NEOPLASIA
www.neoplasia.com

Volume 20 Number 9 September 2018 pp. 943–950 943

Characterizing Microsatellite Instability and Chromosome Instability in Interval Colorectal Cancers

Amy L. Cisyk[*,†], Zoann Nugent[†], Robert H. Wightman[‡], Harminder Singh[†,§] and Kirk J. McManus[*,†]

[*]Department of Biochemistry & Medical Genetics, University of Manitoba, Winnipeg, Canada; [†]Research Institute in Oncology & Hematology, CancerCare Manitoba, Winnipeg, Canada; [‡]Department of Pathology, University of Manitoba, Winnipeg, Canada; [§]Department of Internal Medicine, University of Manitoba, Winnipeg, Canada

Abstract

There are a substantial portion of colorectal cancers (CRCs), termed interval CRCs (I-CRCs), that are diagnosed shortly after a negative colonoscopy (i.e., no detectable polyps or CRC) and before recommended follow-up screening. The underlying cause(s) accounting for I-CRCs remain poorly understood, but may involve aberrant biology that drives genome instability. Genetic defects inducing genome instability are pathogenic events that lead to the development and progression of traditional sporadic (Sp-) CRCs. Classically, there are two genome instability pathways that give rise to virtually all Sp-CRCs, chromosome instability (CIN; ~85% of Sp-CRCs) and microsatellite instability (MSI; ~15% of Sp-CRCs); however, the contribution MSI and CIN have in I-CRCs is only beginning to emerge. To date, no study has simultaneously evaluated both MSI and CIN within an I-CRC cohort, and thus we sought to determine and compare the prevalence of MSI and/or CIN within population-based I-CRC and matched Sp-CRC cohorts. MSI status was established using a clinically validated, immunohistochemical approach that assessed the presence or absence of four proteins (MLH1, MSH2, MSH6 and PMS2) implicated in MSI. By combining the MSI results of the current study with those of our previous CIN study, we provide unprecedented insight into the prevalence of MSI and/or CIN between and within Sp- and I-CRCs. Our data show that MSI^+ tumors are 1.5-times more prevalent within I-CRCs than Sp-CRCs in a population-based setting and further show that CIN^+/MSI^+ I-CRCs occur at similar frequency as CIN^+/MSI^+ Sp-CRCs.

Neoplasia (2018) 20, 943–950

Introduction

CRC is the second leading cause of cancer-related deaths in North America. In 2017, ~162,250 North Americans were newly diagnosed with colorectal cancer (CRC), while an additional ~59,700 succumbed to the disease [2,3]. Further, of those newly diagnosed, ~80% to 85% are sporadic, or randomly occurring (i.e., there is no evidence for any predisposing hereditary cancer syndromes and/or family history) [4]. The high morbidity and mortality rates due to CRC underscore the need for accurate screening and diagnostic strategies [5].

Colonoscopy is an effective diagnostic and screening modality for CRCs and its use correlates with reduced CRC incidence and mortality primarily due to its ability to identify precursor lesions (i.e., polyps) and early stage disease (i.e., I and II) [6,7]. However, even with colonoscopies, there remain a portion of CRCs, termed interval CRCs (I-CRCs) that are diagnosed shortly after a negative colonoscopy (i.e., no detectable polyps or CRC) and before the recommended follow-up CRC screening. A meta-analysis determined that the pooled prevalence of I-CRCs is ~3.7% [8], which represents

Abbreviations: CIMP, CpG island methylator phenotype; CIN, Chromosome instability; CRC, Colorectal cancer; CS, CIN score; FISH, Fluorescent in situ hybridization; I-CRC, Interval CRC; IHC, Immunohistochemistry; MMR, DNA mismatch repair; MPE, Molecular pathologic epidemiology; MSI, Microsatellite instability; SD, Standard deviation; Sp-CRC, Sporadic CRC; TMA, Tumor micro-array.

Address all correspondence to: Research Institute in Oncology & Hematology, ON6010 – 675 McDermot Avenue, Winnipeg, MB, R3E 0V9. E-mail: Kirk.McManus@umanitoba.ca

Received 18 May 2018; Revised 31 July 2018; Accepted 31 July 2018

1476-5586
https://doi.org/10.1016/j.neo.2018.07.007

~6000 North Americans annually [2,3]. I-CRCs could represent missed sporadic CRC (Sp-CRC) cases that present as false-negative colonoscopies [9–15], or a unique subtype of Sp-CRC that harbors distinct biological characteristics [1,8,13,15–23] that enable rapid tumor growth and development (reviewed in [24]).

Genome instability is an enabling feature of virtually all cancer types [25] and is perhaps best understood in CRC. Genome instability typically arises through three pathways: 1) CpG island methylator phenotype (CIMP); 2) chromosome instability (CIN); and 3) microsatellite instability (MSI). While CIMP is an epigenetic phenomenon whereby hypermethylation of CpG islands on gene promoters correlates with gene silencing [26], CIN is defined as an increase in the rate at which chromosomes, or large chromosomal fragments, are gained or lost [27,28]. Finally, MSI arises from defects in the DNA mismatch repair (MMR) pathway [29]. Traditionally, MSI and CIN are proposed to be mutually exclusive pathways giving rise to Sp-CRCs [27], whereas it has been suggested that CIMP may underlie the development of MSI and/or CIN [30].

Currently, there is limited information detailing the etiological origins of I-CRCs. Of the few studies conducted, most have focused on the prevalence of MSI [18–20] and/or CIMP [17,19], with even fewer evaluating specific genes like *BRAF* [22], *KRAS* [21], or *CTNNB1* [16]. Collectively, these studies do support divergent biology relative to Sp-CRCs, as I-CRCs typically exhibit a 1.5- and 3.0-fold increase in the prevalence of CIMP and MSI, respectively. Given the traditional view that MSI and CIN are mutually exclusive [27,28], these observations imply that the prevalence of CIN is likely to be lower in I-CRCs (e.g. ~55%) versus Sp-CRCs (~80%-85%). However, we recently determined that the prevalence of CIN in I-CRCs (~85%) was statistically indistinguishable from Sp-CRC controls [1], raising the possibility that MSI and CIN may co-occur within I-CRCs.

