be hypermethylated[103]. A modest increase in DNA methyltransferase activity has been observed with CRC progression, however the mechanisms underlying CIMP and aberrant hypermethylation remain poorly understood[104]. For instance, silencing of cancer related genes via hypermethylation has been shown to occur independently of heterochromatic DNA compaction[105]. It should be noted that DNA methylation is often permanent and this transcriptional inactivation is maintained through cellular divisions[99]. It is probable that certain patterns of hypermethylation are selected due to a growth advantage resulting in clonal expansion[106].

Approximately 1/3rd of sporadic CRCs harbour the CIMP phenotype, and similar to MSI these tumors occur disproportionally in the proximal colon[94]. Curtin *et al*[103] pooled data from numerous independent CRC studies and found 30-40% of proximal and 3-12% of distal sporadic CRCs possess the CIMP phenotype. Unfortunately, determining whether or not a CRC has developed along the CIMP pathway is not as straightforward as detecting MSI. There are a variety of different detection methods, including quantitative real-time PCR evaluating small sets of CIMP markers, quantitative real-time PCR followed by unsupervised hierarchical clustering analysis, sensitive non-quantitative methylation specific PCR, and next generation sequencing to name a few. In addition, the minimum number and choice of genes to be analysed differs between panels[103]. The

traditional panel of CIMP markers assesses methylation at five CpG sites, including *MLH1*, *CDKN2A*, *MINTS1*, *MINTS2*, and *MINTS31*, and classifies the tumors as either CIMP positive or negative[107]. Many other panels of CIMP markers do exist, some of which further categorise CIMP sporadic tumors as either CIMP high, relating to *BRAF* and *MLH1* methylation; or CIMP low, relating to *KRAS* methylation[103]. Hypermethylation leading to silencing of the MLH1 gene has revealed a link between CIMP and MSI status, and *MLH1* is one of the most common CIMP markers analysed. Almost 50% of CIMP positive CRCs acquire *MLH1* hypermethylation and consequently develop MSI[102].

1.6 GENOMIC INSTABILITY PATHWAYS IN COLORECTAL CANCER ARE NOT MUTUALLY EXCLUSIVE

The three pathways of tumorigenesis have overlapping contributions to the progression of sporadic CRCs. By prevalence alone (CIN 85%[108], MSI 15%[77] and CIMP ~30%[109]) it is evident that these pathways are not mutually exclusive. Most extensively studied in sporadic CRC is the relationship between MSI and CIMP. As reviewed by Issa[99], 70-80% of MSI positive sporadic CRCs can be attributed to CIMP and its effects of hypermethylation and consequent silencing of *MLH1* (Fig. 2). Interestingly, CIMP-positive MSI-positive sporadic CRCs have a relatively good outcome when compared to CIMP cancers without MSI, which tend to have a much poorer prognosis[110]. It has been proposed that MSI and CIN are mutually exclusive pathways[51, 52] and that CIMP may contribute to the development of either MSI or CIN[111].

A progressive increase in colonic epithelial dysplasia is responsible for the progression of an adenoma or adenomatous polyp to an adenocarcinoma, and up to 90%

of CRCs are adenocarcinomas[112]. CIMP is thought to occur early on in this tumorigenic process, as abnormal methylation has been noted in colorectal adenomas[113, 114]. It has been shown that approximately 1/3rd of CIMP positive tumors exhibit a high degree of chromosomal abnormities[55]. In CIMP positive CRCs, CpG island methylation has been observed in the promoter region of genes involved in processes necessary for chromosome stability, including DNA repair and cell cycle control[111].

On the other hand, the conventional view on the discordance between CIN and MSI has recently been disputed. While MSI tumors are typically diploid, one study aimed to identify the extent of CIN in MSI positive and microsatellite stable CRCs, and found chromosomal abnormalities in both[54]. Additionally, in 2007, Shen *et al*[53] identified as many as 12% of CIN-positive tumors to also possess high levels of MSI. As further insight into the individual pathways is gained, it becomes clear that these pathways to genomic instability are intertwined in the development of this disease. Furthermore, an alteration in any of these pathways in Interval CRC would contribute to aberrant tumor biology and possibly rapid tumorigenesis.

1.7 IDENTIFYING THE MOLECULAR TUMORIGENIC PATHWAYS IN INTERVAL COLORECTAL CANCER

There is clearly a wide-range of research on the molecular pathways through which sporadic CRCs develop, however very little is known about the progression of Interval CRCs. Nevertheless, researchers have recently begun to evaluate the molecular pathways that contribute to the development of Interval CRCs, and results indicating altered tumor biology relative to sporadic CRC have emerged. The prevalence of genomic instability in Interval CRCs, and how the contributions of the different pathways differ to those of sporadic CRCs, is presented in Table 2. Below is a detailed review of the advances made in the understanding of Interval CRC molecular genetics.

Table 2. Prevalence of Tumorigenic Pathways and Gene Mutations in Colorectal Cancer.

	CRC Tumor Types (%)			
Pathway/*Gene*	**Sporadic**	**Interval**	**SSA/P[A]**	**Reference**
CIN	85	ND[B]	37.5(APC loss)	108, 115
MSI	10-15	29-30	0	15, 16, 77, 116-118
CIMP	30-33	57	69-76.8	15, 109, 115, 116, 119
BRAF	5-22	28	60.7-82.8	27, 109, 115-117, 120-130
KRAS	32.8-51	12.9	0-8	17, 95, 116, 117, 119, 120, 123-126, 131, 132

[A]SSA/P = Sessile Serrated Adenoma/Polyp
[B]ND = Not determined

The first indication that Interval CRCs harboured an altered molecular biology came from Sawhney *et al*[16] in 2006, when they identified the prevalence of MSI in Interval CRCs. Tumors diagnosed within 5 years after a negative colonoscopy were considered Interval CRCs, and their study cohort consisted of 51 Interval CRCs, and 112 non-interval CRCs. Interval to non-interval cases were both age-matched (i.e. a mean age of 73.6 and 72.5 years respectively), and sex-matched. However, MSI analysis was not available for 5 Interval and 15 non-interval subjects. Patients with FAP and/or IBD were excluded from the selection process, and this is important to note as these traits can significantly increase the risk of CRCs, thus increasing the risk of false negative colonoscopies. As expected, the Interval tumors presented more often in the proximal colon ~61%, compared to the non-interval tumors ~40%. MSI was identified in 30.4% of the Interval CRCs compared to 10.3% of the non-interval CRCs. Out of the total 46 Interval samples, 14 were MSI-high and an additional 12 were MSI-low. For the 97 non-interval patients, 10 were MSI-high and 17 were MSI-low. After adjusting for age the Interval CRCs were found to be 3.7 times more likely to harbour a MSI phenotype. Interestingly, this relation was further amplified when they divided the cohort into proximal verses distal tumors. In the proximal samples, the Interval CRCs were 1.75 times more likely to be MSI positive compared to the non-interval CRC controls, but this was not statistically significant. In the distal colon however, they found a 17.5-fold increase in the prevalence of MSI in the Interval CRCs. Sawhney *et al*[16] suggested that false-negative colonoscopies are not the only contributing factor to the development of Interval CRCs, and that the atypical prevalence of MSI is evidence of an altered tumor biology (particularly in the distal colon) which may contribute to a rapid tumor

development. It is important to note that 98% of the total cohort was male and that may impact the significance of the results, as Interval CRCs are more common among females. Therefore, additional MSI studies aimed at broader Interval CRC cohorts, with respect to both gender and ethnicity, are highly warranted.

In 2009, Arain *et al*[15] identified an increase in the prevalence of CIMP in Interval CRCs. They tested the presence of methylation in the promoter regions of *MINT1*, *MINT2*, *MINT31*, *p16INK4*, *MGMT*, and *MLH1*, and classified samples as CIMP positive when methylation was observed in ≥3 genes. CIMP was identified in 57% of Interval CRCs compared to 33% of the non-interval CRCs, and this represented a significant 2.5-fold increase. Both Interval as well as non-interval CIMP-positive CRCs were more likely to occur in the proximal colon; however this is not surprising as CIMP-positive CRCs occur more frequently in the proximal colon[94, 103]. Interval CRCs were more likely to present in the proximal colon, and therefore the increase that was found in the prevalence of CIMP in Interval CRCs may be a result of missed sporadic CRCs, and not a true phenotype of newly developed Interval CRCs. Unlike the Interval CRC and MSI analysis published by Sawhney *et al*[16], this study did not divide the results of the prevalence of CIMP in Interval CRCs by location of presentation. Had this been done, it would have decreased bias of missed sporadic CRCs in the identification of the prevalence of CIMP in Interval CRCs. If the increase in CIMP was observed in distal Interval CRCs, one could conclude Interval CRCs have altered tumor biology with respect to CIMP.

1.8 ADDITIONAL GENETIC MUTATIONS IDENTIFIED IN INTERVAL COLORECTAL CANCER

Proto-oncogenes, *BRAF* and *KRAS*, encode important components of the Ras/Raf/MEK/ERK signalling pathway which regulates cellular proliferation, differentiation and survival[133], and the mutation rates of these two genes have been identified in both sporadic and Interval CRCs. *KRAS* encodes a signalling kinase that can be transiently activated in response to external growth factors, and is responsible for signal transduction via the RAF family of serine/threonine kinases. *BRAF* encodes one of three RAF kinases, and BRAF activation promotes both cellular proliferation and antiapoptosis. Mutations in either of these two genes have been attributed to unregulated cell growth and proliferation and are frequently observed in cancer, including CRCs[120]. *BRAF* and *KRAS* oncogenic mutations have been characterised in Interval CRCs, adding to the molecular profile of these tumors and providing further evidence that Interval CRCs are a unique subset of sporadic CRC.

An oncogenic mutation in *BRAF* encoding the most predominant V600E mutation leads to elevated kinase activity and has been observed in approximately 5-22% of sporadic CRCs[120-122, 127-130]. Mutation of this gene is strongly associated with MSI positive sporadic CRCs[120, 121, 128], particularly those also positive for CIMP[123, 127, 130]. Microsatellite stable CRCs which harbour the *BRAF* mutation are associated with a much poorer prognosis[130]. A *BRAF* mutation is more predominant in women[123, 129, 130], tumors of the proximal colon[123, 129, 130], and for the most part, is mutually exclusive with the *KRAS* mutations[120-123, 129, 130]. The increase in MSI phenotype observed in Interval CRC is suggestive that these tumors would also harbour an increase in the *BRAF* mutation.

Shaukat *et al*[27] aimed to detect the percentage of *BRAF* mutations in Interval CRCs via PCR DNA amplification followed by a shifted termination assay, to enrich PCR amplification of the mutant DNA, and subsequent capillary electrophoresis detection. They expanded the cohort of tumors used by Sawhney *et al*[16] to identify the prevalence of MSI with the addition of 12 Interval and 19 non-interval tumors, to total 63 Interval and 131 non-interval samples. They reconfirmed the increase of MSI in Interval (29%) compared to non-interval CRCs (11%) and surprisingly, while they observed an increase in the *BRAF* mutation in Interval CRCs, this increase was considerably less significant than that of MSI. The *BRAF* mutation was identified in 28% of Interval CRC samples, and 19% of the non-interval controls, representing less than a 2-fold increase. Although the *BRAF* mutations were more likely to occur in the proximal colon, this study concluded that their results did not support a direct role of *BRAF* in Interval CRCs.

The *KRAS* oncogenic mutation rate, while significant in sporadic CRC at 30-50%[95, 120, 131, 132, 134], is observed less frequently in Interval CRCs and is thought to be inversely associated with the development of this subset of sporadic tumors. Still, there are mixed views on the association of *KRAS* mutation and MSI phenotype. Some studies show that they are negatively correlated[78, 95, 131, 132, 134-136], while others show *KRAS* mutation as independent of MSI[137, 138]. Shaukat *et al*[17] measured the percentage of Interval verses non-interval CRCs that possessed the *KRAS* mutation. They identified the *KRAS* mutation to be inversely associated with Interval CRCs, as only 13% of Interval CRCs possessed the *KRAS* mutation compared to 29% of the non-interval CRCs. None of the *KRAS* mutated cancers exhibited MSI, and only 5% contained the *BRAF* V600E substitution. The CRC cohort used in this study was previously used to identify the

CIMP status[15] and percentage of *BRAF* mutations[27] in Interval CRCs. The Interval CRCs in this cohort exhibited a 3-fold increase in the prevalence of MSI compared to the non-interval CRCs. The decrease in *KRAS* mutations in the Interval CRCs may partially be explained by the increase in MSI tumors. Shaukat *et al*[17] suggested that the *KRAS* mutation is a marker of the CIN pathway due to the mutual exclusion of the *KRAS* mutation and MSI phenotype in their cohort, and implied that Interval CRCs have a decreased prevalence of CIN. The usefulness of *KRAS* as a molecular signature of CIN is unclear, as the *KRAS* mutated cancers were almost equally distributed throughout the colonic segments. This lack of regional differences in the *KRAS* mutation has been shown in other studies as well[72, 95] and is suggestive that this mutation is relevant to the progression of both proximal and distal CRCs. Additionally, *KRAS* is only mutated in 30-50% of sporadic CRCs[95, 120, 131, 132, 134], and thus is not always mutated in CRCs exhibiting CIN. As previously mentioned, the CIN phenotype is predominately observed in distal sporadic CRCs[71, 72], and therefore, the evidence of decreased *KRAS* mutation rate in Interval CRCs is inconclusive in regards to the prevalence of the CIN phenotype in these tumors. Accordingly, studies designed to assess CIN in Interval CRCs are highly warranted.

1.9 SESSILE SERRATED ADENOMAS A POSSIBLE PRECURSOR TO INTERVAL COLORECTAL CANCER

The emerging aspects of Interval CRC molecular biology, combined with the unique physical properties, have transpired the goal of identifying a precursor lesion. The study of colonic polyps as precursors to CRC has evolved greatly since the classification of hyperplastic polyps by Williams *et al*[139] in 1980. At that time, colonic polyps were categorized as either hyperplastic, considered innocuous and lacking neoplastic potential, or adenomatous, recognised as neoplastic precursor lesions for CRC. The hyperplastic polyps were identified by their histology, most notably by their sawtooth or "serrated" margins. In a subsequent study by Urbanski *et al*[140], the term hyperplastic-adenomatous polyp was proposed to classify lesions with characteristics of both histological types. In 1990, accompanied by extensive research, these lesions were renamed serrated adenomas, stressing their neoplastic capabilities. It was recognised that serrated adenomas encompassed a broad morphological range between hyperplastic polyps and adenomas[141]. To date, serrated adenomas encompass four broad categories of colonic polyps[31], much like the histological classification system described by Torlakovic *et al*[142] in 2003. According to the 4th edition of the World Health Organization Classification of Tumours of the Digestive System published in 2010, the classification scheme of serrated polyps includes; 1) hyperplastic polyp, 2) sessile serrated adenoma/polyp (SSA/P), 3) SSA/P with cytological dysplasia, and 4) traditional serrated adenoma (Fig. 3)[31].

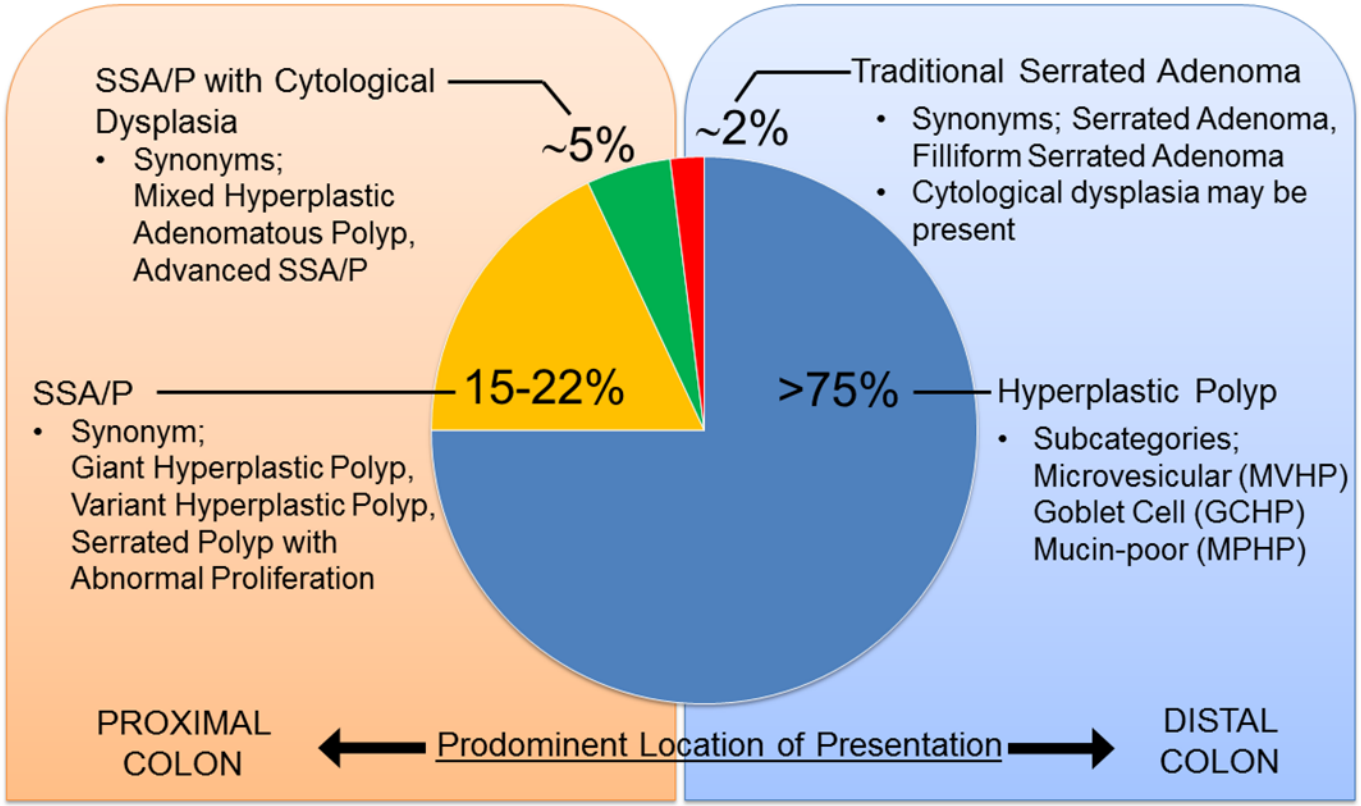

Figure 3. Classification of Serrated Polyps.

The World Health Organization classifies serrated polyps of the colorectum into four morphological categories; SSA/P, SSA/P with Cytological Dysplasia, Traditional Serrated Adenomas, and Hyperplastic Polyps. This figure details the relative frequencies, synonyms, subcategories and predominant location of presentation for each category of serrated polyp.

The hyperplastic polyp comprises >75% of all serrated lesions[31] and with relatively normal proliferation these polyps pose little-to-no risk of neoplasia (Fig. 3). Based on the mucin content of the epithelial cells, hyperplastic polyps can be further divided into three categories, the most common of which is the microvesicular hyperplastic polyp. Traditional serrated adenomas on the other hand, exhibit proliferation that often results in dysplasia. While the traditional serrated adenomas encompass several disparate lesions (e.g. filiform serrated adenomas, conventional adenomas with serration) they are uncommon making up <1% of all colorectal lesions[31]. Both hyperplastic polyps and traditional serrated adenomas are most predominately distributed in the distal colon and rectum (Fig. 3). Compared to the other two subcategories of serrated lesions, hyperplastic polyps and traditional serrated adenomas have a less significant impact on the development and therefore incidence of CRC. This reduced neoplastic potential makes them unlikely candidates of Interval CRC precursor lesions.

SSA/P were first reported by Torlakovic *et al*[143] in 1996 as a serrated lesion which differed morphologically from hyperplastic polyps. In a subsequent study, a larger cohort of serrated polyps were analysed and Torlakovic *et al*[142] established a more complete profile for SSA/Ps. The abnormal morphology of SSA/Ps included, abnormal proliferation, crypt distortion and dilation, and decreased number of endocrine cells. In 2006, it was reported that SSA/Ps accounted for 22% of serrated adenomas, and 9% of all polyps[124]. It is recognised that the CRCs that progress from the SSA/P, develop along a specific tumorigenic course designated the Serrated Pathway[144] (Fig. 2). It has been proposed that the serrated pathway is responsible for the proportion of CRCs that harbour

the CIMP phenotype (approximately 35% of all CRCs). In this pathway, which displays similarities to the more traditional molecular pathways of tumorigenesis discussed above, activation mutations in specific oncogenes have been identified. For example, initiation of the pathway is often random or slowly progressing and requires an activating *BRAF* mutation. This results in normal mucosa developing into either a microvesicular hyperplastic polyp or a SSA/P. At this stage, these lesions are prone to CpG island methylation, and if the *MLH1* gene is methylated, the further development of MSI results in SSA/P with cytological dysplasia. Mutation and methylation of other genes becomes accelerated, and the SSA/P with cytological dysplasia can quickly develop into a CRC possessing both the CIMP and MSI phenotypes.

SSA/Ps exhibit similar genomic and colonic site location characteristics to that of Interval CRCs. The possibility that SSA/Ps are precursor lesions to Interval CRCs has been suggested by studies of serrated adenomas[116, 145-147] as well as those particularly focused on Interval CRCs[15]. Firstly, SSA/Ps are reported to present most often in the proximal colon[115, 116, 124]. They tend to be pale, flat polyps, with minimal changes to the vascular network, and are often covered with a yellow colored mucus[145, 146]. Due to these characteristics they can be difficult to detect by colonoscopy[147], similar to Interval CRCs, and therefore proper bowel preparation is critical (see above). However, unlike Interval CRCs, the risk of developing SSA/Ps does not appear to be influenced by gender[125]. In fact in Korea, the risk of developing a SSA/Ps is ~2-fold higher in men compared to women[115]. Yet, when the same group looked at SSA/P patients from the United States, it was the opposite, as 63% of the SSA/P patients were female. This discordance in gender with SSA/Ps may be a result of ethnic or environmental factors including diet.

The difficulties that SSA/Ps pose during colonoscopy and their potential relationship with Interval CRCs has peeked general interest in these lesions, and as a result, there has been an increase in the number of studies aimed at identifying the aberrant molecular biology associated with SSA/Ps. The serrated pathway of tumorigenesis proposes that an activating *BRAF* mutation is responsible for the development of an SSA/P, either from normal mucosa or a preceding hyperplastic polyp (likely a microvesicular hyperplastic polyp). The *BRAF* mutation is associated with 60% to 83% of all SSA/Ps[115-117, 123-126], a much higher frequency than the 28% of Interval CRCs reported to harbour the *BRAF* mutation[27]. On the other hand, similar to the Interval CRCs, the frequency of the *KRAS* mutation is much lower in SSA/Ps than sporadic CRCs (reported to be anywhere between 0% and 12.5%)[115-117, 119, 123-126]. Similar to sporadic and Interval CRCs, the *KRAS* mutation in SSA/Ps appears to be mutually exclusive to that of the oncogenic *BRAF* mutation[123, 124]. Due to the fact that the *BRAF* mutation is strongly associated with MSI positive CRCs, particularly those also harbouring the CIMP phenotype, SSA/Ps are suggested to be the precursor lesion to CIMP+MSI+ CRCs. MSI is thought to be a major event in the malignant transformation of SSA/P and occurs later on in the serrated pathway to CRC[144]. In fact, one study identified SSA/P to be the precursor polyp to 21% of MSI+ CRCs compared to only 3% of microsatellite stable CRCs[148]. The proportion of SSA/P with CIMP appears to be universally recognised at around 70-75%[115, 116, 119], and it is not surprising that MSI does not appear to be associated with SSA/P, as it occurs on malignant transformation of SSA/P[116-118]. Therefore, while SSA/Ps show elevated levels of CIMP similar to Interval CRCs, where the prevalence of CIMP was identified around 60%[15], both the *BRAF* and

MSI profiles of SSA/Ps compared to Interval CRCs are significantly different. Ultimately, a greater understanding of Interval CRCs, especially at the molecular level, is necessary to identify precursor lesions and/or possible subcategories within Interval CRCs, one of which may be the SSA/Ps.

