Association of the Colorectal CpG Island Methylator Phenotype with Molecular Features, Risk Factors, and Family History

Daniel J. Weisenberger1,2, A. Joan Levine3, Tiffany I. Long4, Daniel D. Buchanan5,6, Rhiannon Walters5, Mark Clendening5, Christophe Rosty7, Amit D. Joshi3, Mariana C. Stern3, Loic Le Marchand8, Noralane M. Lindor9, Darshana Daftary10, Steven Gallinger10, Teresa Selander10, Bharati Bapat10, Polly A. Newcomb11, Peter T. Campbell12, Graham Casey3, Dennis J. Ahnen13, John A. Baron14, Robert W. Haile3, John L. Hopper6, Joanne P. Young7,15, Peter W. Laird1,2,4, and Kimberly D. Siegmund3, for the Colon Cancer Family Registry

Abstract

Background: The CpG island methylator phenotype (CIMP) represents a subset of colorectal cancers characterized by widespread aberrant DNA hypermethylation at select CpG islands. The risk factors and environmental exposures contributing to etiologic heterogeneity between CIMP and non-CIMP tumors are not known.

Methods: We measured the CIMP status of 3,119 primary population-based colorectal cancer tumors from the multinational Colon Cancer Family Registry. Etiologic heterogeneity was assessed by case–case study comparing risk factor frequency of colorectal cancer cases with CIMP and non-CIMP tumors using logistic regression to estimate the case–case odds ratio (ccOR).

Results: We found associations between tumor CIMP status and MSI-H (ccOR = 7.6), BRAF V600E mutation (ccOR = 59.8), proximal tumor site (ccOR = 9; all P < 0.0001), female sex [ccOR = 1.8; 95% confidence interval (CI), 1.5–2.1], older age (ccOR = 4.0 comparing over 70 years vs. under 50; 95% CI, 3.0–5.5), and family history of CRC (ccOR = 0.6; 95% CI, 0.5–0.7). While use of NSAIDs varied by tumor CIMP status for both males and females (P = 0.0001 and P = 0.02, respectively), use of multivitamin or calcium supplements did not. Only for female colorectal cancer was CIMP status associated with increased pack-years of smoking (P_trend < 0.001) and body mass index (BMI; P_trend = 0.03).

Conclusions: The frequency of several colorectal cancer risk factors varied by CIMP status, and the associations of smoking and obesity with tumor subtype were evident only for females.

Impact: Differences in the associations of a unique DNA methylation–based subgroup of colorectal cancer with important lifestyle and environmental exposures increase understanding of the molecular pathologic epidemiology of this heavily methylated subset of colorectal cancer. Cancer Epidemiol Biomarkers Prev; 24(3): 1–8. ©2015 AACR.

Introduction

Human colorectal cancer is a worldwide health concern through being a substantial cause of morbidity and mortality. In 2014, there will be an estimated 136,830 new cases of colon and rectal cancers in the United States and about 50,000 deaths (1). People with Lynch syndrome carry germline mutations in mismatch repair genes, primarily MLH1, MSH2, MSH6, and PMS2, and are predisposed to colorectal cancer. However, Lynch syndrome only

1USC Epigenome Center, University of Southern California, Los Angeles, California. 2Department of Biochemistry and Molecular Biology, University of Southern California, Los Angeles, California. 3Department of Preventive Medicine, University of Southern California, Los Angeles, California. 4Department of Surgery, University of Southern California, Los Angeles, California. 5Oncogenomics Group, Genetic Epidemiology Laboratory, Department of Pathology, The University of Melbourne, Parkville, Victoria, Australia. 6Centre for Epidemiology and Biostatistics, The University of Melbourne, Parkville, Victoria, Australia. 7Queensland Institute of Medical Research, Herston, Queensland, Australia. 8Cancer Research Center of Hawaii, University of Hawaii, Honolulu, Hawaii. 9Department of Health Science Research, Mayo Clinic, Scottsdale, Arizona. 10Department of Pathology and Laboratory Medicine, Samuel Lunenfeld Research Institute, Mt. Sinai Hospital, Toronto, Ontario, Canada. 11Epidemiology Department, University of Washington and Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, Washington. 12Epidemiology Research Program, American Cancer Society, Atlanta, Georgia. 13Division of Gastroenterology, University of Colorado School of Medicine, Denver, Colorado. 14Department of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina. 15The Queen Elizabeth Hospital, Woodville, Australia.

Note: Supplementary data for this article are available at Cancer Epidemiology, Biomarkers & Prevention Online (http://cebp.aacrjournals.org/).