No prior study has simultaneously assessed MSI and CIN within the same I-CRC cohort. We previously determined the prevalence of CIN within a tissue microarray (TMA) comprised of 95 Sp-CRC (control) and 46 I-CRC samples [1]. In the current study, we determined the prevalence of MSI within this same cohort and have now correlated these findings with the CIN findings of the previous study.

Materials and Methods

Ethics Statement

Ethics for this study, including the collection and use of archived clinical CRC tissue samples was approved by the University of Manitoba Research Ethics Board (REB Registry Number: H2010:237 [HS11032]) and Pathology Access Committee for Tissue and Manitoba's Health Information Privacy Committee.

Patient Identification

CRCs were identified from the population-based Manitoba Cancer Registry and linked to patient colonoscopy records through Manitoba Health databases as detailed elsewhere [1].

CRC Cohort

The CRC cohort is described elsewhere [1]. Briefly, I-CRCs were defined as CRCs diagnosed between 6 and 36 months following a colonoscopy, while CRCs detected upon initial colonoscopy (i.e. CRC diagnosis within a month of colonoscopy) were classified as Sp-CRCs and employed as controls. CRCs diagnosed between 1 to 6 months of colonoscopy were excluded from the analysis. The Sp-CRCs were matched 2:1 with I-CRCs based on gender, age and tumor location (proximal vs. distal based on location at or proximal to splenic flexure vs. more distally) [1]. Individuals with history of inflammatory bowel disease were excluded from both groups. Archived clinical formalin-fixed and paraffin-embedded tumor samples were supplied by the Department of Pathology in an anonymized, double-blinded fashion, with the I-CRC status only revealed following completion of the MSI analyses. A total of 141 samples, including 95 Sp- and 46 I-CRCs were evaluated. Minor sample attrition (5 Sp-CRCs and 1 I-CRC) occurred due to lack of informative CIN or MSI status stemming from too few cells for the CIN analyses, or lack of tumor tissue within the TMA cores based on routine hematoxylin and eosin staining, respectively.

CRC Tissue Micro-Array (TMA)

CRC samples were arrayed in duplicate as detailed previously [1].

Immunohistochemistry and Microsatellite Instability (MSI)

Immunohistochemistry (IHC) was performed using a Dako Autostainer Link 48 (Dako; Agilent) and clinically-validated monoclonal antibodies recognizing MLH1 (ES05 at 1:50; Dako), MSH2 (FE11 at 1:150; Dako), MSH6 (EP49 at 1:300; Dako) and PMS2 (EP51 at 1:50; Dako). Briefly, serial sections of the TMA (6 μm) were deparaffinized, subjected to an antigen retrieval and incubated with primary antibodies. Slides were mounted and scored in a double-blinded fashion for the presence (+) or absence (–) of each epitope interrogated. Samples lacking antibody labeling for 1 or more epitopes were considered MSI^+. To assess MSI within the CRC cohort, the presence (+) or absence (–) of four proteins (MLH1, MSH2, MSH6 and PMS2) with essential roles in MMR and causally linked with MSI were immunohistochemically assessed using clinically-validated antibodies. Briefly, serial sections of the CRC tissue microarray (TMA) were independently labeled with antibodies, and samples were qualitatively assessed for the presence or absence of each protein in a double-blinded manner (Figure 1). In agreement with standard clinical practice, samples exhibiting positive antibody labeling for each of the four proteins interrogated were defined as MSI^-, while samples lacking labeling for ≥1 targeted proteins were defined as MSI^+.

Chromosome Instability (CIN) Analysis

Briefly, the previous CIN study employed a FISH-based approach to assess gains and/or losses of three specific chromosomes implicated in CRC pathogenesis (i.e., chromosomes 8, 11, and 17) [1]. Chromosome enumeration probes recognizing each chromosome were quantitatively assessed within each CRC sample. To identify CIN^+ CRCs, we devised a metric, called a mean CIN Score (CS) that reflects both the gains and/or losses of each FISH probe (i.e., chromosome) within a given sample. A mean CS = 0 defines the diploid state and deviations from 0 identifies samples with gains and/or losses of FISH probes. As CIN and MSI typically occur in ~85% [27] and ~15% [29] of Sp-CRCs, respectively, we operationally defined the 15th percentile of the Sp-CRCs as the minimum threshold required to identify CIN^+ CRCs (i.e., mean CS ≥1.68) and determined that ~82% of I-CRCs (36/44 samples) were defined as CIN^+ tumors and were statistically indistinguishable from the 85% of Sp-CRCs (80/94 samples) defined as CIN^+ [1].

Statistical Analyses

Data were described using standard descriptive statistical analyses. Wilcoxon two sample tests were performed for continuous data such as comparing ages. Proportions were compared using Fisher's Exact

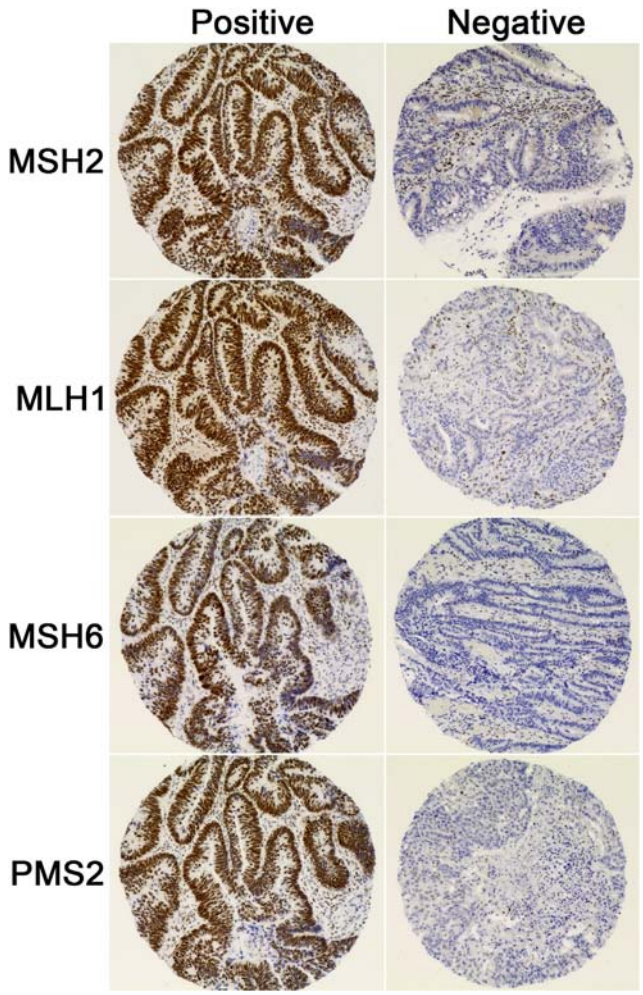

Figure 1. Immunohistochemical Evaluation of CRC Samples.Representative low resolution images of CRC cores from the TMA that are positive or negative for the protein indicated on the left. Note the number of cells labeled and labeling intensity within the left-hand panels (positive labeling; MSI^-) relative to the right-hand panels (negative labeling; MSI^+).