CHAPTER 2: HYPOTHESIS AND RESEARCH AIMS

2.1 HYPOTHESIS

We hypothesize that Interval CRCs may progress through altered tumorigenic pathways contributing to a rapid tumor development. We further hypothesize that CIN contributes to the etiology of Interval CRCs and that Interval CRCs likely either possess an aggressive CIN phenotype or a higher prevalence of CIN compared to sporadic CRCs.

2.2 RESEARCH AIMS

The overall objective of this study is to determine the prevalence of CIN in Interval CRCs, however it was first necessary to develop and optimize a multiplexed, fluorescence-based approach that could be employed to evaluate CIN in formalin-fixed and paraffin embedded (FFPE) tumor samples.

2.2.1 To Establish and Optimize a Protocol that can Detect Changes in Chromosome Numbers

Aneusomy, or abnormal chromosome numbers, is often used as a metric for CIN. A FISH-based approach to enumerate multiple chromosomes can be used to identify cells exhibiting aneusomy and thus CIN. A DNA FISH probe set that multiplexes three specific chromosomes, namely 8, 11 and 17 has been employed. However, before CRC tissues can be evaluated, it was first essential to develop and optimize a FISH protocol, and confirm the specificity of the FISH probes using a tissue culture model system.

2.2.2 Analyze Chromosomal Instability in Colorectal Cancer Samples through Fluorescence *in situ* Hybridization (FISH) Probe Enumeration

Currently, very little is known about the aberrant biological pathways that contribute to the development of Interval CRCs. In fact, Interval CRCs have never been evaluated for the presence of CIN. By assembling a cohort of Interval and sporadic CRCs to which we apply the protocol to enumerate chromosomes designed in Aim 2.2.1, we can determine the prevalence of CIN in Interval CRCs.

2.2.3 Significance

Our current understanding about the etiological origins of Interval CRC is extremely limited, and the mechanism(s) that drives the genomic instability within these tumors needs to be identified. This study provides novel information about the tumorigenic pathways that are associated with Interval CRC, and begins to characterize the potential role CIN may have in that process. Only once we understand how these tumors develop, can we begin to modify screening approaches or develop novel treatments designed to combat this form of CRC.

2.3 ETHICS APPROVAL STATEMENT

This study, including the collection and use of archived clinical CRC tissue samples, was approved by the University of Manitoba Research Ethics Board. Prior to starting this study, I completed the Tri-Council Policy Statement (TCPS 2) Course on Research Ethics.

CHAPTER 3: MATERIALS AND METHODS

3.1 REAGENTS

Appendix A contains a list of the solutions and the reagents used throughout this study (pg. 119). In general, all reagents were purchased through Corning, Fisher Scientific, Fisherbrand, Gibco, Invitrogen, Sarstedt, Sigma-Aldrich and VWR.

3.2 CELL CULTURE

Adherent human cell lines of both immortalized and cancer tissue were utilized in the validation of the DNA FISH probe set. The three cell lines employed for probe evaluation included hTERT, HCT116, and HeLa and Table 3 presents general properties of each line. Both hTERT and HCT116 cells were cultured in McCoy's 5A (Corning) medium supplemented with 10% FBS (fetal bovine serum) (Sigma-Aldrich), while HeLa cells were cultured in DMEM (Dulbecco's modified Eagle's medium) (Thermo Scientific) supplemented with 10% FBS. All cells were grown on 10 cm tissue culture dishes (Sarstedt) and initially seeded at 20-40% confluence. Cells were maintained in a humidified Sanyo CO_2 incubator, at 37°C and 5.0% CO_2, until confluency reached ~80%. Generally, cells were detached (see below) and reseeded every 2-3 days.

All cell culture passagings and manipulations were performed in a biosafety cabinet. For passaging, cells were washed with 1xPBS (phosphate buffered saline), followed by the addition of 1.6 mL of 0.05% trypsin (Fisher Scientific) containing EDTA for 1-5 minutes (min). Cells were monitored using an inverted ID03 microscope (Zeiss) and 10x objective (Zeiss) to ensure cells had rounded up and detached from the culture dish prior to transferring to a 15 mL conical tube (Sarstedt). After the trypsin treatment, 5 mL of media supplemented with FBS was used to inactivate the trypsin. Cells were

transferred into a 15 mL conical, and washed with 5 mL of 1xPBS. A total suspended volume of ~12 mL was centrifuged at 140xg for 5 min in the Thermo Scientific Legend XFR centrifuge. Supernatant was aspirated, and the cell pellet was resuspended by titration in 3 mL of 1xPBS, and subsequently dispensed into a 10 cm culture dish containing 10 mL of fresh medium.

Table 3. Common Properties of the Three Cell Lines Employed in this Study.

Properties	**hTERT**	**HCT116**	**HeLa**
Organism	Human	Human	Human
Tissue	Foreskin	Colon	Cervix
Cell Type/disease	Fibroblast, immortalized with hTERT	Epithelial, CRC	Epithelial, Adenocarcinoma
Culture Properties	Adherent	Adherent	Adherent
Gender	Male	Male	Female
Culture Medium	McCoy's 5A + 10% FBS	McCoy's 5A + 10% FBS	DMEM + 10% FBS
Approximate Doubling Time	36 h	22 h	24 h
Karyotype	Diploid 46, XY	Near Diploid 45, XY	Hypotetraploid 82, XX

3.3 ASSESSING CHROMOSOMAL INSTABILITY IN CELLS

3.3.1 Fluorescence *in situ* Hybridization (FISH) Probes

DNA FISH probes that specifically recognize the α satellite DNA within the pericentromeric regions of chromosomes 8, 11 and 17 (Fig. 4), were purchased from Vysis (Abbott Molecular Inc.). Each FISH probe harbours unique fluorescence properties which allows them to be combined in a multiplexed manner and thus they are collectively referred to as the "DNA probe set". The common properties of the DNA probe set are indicated in Table 4. Chromosomes 8, 11 and 17 were specifically selected, as genes encoded within these chromosomes have been found altered in CRCs. Examples of such genes include the oncogenes *MYC* and *FGFR1* (chromosome 8), and tumor suppressor genes *MRE11A*, *ATM* (chromosome 11) and *p53* (chromosome 17)[57, 59, 60, 149, 150].

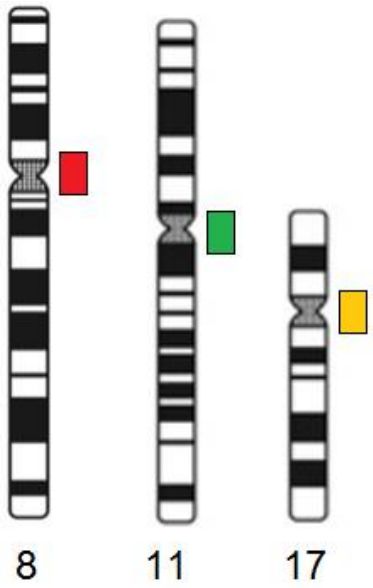

Figure 4. Ideogram Depicting the Location and Pseudocolors of the DNA Probe Set. FISH probes that hybridize to the pericentromeric α satellite DNA of chromosomes 8, 11, and 17 will be employed and are represented by the red, green, and yellow boxes, respectively.

Table 4. Centromeric α Satellite DNA Probes.

Properties	Chromosome 8	Chromosome 11	Chromosome 17
Probe	Vysis CEP 8	Vysis CEP 11 (D11Z1)	Vysis CEP 17 (D17Z1)
Reference Number	07J20-008	06J37-021	06J38-027
Band Region	8p11.1-q11.1	11p11.11-q11	17p11.1-q11.1
Probe Allotment	Probe premixed with Hybridization Buffer 220µL	20µL	20µL
Vysis Probe Spectrum	SpectrumOrange	SpectrumGreen	SpectrumAqua
Excitation Peak, Emission Peak	559, 588	509, 538	433, 480
Chroma Filter, Excitation, Emission	Cy3, 530±30, 570	FITC, 493±16, 525	CFP, 436±20, 476
Pseudo Colored	RED	GREEN	YELLOW

3.3.2 Cell Fixation

Prior to performing FISH, it was first necessary to fix cells so that slides could be prepared and processed. In general, slides were generated for hTERT, HCT116 and HeLa. First, cells were harvested for fixation by removing the medium and washing with 5 mL of 1xPBS. Similar to cell culturing protocol detailed above, cells were trypsinized and transferred with 5-10 mL of 1xPBS to a 15 mL conical tube. The cells were centrifuged (140xg, 5 min), the supernatant was aspirated, and the pelleted cells were resuspended in residual supernatant by gently tapping the centrifuge tube. Cells were resuspended in aliquots of 1-2 mL of hypotonic solution (75 mM KCl) added slowly, up to a total volume of 5 mL. The cells were mixed between increments by gently flicking the tubes to minimize cell clumping. Optimal hypotonic incubation times were experimentally determined for all three cells lines employed in the DNA probe set analysis at various time points, and were 10 min for the hTERT and HCT116 cells, and 25 min for HeLa. Following the hypotonic treatment, 1 mL of freshly prepared 3:1 methanol:acetic acid fixative (Fisher Scientific) was added and the cells were immediately centrifuged. The supernatant was aspirated and the pelleted cells were resuspended in residual supernatant by gently flicking the tubes. Fresh fixative (5 mL) was added to the cells with gentle tapping to mix. The cells were centrifuged as above and the supernatant was aspirated. This process was repeated and cells were treated a second time in fixative as detailed above. For the third and final cycle of fixative, the cells were resuspended in 3 mL of fixative and finally stored at 4°C.

3.3.3 Generating Mitotic Spreads

To confirm the specificity of the DNA probe set, fixed cells were utilized to generate mitotic spreads. Prior to fixation however, the cells were enriched in mitosis through the addition of mitotic arresting agents (e.g. colcemid or nocodazole) to the growth medium. Colcemid depolymerizes microtubules thus disrupting the spindle fibers, while nocodazole (Section 3.4.1) inhibits microtubule polymerization and the generation of spindle fibers. Colcemid (Invitrogen, 10 µg/mL) was added to the medium of asynchronously growing cells at a working concentration of 0.1 µg/mL. Cells were approximately 70% confluent at the time of treatment and were treated for 1.5 hours (h) prior to harvesting. Empirical tests determined that 1.5 h was the optimal time for colcemid treatment in all three cell lines.

After the cells were harvested and fixed, they were used immediately to generate mitotic spreads or stored for up to two weeks in fixative at 4°C. To generate mitotic spreads, cells were centrifuged and resuspended in minimal fresh fixative (~50-500 µL). Surgipath glass slides were cleaned in 70% ethanol (Fisher Scientific), air dried and labelled. 25-30 µL of fixed cells were dropped at arm's length from the slide, and the slide was immediately placed on a hotplate (~37°C) for ~3 seconds and immediately submerged in 50% acetic acid to fix.

To identify and mark regions on the slides for FISH, the mitotic chromosome spreads were stained with Giemsa (Gibco). Giemsa binds to the phosphate groups of DNA and is removed during a standard FISH protocol, making it an ideal stain to visualize the DNA when generating mitotic spreads. After the fixed cells on the slides were air-dried, 1-2 mL of Giemsa was administered to the surface of the slide using a

glass pipette. Following a 3 min incubation period, excess stain was removed by washing with distilled H_2O and the slide was placed on the hotplate to dry. The mitotic spreads were visualized using an inverted ID03 microscope with 10x objective. If the mitotic spreads were inadequate, steps outlined in Table 5 were taken to achieve sufficient spreads for FISH and karyotyping. When a region was identified to have sufficient, high quality mitotic spreads, it was marked using a diamond pen (Fisher Scientific) prior to performing FISH.

Table 5. Trouble Shooting Approaches Employed to Generate Mitotic Spreads.

Problem	Solution
Excessive Chromosome Spreading	Mitotic cells bursting in fixative, shorten hypotonic treatment
Overlapping/ill-defined mitotic spreads	Increase the amount of fixative
	Increase the time on the hotplate after cells are dropped
Too few mitotic spreads	Increase the length of time of the colcemid/nocodazole treatment
	Mitotic cells not bursting, increase hypotonic treatment
	Decrease the amount of fixative
Overlapping chromosomes	Decrease/eliminate the time on the hotplate after cells are dropped
	Chromosomes are too long, increase the length of time of the colcemid/nocodazole treatment
Excessive chromosome compaction	Decrease the length of time of the colcemid/nocodazole treatment

3.3.4 Fluorescence *in situ* Hybridization (FISH) Protocol for Human Cell Lines

Interphase cells and mitotic chromosome spreads from hTERT, HCT116 and HeLa cell lines were used to validate the 3 FISH probes employed in this study (Table 4). Slides were generated as detailed in Section 3.3.3, and were rinsed twice with 2x saline-sodium citrate buffer solution (2xSSC) (Vysis) for 5 min, with shaking. Next pepsin (Sigma-Aldrich), a proteolytic enzyme that digests proteins including histones and other DNA-associated proteins, was added to pre-warmed 0.01M HCl (Fisherbrand) (37°C) at a working concentration of 50 µg/mL. The acidic pH of the HCl solution was matched to the optimal pH for pepsin activity (pH 2.0). The optimal pepsin incubation time was empirically determined for each cell line through a time series and was determined to be 7 min for HCT116, 10 min for HeLa, and 15 min for hTERT all of which were maintained at 37°C. Following the pepsin treatment, the samples were washed twice with 1xPBS and once with 1xPBS/50mM $MgCl_2$ for 5 min each. Next, slides were submerged into freshly prepared 1% formaldehyde/1xPBS/50mM $MgCl_2$ (Sigma-Aldrich) for 10 min. The samples were washed for 5 min in 1xPBS, and dehydrated in serial washes of 70%, 90% and 100% ethanol and air dried for 5 min. The DNA was further denatured by heating the slides at 70°C in an oven (1500 E Incubator, VWR Scientific) (5 min) and immediately transferring them to pre-warmed 70% formamide/2xSSC (Sigma-Aldrich) (70°C water bath, 2 min). Samples were fixed using ice cold 70%, 90% and 100% ethanol and air dried at room temperature. The samples were ready for the addition of the FISH probes.

The FISH probes that make up the DNA probe set (Section 3.3.1) are supplied pre-denatured by the manufacturer (Vysis). To prepare the multiplexed cocktail of

probes, the probes were first thawed on ice and protected from the light to minimize bleaching. The chromosome 8 FISH probe was purchased pre-diluted in hybridization buffer, and was the solution into which the chromosome 11 and 17 FISH probes were added. A total of 1 µL of the chromosome 11 and 17 FISH probes were added to 8 µL of the chromosome 8 FISH probe, mixed and transferred to a 0.5 mL microcentrifuge tube (Fisher Scientific). The multiplexed cocktail was subsequently mixed by vortexing and briefly centrifuged (Biofuge Fresco) to collect the probes. A total of 10 µL of the DNA probe set was transferred to an 18x18mm coverslip (Fisherbrand 1.5) before gently placing the coverslip on the marked region of the denatured slide. The edges of the coverslip were sealed with rubber cement (Elmers) to prevent desiccation, and the samples were incubated overnight at 37°C in the dark, in a ThermoBrite slide processing system.

The following day, excess FISH probes were removed through three 5 min washes in 50% formamide/2xSSC at 42°C followed by two additional rinses in 2xSSC (5 min, 42°C). As a final step, the DNA was counterstained through the addition of 30µL of DAPI (Sigma-Aldrich, 0.1 µg/mL), a DNA intercalating dye, which was added to a coverslip and placed on the slide for 3 min. The DAPI coverslip was removed and a second coverslip supporting a drop of Vectashield anti-fade mounting medium (Vector Laboratories) was placed on the marked region of the slide. The slides were incubated at room temperature for 3-5 h and subsequently transferred and stored at -20°C, prior to imaging (see below).

3.4 CELL SYNCHRONIZATION

To ensure that changes in the FISH probe numbers does not simply reflect differences in DNA content due to DNA replication and cell cycle stages, it was first imperative to evaluate the probes in both pre- and post-S-phase. To analyze the FISH probe signals in cells pre- and post-DNA replication, cell synchronization was employed to enrich cells at the G_1/S-phase boundary through a double thymidine arrest (Section 3.4.1), or at G_2/M via double thymidine arrest followed by a release into nocodazole (Section 3.4.1). Importantly, cells were simultaneously harvested and evaluated by both flow cytometry for cell cycle analysis, and FISH for probe enumeration.

3.4.1 Double Thymidine Arrest

To enrich a population of cells at the G_1/S-phase boundary a double thymidine arrest was performed (Fig. 5). Asynchronously growing HCT116 cells were seeded into two 20 cm culture dishes at 20% confluency, 24 h prior to thymidine treatment. Thymidine is pyrimidine deoxynucleoside that enriches cells at the G_1/S-phase boundary by activating a negative feedback loop and restricting the cell from producing sufficient nucleosides to complete DNA replication. The cells were treated with 2.5 mM thymidine (Sigma-Aldrich) for 12 h at 37°C. After the first thymidine treatment, cells were released from the initial arrest by rinsing twice with pre-warmed 1xPBS (37°C) and once with pre-warmed McCoy's 5A medium (37°C). Cells were allowed to progress through the cell cycle for 10 h in fresh McCoy's 5A medium (37°C) without thymidine, prior to a second thymidine arrest (as above).

To generate a G_2/M-enriched population of cells, cells were released from the double thymidine arrest (Fig. 5), and permitted to grow for 7 h, a time point

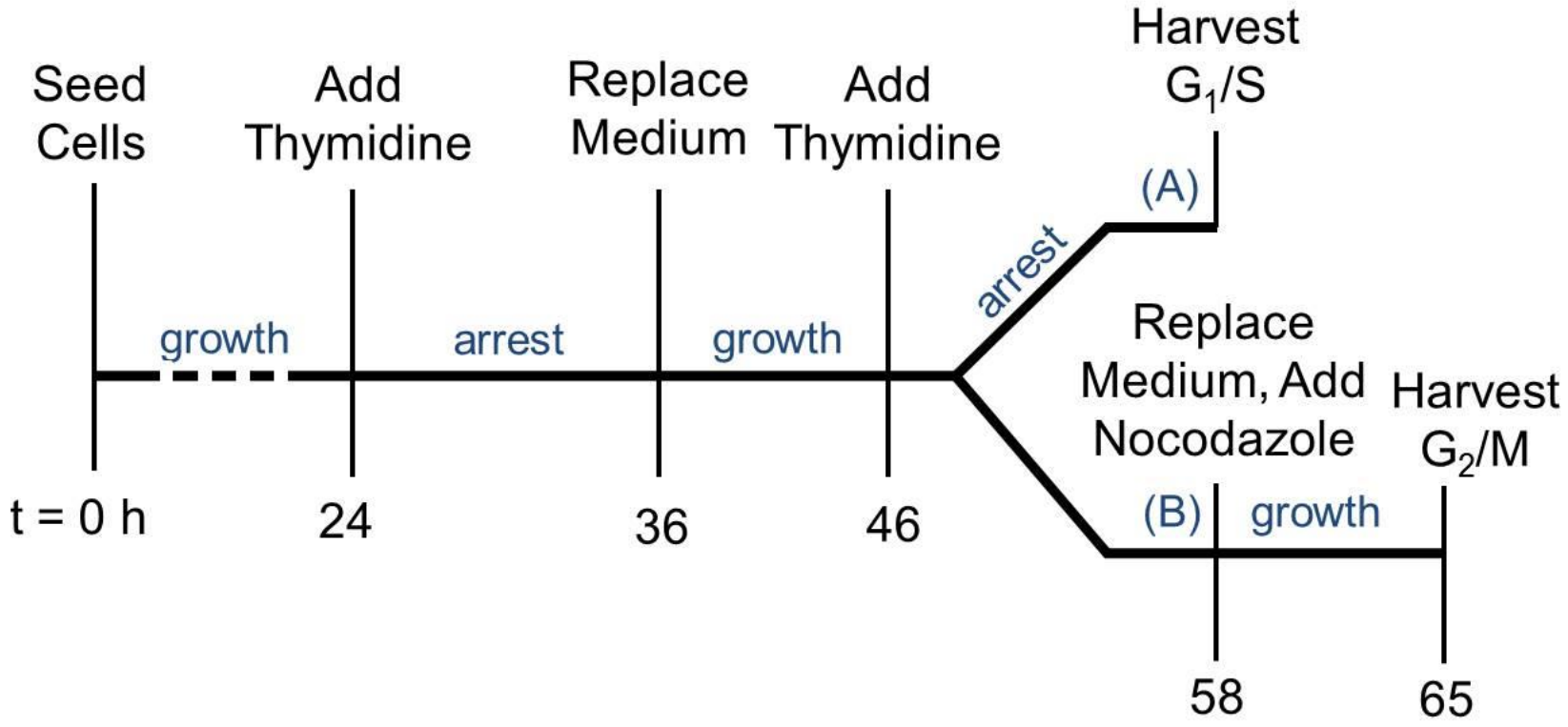

Figure 5. Cell Synchronization Scheme and Timeline.

Schematic depiction of the time line employed to generate either G_1/S-phase or G_2/M-enriched cellular populations. During thymidine treatments cells are incubated for 12 h, while between treatments cells are incubated in thymidine-free medium. At the end of the second thymidine treatment, cells are enriched at the G_1/S-phase boundary and are either harvested ***(A)*** or released into fresh medium containing nocodazole ***(B)*** to be harvested in G_2/M.

established to contain a high percentage of G_2/M cells (McManus KJ, unpublished observations). To prevent cells from progressing beyond G_2/M into G_1, nocodazole (Sigma-Aldrich) was added at working concentration of 1 µg/mL. Cell cycle staging and enrichment was confirmed through standard DNA labeling with propidium iodide (PI) followed by flow cytometric analysis (see below).

3.4.2 Cell Cycle Analysis by Flow Cytometry

To determine the relative proportion of cells in various cell cycle stages (e.g. G_0/G_1, S-phase, or G_2/M) standard flow cytometry was performed. Three populations of cells (asynchronous, G_1/S-phase enriched, and G_2/M enriched) were harvested and labelled with PI in preparation for flow cytometry and cell cycle analysis. PI is a DNA/RNA intercalating agent and by staining fixed cells with PI, DNA content, and thus the stage of the cell cycle, can be determined.

3.4.2.1 *Ethanol Fixation*

Three populations of HCT116 cells, namely asynchronous, G_1/S-phase enriched, and G_2/M enriched were grown on 20 cm tissue culture dishes and detached using trypsin as detailed above. Once detached, cells were pelleted by centrifugation (140xg, 5 min) and resuspended in 4-5 mL of 1xPBS. To determine cell concentrations, 40 µL of cells were transferred to a microcentrifuge tube, combined with 40 µL of Trypan Blue Stain (Invitrogen), and the concentration was determined using the Cedex XS Cellular Analyzer (Roche). A volume containing $2x10^6$ cells was transferred to a 5 mL BD/Falcon fluorescence activated cell sorter (FACS) tube and the remaining cells were used to generate samples for FISH analysis (Section 3.3). Samples reserved for flow cytometry were centrifuged at 315xg (5 min) and the supernatant was gently poured off.

Cells were fixed and permeablized by adding 1 mL of 70% ice cold ethanol, drop wise, while continuously vortexing. Cells were fixed overnight at 4°C prior to counterstaining the DNA with PI.

3.4.2.2 *Propidium Iodide Labelling*

Following ethanol fixation, cells were pelleted by centrifugation (315xg, 1200 rpm, 5 min), ethanol was removed and cells were resuspended in 1 mL of 1xPBS. 10 µL of RNase A (10 mg/mL) (Fisher Scientific), an RNA endonuclease, was added to degrade RNA to ensure an accurate assessment of only the DNA content. In addition, 40 µL of PI (Fisher Scientific) was added and the cells were incubated in the dark for 30 min at 37°C. The cells were handled with minimal exposure to light, centrifuged (315xg, 5 min) and resuspended in 500 µL of 1xPBS and immediately analyzed by flow cytometry.