D.J. Weisenberger and A.J. Levine contributed equally to this article.

Current address for A.J. Levine and R.W. Haile: Stanford Cancer Institute, Stanford University, Palo Alto, CA.

Current address for A.D. Joshi: Program in Genetic Epidemiology and Statistical Genetics, Department of Epidemiology, Harvard School of Public Health, Boston, MA.

Current address for P.W. Laird: Laboratory of Cancer Epigenetics, Van Andel Research Institute, Grand Rapids, MI.

Corresponding Author: Kimberly D. Siegmund, Department of Preventive Medicine, University of Southern California, 2001 N Soto Street, Los Angeles, CA 90089. Phone: 323-442-1310; Fax: 323-442-2993; E-mail: kds@usc.edu

doi: 10.1158/1055-9965.EPI-14-1161

©2015 American Association for Cancer Research.
accounts for 2% to 5% of all colorectal cancer (reviewed in ref. 2). Most colorectal cancer are thought to result from the accumulation of somatic genetic (3–5) and epigenetic alterations (reviewed in refs. 6, 7) often associated with gender, age, diet, lifestyle habits, and environmental exposures (8–15). The majority of non-Lynch syndrome colorectal cancer are located in the distal (descending left) colon and rectum and are enriched for KRAS mutations. In contrast, approximately 15% of colorectal cancer are predominantly located in the proximal (ascending right colon) of older age females with enrichment for BRAFV600E mutations, high levels of microsatellite instability (MSI-H), MLH1 epigenetic silencing, and the CpG island methylator phenotype (CIMP; refs. 16–23).

CIMP tumors were first identified in 1999 by Toyota and colleagues (22) and are thought to develop via the serrated neoplasia pathway (17, 24). Using MethyLight technology, we identified CIMP from a screen of 195 gene loci and presented a 5-gene diagnostic panel to identify CIMP tumors: CACNA1C, IGF2, NEUROG1, RUNX3, and SOCS1 (23). Using this panel, we showed that CIMP tumors are preferentially located in the proximal colon and are associated with the BRAFV600E mutation, MSI-H, increasing age, female gender, and overall improved patient outcome (23). CIMP has also been described in recent reports using genome-scale technologies (25–28).

The associations of colorectal cancer with environmental exposures are well documented. The risk of colorectal cancer is positively associated with smoking, alcohol use, obesity, and physical inactivity. A recent report of genome-scale DNA methylation in normal colorectal tissues suggests that in women, obesity and smoking increase DNA methylation at genes hypermethylated in cancer but that the use of aspirin and hormone replacement therapies is correlated with a reduction in DNA hypermethylation (29).

In this study, we sought to confirm previous associations for colorectal CIMP tumors and evaluate whether the distributions of known colorectal cancer risk factors differ in CIMP and non-CIMP tumors, including family colorectal cancer history, physical activity, smoking history, history of alcohol use, use of NSAID, and body mass index (BMI). We used the resources of the Colon Cancer Family Registry, an international, multi-institutional consortium, and performed CIMP assays on 3,119 population-based primary colorectal cancer. Accompanying these samples are a rich data resource of family history and the level of use/intake of the known colorectal cancer risk factors. We evaluated etiologic heterogeneity of these risk factors using a case–case study, directly comparing the distribution of known colorectal cancer risk factors between CIMP and non-CIMP tumor subtypes.

Materials and Methods

Study population

Data for this study were obtained through the Colon Cancer Family Registry (C-CFR), a National Cancer Institute–funded registry of colorectal cancer cases, family members, and population-based controls, which used standardized methods for data collection and genotyping. Detailed information about the C-CFR can be found elsewhere (30) and at coloncfr.org. Recruitment at individual C-CFR sites was described previously (30). Participants for this study were recruited from 6 centers: the University of Southern California (USC) Consortium (Arizona, Colorado, New Hampshire, Minnesota, North Carolina, and Los Angeles, California), University of Hawaii (Honolulu, HI), Fred Hutchinson Cancer Research Center (FHCRC; Seattle, WA), Mayo Clinic (Rochester, MN), Cancer Care Ontario (Toronto, Canada), and University of Melbourne (Victoria, Australia) using population-based ascertainment strategies. All centers except FHCRC oversampled case probands with first-degree relatives reporting colorectal cancer or colorectal cancer case probands diagnosed younger than 50 years to target families with increased colorectal cancer risk. First-degree and some second-degree relatives with colorectal cancer were also recruited from families with multiple colorectal cancer cases. In this study, we included only colorectal cancer cases recruited from 1997 to 2002 (30), who signed a written informed consent and completed the risk factor questionnaire (RFQ) within 5 years of their colorectal cancer diagnosis.