Table 1. Results of MSI analyses

	Total N^*	Sp-CRC $n^\dagger$ (%)	I-CRC $n^\dagger$ (%)	P value‡
MSI^+ Samples	28	16 (17.4)	12 (27.3)	0.26
MSI^- Samples	108	76 (82.6)	32 (72.7)	
MSI^+ Samples		16 (100)	12 (100)	
Sex				
Female		10 (62.5)	6 (50.0)	0.70
Male		6 (37.5)	6 (50.0)	
Age, mean ± SD		69.9 ± 8.5	72 ± 6.9	0.49
Tumor Location				
Proximal colon		14 (87.5)	12 (100.0)	0.49
Distal colon		2 (12.5)	0 (0.0)	
Stage§				
0/I		3 (18.8)	0 (0)	0.04
II		9 (56.3)	4 (33.3)	
III		2 (12.5)	6 (50.0)	
IV		1 (6.3)	1 (8.3)	
ND¶		1 (6.3)	1 (8.3)	
Grade (clustered)				
1 and 2		8 (50.0)	7 (58.3)	0.70
3 and 4		8 (50.0)	4 (33.3)	
ND		0 (0)	1 (8.3)	
MSI^- Samples		76 (100)	32 (100)	
Sex				
Female		52 (68.4)	21 (65.6)	0.82
Male		24 (31.6)	11 (34.4)	
Age, mean ± SD		70.0 ± 8.0	69.8 ± 8.3	0.91
Tumor Location				
Proximal colon		54 (71.1)	23 (71.9)	1.0
Distal colon		22 (28.9)	9 (28.1)	
Stage				
0/I		7 (9.2)	6 (18.8)	0.30
II		21 (27.6)	8 (25.0)	
III		26 (34.2)	8 (25.0)	
IV		16 (21.1)	6 (18.8)	
ND§		6 (7.9)	4 (12.5)	
Grade (clustered)				
1 and 2		64 (84.2)	25 (78.1)	0.71
3 and 4		6 (7.9)	3 (9.4)	
ND		6 (7.9)	4 (12.5)	

* N = total number of samples in the cohort.
† n = number of samples in a subcategory.
‡ $P < .05$ is statistically significant.
§ American Joint Commission on Cancer (AJCC); staging system version 6.
¶ ND = not determined.

tests, including comparisons of sex distributions, tumor location, and grade. Mantel–Haenszel tests were used to compare multi-categorical ordinal data such as stage. Samples with missing data were omitted from the analyses. Multivariable logistic regression analysis model was used to assess independent association with I-CRC status; model covariates included age, sex, site of CRC in colon, grade of CRC, stage of CRC, CIN + /–, MSI +/– and physician specialty of endoscopy physician performing the initial colonoscopy.

Results

Evaluating MSI in Sp- and I-CRCs Stratified by Sex, Age and Tumor Location

The physical and clinical characteristics of the CRC study cohort were detailed previously (Table 1 therein) [1]. Briefly, the original cohort was selected using a ~2:1 ratio of Sp-CRC to I-CRC samples, matched based on key parameters of sex, age and tumor location distributions. The tumor grade and stage was comparable (80–85% of each group was grade 1 or 2). Further, there were no significant differences in colonoscopy completion rates between groups. The quality of bowel preparations was only recorded in one third of the cohort with no significant differences observed between groups. A higher proportion of I-CRC (79%) than Sp-CRC (60%) had their colonoscopy with a surgeon rather than a gastroenterologist (P = .04). To assess MSI within the CRC cohort, the presence (+) or absence (–) of four proteins (MLH1, MSH2, MSH6 and PMS2) with essential roles in MMR and causally linked with MSI were assessed using IHC and clinically validated antibodies (Figure 1). Overall, MSI (i.e., MSI^+ samples) was ~1.6-times more prevalent within the I-CRC (27%) than the Sp-CRC control group (17%) (Figure 2*A* and Table 1).

Studies have shown that differences exist in the prevalence of MSI based on sex, age, stage and anatomic location [31]. Accordingly, each of these criteria were independently assessed within the CRC cohort (Table 1). First, potential differences in the prevalence of MSI within Sp-and I-CRC were assessed based on sex. As shown in Figure 2*B*, MSI was 1.3- and 1.5-times more prevalent within I-CRCs than Sp-CRCs for males (I-CRC, 21%; Sp-CRC, 16%) and females (I-CRC, 35%; Sp-CRC, 23%), respectively. Furthermore, MSI was also more prevalent in females than males within each group (1.6- and 1.4-times more prevalent in I- and Sp-CRCs respectively). However, overall there was no statistically significant difference noted between sexes for