3.4.2.3 *Flow Cytometry*

Flow cytometry was performed using the Becton Dickinson FACS Calibur (BD Cell Quest Pro software, version 0.4.cf2b) on three populations of HCT116 cells, specifically asynchronous, G_1/S enriched, and G2/M enriched. Instrument settings were optimized utilizing the asynchronous population of HCT116 cells stained with PI. Each flow acquisition collected a minimum of 10,000 gated events, and three replicates were collected for each sample. Cells were gated based on FSC-H and FL2-W. The data was processed with flow cytometry software (FlowJo, version 10.0.5) by gating out doublets and plotting the populations against PI intensity (FL2-A). Finally, the proportion of cells in the different stages of the cell cycle was analyzed for the G_1/S-phase enriched and G_2/M enriched populations using the cell cycle analysis option contained within FlowJo.

3.5 CHROMOSOMAL INSTABILITY ANALYSIS IN COLORECTAL CANCER TUMOR SAMPLES

A FISH based approach was employed to determine the prevalence of CIN within a Manitoba cohort of Interval CRC tumor samples and compare it with that observed within a Manitoba cohort comprised of sporadic CRC samples. Archived clinical FFPE tumor tissue blocks were supplied by various hospitals in Winnipeg, including the Concordia, Grace, Health Sciences, Seven Oaks, St. Boniface, and Victoria General Hospitals. A total of 67 samples including 21 Interval and 46 sporadic (control) CRC samples were specifically selected and obtained for CIN analysis and are collectively referred to as the Manitoba cohort of Interval and sporadic CRCs (detailed below). The collection and use of these samples for this study was approved by the University of Manitoba Research Ethics Board.

3.5.1 Manitoba Cohort of Interval and Sporadic Colorectal Cancer Samples

The Manitoba cohort of Interval and sporadic CRCs consisted of a 1:2 ratio of Interval to sporadic CRC tumor samples. The samples were provided in an anonymized, double-blinded fashion and thus the type of each CRC tumor (i.e. Interval or sporadic) was unknown and only revealed once all samples had been analyzed. Additional clinical information associated with each sample, including gender and age was only revealed after all samples had been evaluated for CIN.

For the purpose of this study, Interval CRCs were defined as CRCs diagnosed between 6-36 months following a negative colonoscopy, as these limits reduce the probability of capturing missed sporadic CRCs (i.e. those missed during the initial colonoscopy) and better ensures the inclusion of those tumor samples that appear to arise

rapidly. Accordingly, the CRCs detected upon initial colonoscopy, or detected before/after the 6-36 month time period were classified as sporadic CRCs and could be included as controls. Whenever possible, sporadic CRCs were matched to each Interval CRC by age (± 5 years), gender and tumor location (i.e. right vs. left). Each CRC tumor sample included in the cohort was assigned a study ID. After the samples were completely analyzed, the study was unmasked and the basic physical characteristics of the samples were charted in a study database. The most vital properties from the database are outlined in Tables 6-8, and includes whether the CRC samples are Interval or sporadic. Additional properties detailed include patient gender and age, location of tumor presentation, and TNM stage.

In addition to the inclusion criteria listed above, several exclusion criteria were also incorporated into the selection of the CRC cohort employed in this study. Exclusion criteria included patients with family history of IBD. In addition, patients diagnosed with CRC before the age of 50 years were excluded due to the higher probability of a genetic predisposition for CRC.

Table 6. Manitoba Cohort of Interval and Sporadic CRC Samples; TMA 24A.

Sample	Interval	Gender	Age	Location*	TNM Stage	TMA Position
1	No	M	61	Rectosigmoid	N1	1a, 1b
2	No	F	58	Ascending	pT3N0MX	2a, 2b
3	Yes	M	72	Transverse	pT3N1	3a, 3b
4	Yes	M	77	Ascending	pT2N0MX	4a, 4b
5	No	M	73	Ascending	pT2N0	5a, 5b
6	Yes	M	61	Ascending	NI	6a, 6b
7	No	M	80	Ascending	pT3N1MX	7a, 7b
8	No	M	78	Rectosigmoid	pT3N0M0	8a, 8b
9	No	F	76	Ascending	pT3	9a, 9b
10	No	F	70	Ascending	pT3N0	10a, 10b
11	No	M	61	Transverse	pT3N1	11a, 11b
12	Yes	F	80	Ascending	pT3N1M0	12a, 12b
13	Yes	M	79	Ascending	NI	13a, 13b
14	No	M	55	Ascending	NI	14a, 14b
15	No	M	65	NI	NI	15a, 15b
16	No	M	64	Ascending	pT2	16a, 16b
17	No	M	75	Rectosigmoid	NI	17a, 17b
18	No	M	80	Ascending	pT2	18a, 18b
19	Yes	M	73	NI	pT3N1M0	19a, 19b
20	No	M	79	Ascending	pT3N0M0	20a, 20b
21	No	M	73	Ascending	pT3N0MX	21a, 21b

*Abbreviation; NI (Not Indicated)

Table 7. Manitoba Cohort of Interval and Sporadic CRC Samples; TMA 24B.

Sample	Interval	Gender	Age	Location	TNM Stage	TMA Position
22	No	F	70	Transverse	pT4N2	1a, 1b
23	No	M	76	Ascending	pT1N0MX	2a, 2b
24	No	F	69	Transverse	pT3N0M0	3a, 3b
25	No	M	70	Ascending	pT3	4a, 4b
26	Yes	F	59	Rectosigmoid	pT4N2M?	5a, 5b
27	No	F	63	Rectosigmoid	NI	6a, 6b
28	Yes	M	80	Ascending	pT4N0Mx	7a, 7b
29	Yes	M	78	Descending	pT2N0MX	8a, 8b
30	No	M	75	Ascending	pT3	9a, 9b
31	No	F	79	Transverse	pT2N0	10a, 10b
32	No	M	73	Transverse	pT3N2	11a, 11b
33	Yes	F	57	Ascending	pT2N0MX	12a, 12b
34	Yes	F	50	Ascending	pT4bN2a	13a, 13b
35	No	F	52	Ascending	pT3N2	14a, 14b
36	Yes	F	64	Ascending	pT3N0MX	15a, 15b
37	No	M	77	Ascending	pT2N1MX	16a, 16b
38	No	M	74	Rectosigmoid	NI	17a, 17b
39	Yes	M	61	Rectosigmoid	pT4N1	18a, 18b
40	No	M	78	Descending	pT4aN2aMX	19a, 19b
41	No	M	59	Descending	pT4bN0	20a, 20b
42	Yes	F	77	Transverse	pT3N0Mx	21a, 21b
43	No	M	55	Ascending	pT3N2MX	22a, 22b

*Abbreviation; NI (Not Indicated)

Table 8. Manitoba Cohort of Interval and Sporadic CRC Samples; TMA 24C.

Sample	Interval	Gender	Age	Location	TNM Stage	TMA Position
44	No	F	76	Rectosigmoid	pT2N2	1a, 1b
45	No	F	67	Ascending	T4N2M1	2a, 2b
46	No	M	62	Rectosigmoid	pT3bN0	3a, 3b
47	No	F	62	Ascending	pT3N0	4a, 4b
48	No	F	69	Rectosigmoid	pT4N2	5a, 5b
49	Yes	F	70	Transverse	pT3b	6a, 6b
50	No	F	77	Ascending	pT4aN2b	7a, 7b
51	No	M	67	Ascending	NI	8a, 8b
52	No	M	79	Ascending	pT3N1M1	9a, 9b
53	Yes	M	78	Transverse	pT3N0	10a, 10b
54	No	M	59	NI	NI	11a, 11b
55	Yes	F	58	Ascending	pT3N1M0	12a, 12b
56	Yes	M	60	Ascending	pT3N0MX	13a, 13b
57	No	F	54	Ascending	pT3N2a	14a, 14b
58	No	M	78	Rectosigmoid	pT3N2MX	15a, 15b
59	No	NI	NI	NI	NI	16a, 16b
60	Yes	M	73	Ascending	pT4N0	17a, 17b
61	No	F	71	Rectosigmoid	pT4bN0	18a, 18b
62	Yes	M	74	Rectosigmoid	NI	19a, 19b
63	Yes	M	79	Rectosigmoid	pT3N0	20a, 20b
64	No	F	76	Rectosigmoid	pT2N0MX	21a, 21b
65	No	F	64	Ascending	pT3N1MX	22a, 22b
66	No	F	74	Ascending	pT3N0MX	23a, 23b
67	No	M	65	Ascending	pT3N0M0	24a, 24b

*Abbreviation; NI (Not Indicated)

3.5.2 Tumor Microarrays

The CRC patient tumor samples that were selected to be a part of the Manitoba cohort of Interval and sporadic CRCs were supplied in the form of archived clinical FFPE tissue blocks with corresponding hematoxylin and eosin (H&E) stained slides. Haematoxylin is a basic dye that stains acidic structures, most predominantly the nucleic acids of DNA, a deep purple/blue color. Eosin is an acidic dye and works by counterstaining basic cellular components, often referred to as eosinophilic structures, shades of pink. All slides from the CRC cohort were reviewed by the study collaborator and pathologist Dr. Robert Wightman (University of Manitoba and Grace General Hospital, Winnipeg). Slides were reviewed and marked with a felt pen by Dr. Wightman for regions containing high cellularity that were not associated with a tumor border or necrotic zone, and 6-10 regions per sample were identified for subsequent coring.

To facilitate the CIN enumeration, three "mini-TMAs" were generated for this study on a fee-for-service basis by the Manitoba Tumor Bank. Mini-TMAs were generated to contain a manageable number of samples to ensure that samples could be labeled and imaged with minimal overall photobleaching. In brief, each mini-TMA contained 21-24 unique patient tumor samples that were arrayed in duplicate as 0.6 mm cores that were selected from the regions identified by our pathology collaborator. Figures 6A, 6B, and 6C depict the maps of TMA 24A, 24B and 24C, respectively. Mouse tissue was included in each mini-TMA (designated L) for orientation purposes during imaging. Each mini-TMA was subsequently sectioned by the Manitoba Tumor Bank at a thickness of 0.5 µm, a thickness determined to include intact (i.e. complete) nuclei and be compatible with the FISH probe labeling protocol (Fig. 4, Table 4).

A. TMA 24A

L	1A	1B	2A	2B	3A	3B	4A	4B	5A	5B
	10A	10B	9A	9B	8A	8B	7A	7B	6A	6B
	11A	11B	12A	12B	13A	13B	14A	14B	15A	15B
	20A	20B	19A	19B	18A	18B	17A	17B	16A	16B
	21A	21B	L	L	L					

B. TMA 24B

L	1A	1B	2A	2B	3A	3B	4A	4B	5A	5B
	10A	10B	9A	9B	8A	8B	7A	7B	6A	6B
	11A	11B	12A	12B	13A	13B	14A	14B	15A	15B
	20A	20B	19A	19B	18A	18B	17A	17B	16A	16B
	21A	21B	22A	22B						
	L	L	L							

C. TMA 24C

L	1A	1B	2A	2B	3A	3B	4A	4B	5A	5B
	10A	10B	9A	9B	8A	8B	7A	7B	6A	6B
	11A	11B	12A	12B	13A	13B	14A	14B	15A	15B
	20A	20B	19A	19B	18A	18B	17A	17B	16A	16B
	21A	21B	22A	22B	23A	23B	24A	24B	L	L
										L

Figure 6. The Position of Colorectal Cancer Tumor Samples within the Mini-TMAs. Schematic map of the 67 unique CRC samples cored and positioned in TMA 24A ***(A)***, TMA 24B ***(B)***, and TMA 24C ***(C)***. The mini-TMAs include CRC samples cored in duplicate (A, B) (blue), and 3 mouse tissue samples (L) used for mini-TMA section orientation (white).

3.5.3 Fluorescence *in situ* Hybridization (FISH) of Formalin-Fixed, Paraffin Embedded Colorectal Cancer Samples

FISH was initially performed and optimized using CRC tissue sections and subsequently performed for each of the three mini-TMAs. First, paraffin was removed from the samples through two washes in xylene (6 – 10 min each) (Fisher Scientific), followed by two washes in 100% ethanol (2 x 10 min), and allowed to air dry. To prepare the DNA for hybridization with the DNA FISH probe set, the samples were incubated in a citric acid buffer for 1 h at 80°C, followed by a 45 min pepsin treatment. Pepsin was added to pre-warmed 0.2N HCl (pH 2.0, 37°C water bath) at a working concentration of 0.5 mg/mL, immediately prior to use. After two washes in 2xSSC and a serial dehydration in 70%, 90% and 100% ethanol, the samples were hybridized with the DNA probe set, which was prepared as detailed above (Section 3.3.4). Next, samples were incubated at 77°C for 10 min in the ThermoBrite slide processing system, followed by an overnight incubation at 37°C.

The following day, excess probe was removed through three 10 min washes in 50% formamide/2xSSC (46°C), followed by 2xSSC for 10 min (46°C), and 2xSSC/0.1% NP-40 (VWR, 46°C) for 5 min. Finally, the samples were washed with 2xSSC (5 min, room temperature), counterstained with DAPI, mounted in Vectashield as detailed above (Section 3.3.4) and stored in the dark at -20°C until imaged.

3.6 MICROSCOPY AND IMAGING

All 3D images were acquired with an Imager.Z1 (Zeiss) widefield fluorescence microscope equipped with a charged couple device HRm (Zeiss) camera. The oil-immersion objectives lenses employed included a 40x EC Plan-Neofluar (numerical aperture 1.30) and a 63x Plan-Apochromat (numerical aperture 1.40). Zeiss Immersol 518F immersion oil (refractive index 1.518) was used in the acquisition of all images. Images of the various fluorophores were collected using fluorescent filters for DAPI, and those specific to the DNA probe set, Cy3, FITC, and CFP (Table 4).

3.6.1 Two-Dimensional Image Acquisition

2D images of mitotic spreads were utilized for karyotype analysis and subsequent examination of the fluorescent probe specificity. 2D images were acquired using an Axioplan 2 microscope (Zeiss). Images were analyzed using the Case Data Manager software from ASI (Applied Spectral Imaging). The identification of chromosomes was automated as well as the generation of the karyotypes. Manual corrections were performed when the mitotic spreads contained overlapping chromosomes. To ensure the mitotic spreads were accurately karyotyped, the results were analyzed and confirmed by a trained cytogeneticist, Ms. Zelda Lichtensztejn.

3.6.2 Three-Dimensional Image Acquisition

All 3D images for either the cell lines or the mini-TMA analyses were acquired using AxioVision 4.8 (Zeiss) software. Images of cell lines (Section 4.1, pg. 67) were acquired with the 63x oil objective, at 400 nm intervals. Slides generated from hTERT, HCT116 and HeLa consisted of a single plane of cells, and therefore the total thickness of the 3D images was considerably less than the tissue section/TMA images. Depending on

the cell type, the slides required 15-30 optical sections per image. All images were saved as 16-bit Carl Zeiss (.zvi) files and exported for image deconvolution as detailed in Section 3.6.3 (see below).

Orientation of the tissue sections and mini-TMAs was performed using a 20x objective for position and image registration, and subsequently changed to the 40x oil-immersion objective for image acquisition. The optical sections of the 3D images were optimized for the visualization of the DNA probe set and both 200 nm and 400 nm intervals were evaluated. The 400 nm optical sections were sufficient to visualize the fluorescent probes while minimizing photo-bleaching of the samples. The total thickness of the 3D images were determined by setting the top and bottom boundaries of the optical sections, and this was independently optimized for each image collected. Due to the 3D orientation of the cells in CRC tissue, each image of the tissue sections/TMAs consisted of approximately 40-46 optical sections. Depending on the cellularity of each cores within the mini-TMAs, 4-8 optical series were acquired. Imaging systematically across the mini-TMAs and within each sample (i.e. from bottom to top, left to right) ensured that no nucleus was imaged more than once. To account for sample variability and autofluorescence, the exposure times for each of the fluorescent probes were independently optimized. Briefly, exposure times were selected such that no image saturation was apparent and the dynamic range for the specific channel (e.g. DAPI, Cy3, FITC and CFP) was set to ~85% of the entire range.

3.6.3 Image Deconvolution

All 3D image series were imported and processed by maximum-likelihood-expectation deconvolution in AutoQuant X3 (Media Cybernetics). The deconvolution software employs a constrained iterative algorithm and theoretical point spread functions for each of the fluorescent channels, DAPI (461nm), CFP (476 nm), FITC (525 nm), and Cy3 (570 nm). Each image was saved as a 16-bit IMS (Bitplane) file that could be imported directly into Imaris (Bitplane) image visualization software for FISH probe enumeration purposes.

3.7 PROBE ENUMERATION

Imaris version 7.6.3 was employed for the manual enumeration of the fluorescent probes. Briefly, each channel (i.e. FISH probe) was independently scored, and 100 complete (i.e. intact) nuclei were analyzed from each core. Surface renderings were generated for the DAPI channel to ensure that only complete nuclei were included in the enumeration process, and only foci (of the DNA probe set) contained within the nuclear boundaries as defined by the DAPI signal intensity were enumerated. A fluorescent signal was considered a focus and enumerated if; 1) it was within the nuclear boundary, 2) it did not overlap with fluorescent signal in the other two channels (i.e. tissue autofluorescence), 3) it was spheroid in shape, and 4) it had similar fluorescent signal intensity to the other foci within the particular channel and image.

To ensure that each nucleus was only analyzed once, each enumerated nucleus was assigned a number and the enumeration data was assembled in Excel (Microsoft). 3D projections (i.e. “maps”) were generated for each image enumerated and opened in Photoshop CS6 (Adobe, Version 13.0.5x64) to record the assigned numbers for each

nucleus that was enumerated. All image panels were also generated in Photoshop CS6 and the figures and maps were saved as JPEG files.

3.7.1 Evaluation of Chromosomal Instability

In order to determine the prevalence of CIN within Interval CRC, the extent of chromosomal gains and losses was compared between the sporadic and Interval CRC samples. In detail, a CIN score was devised to measure the amount of chromosome gains and losses within each sample enumerated and was defined as the average number of chromosome gains and losses per nucleus. Chromosomes 8, 11 and 17 were enumerated in ~100 nuclei per core. Any deviation from the expected 2 foci/chromosome/nucleus was determined, and the sum of these aberrations (i.e. chromosome gains and/or losses) was divided by the total number of nuclei enumerated in a sample. For example, if a nucleus lost a copy of a particular chromosome, it was recorded as a single aberration, if it lost two copies of a single chromosome or lost a single copy of two chromosomes, it was recorded as having two aberrations, *etc*. Similarly, gains in chromosome foci were also recoded as aberrations, and thus with this calculation, nuclei harbouring both chromosome gains and losses were not scored as diploid/normal. Finally, the data were imported to Prism (GraphPad, Version 6.0a), and bar graphs and scatterplots were generated to compare both CIN scores as well as individual chromosomal gains/losses between the Interval and sporadic CRC samples.

CHAPTER 4: RESULTS

4.1 AIM 1: TO ESTABLISH AND OPTIMIZE A PROTOCOL THAT CAN DETECT CHANGES IN CHROMOSOME NUMBERS

To assess aneusomy (i.e. a surrogate marker for CIN) a FISH-based approach was employed that could simultaneously evaluate three different chromosomes, namely chromosomes 8, 11 and 17. The pericentromeric FISH probes were selected as they hybridize to regions of DNA that surround the centromere, an essential element of a chromosome that is required for accurate segregation. Thus gains/losses in FISH signals are predicted to reflect increases/decreases respectively, in chromosome numbers. Accordingly, deviations from the expected number of 2 FISH signals per chromosome per nucleus are suggestive of a CIN phenotype.

4.1.1 Optimizing the Pepsin Treatment for Fluorescence *in situ* Hybridization (FISH) in Human Cell Lines

Prior to evaluating the FISH probe specificity it was first necessary to optimize several of the procedures used to generate the chromosome spreads and to optimize the DNA probe set hybridization to the various human cell lines. Optimizing the pepsin treatment in the FISH protocol for cell lines was required prior to the analysis of the DNA probe set in the G_1/S-phase and G_2/M-enriched HCT116 populations. Four samples of HCT116 cells were treated with pepsin for various incubation times (1, 5, 10 and 15 min) prior to hybridization with the DNA probe set. High-resolution (63x) 3D images were acquired of 15 nuclei per sample and subjected to image deconvolution. Both the fluorescent exposure times as well as the compaction of the FISH probe foci were evaluated for the final selection of the optimal pepsin incubation treatment. While 1

and 5 min pepsin treatments resulted in compact foci compared to the slightly more dispersed signals observed in the 10 and 15 min incubations, these samples required longer exposure times for image acquisition and consequently the signal intensities bleached more rapidly. In general, it was also noted the chromosome 11 foci appeared more diffuse (i.e. less punctate) than the other two chromosome FISH foci, yet were still easily enumerated. Based on these observations a 7 min treatment was selected as the optimal pepsin incubation time and was employed for all subsequent experiments (Fig. 7).

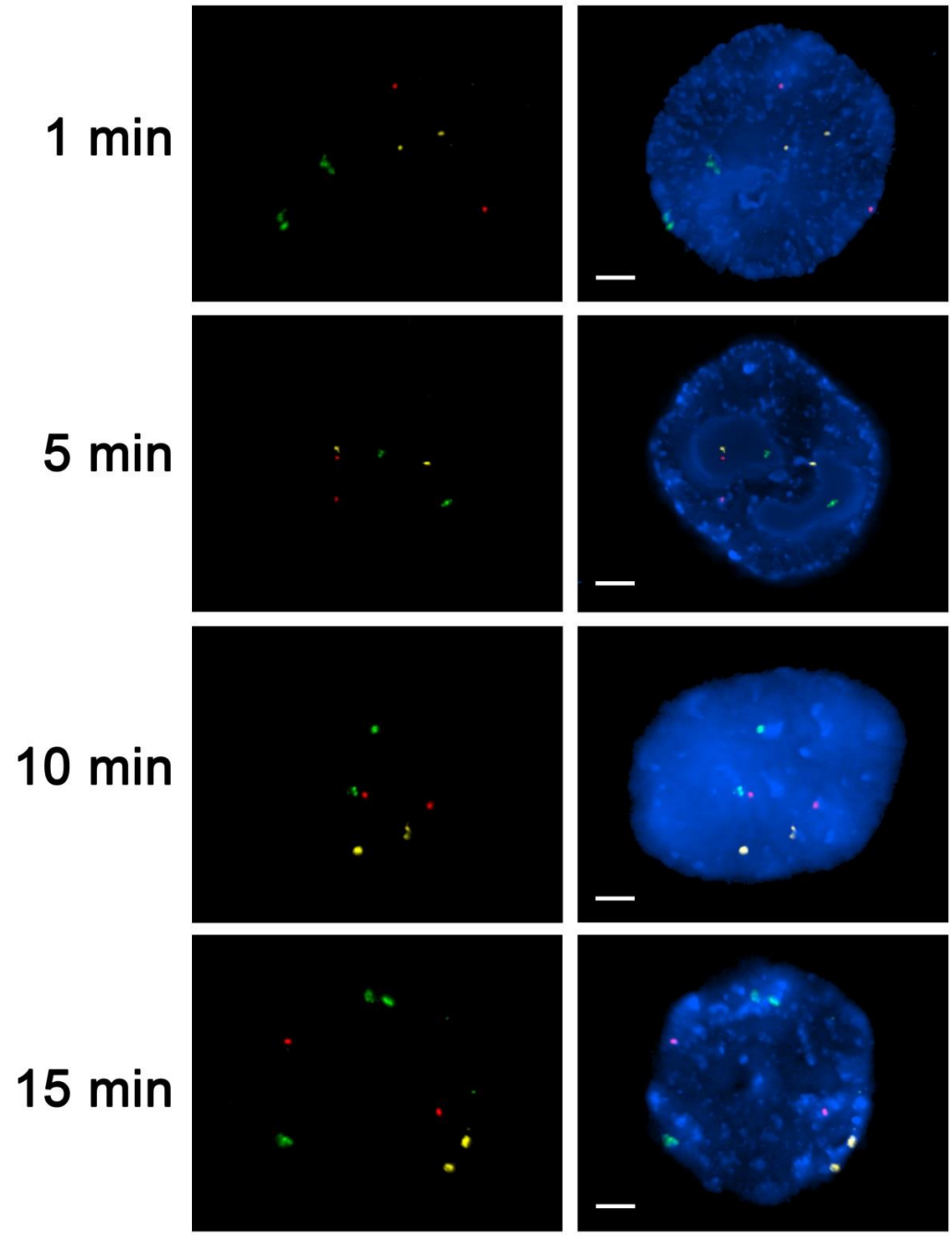

Figure 7. The Optimization of Pepsin Treatment in Human Cell Lines.
Representative 3D projections of individual HCT116 nuclei incubated with pepsin for 1, 5, 10 and 15 min each (indicated) and hybridized with the DNA probe set. The FISH probes for chromosomes 8, 11 and 17 are pseudocolored red, green and yellow respectively. As pepsin incubation times increased, the foci became less punctate, but required shorter exposure times for image acquisition. The optimal pepsin-treatment was determined to be 7 min based on signal shape and intensity (see Fig. 10 for example). Scale bar is equal to 7 um.