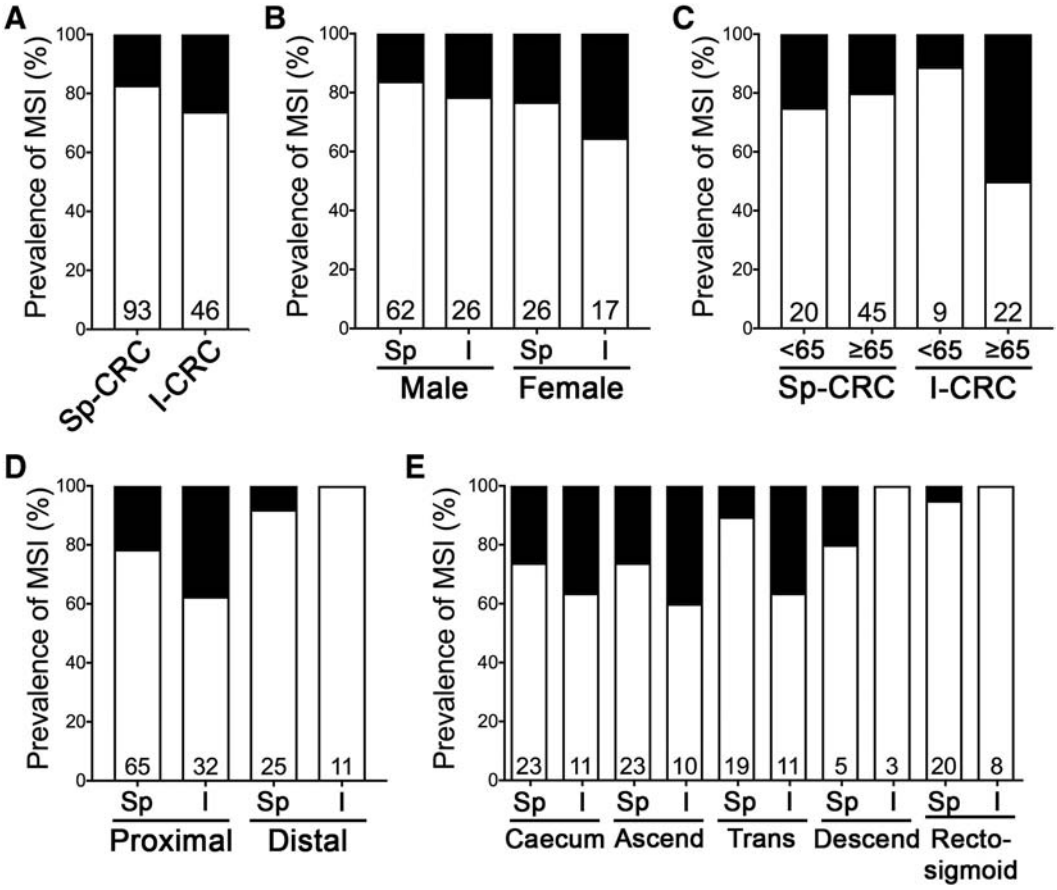

Figure 2. The Prevalence of MSI in Sp- and I-CRCs Stratified by Sex, Age and Anatomic Location. *(A)* Column graph depicting the prevalence of MSI within Sp- and I-CRCs, with MSI^+ samples indicated in black. Numbers at the base of each column identify the total number of patient samples within each group. *(B)* Bar graph presenting the percentage of MSI^+ (black) Sp- and I-CRCs stratified by sex. Numbers (*x*-axis) identify the total number of samples within each group. *(C)* The prevalence of MSI^+ Sp- (black) and I-CRCs in individuals <65, or ≥65 years of age. Numbers (*x*-axis) indicate total samples in each category. *(D)* Prevalence of MSI^+ (black) Sp- and I-CRCs located within the proximal and distal colon; total sample numbers indicated at the base of each column. *(E)* Column graph presenting the overall prevalence of MSI^+ (black) Sp- and I-CRCs within five anatomic locations; caecum, ascending, transverse, descending and recto-sigmoid colon.

Sp- and I-CRCs occurrence (Table 1). Next, we assed potential differences in the prevalence of MSI based on age, and due to the limited sample size, the patient cohort was stratified into two age categories, <65 or ≥65 years of age. In addition, age comparisons were restricted to samples arising within the proximal colon as MSI did not occur within the distal colon. Thus, comparisons of samples isolated from the proximal colon show that the prevalence of MSI within Sp-CRCs is similar for both age categories (25% for <65 versus 20% for ≥65 years of age); however, within the I-CRCs, MSI was 4.5-times more prevalent within the older age category (50%) than the younger category (11%).

Next, the prevalence of MSI was evaluated in Sp- and I-CRCs stratified based on anatomic location and stage. Overall, MSI^+ samples were restricted to tumors isolated form the proximal colon (Figure 2*D*), but were statistically indistinguishable between Sp- and I-CRCs (Table 1). More specifically, 22% and 8% of Sp-CRCs occurred in the proximal and distal colon, respectively, while all 12 MSI^+ I-CRCs (38% of all I-CRCs) occurred within the proximal colon. When further subdivided into the five anatomic locations presented within Figure 2*E*, MSI^+ Sp-CRCs occurred more frequently in the caecum (26%) and ascending colon (26%), while MSI^+ I-CRCs are roughly equally distributed between the caecum (36%), ascending (40%) and transverse (36%) colon – no MSI^+ I-CRCs occurred within the descending or recto-sigmoid colon. Interestingly, a significant difference ($P = .04$) was revealed when comparing stage (Table 1); MSI was most frequently observed in stage II for Sp-CRCs (56%) and stage III for I-CRCs (50%). No significant differences were observed when comparing tumor grade (Table 1). Collectively, the data presented above show that the prevalence of MSI is ~1.6-fold greater in Sp- and I-CRCs, but also show that there is no statistically significant difference between Sp- and I-CRCs with respect to sex, age, tumor location or grade.

MSI and CIN Co-occur in a Subset of Sp- and I-CRCs

To date, no study has simultaneously evaluated the prevalence of CIN and MSI within a single CRC cohort, let alone in I-CRCs. Having previously characterized CIN within this study cohort [1], we are strategically positioned to gain unprecedented insight into their individual and co-occurrence within Sp- and I-CRCs. To accomplish this, we used the CIN data gleaned from the previous CIN study [1] and combined it with the new MSI data generated in the current study. Initially, we sought to identify and compare the prevalence of CIN^+/MSI^-, CIN^-/MSI^+, CIN^+/MSI^+ and CIN^-/MSI^- tumors within the Sp- and I-CRC groups. As shown in Figure 3*A*, 73.3%

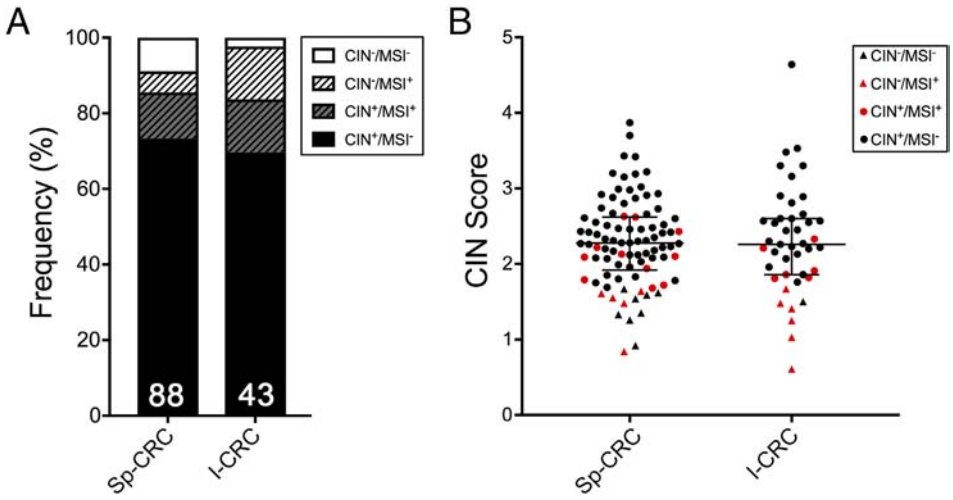

Figure 3. Frequency of CIN and MSI in Sp- and I-CRCs. *(A)* Bar graph presenting the frequency of Sp- and I-CRCs defined as CIN^+/MSI^-, CIN^+/MSI^+, CIN^-/MSI^+ and CIN^-/MSI^- (see key for details). Numbers at the base of each column are the total number of samples within each group. *(B)* Dot plot presenting the mean CS values for each Sp- and I-CRC evaluated. CRCs defined as CIN^+ (CS value ≥1.68) are represented by circles, while CIN^- CRCs are identified with triangles. MSI^+ samples are shown in red irrespective of CIN status. Bars define the interquartile range (25th, 50th and 75th percentiles) of each group.

(66/90 samples) of Sp-CRCs were CIN^+/MSI^-, 5.6% (5) CIN^-/MSI^+, ~12.2% (11) were both CIN^+/MSI^+ and 8.9% (8) were CIN^-/MSI^-. By comparison, of I-CRCs 69.8% (30/43 samples) were CIN^+/MSI^-, 14.0% (6) CIN^-/MSI^+, 14.0% (6) CIN^+/MSI^+ and 2.3% (1) CIN^-/MSI^-. Further, as shown in Figure 3*B*, the majority of MSI^+ Sp- and I-CRCs occurred within tumors with lower CS values. More specifically, 81.3% (13/16) of the MSI^+ Sp-CRCs occurred below the median CS value, while 91.7% (11/12) of the MSI^+ I-CRCs were below the median, with no MSI^+ samples identified within the top CS quartile for either Sp- or I-CRCs. Collectively, these data show that the prevalence of CIN^+/MSI^+ tumors is similar in Sp- and I-CRCs, and that CIN^+/MSI^+ tumors do not exhibit high CS values (i.e., high levels of CIN), but rather, harbor CS values that are typically below the median CS values for each group.

CIN^+/MSI^+ Sp- and I-CRCs Stratified by Sex, Age and Tumor Location

We next sought to assess the prevalence of CIN and MSI in Sp- and I-CRCs based on sex, age and anatomic location. In general, there were no significant differences in the prevalence of CIN^+, MSI^+, or CIN^+/MSI^+ samples between males and females in both Sp- and I-CRCs. As shown in Figure 4*A*, 71% (44/62 samples) of male Sp-CRCs were CIN^+/MSI^-, 4.8% (3) were CIN^-/MSI^+, while 11.3% (7) were CIN^+/MSI^+ and 12.9% (8) were CIN^-/MSI^-. Similarly, 78.6% (22/28 samples) of female Sp-CRCs were CIN^+/MSI^-, 7.1% (2) were CIN^-/MSI^+, while 14.3% (11) were CIN^+/MSI^+ and 0.0% being CIN^-/MSI^-. With respect to the I-CRCs, 73.1% (19/26 samples) of male samples were CIN^+/MSI^-, whereas CIN^-/MSI^+ and CIN^+/MSI^+ samples each occurred in 11.5% of samples (3 each), with only 3.8% (1) sample being CIN^-/MSI^-. Similar trends were also noted for female I-CRCs as 64.7% (11/17 samples) were CIN^+/MSI^-, while CIN^-/MSI^+ and CIN^+/MSI^+ samples were prevalent in 17.6% (3 each) of samples. Interestingly, none of the 45 total female samples were CIN^-/MSI^-. Further, and as shown in Figure 4*B*, the majority of MSI^+ CRC samples (irrespective of sex) occurred within the bottom 50% of CS values for both Sp- or I-CRCs. In fact, all Sp- and I-CRCs, with the exception of only one male Sp-CRC, had CS values that were below the 75th percentile, indicating that MSI^+ is not associated with extreme CS values (i.e. top 25th percentile) in either group or sex.

Next, the prevalence of CIN^+/MSI^+ Sp- and I-CRCs was further scrutinized based on age (Figure 4*C*). Overall, there was 1.8-fold increase in the prevalence of CIN^+/MSI^+ Sp-CRC samples within the <65 (20%; 4/20 samples) relative to the ≥65 (11%; 5/45 samples) years of age group. Interestingly, this ratio was inverted within the I-CRCs and there was a 2.0-fold increase within ≥65 (21.7%; 5/23 samples) relative to the <65 (11.1%; 1/9) years of age group. Although limited sample sizes, these data suggest CIN^+/MSI^+ samples occur more frequently within the younger age group (<65) in Sp-CRCs and more frequently within the older age group (≥65 years of age) of the I-CRCs.

Finally, the prevalence of CIN^+/MSI^+ was assessed in Sp- and I-CRCs stratified by anatomic location. As shown in Figure 4*D*, the majority of CIN^+/MSI^+ Sp- and I-CRCs occur in the proximal colon, with only two CIN^+/MSI^+ Sp-CRCs found within the distal colon; no CIN^+/MSI^+ I-CRCs were observed within the proximal samples. Upon further subdivision of the proximal colon, it was noted that the majority of CIN^+/MSI^+ occur within the ascending colon for both Sp-CRCs (17.4%; 4/23 samples) and I-CRCs (30.0%; 3/10 samples). Collectively, the above findings indicate that minor differences may exist in the prevalence of CIN^+/MSI^+ status based on age, sex and anatomic location.

In the multivariate logistic model, the only factor statistically significantly associated with I-CRC status was the specialty of the physician performing the initial colonoscopy (gastroenterologist vs. non-gastroenterologist odds ratio: 0.25; 95% confidence interval: 0.09–0.70; P = .008).