4.1.2 Confirming the Specificity of the DNA Probe Set

The specificity of the individual FISH probes was confirmed utilizing three human cell lines, namely hTERT (diploid), HCT116 (near diploid) and HeLa (hypotetraploid) (see Table 3 for additional cell line properties, pg. 43). Mitotic cells were used to confirm correct FISH probe localization and position (i.e. pericentromeric regions) within the corresponding chromosomes. Mitotic spreads of hTERT fibroblast cells were hybridized with the DNA probe set, and high resolution (63x) imaging and karyotype analysis revealed that the three FISH probes hybridize to the expected pericentromeric regions of chromosomes 8, 11 and 17 with high specificity (Fig. 8). Similar results were obtained in HCT116 and HeLa.

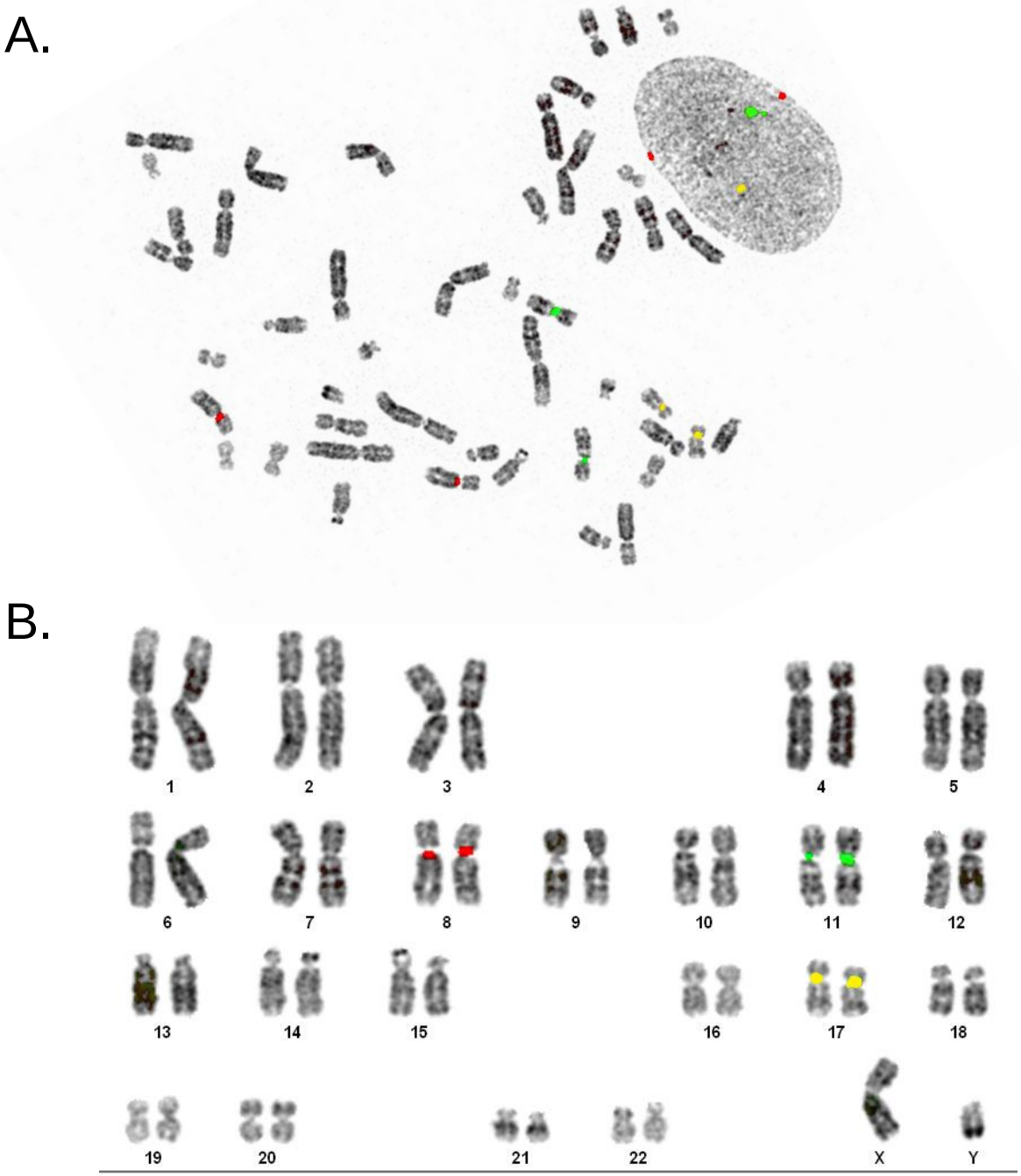

Figure 8. The DNA Probe Set Specifically Hybridizes to Chromosomes 8, 11 and 17. The DNA probe set that contains FISH probes for chromosomes 8 [red], 11 [green], and 17 [yellow] was hybridized to mitotic chromosome spreads generated from hTERT cells. ***(A)*** A representative mitotic chromosome spread for which karyotypic analysis ***(B)*** was performed and confirmed the expected pericentromeric chromosome localization. The karyotype of hTERT is 46, XY.

4.1.3 Evaluating the DNA Probe Set to Detect Changes in Chromosome Numbers

Although the above data confirm the DNA probe set correctly recognizes the three chromosomes in mitotic spreads, it was imperative to evaluate their ability to reflect changes in chromosome numbers in interphase cells, as the CRC tumor samples to be evaluated are predominately comprised of interphase cells. Consequently, the FISH protocol was performed on both mitotic spreads and interphase cells from both hTERT fibroblast (diploid) and HeLa (hypotetraploid) cell lines. When hybridized with the DNA probe set, interphase hTERT cells produced the expected two fluorescent foci/chromosome/nucleus (Fig. 9). The foci within the interphase cells were spheroid (punctate) and reflected similar numbers as observed within the mitotic chromosome spreads indicating that this approach is capable of accurately detecting chromosome (i.e. pericentromeric) copy numbers for each of the three chromosomes evaluated. HeLa cells were also evaluated using the DNA probe set and provided additional support that the protocol can be employed to enumerate abnormal chromosome numbers and visualize aneusomy as deviations from the diploid expectation of 2 foci/chromosome/nucleus were reproducibly detected (Fig. 9). In fact, most interphase HeLa cells evaluated had 4 copies of chromosome 8, and 3 copies each of chromosomes 11 and 17, which correlated with what was observed within the mitotic chromosome spreads. Finally it should be mentioned that there was no bleed-through of the various fluorophores or significant background hybridization. Collectively, the above data indicate that this approach is ideally suited to the analysis of CIN within cell lines and likely within interphase CRC tumor samples.

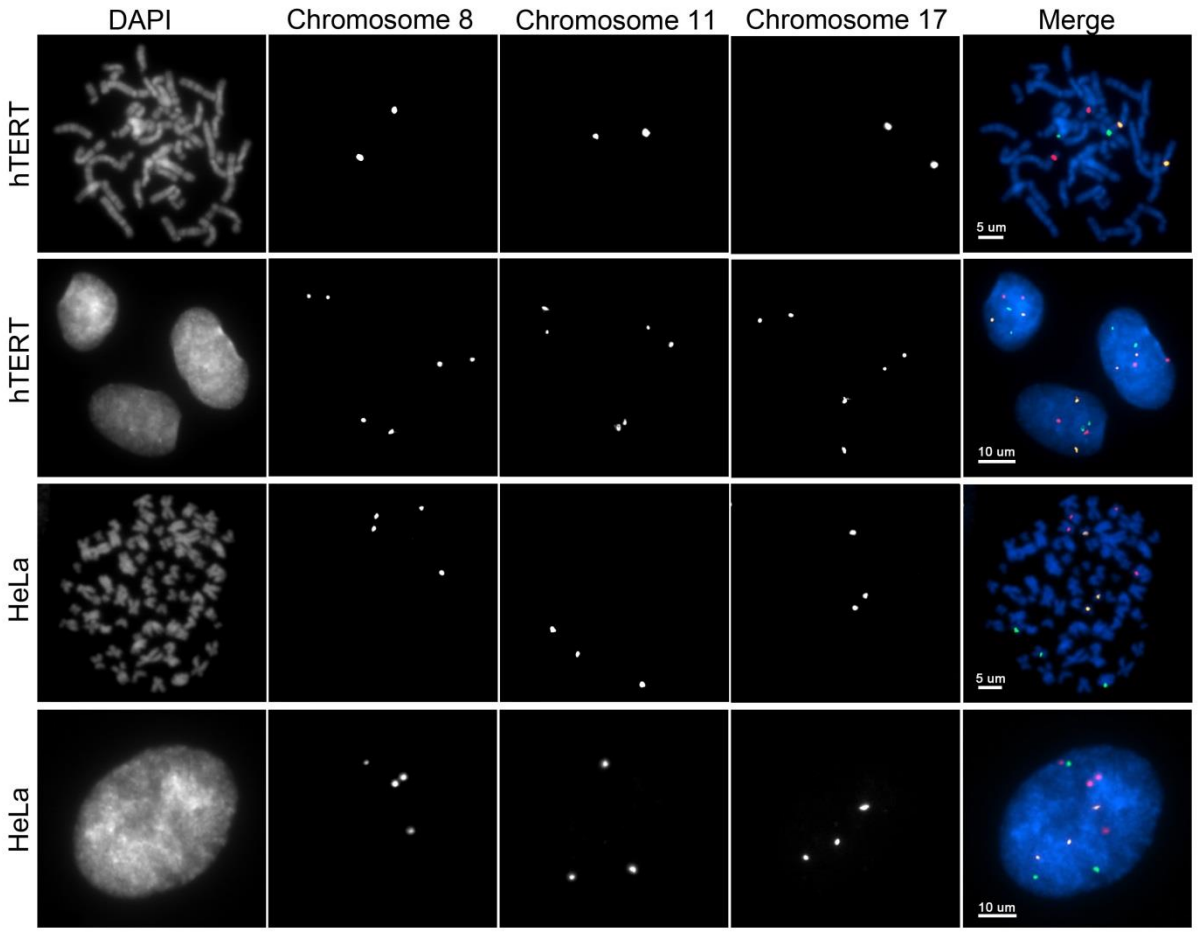

Figure 9. Mitotic and Interphase FISH Staining Patterns in Human Cell Lines.

Representative images depicting interphase and mitotic cells hybridized with the DNA probe set [chromosome 8, 11 and 17 FISH probes, pseudocolored red, green, yellow respectively within Merge]. The hTERT cells represent a diploid cell line, and produced the expected 2 FISH signals/probe per cell, while the HeLa cells are aneuploid (hypotetraploid) and thus harbor additional FISH signals for the three chromosomes evaluated.

4.1.4 Evaluating the DNA Probe Set Pre and Post S-Phase

To address concerns over the possibility that DNA replication may impact the ability to accurately enumerate FISH signals in interphase cells, the DNA probe set was evaluated in both pre and post S-phase cells. Recall that the FISH probe set was purposefully selected as it specifically recognizes pericentromeric DNA, a chromosomal region that remains tethered closely together throughout G2 and the early stages of mitosis due to sister chromatid cohesion[151]. Thus, signal 'duplication' due to DNA replication is less likely to occur within this region as compared to a chromosomal locus located along an arm that would be more likely to produce a double-dot FISH signal if within the resolution capabilities of the wide-field fluorescence microscopy employed.

To ensure DNA replication does not impact the ability to accurately enumerate chromosomes, cells at different stages of interphase were analyzed using the DNA probe set. Briefly, the DNA probe set was evaluated in HCT116 cells enriched at the G_1/S-phase boundary through a standard double thymidine arrest, and compared to those enriched in G_2 generated via a standard double thymidine arrest and release into nocodazole. The two cell cycle enriched populations and asynchronous controls were evaluated by flow cytometry and were determined to be 83.3% for the G_1/S-phase enriched population, and 89.5% for the G_2/M enriched population (Fig. 10A). Samples from the two cell cycle enriched populations of cells were processed using the standard FISH, imaged (Fig. 10B) and the fluorescent foci were enumerated. Overall, the enumeration (Table 9) revealed highly similar numbers of foci within the pre and post S-phase enriched populations, and were consistent for all three chromosomes evaluated.

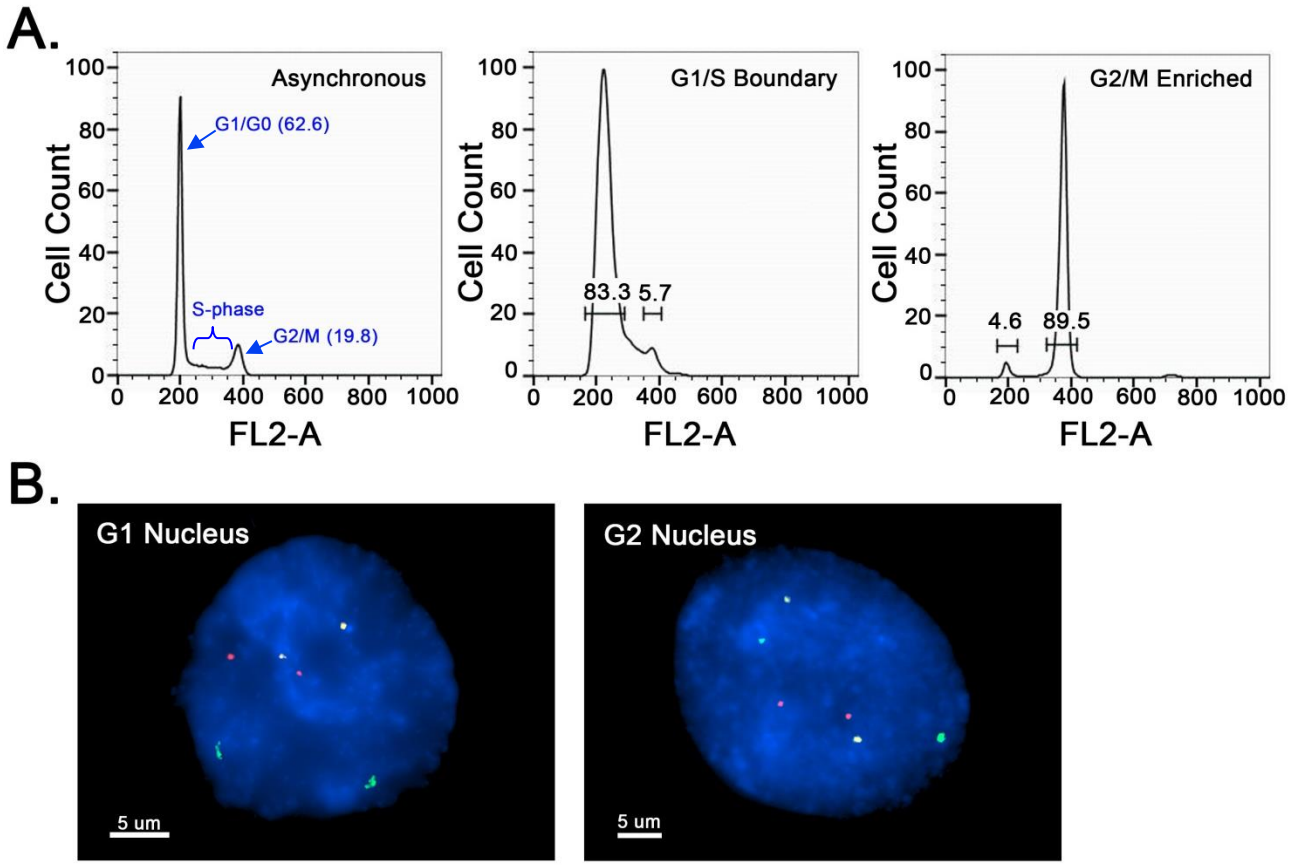

Figure 10. DNA Replication Does Not Impact Enumeration of FISH Probes.

(A) HCT116 cell cycle analysis, depicting the enrichment at the G_1/S-phase boundary (***A***; middle panel), and the G_2/M-enrichment (***A***; right panel) following treatments relative to an asynchronous control (***A***; left panel). The DNA probe set was hybridized to the two enriched populations of HCT116 cells, and the foci were analyzed in 25 nuclei for each population. ***(B)*** 3D projections of a representative nucleus from each population. FISH signals for chromosomes 8 (red), 11 (green) and 17 (yellow) remained compact in the G2 nuclei, confirming DNA replication will not impact chromosome enumeration

Table 9. DNA Probe Set Enumeration Pre and Post S-Phase.

G_1 Nucleus	Foci Count			G_2 Nucleus	Foci Count		
	Chr 8	Chr 11	Chr 17		Chr 8	Chr 11	Chr 17
1	2	2	2	1	2	2	2
2	2	2	2	2	2	2	2
3	2	2	2	3	2	2	2
4	2	2	2	4	2	2	2
5	2	2	2	5	2	2	2
6	2	2	2	6	2	2	2
7	2	2	2	7	2	2	2
8	2	2	2	8	2	2	2
9	2	2	2	9	2	2	2
10	2	2	2	10	2	2	2
11	2	2	2	11	2	2	2
12	2	2	2	12	2	2	2
13	2	2	2	13	2	2	2
14	2	2	2	14	2	2	2
15	2	2	2	15	2	2	2
16	2	2	2	16	2	2	2
17	2	2	2	17	2	2	2
18	2	2	2	18	2	2	2
19	2	2	2	19*	4	4	4
20	2	2	2	20	2	2	2
21	2	2	2	21	2	2	2
22	2	2	2	22	2	2	2
23	2	2	2	23	2	2	2
24	2	2	2	24	2	2	2
25	2	2	2	25	2	2	2

*Abnormal foci counts not due to DNA replication.

Interestingly however, one G_2 nucleus (number 19) contained four distinct FISH signals for each of the three chromosomes evaluated. DNA replication does not appear to be the underlying cause of the extra two copies of each FISH signal as the foci for any given chromosome evaluated were not in close proximity within the nucleus, which is what would be expected if the extra signals were from the replicated sister chromatids. Thus, the results presented above strongly suggest that DNA replication did not impact the ability to enumerate the FISH probes.

4.2 AIM 2: TO EVALUATE CHROMOSOMAL INSTABILITY IN INTERVAL AND SPORADIC COLORECTAL CANCER TUMOR SAMPLES

4.2.1 Extrapolating and Optimizing the FISH-Based Protocol for Colorectal Cancer Tissue Sections

Having validated the specificity of the DNA probe set in cell lines (Aim 1), it was now necessary to optimize the approach for CRC tissue sections prior to extrapolating this work to the mini-TMAs. Overall the general FISH protocol for the FFPE samples was similar to that performed on the cell lines above. However, slight modifications were performed to enhance signal strength and detection, which included an enhanced pepsin treatment, namely an increased concentration and extended incubation time, due to the 3D tissue structure and architecture compared to the 2D orientation of the cell lines. Pepsin treatment was optimized for CRC tissue by administering different pepsin incubation treatments (e.g. 30 min, 60 min) to two CRC tissue sections and analyzing the ability to identify and enumerate the FISH foci. Similar to the pepsin optimization in the FISH protocol for cell lines, the DNA probe set foci were more punctate with the shorter pepsin treatment, and also required extended acquisition times which resulted in greater tissue auto-fluorescence, particularly within the green fluorescent channel. A 45 min pepsin treatment, at a working concentration of 0.5 mg/mL, was determined to be the optimal incubation time.

Based on the 3D size and shapes of interphase nuclei within the CRC tumor samples and the necessity to only analyze intact nuclei, tissues samples were sectioned at 5 µm intervals. This thickness was confirmed to be adequate within the CRC tumor samples and was employed in all subsequent analyses. By generating nuclear surface

renderings within each image, complete/intact nuclei could be easily identified and evaluated within the mini-TMA sections (Fig. 11). Subsequently, only complete nuclei were enumerated in the mini-TMA sections. Foci were also identified throughout the 3D CRC tumor sample, indicating the section thickness was appropriate to ensure adequate probe hybridization throughout the section (i.e. in 3 dimensions).

Next, it was imperative to identify the optimal optical section thickness (i.e. z-series step size) to best capture the FISH signals for enumerations purposes. As with the cell lines, the optical section thickness was compared between images acquired from the tissue sections at 200 nm and 400 nm intervals. The 400 nm images was determined to be optimal as the FISH foci enumeration results were identical between images acquired at 200 nm and 400 nm intervals and the reduced number of optical sections collected had the added benefits of minimizing photobleaching and image processing (deconvolution) time.

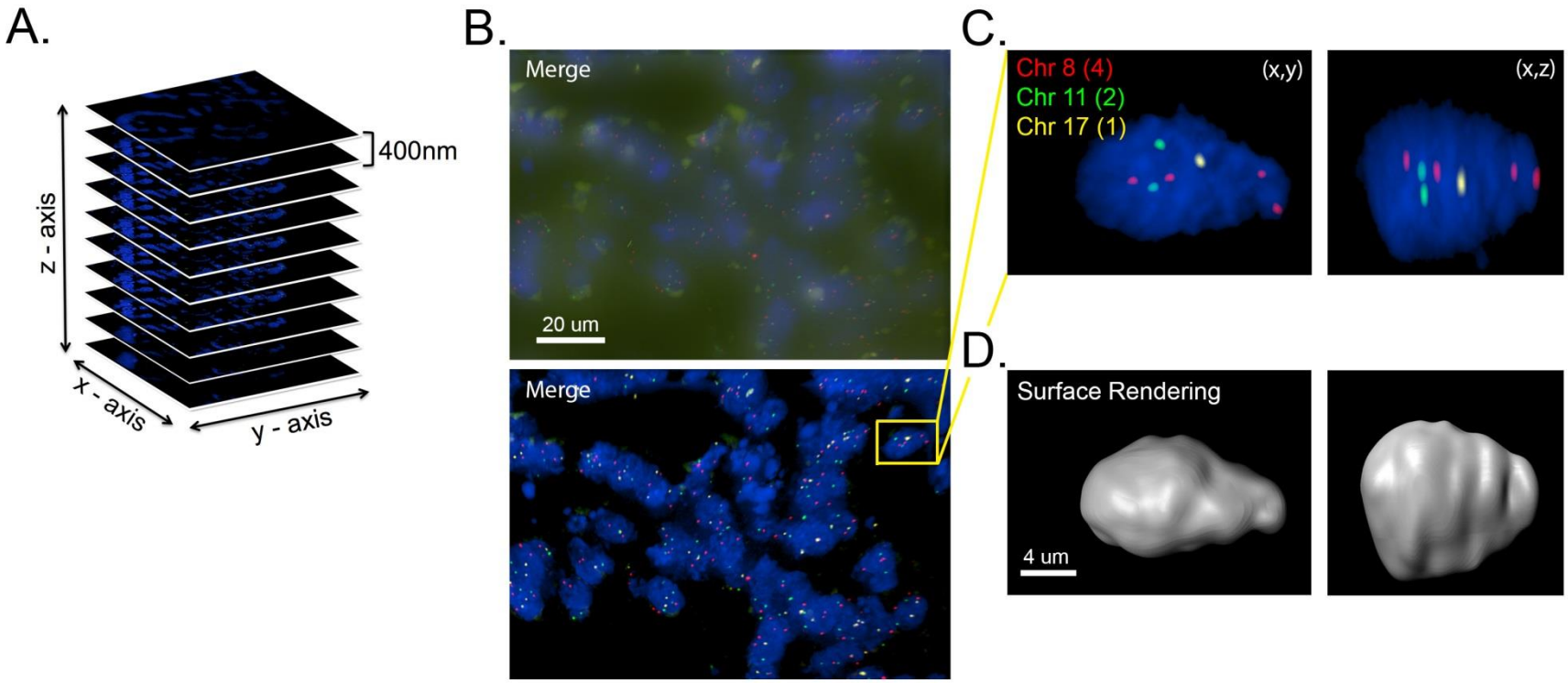

Figure 11. Chromosome Enumeration Protocol; Image Acquisition and Analysis.
Schematic depicting the steps involved in the chromosome enumeration process after the CRC samples were hybridized with the DNA probe set. ***(A)*** A schematic depicting 3D image acquisition at 400 nm intervals. ***(B)*** 3D projection of a representative image, collected from a single TMA core, before (top) and after (bottom) image deconvolution. ***(C)*** A representative high magnification of a single complete nucleus exhibiting aneusomy and is suggestive of a CIN phenotype. ***(D)*** A surface rendering generated from the DAPI channel of the nucleus in ***C***.