Discussion

In this study, we evaluated MSI status within a Manitoban CRC cohort comprised of Sp-CRCs and I-CRCs in which we previously assessed CIN [1]. Using a clinically validated, IHC approach, we determined that MSI is ~1.5-times more prevalent within I-CRCs

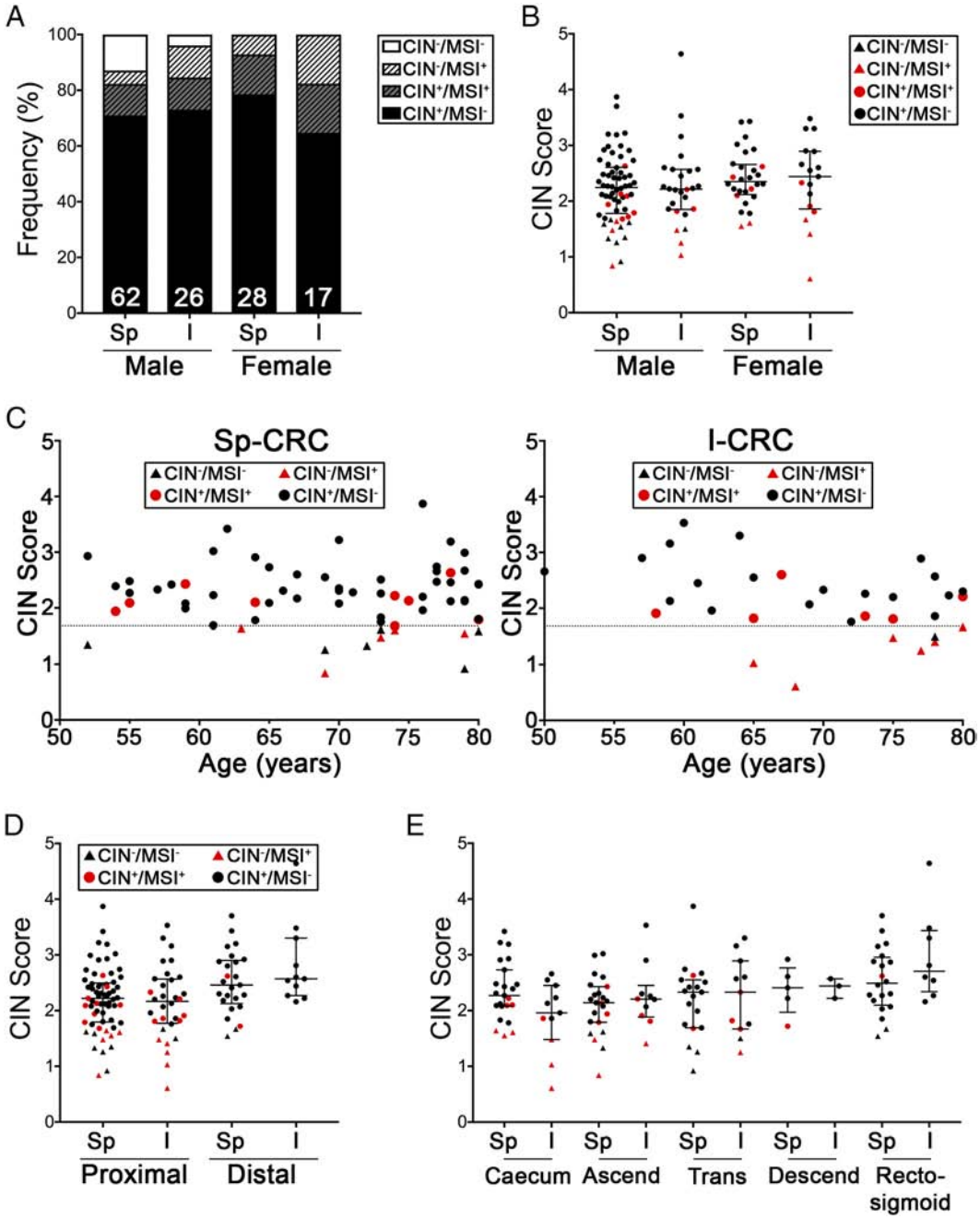

Figure 4. The Prevalence of CIN and MSI within Sp- and I-CRCs Stratified by Sex, Age and Anatomic Location. *(A)* Bar graph presenting the frequency of Sp- and I-CRCs stratified by sex that are CIN^+/MSI^-, CIN^+/MSI^+, CIN^-/MSI^+ and CIN^-/MSI^-. Numbers at the base of each column indicate the total number of samples in each group. *(B)* Dot plot presenting the CS values for Sp- and I-CRCs stratified by sex. CIN^+ (CS value ≥1.68) and CIN^- CRCs are represented by circles and triangles, respectively, while MSI^+ samples are shown in red. Bars identify the interquartile range (25th, 50th and 75th percentiles) within each group. *(C)* Scatter plots presenting the CS values versus age for Sp- (left) and I-CRCs (right). Dotted horizontal lines identify CS = 1.68. *(D)* CS values for Sp- and I-CRCs within the proximal and distal colon. Bars define the interquartile ranges. *(E)* Overall distribution and interquartile ranges of Sp- and I-CRCs isolated from the caecum, ascending, transverse, descending and recto-sigmoid colon.

than Sp-CRC controls. We subsequently showed that MSI occurs more frequently in females than males in both I- and Sp-CRCs groups, and further show that the majority MSI^+ I-CRCs occur in individuals ≥65 years of age. Further, MSI^+ I-CRCs occurred exclusively within the proximal colon, and were roughly equally distributed between the caecum, ascending and transverse colon. Having previously established the prevalence of CIN within this cohort [1], we were uniquely positioned to gain unprecedented insight into the relationship between CIN and MSI within this cohort. Although CIN occurred in ~80% to 85% of Sp-CRCs and I-CRCs [1], the frequencies of CIN^+/MSI^+ tumors are statistically indistinguishable in Sp-CRCs (12%) and I-CRCs (14%). Medical specialty of the physician performing the initial colonoscopy is an independent factor associated with I-CRCs. Collectively, these data suggest that most I- and Sp-CRCs arise through similar pathways. These data further imply that underlying reason accounting for the

high prevalence of I-CRCs in Manitoba is unlikely to be based on distinct biology, but perhaps an issue associated with thoroughness of colonoscopy examination.

Prior to the current study, only three previous studies assessed MSI within I-CRCs [18–20] and no study had ever simultaneously assessed both MSI and CIN within the same cohort. Accordingly, the current study was designed to assess both MSI and CIN within both Sp- and I-CRCs to identify any potential relationship existing between these traditionally distinct pathways [27–29]. A critical component of this study were the criteria used to define I-CRC samples; too long an interval following colonoscopy could inadvertently include Sp-CRCs that arose through traditional means, whereas too short an interval could inappropriately include Sp-CRCs whose date of diagnosis was delayed due to delays in processing and reporting pathology specimens (reviewed in [24]). Thus, we purposefully excluded CRCs diagnosed 1 and 6 months after a colonoscopy from both the Sp-CRC and I-CRC cohorts, and only designated those that occurred 6 to 36 months following a clearing colonoscopy as our I-CRC cohort samples.

Previous studies have shown that MSI is ~3 to 4-times more prevalent in I-CRCs (~45%) than Sp-CRCs (~15%) [18–20]. Our findings also show that MSI is more frequently observed in I-CRCs, but to a lesser extent – only ~1.5-times more frequently in I-CRCs (26%) than Sp-CRCs (17%). The underlying reason(s) accounting for these differences is not readily apparent, although it may be due to differences in the definition of I-CRCs (e.g., 3 years for the current study versus 5 years for other studies; inclusion of CRC occurring within 1–6 months of colonoscopy within different groups in the previous studies), sample selection, or the methodologies used to assess MSI status. For example, the Sawhney et al. [20] evaluated a cohort that was almost exclusively white (99%) male (98%) veterans and Lee et al. [18] evaluated a Korean cohort, while our cohort was population-based (therefore more likely to be reflective of usual clinical practice) and included both males (66%) and females (34%) and although ethnicity was not recorded, it is known to be diverse in Manitoba [1]. Furthermore, while our study employs IHC to assess the presence of four key proteins involved in MMR, others have employed PCR-based approaches of established microsatellites contained within the DNA extracted from fixed tumor samples [18–20]. However, since both approaches are employed clinically to screen for Lynch Syndrome (for which they are equivalent), the fundamental differences in MSI frequencies observed in the different studies [18–20] may simply reflect differences in the cohort and/or the biological material employed for the assessment (i.e., protein versus DNA). We believe there are differences in rates and etiology of I-CRCs occurring in different jurisdictions and practices, with higher rates and population-based samples (usual clinical practice) correlating with less biological differences from Sp-CRCs.

While there is limit knowledge pertaining to the pathogenic origins of I-CRCs, it is now clear that genome instability is highly prevalent within tumor samples suggesting it may be a pathogenic event in most I-CRCs. Traditionally, MSI and CIN were considered mutually exclusive pathways [27,28], though more recent findings show these pathways can co-occur in Sp-CRCs. The results of the current study support these observations by showing that MSI and CIN do co-occur in I-CRCs, albeit in only a small subset of tumors. A further strength of this study is the coupling of pathologic approaches (e.g. CIN analyses and the IHC evaluation of MMR proteins) with clinical, epidemiological and outcome analyses in an emerging transdisciplinary science referred to as molecular pathologic epidemiology (MPE; reviewed in [32]). Briefly, MPE seeks to integrate both molecular and population-based health information to aid in the identification of factors that contribute to the pathology of diseases, including I-CRC. Accordingly, this study has for the first time, combined two epidemiological exposure variables (i.e. CIN and MSI datasets) that can be incorporated into future MPE studies. Thus, greater insight into the molecular factors associated with, and driving I-CRCs may reveal important sub-groups contained within the larger I-CRC context, such as missed Sp-CRCs, synergistic growth conditions, or additional contributors that have yet to be determined. We found interesting differences in I-CRC biology by age, which need to be evaluated in subsequent studies. Our study suggests I-CRC maybe a heterogeneous group, which needs through evaluation of individual CRCs. Collectively, these data along with additional insight they provide, may impact screening programs that would facilitate earlier identification of precursor lesions and improve outcomes for those with I-CRCs.

In conclusion, our study data suggest that MSI^+ tumors are only 1.5-times more prevalent within I-CRCs than Sp-CRCs in a population-based cohort and further show that CIN^+/MSI^+ I-CRCs occur at similar frequency as in Sp-CRCs. Importantly, these data suggest biological differences are unlikely to account for the vast majority of I-CRCs in usual clinical practices, emphasizing the need for enhanced colonoscopy examination, which could lead to reduction of I-CRCs.

Acknowledgements

This study was supported by the Manitoba Tumor Bank, funded in part by the CancerCare Manitoba Foundation and CIHR. We acknowledge the strong support of the Research Institute in Oncology and Hematology, and the CancerCare Manitoba Foundation. The results and conclusions are those of the authors, and no official endorsement by Manitoba Health is intended or should be inferred.

Funding

This work was supported by the Canadian Institutes of Health Research (MOP 115179), CancerCare Manitoba (KJM/HS), and Colon Cancer Canada (HS/KJM).

References

[1] Cisyk AL, Penner-Goeke S, Lichtensztejn Z, Nugent Z, Wightman RH, Singh H, and McManus KJ (2015). Characterizing the prevalence of chromosome instability in interval colorectal cancer. *Neoplasia* **17**, 306–316.

[2] American Cancer Society (2017). Cancer Facts & Figures 2017. Atlanta, GA: American Cancer Society, Inc.; 2017 [Vol.].

[3] Canadian Cancer Society's Advisory Committee on Cancer Statistics (2017). Canadian Cancer Statistics 2017. Toronto, ON: Canadian Cancer Society; 2017 [Vol.].

[4] Markowitz SD and Bertagnolli MM (2009). Molecular origins of cancer: Molecular basis of colorectal cancer. *N Engl J Med* **361**, 2449–2460.

[5] Grady WM and Markowitz SD (2015). The molecular pathogenesis of colorectal cancer and its potential application to colorectal cancer screening. *Dig Dis Sci* **60**, 762–772.

[6] Meissner HI, Breen N, Klabunde CN, and Vernon SW (2006). Patterns of colorectal cancer screening uptake among men and women in the United States. *Cancer Epidemiol Biomark Prev* **15**, 389–394.