4.2.2 Assessing the Stability of the DNA Probe Set Fluorescent Signals

Although the mini-TMAs were specifically designed to house a manageable number of samples (20-25 samples in duplicate) for FISH labeling and image acquisition, the 3D image acquisition process was extensive and required a significant amount of time to complete (~2-weeks per core). Consequently, even after storing samples in the dark and at -20°C, it was critical to address concerns that the fluorescent FISH signals may diminish over time.

Four random cores within TMA 24B were identified and subjected to a second round of imaging and analyzed to determine if signal decay was significant over time. The second round of imaging occurred ~7 months after the initial image acquisition and was performed within the identical regions previously imaged. This way direct comparisons could be made for each chromosome evaluated on a nucleus-by-nucleus basis. Images were acquired in an analogous fashion to the original round, subjected to image deconvolution and the foci were enumerated from ~25 nuclei from each core (N=95 total nuclei). The DAPI channel was used to identify the identical regions originally imaged for image registration purposes, so that FISH signal comparisons could be performed. It should be noted that for the second round of imaging, longer exposure times were required and is likely attributed to a minimal amount of photo-bleaching that occurred during the initial image acquisition. Following image deconvolution the FISH foci of all three chromosomes remained clearly identifiable (Fig. 12) and were easily enumerated. Interestingly, these results demonstrate that 98.3% of the original foci were identified and reproducibly enumerated in the 95 nuclei enumerated (Table 10).

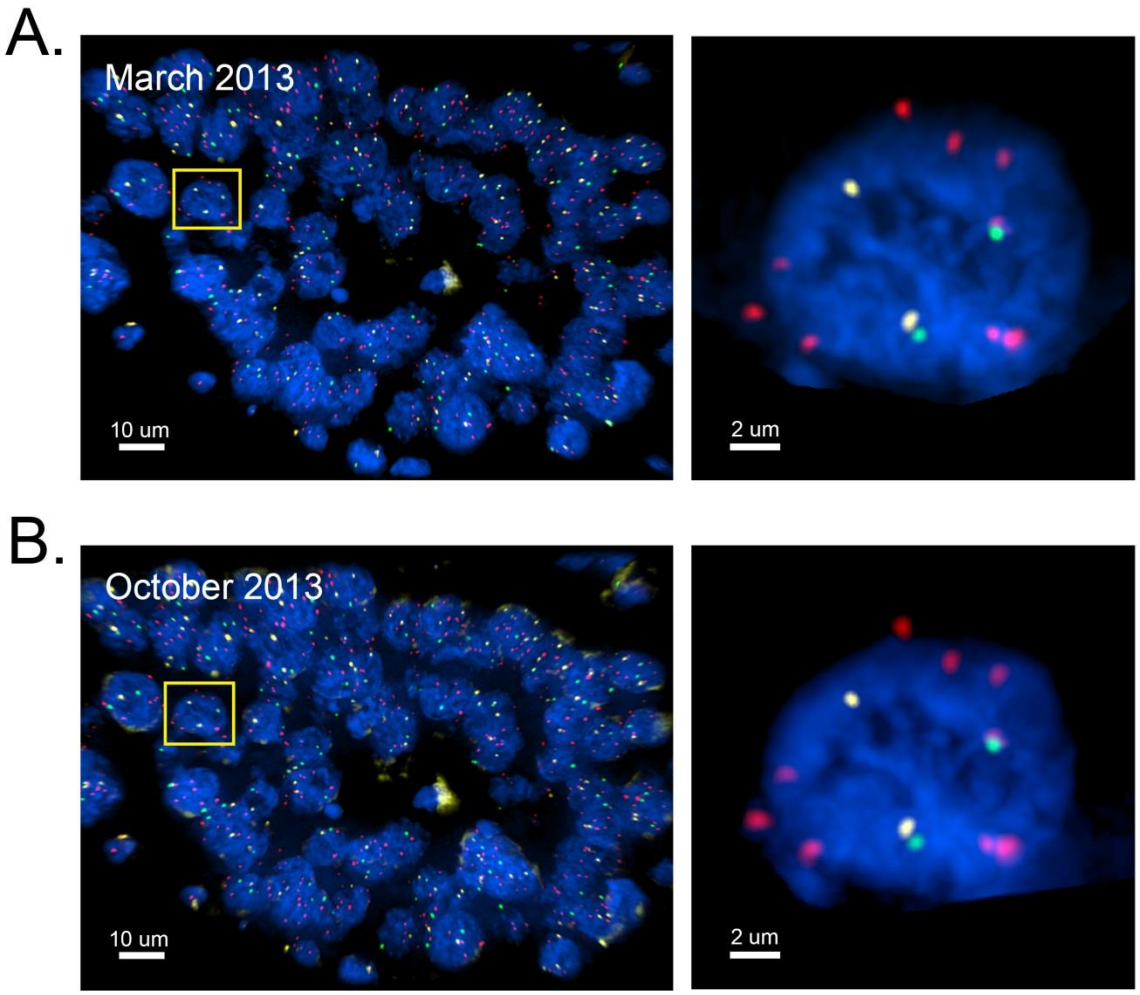

Figure 12. FISH Foci Signals are Maintained and Persist Over Time.
(A) A 3D image projection of a representative image initially collected from TMA 24B. The yellow bounding box within the low-resolution field image identifies the nucleus presented at a higher magnification on the right. ***(B)*** A 3D image projection of the identical region of TMA 24B acquired 7 months after the initial image acquisition. Note that similar numbers of FISH signals for each of the 3 chromosomes evaluated in the first image acquisition are maintained within the second image acquisition.

Table 10. Foci Enumeration Results Following 7 Months of Storage at -20°C.

		Enumeration Results (7 Mth)						
Core	**# of Nuclei**	**Chr 8**		**Chr 11**		**Chr 17**		**Total**
4B	22	66	(66)	36	(35)	35	(34)	
5B	25	102	(102)	33	(32)	35	(35)	
17B	24	47	(45)	40	(38)	31	(31)	
18B	24	51	(51)	52	(49)	50	(50)	
Initial Foci Count	95		266		161		151	578
% Identically Scored			99.2		95.7		99.3	98.3%

4.2.3 Applying the Chromosome Enumeration Protocol to the Manitoba Cohort of Interval and Sporadic Colorectal Cancers

The Manitoban cohort of Interval and sporadic CRCs was composed of 21 Interval and 46 sporadic (control) CRC samples. The CRC tumors were cored in duplicate and arrayed across three mini-TMAs, as outlined in Section 3.5. The overall goal of this study was to enumerate 100 nuclei per core, and thus 200 nuclei per sample. In 10 cores, circumstances involving high tissue auto-fluorescence or low cellularity resulted in less than 100 nuclei being enumerated. However, since the samples were cored in duplicate, it was possible to enumerate 100 nuclei in the duplicate core for all 10 samples. Nevertheless, one sample from TMA 24C (Sample 59) could not be enumerated in either core due to extensive tissue auto-fluorescence. The remaining 122 total cores were completely enumerated. Finally, it should be noted that only once the enumeration process was complete were the study samples unmasked and their status (i.e. Interval vs. sporadic) or various features (e.g. age, sex, tumor location, etc.) revealed (Tables 6-8, pg. 57-59).

4.2.4 Description of the Manitoba Colorectal Cancer Cohort

As indicated above, the details of the CRC samples were only made available following the completion of the study and are presented in Tables 6-8 (pg. 57-59). The sex, age and tumor location distributions were well-matched between the Interval and sporadic CRC samples (Fig. 13) as the sporadic CRC cohort was 60% male (27 males, 18 females), and similarly the Interval CRC cohort was 61.9% male (13 males, 8 females).

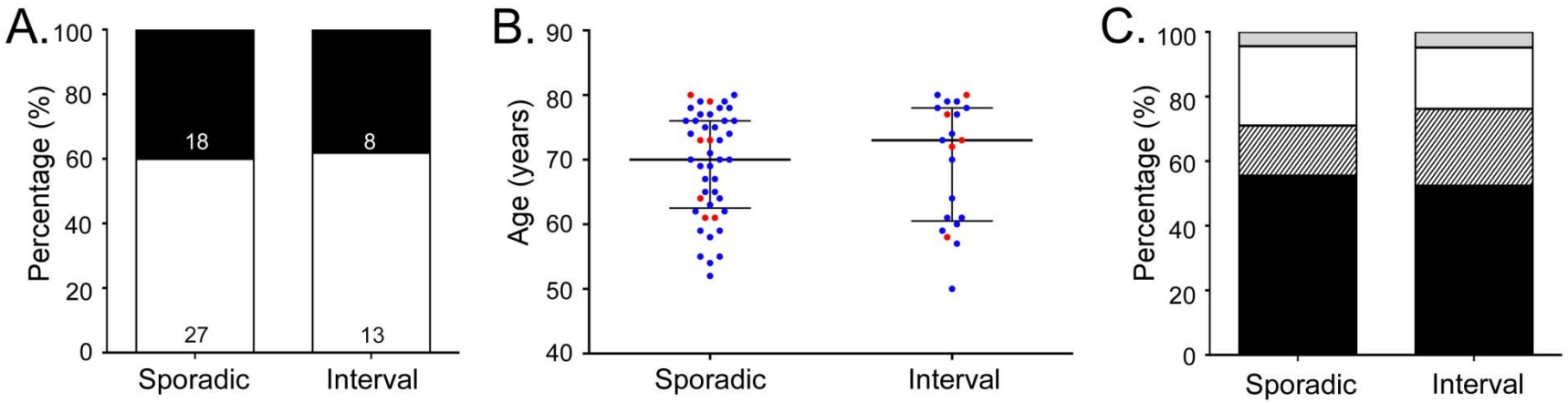

Figure 13. Properties of the Manitoban Cohort of Interval and Sporadic Colorectal Cancers.

The CRC cohort consisted of 21 Interval and 45 sporadic CRC samples. ***(A)*** Sex distributions are nearly identical in sporadic and Interval CRC samples; approximately 60% male (white) and 40% female (black). ***(B)*** A scatter plot depicting patient age with median, 25th and 75th percentiles. Patients with tumor samples exhibiting CIN are represented in blue. ***(C)*** Similar location of tumor presentation (ascending [black], transverse/descending[cross-hatch], rectosigmoid [white], unidentified [grey]) within the sporadic and Interval CRC samples.

As for age distributions, the 25th, median and 75th percentiles were 62.5, 70, and 76 years for the sporadic CRCs while the Interval CRCs displayed a slightly wider, yet comparable, distribution of 60.5, 73, and 78 years. Finally, the sporadic CRC samples were predominately located within the ascending colon (55.5%), followed by the transverse/descending (15.5%) and rectosigmoid (24.6%). The location was not specified for two of the sporadic CRC samples (Samples 15 and 54). Similar distributions were also observed within the Interval CRC samples; ascending colon (52.4%), transverse/descending (23.8%), and rectosigmoid (19.0%). The location for one Interval CRC (Sample 19) was not specified.

4.2.5 Determining the Prevalence of Chromosomal Instability in Interval Colorectal Cancers

To determine the prevalence of CIN within each tumor a "CIN score" was developed and calculated so that comparison between Interval CRCs and sporadic CRCs could be made (Section 3.7.1, pg. 66). Briefly, the CIN score is defined as the sample's average chromosome gain and/or loss per nucleus. Each CIN score was calculated by summing the aberrations (i.e. chromosome gains and losses) in chromosomes 8, 11 and 17 within a sample and normalizing it to the total number of nuclei enumerated within each given core. An example of such calculations is presented in Table 11. The aberrations for each CRC sample enumerated and the corresponding CIN scores are presented in Appendix B (Supplementary Table 1, pg. 125).

Table 11. Calculating the CIN Score for Sample 19.

Core	N	Chr 8 Gains	Chr 8 Losses	Chr 11 Gains	Chr 11 Losses	Chr 17 Gains	Chr 17 Losses	Total Aberrations
A	50	9	12	20	19	15	13	88
B	100	77	19	45	30	61	19	251

CIN Score = total aberrations / total nuclei enumerated
= (88 + 251) / (50 + 100)
= 2.26

N = number of nuclei enumerated in core.

Overall the CIN scores were comparable but slightly elevated within the Interval CRCs relative to the control sporadic CRCs. They ranged from 0.9 – 3.7 in the sporadic CRC samples and 1.2 – 4.6 in the Interval CRC samples (Fig. 14A). Plotting the CIN scores within a scatterplot revealed a slightly larger distribution within the interval CRC samples compared to the sporadic control CRC samples. While there appeared to be variation within the minimum and maximum CIN scores between the two cohorts, 1.2, 4.6 for Interval compared to 0.9, 3.7 for sporadic, the median CIN scores was very similar between the Interval and sporadic CRC samples, and calculated to be 2.60, and 2.50 respectively.

In this study, the sporadic (control) CRCs were used for comparison purposes to determine if the prevalence of CIN differed within the Interval CRCs. Recall it is generally accepted that ~85% of sporadic CRCs possess a CIN phenotype[51] (Section 1.5.1, pg. 19). Thus, we identified that a CIN score value of 2.0 represents 85% of the control sporadic cases and set this as the threshold above which defines a CIN phenotype.

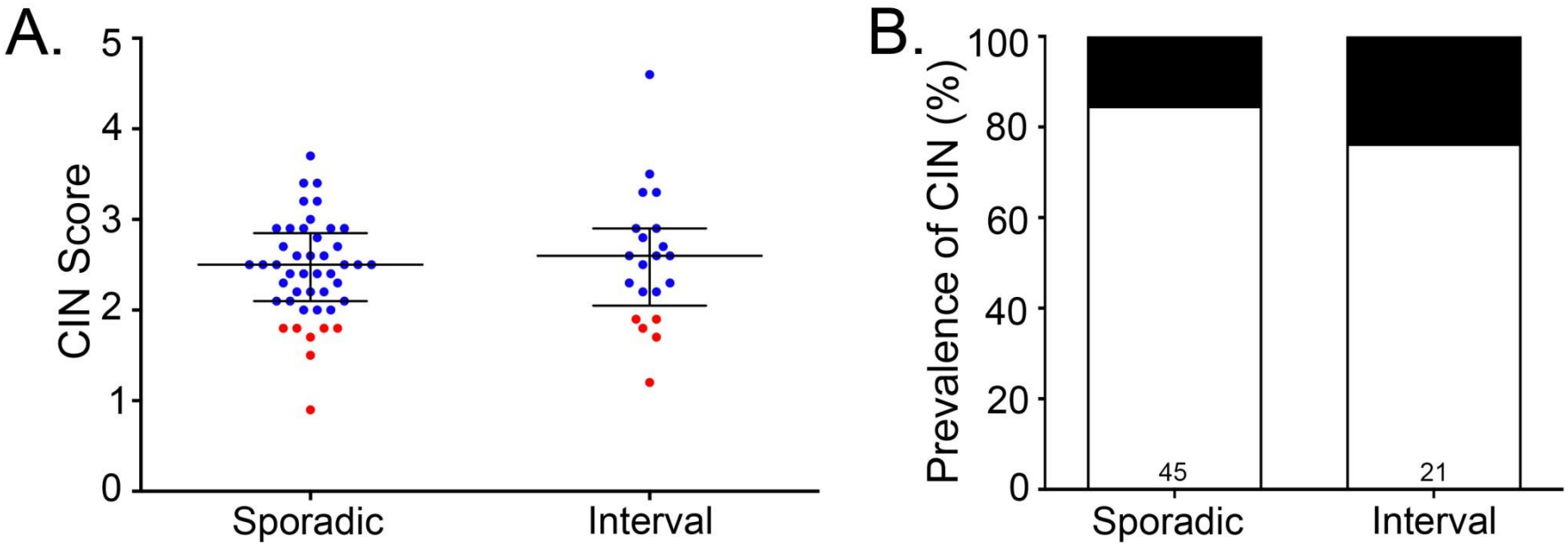

Figure 14. Sporadic and Interval Colorectal Cancers Display Similar Frequencies of Chromosomal Instability.
Chromosomes 8, 11 and 17 were enumerated in sporadic and Interval CRC samples. ***(A)*** Scatterplot depicting the CIN scores (average chromosome gain and loss per nucleus) for each sample, with median, 25th and 75th percentiles. 85% of sporadic and 76% of Interval CRCs had a CIN score ≥ 2 (Blue), indicating the sample exhibits a CIN phenotype. The remaining samples displayed more subtle chromosome gains/losses resulting in a CIN score < 2 (Red). ***(B)*** The prevalence of CIN (white) in Interval (n=21) compared to sporadic (n=45) CRC samples, as determined by chromosome enumeration.

For example, within the sporadic CRC samples enumerated in this study, ~85% (38/45) had a CIN score ≥ 2.0 and thus a CIN score of ≥ 2.0 was used to define those samples exhibiting a CIN phenotype. Out of the 21 Interval CRC samples evaluated, 16 (76.2%) had a CIN score ≥ 2 and were considered to exhibit a CIN phenotype (Fig. 14B). Thus, the prevalence of CIN is slightly lower within the Interval CRCs relative to the sporadic CRC samples.

The CIN scores were further scrutinized by determining the average chromosome gains per nucleus and the average chromosome losses per nucleus for each of the three chromosomes evaluated (Appendix B, pg. 125). In addition to the similar CIN score distributions, the Interval and sporadic CRCs also exhibited very similar chromosomal gains and losses, as indicated by the mean, 25th and 75th percentiles in the chromosome 8, 11 and 17 gain and loss scatterplots (Fig. 15). The only exception was the difference observed between the Interval and sporadic CRC samples with regards to chromosome 17. The chromosome 17 enumeration results revealed a greater distribution in gains within the Interval CRC cohort (Fig. 15C). This is particularly evident by the variation observed between the 75th percentile and maximum CIN scores between the two cohorts compared to the very similar median gain score values. The differences between the median, 75th percentile, and maximum chromosome 17 gain scores between the Interval and sporadic CRC samples were 0.01, 0.25, and 0.27 respectively.

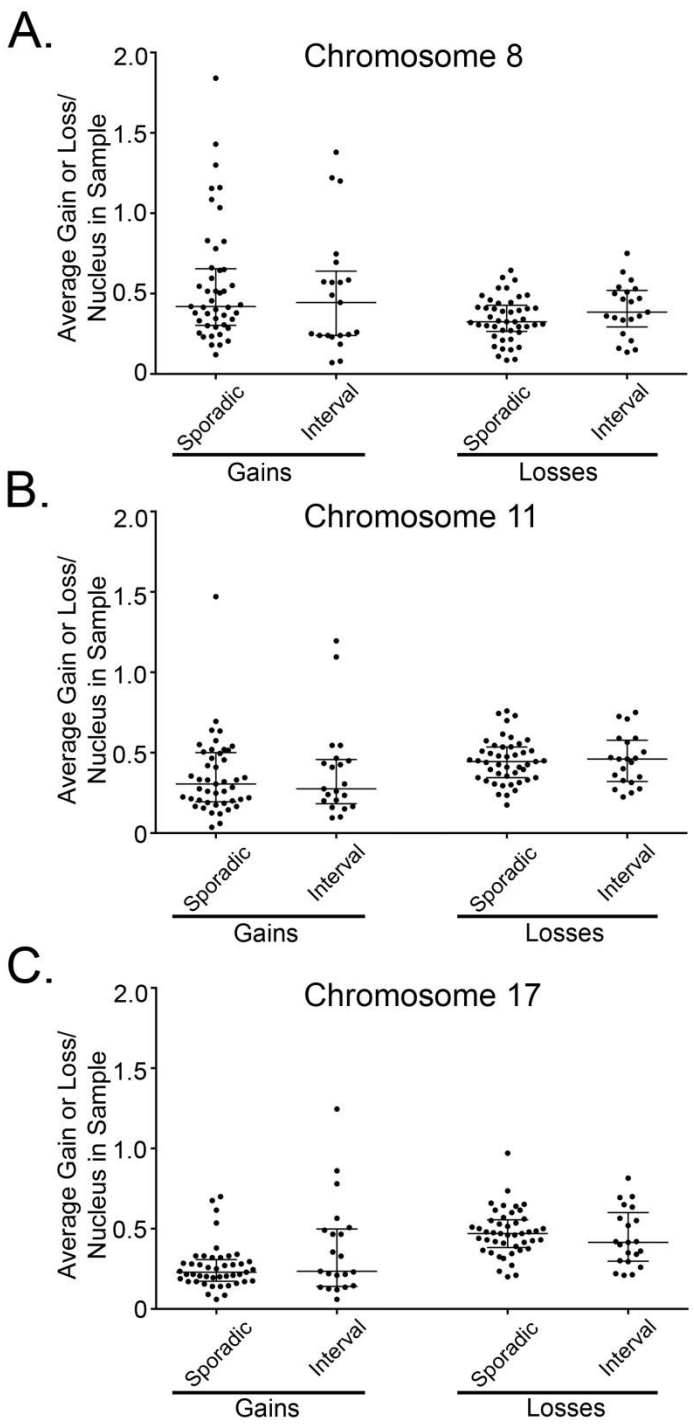

Figure 15. Comparing the Gains and Losses in Chromosome Numbers in Interval and Sporadic Colorectal Cancer Tumors.

Scatter plots depicting the average gain and loss of chromosome 8 ***(A)***, 11 ***(B)*** and 17 ***(C)***, in the Interval and Sporadic CRC samples. The median, 25th and 75th percentiles are indicated and are comparable between the Interval and Sporadic CRCs for all chromosome gains and losses, except for that of chromosome 17, which shows a greater distribution for gains in the Interval CRC samples.

Further examination of the chromosome gain and loss distributions revealed that chromosome 8 was more frequently gained while chromosomes 11 and 17 were more frequently lost. This was evident by the median gain compared to median loss scores for each of the chromosomes. For chromosome 8 the median gain score was greater than that of the median loss score for both the Interval (0.45 [gain], 0.39 [loss]) and the sporadic CRC samples (0.42 [gain], 0.33 [loss]). This was opposite to that of chromosomes 11 and 17, which had greater median loss scores for both the Interval (11; 0.28 [gain], 0.46 [loss], 17; 0.24 [gain], 0.42 [loss]) and sporadic CRC samples (11; 0.31 [gain], 0.45 [loss], 17; 0.30 [gain], 0.47 [loss]).

4.2.6 Unique Chromosomal Instability Profiles in Interval and Sporadic Colorectal Cancers

The definition of CIN is an increase in the rate of which whole chromosomes or large parts thereof are *gained* or *lost* (Section 1.5.1, pg. 19) and therefore the calculation of the CIN score included both chromosome *gains* and *losses*. While the largest CIN scores were attributed to significant chromosome gains, a few samples displayed substantial chromosomal losses with a relatively fewer numbers of chromosome gains. Figure 16 displays two Interval CRC samples that have CIN scores > 2; one sample with a large increase in chromosome foci and the other with a substantial decrease in chromosome foci. These extreme CIN phenotypes were observed periodically in both the Interval and the sporadic CRC samples.

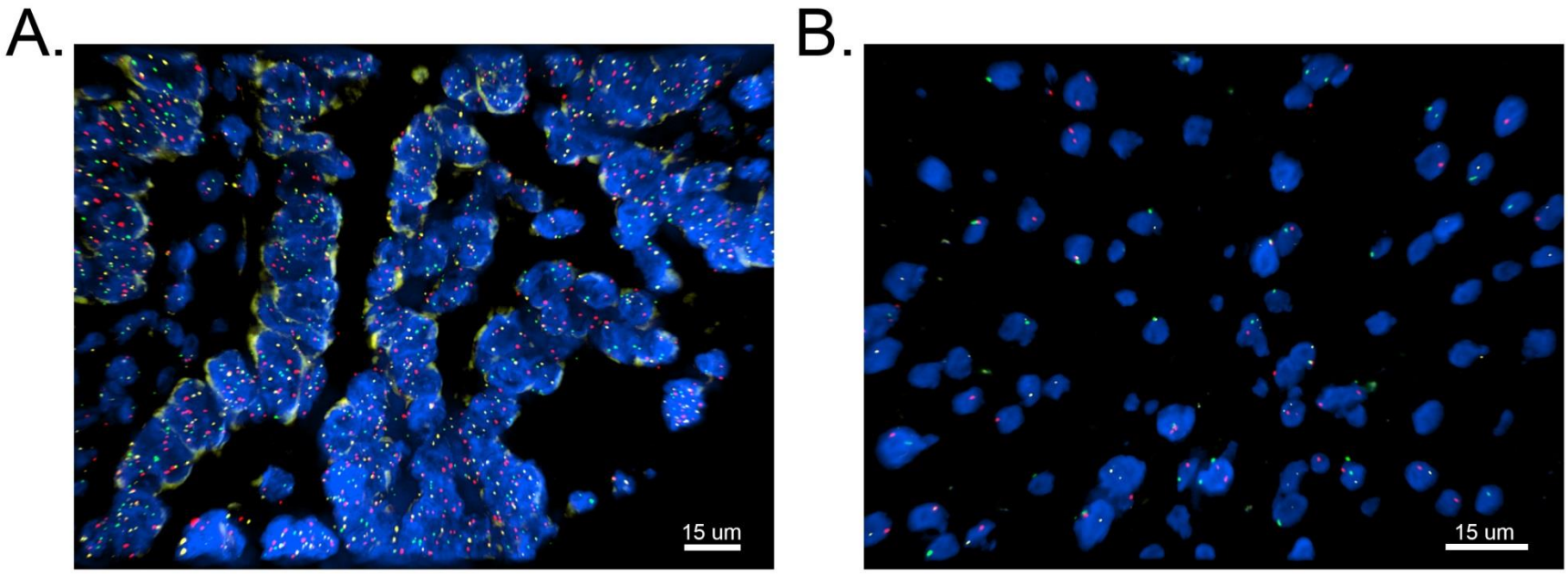

Figure 16. Interval Colorectal Cancers Exhibiting Extreme Chromosomal Instability Phenotypes.
Representative 3D projections of two different Interval CRCs both displaying aneusomy of chromosomes 8 (red), 11 (green) and 17 (yellow). ***(A)*** An Interval CRC sample with a CIN score of 4.6, exhibits a visually striking increase in chromosome foci, while another Interval sample exhibiting CIN ***(B)*** (CIN score 2.6), displays extreme loss in chromosome foci.

Periodically, chromosome 8 or 11 was more substantially gained in a sample over the other two chromosomes enumerated. The samples that showed visually striking increases in chromosome 8 or 11 were also aneuploid for the other two chromosomes enumerated. Figure 17 illustrates an Interval CRC sample with a significant increase in chromosome 8 accompanied with decreases in chromosomes 11 and 17. Figure 18 on the other hand depicts a sporadic CRC with an extensive increase in chromosome 11. In this example however, compared to Figure 17 the other chromosomes (i.e. 8, 17) also exhibited chromosomal gains, making the differences between the three chromosomes slightly less visually apparent. These unique CIN profiles were not restricted to the Interval CRC samples, and were also observed periodically within the sporadic CRC samples. As mentioned previously, it was rare to find a sample with increased chromosome 17 gains, and in fact, there was not a single instance where chromosome 17 was preferentially gained over chromosome 8 and 11. These results suggest that there may be particular genes encoded within chromosomes 8 and 11, that when these chromosomes are gained, confer a selective advantage for the cell, and vice versa for chromosome 17.

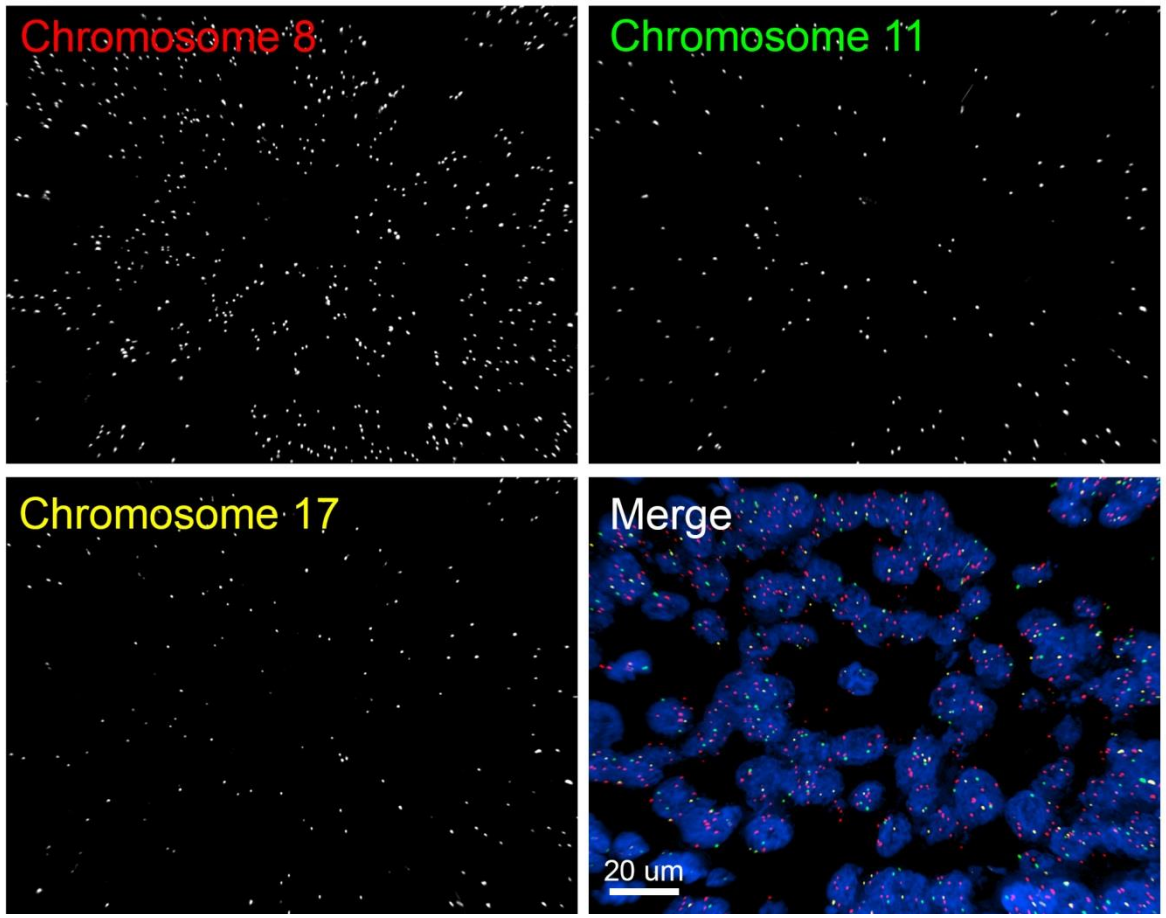

Figure 17. Interval Colorectal Cancer with Selective Gains of Chromosome 8.
Representative 3D projections of an Interval CRC aneuploid for chromosomes 8, 11 and 17, however with increased gain in chromosome 8 (CIN scores of 2.4, 0.9, and 0.9 respectively). There is likely a selective growth advantage for those cells harbouring multiple copies of chromosome 8.

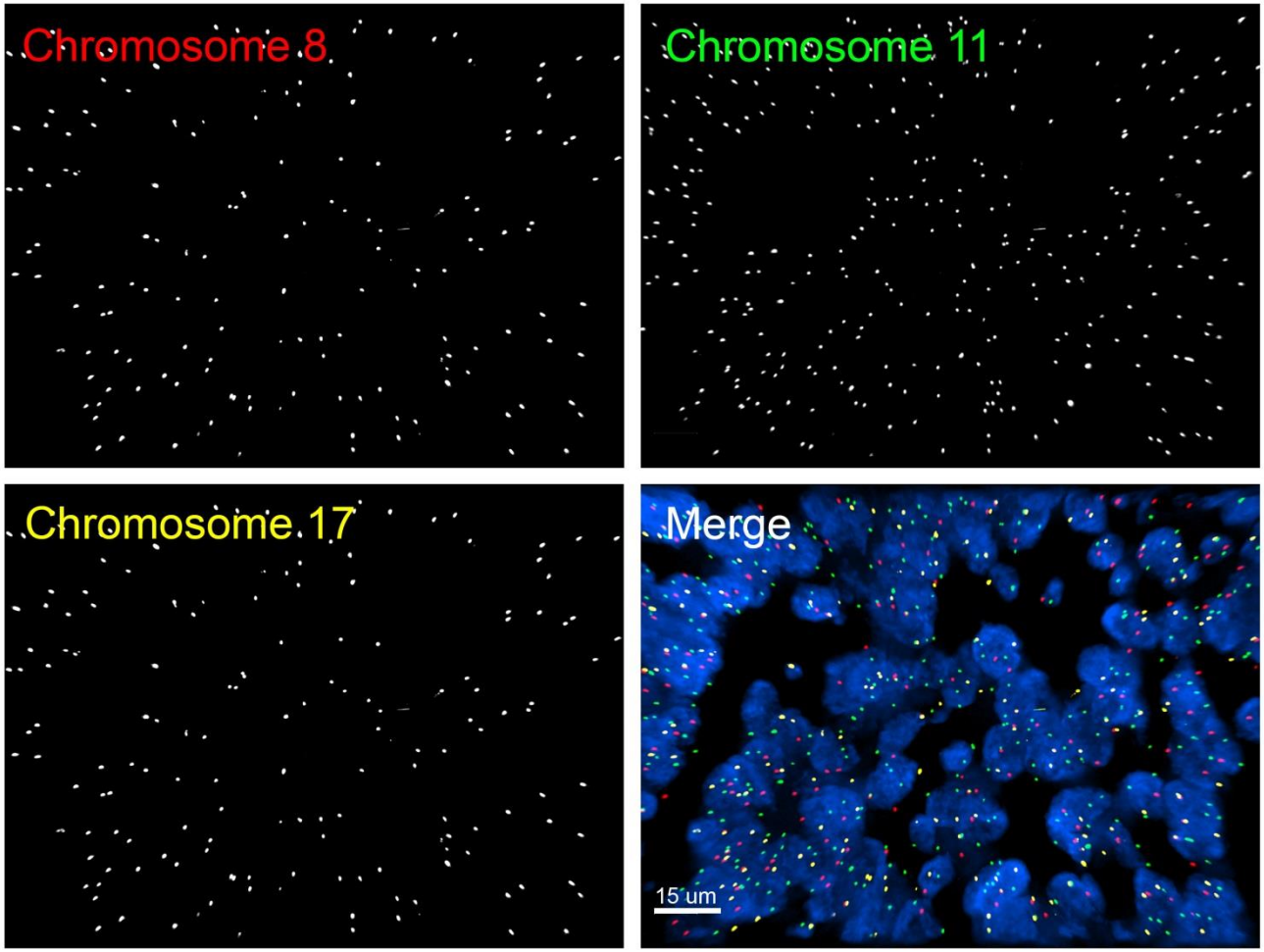

Figure 18. Sporadic Colorectal Cancer with Selective Gains of Chromosome 11.
Representative 3D projections of a sporadic CRC aneuploid for chromosomes 8, 11, and 17, however with increased gain in chromosome 11 foci (CIN Scores of 0.6, 1.8, 0.7 for chromosomes 8, 11 and 17 respectively).

CHAPTER 5: CONCLUSION

We have established and optimized a FISH-based chromosome enumeration protocol to evaluate the presence of aneusomy (i.e. a metric for CIN). This approach utilizes a DNA FISH probe set to specifically evaluate gains or losses of three specific chromosomes (8, 11 and 17) within interphase nuclei. We have confirmed the specificity of the FISH probes and validated the capacity of the DNA probe set to accurately visually display aneusomy within tissue culture model systems that is independent of cell cycle stage.

The central goal of this study was to extrapolate the FISH-based approach and evaluate CIN within a Manitoba cohort comprised of both Interval and control sporadic CRCs. Accordingly, a Manitoba cohort of 21 Interval and 46 sporadic CRC samples was established and evaluated for the prevalence of CIN. We assembled three mini-TMAs that were purposefully designed to contain ~25 CRC samples cored in duplicate. Each TMA was processed and imaged and 100 nuclei/core were evaluated. A CIN score metric (i.e. the average chromosomal gains and losses per nucleus) was developed and was calculated for each sample so that comparisons between the Interval and control sporadic CRC samples could be made. We set the threshold for the CIN score to be 2.0 based on the 15^{th} percentile of the sporadic CRC CIN scores (i.e. 85% of the sporadic CRCs had a CIN score $\geq$ 2.0). After evaluating the Interval CRC CIN scores, our data indicate that ~75% of Interval CRCs within the Manitoba cohort exhibit a CIN phenotype.

Finally, we analyzed and compared the enumeration results, including the CIN scores, and the individual chromosomal gains and losses, between the Interval and sporadic CRC samples. The CIN scores displayed similar distributions within the

Interval (1.2 – 4.6) and sporadic (0.9 – 3.7) CRCs cohorts. Additionally, similar gains and losses were observed between the Interval and sporadic CRCs for both chromosomes 8 and 11. Gains of chromosome 17 revealed a greater distribution within the Interval (0.06 – 1.25) compared to the sporadic (0.06 – 0.70) CRC samples, however chromosome 17 was lost more frequently than gained in both the sporadic and Interval CRCs. Similarly, chromosome 11 was also more frequently lost, however the opposite (more gains) was observed for chromosome 8. These results suggest that a selective growth advantage for cells harbouring increased copy number of chromosome 8, and a loss of chromosome 11 and 17.

The results of this study indicate that 75% of Interval CRCs exhibit a CIN phenotype, making CIN the most prevalent contributor to genomic instability in Interval CRCs. These results expand our limited understanding of the molecular biology of Interval CRCs. Only once we can better characterize the tumorigenic pathways through which Interval CRCs develop, will we be able to tailor screening strategies and treatment options to specifically identify and combat this aggressive subset of sporadic CRC.

CHAPTER 6: DISCUSSION

6.1 DEVELOPING THE FLUORESCENCE IN SITU HYBRIDIZATION (FISH)-BASED CHROMOSOME ENUMERATION PROTOCOL

In this study, we established and employed a FISH-based approach to evaluate CIN within a Manitoban cohort comprised of Interval and sporadic CRCs. Three specific chromosomes (i.e. chromosomes 8, 11 and 17) were selected for the analysis using the DNA FISH probe set. These three chromosomes were specifically selected for evaluation purposes, as genes encoded within these chromosomes have been found altered in CRCs and/or are thought to play a role in CRC tumorigenesis. Examples of such genes include the oncogenes *MYC* and *FGFR1* (chromosome 8) and tumor suppressor genes *ATM*, *MRE11A* (chromosome 11) and *p53* (chromosome 17)[57, 59, 60, 149, 150]. Vogelstein *et al*[51] identified chromosome 17 to exhibit more variation between passages of aneuploid CRC derived cell lines over 8 other chromosomes analysed (1, 3, 4, 7, 8, 15, 16, 18). Additionally, these FISH probes were selected based on the pericentromeric location of hybridization, as the centromere is an essential component of all chromosomes and a gain/loss from the expected 2 foci/chromosome/nucleus is indicative of aneusomy (i.e. the metric for CIN).

The FISH-based chromosome enumeration protocol was validated (i.e. probe specificity, enumeration in diploid verses aneuploid cells) through tissue culture model systems and further extrapolated to CRC tissue sections (Section 4.1, pg. 67). To address the possibility that DNA replication may impact the ability to accurately enumerate the DNA probe set, we evaluated the probes in cells in pre- and post-S-phase. Due to the pericentromeric location of the FISH probes, and the fact that sister chromatid cohesion is

maintained at the centromere throughout DNA replication up to the metaphase to anaphase transition[151], it was expected (and confirmed) that the foci would remain punctate and of correct quantity within the G_2 cells. While it is also expected that these foci would be larger or of greater intensity, for the purpose of our study only shape and copy number were assessed. Using CRC tissue sections we were able to optimize the FISH protocol, as well as the CRC tissue section thickness and optical section thickness for 3D image acquisition, prior to extrapolating to the mini-TMAs (Section 4.2.1, pg. 78).

6.2 CHROMOSOMAL INSTABILITY ANALYSIS IN INTERVAL COLORECTAL CANCERS

6.2.1 Examining the Molecular Biology of Interval Colorectal Cancers

Prior to our study, only a single cohort of Interval CRCs had been studied with regards to genomic mutations and the tumorigenic pathways MSI[16] and CIMP[15], and CIN had yet to be investigated. The previous studies provided data indicating that Interval CRCs possessed an altered molecular biology compared to sporadic CRC. They showed a 1.5- and 3-fold increase in the prevalence of CIMP and MSI respectively, within the Interval CRCs, and suggested greater prevalence of these genomic instabilities may contribute to an increased rate of tumor progression. An alternative explanation for a rapid CRC development, may be the result of a synergistic growth advantage conferred by all three pathways, namely MSI, CIMP and CIN, and therefore we aimed to identify the prevalence of CIN within the Interval CRCs.

6.2.2 Identifying the Prevalence of Chromosomal Instability in Interval Colorectal Cancers

Through our analysis of the Manitoba cohort of Interval and sporadic CRCs, we have identified the prevalence of CIN within Interval CRCs. Specifically, 75% of Interval CRCs were identified as exhibiting a CIN phenotype, a slightly lower prevalence compared to 85% of sporadic CRCs (Section 4.2.5, pg. 86). These findings provide novel insight into the relationship between the genomic instability pathways in Interval CRCs. To clarify, within sporadic CRC, MSI and CIN are thought to be mutually exclusive pathways[51, 52] and it has been proposed that CIMP may contribute to the development of either MSI or CIN[111]. If MSI and CIN were to remain mutually exclusive in the Interval CRCs then we would expect these tumors to exhibit a decreased prevalence of CIN, from 85% observed in the sporadic CRCs to 55%, due to the increased prevalence of MSI in Interval CRC identified by Sawhney *et al*[16] (i.e. form 15% in sporadic CRC to 45% in Interval CRC). However, our study identified 75% of Interval CRC to exhibit a CIN phenotype, and assuming our cohort exhibits the same prevalence of MSI, then there must be an interaction between these two pathways within the Interval CRCs. This could possibly be a synergistic interaction resulting in the increased rate of CRC tumorigenesis in this subset of sporadic CRC, and therefore additional studies are warranted to evaluate both CIN and MSI within the same cohort of Interval CRCs.

6.2.3 Identifying Precursor(s) and/or Subcategories of Interval Colorectal Cancers

It has been suggested that SSA/Ps may be the precursor lesion to Interval CRCs[15, 116, 145-147], due to the physical and molecular similarities between SSA/Ps and Interval CRCs (Section 1.9, pg. 34). However, it is important to acknowledge that a portion of Interval CRCs are attributed to false negative colonoscopies. It is possible that SSA/Ps contribute to the missed sporadic CRC portion of Interval CRCs, as the physical characteristics of SSA/Ps make them more challenging to detect by colonoscopy. Approximately 75% of SSA/Ps possesses the CIMP phenotype[115, 116, 119], and this could explain the increased prevalence of CIMP in Interval CRCs[15]. Additionally, even though CIN has not been directly analyzed in SSA/Ps, 37.5% of SSA/Ps exhibit APC loss[115], a mutation strongly associated with the CIN phenotype within CRCs[57]. Our study revealed a higher prevalence of CIN within the Interval CRCs (75%) than that within the SSA/Ps (37.5%) further suggesting that SSA/Ps may only contribute to the development of a portion of Interval CRCs. It would be advantageous to test the prevalence of both SSA/Ps and Interval CRCs in multiple cohorts/countries to determine if there is a positive correlation, and to what extent, between the two lesions.

Even though false-negative colonoscopies undoubtedly contribute to the development of Interval CRCs, aberrant pathways and players are proposed to have a more substantial role in driving the progression of Interval CRCs, as indicated by the increased prevalence of MSI and CIMP in these tumors[15, 16]. It is possible that a greater understanding of Interval CRCs, especially at the molecular level, will uncover subcategories within Interval CRCs. Such categories may include missed sporadic CRCs, SSA/Ps, synergistic growth advantage CRCs (i.e. CRC possessing all three

genomic instability phenotypes [CIN, MSI, CIMP]) and other contributors that are currently unknown. It is necessary to separate and determine the extent of missed sporadic CRCs from true Interval CRCs, and this will come as we further classify the Interval CRC sub-categories. This evidence (i.e. portion of missed sporadic CRCs and Interval CRC sub-categories) will help to conclusively identify precursor lesions, and may impact screening programs with regards to the recommended time interval between screening colonoscopies.

6.3 A MODEL INTERVAL COLORECTAL CANCER STUDY COHORT

6.3.1 Selecting the Manitoba Cohort of Interval and Sporadic Colorectal Cancers

We designed an Interval CRC study cohort to accurately compare and contrast Interval CRCs to that of control sporadic CRCs, while reducing the inclusion of missed sporadic CRCs. This cohort was specifically designed to include a variety of Interval CRCs (e.g. location, gender, age) to grasp a more complete understanding of this cancer in the general population. Before selecting the tumor samples for the study cohort, we first designed an ideal definition for Interval CRCs. Interval CRCs were classified as CRCs detected between 6-36 months after a negative colonoscopy. The theory behind this time interval, as well as the other inclusion/exclusion criteria incorporated into the definition is discussed below (Section 6.3.3). After narrowing down the selection of Interval CRCs, the samples were matched to control sporadic CRC samples, in a 1:2 ratio, by numerous factors, including patient age, sex and location of tumor presentation.

We compared our cohort of Interval and sporadic CRCs to the ones developed in Minnesota, USA for the analysis of MSI[16] and CIMP[15] (Section 1.7, pg. 28). Similar to our study cohort, their Interval CRC samples were match to the sporadic CRCs by age,

either <70 or ≥70, and collected in a 1:2 ratio. Exclusion criteria, was also similar (i.e. patients with IBD) however they did not exclude patients based on age, or prior history of colon disease/polyps (provisions incorporated into our selection criteria). Location of tumor presentation was not considered when selecting their cohort, as evident by the uneven distribution between the Interval and sporadic CRCs (i.e. 63% of Interval and only 39% of sporadic samples in their cohort were located in the proximal colon). While their cohort was larger, consisting of 63 Interval CRCs compared to our 21 Interval CRCs, we have fortunately acquired an additional ~20 Interval and ~40 sporadic CRCs, doubling our cohort size for future studies. One major advantage of our cohort was the inclusion of samples from both sexes. This was the first Interval CRC study to assess the molecular biology of both male and female Interval CRCs, as previous studies only included CRCs from male patients. Other sample information that was identified in their cohorts included race, tumor differentiation, tumor size and mucinous history, all of which could be collected for our cohort utilizing patient records, if required.