[7] Wernli KJ, Hubbard RA, Johnson E, Chubak J, Kamineni A, Green BB, and Rutter CM (2014). Patterns of colorectal cancer screening uptake in newly eligible men and women. *Cancer Epidemiol Biomark Prev* **23**, 1230–1237.

[8] Singh S, Singh PP, Murad MH, Singh H, and Samadder NJ (2014). Prevalence, risk factors, and outcomes of interval colorectal cancers: a systematic review and meta-analysis. *Am J Gastroenterol* **109**, 1375–1389.

[9] Gorski TF, Rosen L, Riether R, Stasik J, and Khubchandani I (1999). Colorectal cancer after surveillance colonoscopy: false-negative examination or fast growth? *Dis Colon Rectum* **42**, 877–880.

[10] Hosokawa O, Shirasaki S, Kaizaki Y, Hayashi H, Douden K, and Hattori M (2003). Invasive colorectal cancer detected up to 3 years after a colonoscopy negative for cancer. *Endoscopy* **35**, 506–510.

[11] le Clercq CM, Bouwens MW, Rondagh EJ, Bakker CM, Keulen ET, de Ridder RJ, Winkens B, Masclee AA, and Sanduleanu S (2014). Postcolonoscopy colorectal cancers are preventable: a population-based study. *Gut* **63**, 957–963.

[12] Leaper M, Johnston MJ, Barclay M, Dobbs BR, and Frizelle FA (2004). Reasons for failure to diagnose colorectal carcinoma at colonoscopy. *Endoscopy* **36**, 499–503.

[13] Pabby A, Schoen RE, Weissfeld JL, Burt R, Kikendall JW, Lance P, Shike M, Lanza E, and Schatzkin A (2005). Analysis of colorectal cancer occurrence during surveillance colonoscopy in the dietary Polyp Prevention Trial. *Gastrointest Endosc* **61**, 385–391.

[14] Pohl H and Robertson DJ (2010). Colorectal cancers detected after colonoscopy frequently result from missed lesions. *Clin Gastroenterol Hepatol* **8**, 858–864.

[15] Robertson DJ, Lieberman DA, Winawer SJ, Ahnen DJ, Baron JA, Schatzkin A, Cross AJ, Zauber AG, Church TR, and Lance P, et al (2014). Colorectal cancers soon after colonoscopy: a pooled multicohort analysis. *Gut* **63**, 949–956.

[16] Ahadova A, von Knebel Doeberitz M, Blaker H, and Kloor M (2016). CTNNB1-mutant colorectal carcinomas with immediate invasive growth: a model of interval cancers in Lynch syndrome. *Familial Cancer* **15**, 579–586.

[17] Arain MA, Sawhney M, Sheikh S, Anway R, Thyagarajan B, Bond JH, and Shaukat A (2010). CIMP status of interval colon cancers: another piece to the puzzle. *Am J Gastroenterol* **105**, 1189–1195.

[18] Lee KW, Park SK, Yang HJ, Jung YS, Choi KY, Kim KE, Jung KU, Kim HO, Kim H, and Chun HK, et al (2016). Microsatellite Instability Status of Interval Colorectal Cancers in a Korean Population. *Gut Liver* **10**, 781–785.

[19] Nishihara R, Wu K, Lochhead P, Morikawa T, Liao X, Qian ZR, Inamura K, Kim SA, Kuchiba A, and Yamauchi M, et al (2013). Long-term colorectal-cancer incidence and mortality after lower endoscopy. *N Engl J Med* **369**, 1095–1105.

[20] Sawhney MS, Farrar WD, Gudiseva S, Nelson DB, Lederle FA, Rector TS, and Bond JH (2006). Microsatellite instability in interval colon cancers. *Gastroenterology* **131**, 1700–1705.

[21] Shaukat A, Arain M, Anway R, Manaktala S, Pohlman L, and Thyagarajan B (2012). Is KRAS mutation associated with interval colorectal cancers? *Dig Dis Sci* **57**, 913–917.

[22] Shaukat A, Arain M, Thaygarajan B, Bond JH, and Sawhney M (2010). Is BRAF mutation associated with interval colorectal cancers? *Dig Dis Sci* **55**, 2352–2356.

[23] Richter JM, Pino MS, Austin TR, Campbell E, Szymonifka J, Russo AL, Hong TS, Borger D, Iafrate AJ, and Chung DC (2014). Genetic mechanisms in interval colon cancers. *Dig Dis Sci* **59**, 2255–2263.

[24] Cisyk AL, Singh H, and McManus KJ (2014). Establishing a biological profile for interval colorectal cancers. *Dig Dis Sci* **59**, 2390–2402.

[25] Hanahan D and Weinberg RA (2011). Hallmarks of cancer: the next generation. *Cell* **144**, 646–674.

[26] Jia M, Gao X, Zhang Y, Hoffmeister M, and Brenner H (2016). Different definitions of CpG island methylator phenotype and outcomes of colorectal cancer: a systematic review. *Clin Epigenetics* **8**, 25.

[27] Lengauer C, Kinzler KW, and Vogelstein B (1997). Genetic instability in colorectal cancers. *Nature* **386**, 623–627.

[28] Lengauer C, Kinzler KW, and Vogelstein B (1998). Genetic instabilities in human cancers. *Nature* **396**, 643–649.

[29] Boland CR and Goel A (2010). Microsatellite instability in colorectal cancer. *Gastroenterology* **138**, 2073–2087.e2073.

[30] Kim YS and Deng G (2007). Epigenetic changes (aberrant DNA methylation) in colorectal neoplasia. *Gut Liver* **1**, 1–11.

[31] Murphy G, Devesa SS, Cross AJ, Inskip PD, McGlynn KA, and Cook MB (2011). Sex disparities in colorectal cancer incidence by anatomic subsite, race and age. *Int J Cancer* **128**, 1668–1675.

[32] Ogino S, Nishihara R, VanderWeele TJ, Wang M, Nishi A, Lochhead P, Qian ZR, Zhang X, Wu K, and Nan H, et al (2016). Review Article: The Role of Molecular Pathological Epidemiology in the Study of Neoplastic and Non-neoplastic Diseases in the Era of Precision Medicine. *Epidemiology* **27**, 602–611.