6.3.2 Limitations of the Study

CIN is defined as an increase in the rate of which whole chromosomes or large parts thereof are gained or lost, resulting in numerical and/or structural aneusomy[56] (Section 1.5.1, pg. 19). Our evaluation of CIN utilized a FISH-based chromosome enumeration protocol which is ultimately an end point analysis of aneusomy, unable to detect a "rate". Additionally, our protocol can only assess numerical aneusomy. Therefore, any CRC sample with structural aneusomy in our study would be identified as CIN-negative, unless the structural aneusomy includes duplications or deletions within the pericentromeric region of one of the three chromosomes analysed. Furthermore, it

would be difficult to identify tandem duplications within those regions. We may be able to see an increase in focal size, however the protocol would have to be modified to include focal size or intensity screens.

Our protocol only assesses the numerical aneusomy of chromosomes 8, 11, and 17, and therefore would miss any CIN-positive CRC sample that remains diploid for the three chromosomes enumerated. While increasing the number of chromosomes analyzed would increase the probability of detecting numerical aneusomy, many factors had to be considered when designing the multiplexed DNA probe set, including cost, time, compatibility of fluorophores, photobleaching, *etc*. Ultimately, the protocol was designed to enumerate three chromosomes, as this was a feasible number of fluorophores to enumerate while tripling the probability of detecting CIN within the tumor samples.

6.3.3 Developing a Universal Definition for Interval Colorectal Cancers

Currently, there is no universal definition for Interval CRCs. The time period after a negative colonoscopy in which Interval CRCs are detected ranges considerably between studies (any time after the negative colonoscopy up to10 years) (Table 1, pg. 3). This does limit our understanding of Interval CRCs, as differences in selection criteria will influence results, and impact our ability to compare findings across studies. It is therefore extremely important that a universal definition is established and adopted or the subtypes better categorised. While Interval CRCs are thought to possess an aberrant biology that contributes to a rapid tumorigenesis, a proportion of these tumors are attributed to false negative colonoscopies and potentially missed sporadic CRCs. To limit the number of these sporadic CRCs included in the definition, a time interval of 6-36 months after a negative colonoscopy may be ideal. Setting the lower limit to 6 months

will reduce the number of included sporadic CRCs that were actually suspected at the initial colonoscopy, but the definite diagnosis was delayed due to circumstances such as obtaining tissue on surgery, pathology reporting or follow-up testing. Establishing a maximum limit of 3 years ensures the inclusion of extremely rapid developing tumors, but also excludes sporadic CRCs that develop through traditional or perhaps slightly accelerated tumorigenic pathways. Additionally, a universal definition should consider the exclusion of patients with a history of diverticular disease, abdominal/pelvic surgery, prior colonoscopies with polypectomy or those with a family history of CRC such as FAP or HNPCC. Finally, as we gain greater understanding of the molecular biology of Interval CRCs, the definition for Interval CRCs should include biological and/or molecular criterion, such as specific genetic mutations or genomic instability phenotypes.

6.4 SIGNIFICANCE

Our study was the first to identify the role that CIN has on the development of Interval CRCs, and prior to our study, only MSI and CIMP had been analyzed in these aggressive CRC tumors. The high prevalence of CIN identified within our cohort (~75%), combined with the previous discoveries of increased prevalence of MSI[16] (45%) and CIMP[15] (60%), suggest there may be a synergistic effect from all three pathways contributing to a rapid tumor development within the Interval CRCs. While our current understanding about the etiological origins of Interval CRCs is limited, the pathways contributing to genomic instability within these tumors have now been identified. The ultimate goal is to understand the altered biology of Interval CRCs, in order to tailor screening strategies and treatment options to specifically identify and combat this aggressive subset of sporadic CRC. Our research has not only provided novel

information about the molecular biology of these tumors, but also identified new concepts about Interval CRC development that require further examination (see below).

6.5 FUTURE DIRECTIONS

Our data indicates that CIN is a frequently observed in Interval CRCs and that it likely contributes to the development of these tumors. We have suggested that the rapid tumor development observed within the Interval CRCs may be a result of a synergistic growth advantage conferred by all three pathways (i.e. MSI, CIMP and CIN). In order to investigate this theory, we would utilize our expanded cohort to assess the prevalence of CIN, MSI and CIMP within the same Interval and sporadic CRC samples. This research will not only be the first to look at CIN and MSI in the same cohort of CRC samples, but will also identify any interactions between these two (seemingly mutually exclusive) genomic instability pathways in addition to the contribution of CIMP, in both Interval and sporadic CRC.

In order to grasp a greater understanding of the molecular biology of Interval CRCs, we would expand the number of Interval CRC and sporadic (control) CRC samples to be included within our study cohort. After accumulating more CRC samples we would analyze these tumors for CIN by employing our FISH-based chromosome enumeration protocol.

To further examine the molecular biology of the study cohort we would test for the prevalence of MSI. Ideally, both PCR analysis of microsatellite markers as well as immunohistochemical analysis against the most commonly mutated/epigenetically silenced DNA MMR proteins (e.g. *MLH1*, *MSH2*, *PMS2* and *MSH6*) would be performed. As detailed in Section 1.5.2 (pg. 22), the PCR analysis specifically detects

expansions/contractions in highly repetitive (microsatellite) regions throughout the genome that likely arise due to defects in DNA MMR proteins. This technique is designed to detect DNA MMR protein function however immunohistochemical analysis would enhance the study by analyzing the expression and localization of the DNA MMR proteins and thus potentially identify the underlying defect accounting for the MSI phenotype. In this regard, an immunohistochemical approach could determine if epigenetic silencing of any of these MMR genes is a contributing factor in the etiology of Interval CRCs.

To complete the analysis of genomic instability pathways and their contributions to Interval CRC development, we would analyse the study cohort with respect to CIMP. Specifically, the promoter methylation status of 6 genes implicated in the pathogenesis of CRC (i.e. *MINT1*, *MINT2*, *MINT31*, *p16INK4*, *MGMT*, and *MLH1*) would be evaluated by Real Time PCR and we would subsequently classify samples as CIMP positive when methylation is observed in ≥3 genes. Utilizing these similar parameters to that of Arian *et al*[15] would allow us to compare our findings to that of the Minnesota cohort. Finally, we would evaluate the prevalence of genomic instability phenotypes, CIN, MSI and CIMP, within the Interval and sporadic CRCs. Only once all three of these classical pathways have been evaluated will we have a greater understanding of the molecular etiology of Interval CRCs, and can begin to establish novel screening and/or treatment strategies to minimize the impact of these tumors.

CHAPTER 7: REFERENCES

1. Committee, C.C.S.s.S. (ed. C.C. Society) (Toronto, ON.; 2013).

2. Society, A.C., Vol. 2013 (American Cancer Society, Inc., 2013).

3. Markowitz, S.D. & Bertagnolli, M.M. Molecular origins of cancer: Molecular basis of colorectal cancer. *N Engl J Med* **361**, 2449-2460 (2009).

4. Hewitson, P., Glasziou, P., Watson, E., Towler, B. & Irwig, L. Cochrane systematic review of colorectal cancer screening using the fecal occult blood test (hemoccult): an update. *Am J Gastroenterol* **103**, 1541-1549 (2008).

5. Elmunzer, B.J. *et al.* Effect of flexible sigmoidoscopy-based screening on incidence and mortality of colorectal cancer: a systematic review and meta-analysis of randomized controlled trials. *PLoS medicine* **9**, e1001352 (2012).

6. Meissner, H.I., Breen, N., Klabunde, C.N. & Vernon, S.W. Patterns of colorectal cancer screening uptake among men and women in the United States. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* **15**, 389-394 (2006).

7. Leddin, D. *et al.* Canadian Association of Gastroenterology and the Canadian Digestive Health Foundation: Guidelines on colon cancer screening. *Can J Gastroenterol* **18**, 93-99 (2004).

8. Ferrucci, J.T. Colonoscopy and barium enema: radiologist's response. *Gastroenterology* **112**, 294-297 (1997).

9. Pabby, A. *et al.* Analysis of colorectal cancer occurrence during surveillance colonoscopy in the dietary Polyp Prevention Trial. *Gastrointest Endosc* **61**, 385-391 (2005).

10. Leaper, M., Johnston, M.J., Barclay, M., Dobbs, B.R. & Frizelle, F.A. Reasons for failure to diagnose colorectal carcinoma at colonoscopy. *Endoscopy* **36**, 499-503 (2004).

11. Gorski, T.F., Rosen, L., Riether, R., Stasik, J. & Khubchandani, I. Colorectal cancer after surveillance colonoscopy: false-negative examination or fast growth? *Dis Colon Rectum* **42**, 877-880 (1999).

12. Hosokawa, O. *et al.* Invasive colorectal cancer detected up to 3 years after a colonoscopy negative for cancer. *Endoscopy* **35**, 506-510 (2003).

13. Robertson, D.J.e.a. Interval Cancer After Total Colonoscopy: Results from a Pooled Analysis of Eight Studies. *Gastroenterology* **134**, A-111-A-112 (2008).

14. Pohl, H. & Robertson, D.J. Colorectal cancers detected after colonoscopy frequently result from missed lesions. *Clin Gastroenterol Hepatol* **8**, 858-864 (2010).

15. Arain, M.A. *et al.* CIMP status of interval colon cancers: another piece to the puzzle. *Am J Gastroenterol* **105**, 1189-1195 (2010).

16. Sawhney, M.S. *et al.* Microsatellite instability in interval colon cancers. *Gastroenterology* **131**, 1700-1705 (2006).

17. Shaukat, A. *et al.* Is KRAS mutation associated with interval colorectal cancers? *Dig Dis Sci* **57**, 913-917 (2012).

18. Kaminski, M.F. *et al.* Quality indicators for colonoscopy and the risk of interval cancer. *N Engl J Med* **362**, 1795-1803 (2010).

19. Bressler, B. *et al.* Rates of new or missed colorectal cancers after colonoscopy and their risk factors: a population-based analysis. *Gastroenterology* **132**, 96-102 (2007).

20. Brenner, H., Chang-Claude, J., Seiler, C.M. & Hoffmeister, M. Interval cancers after negative colonoscopy: population-based case-control study. *Gut* (2011).

21. Steele, R.J. *et al.* Interval cancers in a FOBT-based colorectal cancer population screening programme: implications for stage, gender and tumour site. *Gut* **61**, 576-581 (2012).

22. Singh, H., Nugent, Z., Demers, A.A. & Bernstein, C.N. Rate and predictors of early/missed colorectal cancers after colonoscopy in Manitoba: a population-based study. *Am J Gastroenterol* **105**, 2588-2596 (2010).

23. Rex, D.K. *et al.* Relative sensitivity of colonoscopy and barium enema for detection of colorectal cancer in clinical practice. *Gastroenterology* **112**, 17-23 (1997).

24. Haseman, J.H., Lemmel, G.T., Rahmani, E.Y. & Rex, D.K. Failure of colonoscopy to detect colorectal cancer: evaluation of 47 cases in 20 hospitals. *Gastrointest Endosc* **45**, 451-455 (1997).

25. Martinez, M.E. *et al.* A pooled analysis of advanced colorectal neoplasia diagnoses after colonoscopic polypectomy. *Gastroenterology* **136**, 832-841 (2009).

26. Robertson, D.J. *et al.* Colorectal cancer in patients under close colonoscopic surveillance. *Gastroenterology* **129**, 34-41 (2005).

27. Shaukat, A., Arain, M., Thaygarajan, B., Bond, J.H. & Sawhney, M. Is BRAF mutation associated with interval colorectal cancers? *Dig Dis Sci* **55**, 2352-2356 (2010).

28. Farrar, W.D., Sawhney, M.S., Nelson, D.B., Lederle, F.A. & Bond, J.H. Colorectal cancers found after a complete colonoscopy. *Clin Gastroenterol Hepatol* **4**, 1259-1264 (2006).

29. Lieberman, D.A. *et al.* Five-year colon surveillance after screening colonoscopy. *Gastroenterology* **133**, 1077-1085 (2007).

30. Singh, H., Turner, D., Xue, L., Targownik, L.E. & Bernstein, C.N. Risk of developing colorectal cancer following a negative colonoscopy examination: evidence for a 10-year interval between colonoscopies. *JAMA* **295**, 2366-2373 (2006).

31. Bosman F. T., C.F., Hruban R. H., Theise N. D., Edn. 4th (International Agency for Research on Cancer (IARC), Lyon, France; 2010).

32. Iacopetta, B. Are there two sides to colorectal cancer? *Int J Cancer* **101**, 403-408 (2002).

33. Shamsuddin, A.M., Phelps, P.C. & Trump, B.F. Human large intestinal epithelium: light microscopy, histochemistry, and ultrastructure. *Hum Pathol* **13**, 790-803 (1982).

34. Arai, T. & Kino, I. Morphometrical and cell kinetic studies of normal human colorectal mucosa. Comparison between the proximal and the distal large intestine. *Acta pathologica japonica* **39**, 725-730 (1989).

35. Stang, A. & Kluttig, A. Etiologic insights from surface adjustment of colorectal carcinoma incidences: an analysis of the U.S. SEER data 2000-2004. *Am J Gastroenterol* **103**, 2853-2861 (2008).

36. Rex, D.K. *et al.* Quality indicators for colonoscopy. *Am J Gastroenterol* **101**, 873-885 (2006).

37. Goncalves, A.R., Ferreira, C., Marques, A., Ribeiro, L.C. & Velosa, J. Assessment of quality in screening colonoscopy for colorectal cancer. *Clin Exp Gastroenterol* **4**, 277-281 (2011).

38. Cirocco, W.C. & Rusin, L.C. Confirmation of cecal intubation during colonoscopy. *Dis Colon Rectum* **38**, 402-406 (1995).

39. Baxter, N.N. *et al.* Association of colonoscopy and death from colorectal cancer. *Ann Intern Med* **150**, 1-8 (2009).

40. Lakoff, J., Paszat, L.F., Saskin, R. & Rabeneck, L. Risk of developing proximal versus distal colorectal cancer after a negative colonoscopy: a population-based study. *Clin Gastroenterol Hepatol* **6**, 1117-1121; quiz 1064 (2008).

41. Baxter, N.N. *et al.* Analysis of administrative data finds endoscopist quality measures associated with postcolonoscopy colorectal cancer. *Gastroenterology* **140**, 65-72 (2011).

42. Harewood, G.C., Sharma, V.K. & de Garmo, P. Impact of colonoscopy preparation quality on detection of suspected colonic neoplasia. *Gastrointest Endosc* **58**, 76-79 (2003).

43. Froehlich, F., Wietlisbach, V., Gonvers, J.J., Burnand, B. & Vader, J.P. Impact of colonic cleansing on quality and diagnostic yield of colonoscopy: the European

Panel of Appropriateness of Gastrointestinal Endoscopy European multicenter study. *Gastrointest Endosc* **61**, 378-384 (2005).

44. Simmons, D.T. *et al.* Impact of endoscopist withdrawal speed on polyp yield: implications for optimal colonoscopy withdrawal time. *Alimentary pharmacology & therapeutics* **24**, 965-971 (2006).

45. Sanchez, W., Harewood, G.C. & Petersen, B.T. Evaluation of polyp detection in relation to procedure time of screening or surveillance colonoscopy. *Am J Gastroenterol* **99**, 1941-1945 (2004).

46. Barclay, R.L., Vicari, J.J., Doughty, A.S., Johanson, J.F. & Greenlaw, R.L. Colonoscopic withdrawal times and adenoma detection during screening colonoscopy. *N Engl J Med* **355**, 2533-2541 (2006).

47. Barclay, R.L., Vicari, J.J. & Greenlaw, R.L. Effect of a time-dependent colonoscopic withdrawal protocol on adenoma detection during screening colonoscopy. *Clin Gastroenterol Hepatol* **6**, 1091-1098 (2008).

48. Singh, H., Nugent, Z., Mahmud, S.M., Demers, A.A. & Bernstein, C.N. Predictors of colorectal cancer after negative colonoscopy: a population-based study. *Am J Gastroenterol* **105**, 663-673; quiz 674 (2010).

49. Hotta, K., Fujii, T., Saito, Y. & Matsuda, T. Local recurrence after endoscopic resection of colorectal tumors. *Int J Colorectal Dis* **24**, 225-230 (2009).

50. Brenner, H., Chang-Claude, J., Seiler, C.M., Sturmer, T. & Hoffmeister, M. Does a negative screening colonoscopy ever need to be repeated? *Gut* **55**, 1145-1150 (2006).

51. Lengauer, C., Kinzler, K.W. & Vogelstein, B. Genetic instability in colorectal cancers. *Nature* **386**, 623-627 (1997).

52. Lengauer, C., Kinzler, K.W. & Vogelstein, B. Genetic instabilities in human cancers. *Nature* **396**, 643-649 (1998).

53. Shen, L. *et al.* Integrated genetic and epigenetic analysis identifies three different subclasses of colon cancer. *Proc Natl Acad Sci U S A* **104**, 18654-18659 (2007).

54. Trautmann, K. *et al.* Chromosomal instability in microsatellite-unstable and stable colon cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research* **12**, 6379-6385 (2006).

55. Cheng, Y.W. *et al.* CpG island methylator phenotype associates with low-degree chromosomal abnormalities in colorectal cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research* **14**, 6005-6013 (2008).

56. Rajagopalan, H., Nowak, M.A., Vogelstein, B. & Lengauer, C. The significance of unstable chromosomes in colorectal cancer. *Nat Rev Cancer* **3**, 695-701 (2003).

57. Pino, M.S. & Chung, D.C. The chromosomal instability pathway in colon cancer. *Gastroenterology* **138**, 2059-2072 (2010).

58. Mai, S. Initiation of telomere-mediated chromosomal rearrangements in cancer. *Journal of cellular biochemistry* **109**, 1095-1102 (2010).

59. Barber, T.D. *et al.* Chromatid cohesion defects may underlie chromosome instability in human colorectal cancers. *Proc Natl Acad Sci U S A* **105**, 3443-3448 (2008).

60. Price, J.C. *et al.* Sequencing of candidate chromosome instability genes in endometrial cancers reveals somatic mutations in ESCO1, CHTF18, and MRE11A. *PLoS One* **8**, e63313 (2013).

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

62. Rowan, A. *et al.* Refining molecular analysis in the pathways of colorectal carcinogenesis. *Clin Gastroenterol Hepatol* **3**, 1115-1123 (2005).

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

64. Lurje, G., Zhang, W. & Lenz, H.J. Molecular prognostic markers in locally advanced colon cancer. *Clinical colorectal cancer* **6**, 683-690 (2007).

65. Choi, S.W. *et al.* Genetic classification of colorectal cancer based on chromosomal loss and microsatellite instability predicts survival. *Clinical cancer research : an official journal of the American Association for Cancer Research* **8**, 2311-2322 (2002).

66. Watanabe, T. *et al.* Chromosomal instability (CIN) phenotype, CIN high or CIN low, predicts survival for colorectal cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **30**, 2256-2264 (2012).

67. Walther, A., Houlston, R. & Tomlinson, I. Association between chromosomal instability and prognosis in colorectal cancer: a meta-analysis. *Gut* **57**, 941-950 (2008).

68. Sheffer, M. *et al.* Association of survival and disease progression with chromosomal instability: a genomic exploration of colorectal cancer. *Proc Natl Acad Sci U S A* **106**, 7131-7136 (2009).

69. Lee, A.J. *et al.* Chromosomal instability confers intrinsic multidrug resistance. *Cancer Res* **71**, 1858-1870 (2011).

70. Chang, S.C., Lin, J.K., Lin, T.C. & Liang, W.Y. Loss of heterozygosity: an independent prognostic factor of colorectal cancer. *World journal of gastroenterology : WJG* **11**, 778-784 (2005).

71. Reichmann, A., Levin, B. & Martin, P. Human large-bowel cancer: correlation of clinical and histopathological features with banded chromosomes. *Int J Cancer* **29**, 625-629 (1982).

72. Delattre, O. *et al.* Multiple genetic alterations in distal and proximal colorectal cancer. *Lancet* **2**, 353-356 (1989).

73. Hagland, H.R., Berg, M., Jolma, I.W., Carlsen, A. & Soreide, K. Molecular Pathways and Cellular Metabolism in Colorectal Cancer. *Digestive surgery* **30**, 12-25 (2013).

74. Bocker, T. *et al.* Genomic instability in colorectal carcinomas: comparison of different evaluation methods and their biological significance. *The Journal of pathology* **179**, 15-19 (1996).

75. Geigl, J.B., Obenauf, A.C., Schwarzbraun, T. & Speicher, M.R. Defining 'chromosomal instability'. *Trends in genetics : TIG* **24**, 64-69 (2008).

76. Chung, D.C. & Rustgi, A.K. DNA mismatch repair and cancer. *Gastroenterology* **109**, 1685-1699 (1995).

77. Boland, C.R. & Goel, A. Microsatellite instability in colorectal cancer. *Gastroenterology* **138**, 2073-2087 e2073 (2010).

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

79. Peltomaki, P. Lynch syndrome genes. *Familial cancer* **4**, 227-232 (2005).

80. Salovaara, R. *et al.* Population-based molecular detection of hereditary nonpolyposis colorectal cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **18**, 2193-2200 (2000).

81. Peel, D.J. *et al.* Characterization of hereditary nonpolyposis colorectal cancer families from a population-based series of cases. *J Natl Cancer Inst* **92**, 1517-1522 (2000).

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

83. Kuismanen, S.A., Holmberg, M.T., Salovaara, R., de la Chapelle, A. & Peltomaki, P. Genetic and epigenetic modification of MLH1 accounts for a major share of microsatellite-unstable colorectal cancers. *The American journal of pathology* **156**, 1773-1779 (2000).

84. Herman, J.G. *et al.* Incidence and functional consequences of hMLH1 promoter hypermethylation in colorectal carcinoma. *Proc Natl Acad Sci U S A* **95**, 6870-6875 (1998).

85. Kane, M.F. *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* **57**, 808-811 (1997).

86. Gafa, R. *et al.* Sporadic colorectal adenocarcinomas with high-frequency microsatellite instability. *Cancer* **89**, 2025-2037 (2000).

87. Elsaleh, H. *et al.* Association of tumour site and sex with survival benefit from adjuvant chemotherapy in colorectal cancer. *Lancet* **355**, 1745-1750 (2000).

88. Ribic, C.M. *et al.* Tumor microsatellite-instability status as a predictor of benefit from fluorouracil-based adjuvant chemotherapy for colon cancer. *N Engl J Med* **349**, 247-257 (2003).

89. Wright, C.M. *et al.* Prognostic significance of extensive microsatellite instability in sporadic clinicopathological stage C colorectal cancer. *The British journal of surgery* **87**, 1197-1202 (2000).

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

91. Bubb, V.J. *et al.* Microsatellite instability and the role of hMSH2 in sporadic colorectalcancer. *Oncogene* **12**, 2641-2649 (1996).

92. Guastadisegni, C., Colafranceschi, M., Ottini, L. & Dogliotti, E. Microsatellite instability as a marker of prognosis and response to therapy: a meta-analysis of colorectal cancer survival data. *European journal of cancer (Oxford, England : 1990)* **46**, 2788-2798 (2010).

93. Vilkin, A. *et al.* Microsatellite instability, MLH1 promoter methylation, and BRAF mutation analysis in sporadic colorectal cancers of different ethnic groups in Israel. *Cancer* **115**, 760-769 (2009).

94. Sugai, T. *et al.* Analysis of molecular alterations in left- and right-sided colorectal carcinomas reveals distinct pathways of carcinogenesis: proposal for new molecular profile of colorectal carcinomas. *J Mol Diagn* **8**, 193-201 (2006).

95. Breivik, J. *et al.* Different genetic pathways to proximal and distal colorectal cancer influenced by sex-related factors. *Int J Cancer* **74**, 664-669 (1997).

96. Boland, C.R. *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* **58**, 5248-5257 (1998).

97. Hatch, S.B. *et al.* Microsatellite instability testing in colorectal carcinoma: choice of markers affects sensitivity of detection of mismatch repair-deficient tumors. *Clinical cancer research : an official journal of the American Association for Cancer Research* **11**, 2180-2187 (2005).

98. Bird, A.P. CpG-rich islands and the function of DNA methylation. *Nature* **321**, 209-213 (1986).

99. Issa, J.P. CpG island methylator phenotype in cancer. *Nat Rev Cancer* **4**, 988-993 (2004).

100. Yoder, J.A., Walsh, C.P. & Bestor, T.H. Cytosine methylation and the ecology of intragenomic parasites. *Trends in genetics : TIG* **13**, 335-340 (1997).

101. Latham, K.E. X chromosome imprinting and inactivation in the early mammalian embryo. *Trends in genetics : TIG* **12**, 134-138 (1996).

102. Toyota, M. *et al.* CpG island methylator phenotype in colorectal cancer. *Proc Natl Acad Sci U S A* **96**, 8681-8686 (1999).

103. Curtin, K., Slattery, M.L. & Samowitz, W.S. CpG island methylation in colorectal cancer: past, present and future. *Pathology research international* **2011**, 902674 (2011).

104. Issa, J.P. *et al.* Increased cytosine DNA-methyltransferase activity during colon cancer progression. *J Natl Cancer Inst* **85**, 1235-1240 (1993).

105. Easwaran, H.P. *et al.* Aberrant silencing of cancer-related genes by CpG hypermethylation occurs independently of their spatial organization in the nucleus. *Cancer Res* **70**, 8015-8024 (2010).

106. Esteller, M. Epigenetic gene silencing in cancer: the DNA hypermethylome. *Human molecular genetics* **16 Spec No 1**, R50-59 (2007).

107. Park, S.J. *et al.* Frequent CpG island methylation in serrated adenomas of the colorectum. *The American journal of pathology* **162**, 815-822 (2003).

108. Grady, W.M. & Carethers, J.M. Genomic and epigenetic instability in colorectal cancer pathogenesis. *Gastroenterology* **135**, 1079-1099 (2008).

109. Samowitz, W.S. *et al.* Evaluation of a large, population-based sample supports a CpG island methylator phenotype in colon cancer. *Gastroenterology* **129**, 837-845 (2005).

110. Ward, R.L. *et al.* Adverse prognostic effect of methylation in colorectal cancer is reversed by microsatellite instability. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **21**, 3729-3736 (2003).

111. Kim, Y.S. & Deng, G. Epigenetic changes (aberrant DNA methylation) in colorectal neoplasia. *Gut and liver* **1**, 1-11 (2007).

112. Fleming, M., Ravula, S., Tatishchev, S.F. & Wang, H.L. Colorectal carcinoma: Pathologic aspects. *Journal of gastrointestinal oncology* **3**, 153-173 (2012).

113. Chan, A.O. *et al.* CpG island methylation in aberrant crypt foci of the colorectum. *The American journal of pathology* **160**, 1823-1830 (2002).

114. Rashid, A., Shen, L., Morris, J.S., Issa, J.P. & Hamilton, S.R. CpG island methylation in colorectal adenomas. *The American journal of pathology* **159**, 1129-1135 (2001).

115. Kim, K.M. *et al.* Molecular features of colorectal hyperplastic polyps and sessile serrated adenoma/polyps from Korea. *Am J Surg Pathol* **35**, 1274-1286 (2011).

116. O'Brien, M.J. *et al.* Comparison of microsatellite instability, CpG island methylation phenotype, BRAF and KRAS status in serrated polyps and traditional adenomas indicates separate pathways to distinct colorectal carcinoma end points. *Am J Surg Pathol* **30**, 1491-1501 (2006).

117. Maeda, T. *et al.* Sessile serrated adenoma shares similar genetic and epigenetic features with microsatellite unstable colon cancer in a location-dependent manner. *Experimental and therapeutic medicine* **2**, 695-700 (2011).

118. Kim, Y.H., Kakar, S., Cun, L., Deng, G. & Kim, Y.S. Distinct CpG island methylation profiles and BRAF mutation status in serrated and adenomatous colorectal polyps. *Int J Cancer* **123**, 2587-2593 (2008).

119. O'Brien, M.J. *et al.* Hyperplastic (serrated) polyps of the colorectum: relationship of CpG island methylator phenotype and K-ras mutation to location and histologic subtype. *Am J Surg Pathol* **28**, 423-434 (2004).

120. Rajagopalan, H. *et al.* Tumorigenesis: RAF/RAS oncogenes and mismatch-repair status. *Nature* **418**, 934 (2002).

121. Yuen, S.T. *et al.* Similarity of the phenotypic patterns associated with BRAF and KRAS mutations in colorectal neoplasia. *Cancer Res* **62**, 6451-6455 (2002).

122. Davies, H. *et al.* Mutations of the BRAF gene in human cancer. *Nature* **417**, 949-954 (2002).

123. Kambara, T. *et al.* BRAF mutation is associated with DNA methylation in serrated polyps and cancers of the colorectum. *Gut* **53**, 1137-1144 (2004).

124. Spring, K.J. *et al.* High prevalence of sessile serrated adenomas with BRAF mutations: a prospective study of patients undergoing colonoscopy. *Gastroenterology* **131**, 1400-1407 (2006).

125. Mohammadi, M., Kristensen, M.H., Nielsen, H.J., Bonde, J.H. & Holck, S. Qualities of sessile serrated adenoma/polyp/lesion and its borderline variant in the context of synchronous colorectal carcinoma. *Journal of clinical pathology* **65**, 924-927 (2012).

126. Fujita, K. *et al.* Sessile serrated adenoma with early neoplastic progression: a clinicopathologic and molecular study. *Am J Surg Pathol* **35**, 295-304 (2011).

127. Weisenberger, D.J. *et al.* CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. *Nature genetics* **38**, 787-793 (2006).

128. Wang, L. *et al.* BRAF mutations in colon cancer are not likely attributable to defective DNA mismatch repair. *Cancer Res* **63**, 5209-5212 (2003).

129. Nagasaka, T. *et al.* Colorectal cancer with mutation in BRAF, KRAS, and wild-type with respect to both oncogenes showing different patterns of DNA methylation. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **22**, 4584-4594 (2004).

130. Samowitz, W.S. *et al.* Poor survival associated with the BRAF V600E mutation in microsatellite-stable colon cancers. *Cancer Res* **65**, 6063-6069 (2005).

131. Liu, X., Jakubowski, M. & Hunt, J.L. KRAS gene mutation in colorectal cancer is correlated with increased proliferation and spontaneous apoptosis. *American journal of clinical pathology* **135**, 245-252 (2011).

132. Samowitz, W.S. *et al.* Inverse relationship between microsatellite instability and K-ras and p53 gene alterations in colon cancer. *The American journal of pathology* **158**, 1517-1524 (2001).

133. Kolch, W. Meaningful relationships: the regulation of the Ras/Raf/MEK/ERK pathway by protein interactions. *The Biochemical journal* **351 Pt 2**, 289-305 (2000).

134. Oliveira, C. *et al.* Distinct patterns of KRAS mutations in colorectal carcinomas according to germline mismatch repair defects and hMLH1 methylation status. *Human molecular genetics* **13**, 2303-2311 (2004).

135. Iacopetta, B.J. *et al.* Mutation of the transforming growth factor-beta type II receptor gene in right-sided colorectal cancer: relationship to clinicopathological features and genetic alterations. *The Journal of pathology* **184**, 390-395 (1998).

136. Konishi, M. *et al.* Molecular nature of colon tumors in hereditary nonpolyposis colon cancer, familial polyposis, and sporadic colon cancer. *Gastroenterology* **111**, 307-317 (1996).

137. Olschwang, S. *et al.* Alternative genetic pathways in colorectal carcinogenesis. *Proc Natl Acad Sci U S A* **94**, 12122-12127 (1997).

138. Kawabata, Y. *et al.* Molecular characteristics of poorly differentiated adenocarcinoma and signet-ring-cell carcinoma of colorectum. *Int J Cancer* **84**, 33-38 (1999).

139. Ferrandez, A., Samowitz, W., DiSario, J.A. & Burt, R.W. Phenotypic characteristics and risk of cancer development in hyperplastic polyposis: case series and literature review. *Am J Gastroenterol* **99**, 2012-2018 (2004).

140. Urbanski, S.J., Kossakowska, A.E., Marcon, N. & Bruce, W.R. Mixed hyperplastic adenomatous polyps--an underdiagnosed entity. Report of a case of adenocarcinoma arising within a mixed hyperplastic adenomatous polyp. *Am J Surg Pathol* **8**, 551-556 (1984).

141. Longacre, T.A. & Fenoglio-Preiser, C.M. Mixed hyperplastic adenomatous polyps/serrated adenomas. A distinct form of colorectal neoplasia. *Am J Surg Pathol* **14**, 524-537 (1990).

142. Torlakovic, E., Skovlund, E., Snover, D.C., Torlakovic, G. & Nesland, J.M. Morphologic reappraisal of serrated colorectal polyps. *Am J Surg Pathol* **27**, 65-81 (2003).

143. Torlakovic, E. & Snover, D.C. Serrated adenomatous polyposis in humans. *Gastroenterology* **110**, 748-755 (1996).

144. Snover, D.C. Update on the serrated pathway to colorectal carcinoma. *Hum Pathol* **42**, 1-10 (2011).

145. Leggett, B. & Whitehall, V. Role of the serrated pathway in colorectal cancer pathogenesis. *Gastroenterology* **138**, 2088-2100 (2010).

146. Snover, D.C. Sessile serrated adenoma/polyp of the large intestine: a potentially aggressive lesion in need of a new screening strategy. *Dis Colon Rectum* **54**, 1205-1206 (2011).

147. Hetzel, J.T. *et al.* Variation in the detection of serrated polyps in an average risk colorectal cancer screening cohort. *Am J Gastroenterol* **105**, 2656-2664 (2010).

148. Patil, D.T., Shadrach, B.L., Rybicki, L.A., Leach, B.H. & Pai, R.K. Proximal colon cancers and the serrated pathway: a systematic analysis of precursor histology and BRAF mutation status. *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc* **25**, 1423-1431 (2012).

149. Goke, F. *et al.* Fibroblast growth factor receptor 1 as a putative therapy target in colorectal cancer. *Digestion* **88**, 172-181 (2013).

150. Forbes, S.A. *et al.* COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic acids research* **39**, D945-950 (2011).

151. Mehta, G.D., Rizvi, S.M. & Ghosh, S.K. Cohesin: a guardian of genome integrity. *Biochimica et biophysica acta* **1823**, 1324-1342 (2012).

APPENDIX A: SOLUTIONS

CELL CULTURE

10xPBS (Stock Solution)

Name	Amount
NaCl	80.0 g
KCl	2.0 g
Na_2HPO_4	14.4 g
KH_2PO_4	2.4 g
ddH_2O	800 mL
Titrate	
ddH_2O	up to 1.0 L total volume
Total Volume	1.0 L

-titrate to pH 7.4

1xPBS

Name	Amount
10xPBS (stock)	100.0 mL
ddH_2O	900.0 mL
Total Volume	1.0 L

CELL FIXATION AND MITOTIC SPREADS

1M KCl (Stock Solution)

Name	Amount
KCl	7.5 g
ddH_2O	up to 100.0 mL total volume
Total Volume	100.0 mL

75mM KCl (Hypotonic Solution)

Name	Amount
1M KCl	750 µL
ddH_2O	up to 10.0 mL total volume
Total Volume	10.0 mL

3:1 Methanol:Acetic Acid (Fixative)

Name	Amount
Methanol	3.0 mL
Acetic Acid	1.0 mL
Total Volume	4.0 mL

70% Ethanol

Name	Amount
95% Ethanol	737.0 mL
ddH_2O	263.0 mL
Total Volume	1.0 L

50% Acetic Acid

Name	Amount
Acetic Acid, Glacial	50.0 mL
ddH_2O	50.0 mL
Total Volume	100.0 L

Giemsa

Name	Amount
Giemsa Stain	1.0 mL
Phosphate Buffer	5.0-6.0 mL
Vortex	
Total Volume	6.0-7.0 mL

FLUORESCENCE IN SITU HYBRIDIZATION OF CELL LINES

20xSSC (Stock Solution)

Name	Amount
20xSSC	132 g
ddH_2O	400.0 mL
Titrate	
ddH_2O	up to 500.0 mL total volume
Total Volume	500.0 mL

-titrate to pH 5.3 with HCl

2xSSC

Name	Amount
20xSSC (stock)	100.0 mL
ddH_2O	850.0 mL
Titrate	
ddH_2O	up to 1.0 L total volume
Total Volume	1.0 L

-titrate to pH 7.0 with NaOH

1M HCl (Stock Solution)

Name	Amount
HCl (concentrated, 37.5%)	8.2 mL
ddH_2O	up to 100.0 mL
Total Volume	100.0 mL

0.01M HCl

Name	Amount
1M HCl (stock)	1.5 mL
ddH_2O	148.5 mL
Total Volume	150.0 mL

Pepsin

Name	Amount
Pepsin (stock, 100 mg/mL)	75 µL
0.01M HCl	150 mL
Total Volume	~150 mL

1M $MgCl_2$

Name	Amount
$MgCl_2$	20.3 g
ddH_2O	up to 100.0 mL total volume
Total Volume	100.0 mL

1xPBS/50 mM $MgCl_2$

Name	Amount
1M $MgCl_2$	50.0 mL
10xPBS (stock)	100.0 mL
ddH_2O	up to 1.0 L total volume
Total Volume	150.0 mL

1% Formaldehyde/1xPBS/50mM $MgCl_2$

Name	Amount
37% formaldehyde	4.05 mL
1xPBS/$MgCl_2$	145 mL
Total Volume	~150.0 mL

90% Ethanol

Name	Amount
95% Ethanol	947.0 mL
ddH_2O	53.0 mL
Total Volume	1.0 L

70% Formamide/2xSSC

Name	Amount
Deionized formamide	35.0 mL
2xSSC	15.0 mL
Total Volume	40.0 mL

-titrate to pH 7.0

100x DAPI

Name	Amount
DAPI	50.0 µg
2xSSC	1.0 mL
Total Volume	1.0 mL

CELL SYNCHRONIZATION

Nocodazole (Stock Solution)

Name	Amount
Nocodazole	5.0 mg
ddH_2O	5.0 mL
Total Volume	5.0 mL

100 mM Thymidine (stock solution)

Name	Amount
Thymidine	24.2 mg
ddH_2O	1.0 mL
Total Volume	1.0 mL

PI (Stock Solution)

Name	Amount
Propidium Iodide	1.0 mg
ddH_2O	1.0 mL
Total Volume	1.0 mL

RNase A (Stock Solution)

Name	Amount
RNase A	10.0 mg
ddH_2O	1.0 mL
Total Volume	1.0 mL

FLUORESCENCE IN SITU HYBRIDIZATION IN FFPE TISSUE

Citric Acid Buffer

Name	Amount
Citric Acid	0.525 g
ddH_2O	200.0 mL
Titrate	
ddH_2O	up to 250.0 mL total volume
Total Volume	250.0 mL

-titrate to pH 6.0 with NaOH

0.2N HCl

Name	Amount
1M HCl	10.0 mL
ddH_2O	40.0 mL
Total Volume	50.0 mL

-titrate to pH 2.0

Pepsin

Name	Amount
Pepsin (stock, 100 mg/mL)	250.0 µL
0.2N HCl pH 2.0	50 mL
Total Volume	~50 mL

50% Formamide/2xSSC

Name	Amount
Formamide	500.0 mL
20xSSC (stock)	50.0 mL
ddH_2O	up to 1.0 L total volume
Total Volume	1.0 L

2xSSC/0.1% NP-40

Name	Amount
20xSSC (stock)	100.0 mL
ddH_2O	850 mL
Titrate	
NP-40	1.0 mL
Mix until dissolved	
ddH_2O	up to 1.0 L total volume
Total Volume	1.0 L

APPENDIX B: SUPPLEMENTARY TABLES

Supplementary Table 1.

Chromosome Enumeration Results with Calculated CIN Score.

Sample	Core	N	Chr 8		Chr 11		Chr 17		CIN
			Gains	Losses	Gains	Losses	Gains	Losses	Score
1	A	53	16	13	7	16	9	15	
	B	100	36	13	31	32	26	21	1.54
2	A	100	38	33	52	40	30	55	
	B	100	38	26	76	19	38	39	2.42
3	A	100	26	24	27	30	12	33	
	B	100	26	43	28	40	15	47	1.76
4	A	100	15	21	14	22	11	36	
	B	100	22	11	19	32	14	32	1.25
5	A	100	15	48	20	43	22	57	
	B	100	21	29	9	42	17	43	1.83
6	A	100	8	75	10	75	12	65	
	B	0	0	0	0	0	0	0	2.45
7	A	100	52	20	42	21	49	24	
	B	100	77	13	61	32	74	18	2.42
8	A	100	59	30	29	55	14	71	
	B	100	106	30	42	40	41	57	2.87
9	A	100	11	55	9	67	7	72	
	B	100	13	43	3	56	5	51	1.96
10	A	100	33	24	30	39	19	37	
	B	100	76	30	52	40	36	45	2.31
11	A	100	17	29	18	31	18	34	
	B	100	24	42	13	44	10	58	1.69
12	A	100	9	54	18	42	21	30	
	B	100	5	63	23	30	26	12	1.67
13	A	100	57	12	32	38	53	29	
	B	100	57	15	53	25	60	15	2.23
14	A	100	39	30	24	35	13	55	
	B	100	117	13	67	30	16	57	2.48
15	A	100	57	16	22	46	13	45	
	B	100	109	2	23	28	29	28	2.09
16	A	100	87	27	43	28	17	36	
	B	100	199	3	58	19	15	49	2.91

Supplementary Table 1 Continued.

Chromosome Enumeration Results with Calculated CIN Score.

Sample	Core	N	Chr 8		Chr 11		Chr 17		CIN
			Gains	Losses	Gains	Losses	Gains	Losses	Score
17	A	100	56	11	23	35	21	34	
	B	100	63	20	15	55	29	44	2.03
18	A	100	24	34	17	49	22	31	
	B	51	20	7	8	32	17	10	1.79
19	A	50	9	12	20	19	15	13	
	B	100	77	19	45	30	61	19	2.26
20	A	100	11	13	5	20	7	21	
	B	100	25	4	20	28	10	19	0.92
21	A	100	22	23	14	29	18	28	
	B	100	24	43	19	53	36	41	1.75
22	A*	66	38	30	0	0	7	58	
	B	100	17	77	6	83	8	103	2.57
23	A	100	46	42	13	64	17	46	
	B	100	40	40	28	43	23	38	2.10
24	A	100	59	38	46	44	30	58	
	B	100	44	27	58	41	36	28	2.55
25	A	100	138	24	28	59	22	52	
	B	100	93	40	71	23	25	68	3.22
26	A	100	37	44	30	45	23	54	
	B	100	207	33	22	73	19	73	3.30
27	A	100	20	76	24	74	14	66	
	B	100	110	41	86	35	50	34	3.15
28	A	100	28	47	35	38	23	50	
	B	100	20	43	26	56	21	54	2.21
29	A	100	44	49	29	36	38	45	
	B	100	73	23	70	14	55	27	2.57
30	A	100	42	32	38	46	29	40	
	B	100	31	36	28	45	14	55	2.18
31	A	100	76	38	44	57	50	50	
	B	100	27	44	25	54	16	53	2.82
32	A	100	16	71	13	86	12	74	
	B	100	45	36	24	54	22	49	2.51
33	A	100	61	47	41	41	52	47	
	B	100	53	46	49	60	46	37	2.90

Supplementary Table 1 Continued.

Chromosome Enumeration Results with Calculated CIN Score.

Sample	Core	N	Chr 8		Chr 11		Chr 17		CIN
			Gains	Losses	Gains	Losses	Gains	Losses	Score
34	A	0	0	0	0	0	0	0	
	B	100	24	51	41	46	78	26	2.66
35	A	100	43	53	26	52	73	39	
	B	100	40	54	26	64	67	49	2.94
36	A	70	8	57	7	57	8	65	
	B	100	119	28	72	39	48	53	3.30
37	A	100	38	37	54	40	30	50	
	B	100	53	44	54	49	36	46	2.66
38	A	100	47	40	10	90	39	45	
	B	100	37	43	14	62	25	52	2.52
39	A	100	47	53	51	61	33	50	
	B	100	42	41	58	32	60	33	2.81
40	A	100	25	56	21	64	13	63	
	B	100	51	36	45	32	46	29	2.41
41	A	29	22	15	10	20	7	21	
	B	100	43	47	33	76	19	64	2.92
42	A	100	46	66	21	77	27	73	
	B	100	52	40	27	65	44	40	2.89
43	A	100	22	65	25	56	23	60	
	B	100	25	55	17	63	11	69	2.45
44	A	100	120	18	39	26	56	34	
	B	100	140	44	65	43	51	49	3.43
45	A	100	39	37	34	52	28	40	
	B	100	30	41	21	50	29	35	2.18
46	A	100	84	28	35	28	34	28	
	B	100	26	31	29	33	22	35	2.07
47	A	100	151	24	91	26	83	33	
	B	100	66	35	48	34	52	40	3.42
48	A	100	53	43	21	65	17	66	
	B	100	22	46	14	81	14	81	2.62
49	A	100	23	62	10	57	4	86	
	B	100	23	46	9	61	8	23	2.33
50	A	100	56	25	33	39	18	52	
	B	100	76	22	28	60	23	62	2.47

Supplementary Table 1 Continued.

Chromosome Enumeration Results with Calculated CIN Score.

Sample	Core	N	Chr 8		Chr 11		Chr 17		CIN
			Gains	Losses	Gains	Losses	Gains	Losses	Score
51	A	100	26	45	34	49	39	57	
	B	100	54	43	50	48	47	37	2.60
52	A	100	31	29	171	13	25	47	
	B	100	20	50	123	22	18	49	2.99
53	A	100	74	22	16	44	21	30	
	B	100	166	8	24	44	25	40	2.57
54	A	100	41	29	38	33	33	23	
	B	100	42	36	19	38	23	42	2.00
55	A	100	25	32	13	42	8	55	
	B	100	24	40	19	50	19	24	1.91
56	A	100	52	36	63	31	73	29	
	B	100	87	34	156	14	99	31	3.53
57	A	100	44	37	45	52	8	58	
	B	100	59	24	70	17	20	44	2.39
58	A	100	176	10	61	37	22	49	
	B	100	192	12	66	29	27	58	3.70
60	A	100	27	34	21	33	19	38	
	B	55	23	34	26	47	24	46	1.86
61	A	100	27	36	24	53	29	50	
	B	100	11	27	9	36	3	51	2.30
62	A	100	138	28	104	30	122	33	
	B	100	138	22	135	25	127	26	4.64
63	A	100	7	100	1	102	8	88	
	B	100	41	43	29	43	21	52	2.60
64	A	100	164	14	67	36	17	60	
	B	100	43	38	26	43	17	50	2.88
65	A	100	30	32	19	57	20	50	
	B	100	27	29	25	31	15	20	1.78
66	A	100	34	31	25	40	22	42	
	B	70	17	52	9	47	10	48	2.22
67	A	100	122	22	27	47	19	44	
	B	100	110	20	25	43	24	42	2.73

N = number of nuclei enumerated in core

*Could not enumerate chromosome 11 due to excessive autofluorescence. Enumerated chromosome 8 and 17 in 100 nuclei.