Risk factor and clinical data

We obtained risk factor data from the completed RFQs. Age at the time of enrollment was categorized as a 3-category variable: <50, 51–69, and ≥70 years. Family history of colorectal cancer was self-reported and was considered positive if the case reported colorectal cancer in one or more first-degree family members (e.g., parents, siblings or children). Cigarette smoking pack-years was estimated by multiplying the average reported cigarettes smoked per day times the total years of smoking and was categorized with never-smokers as the referent group. BMI was categorized into 3 groups based on WHO criteria for overweight and obesity: 18–24.9, 25–29.9, and ≥30 kg/m².

The average weekly hours of physical activity was derived for each of 10 common activities within 3 age periods during adulthood (20–29, 30–49, and ≥50 years). Average mode-specific minutes per week, computed using responses to total number of years and months the activity was conducted and its typical duration per week, was multiplied by the mode’s average MET cost per activity, transverse colon, and splenic flexure. Tumors were labeled as distal colon if located in the descending colon, sigmoid colon, and the region overlapping the colon and rectum. Tumors were labeled as rectal if located in the rectum or rectosigmoid junction.

Supplement intake was a 3-level variable (current user/former user/non-user) with a “user” answer indicating ever use ≥2 times/week for more than a month and use within 1 year before cancer diagnosis. NSAID use was coded as “user” if the subject used either aspirin or ibuprofen over the same time period and ‘non-user’ if neither was used. Former users were users who had stopped using supplement or NSAID more than 1 year before cancer diagnosis.

Hormone replacement therapy (HRT) use was coded as yes if the subject answered “yes” to the question “have you ever used a pill or patch form of hormone replacement therapy for 6 months or longer” for any hormone replacement preparation (estrogen only or estrogen + progesterone).

Tumor site was abstracted from pathology reports and/or state or provincial cancer registries and coded using International Classification of Diseases for Oncology, third edition codes. Tumors were labeled as proximal colon if located in the cecum, ascending colon, hepatic flexure, transverse colon, and splenic flexure. Tumors were labeled as distal colon if located in the descending colon, sigmoid colon, and the region overlapping the colon and rectum. Tumors were labeled as rectal if located in the rectum or rectosigmoid junction.
Sample receipt and processing
We requested colorectal tumor specimens from all population-based case probands recruited in 1997–2002 as well as their colorectal cancer–affected first-, second-, and third-degree relatives. This provided a total of 3,970 specimens, out of which we received 3,732 (94%) formalin-fixed, paraffin-embedded (FFPE) tissues. Specifically, we received 2 unstained 5-μm tissue sections embedded in paraffin from each tumor on positively charged ‘plus’ glass slides without c overslips.

Slides were randomized to avoid batch effects attributed to source site and reagents. We deparaffinized each slide, micro-dissected tumor tissues, and extracted genomic DNA as previously described (32). Proteinase K was inactivated by heating at 100°C for 10 minutes. An aliquot was then removed for bisulfite conversion using the Zymo EZ-96 DNA Methylation Kit (Zymo Research) as specified by the manufacturer. CIMP status in each sample was determined using a 5-gene MethyLight assay (as described previously (23)). All MethyLight CIMP assays were performed using a control reaction specific for ALU repeats as a means of normalizing for input bisulfite DNA amounts. MethyLight data were organized as percent of methylated reference (PMR) value. Tumors were classified as CIMP if ≥3 of 5 genes gave PMR ≥ 10, and non-CIMP if ≤2 genes gave PMR ≥ 10, as described previously (23). Out of the 3,732 samples processed, 46 (1.2%) failed the assay. For a subset of 25 tumors with 2 independent samples analyzed, 24 pairs were concordant for non-CIMP and 1 pair was discordant. In later analyses, the tumor with discordant results was classified as CIMP.

The processed samples yielded a total of 3,660 colorectal cancer with CIMP results: 3,544 primary colorectal cancer from case probands and 116 colorectal cancer from affected relatives. Associations between tumor CIMP status and demographic, molecular, and environmental risk factors were performed using the population-based colorectal cancer samples from case probands. Of these primary colorectal cancer, 108 case probands (3.0%) were excluded for having been interviewed more than 5 years after diagnosis, 203 (5.7%) for missing RFQ data, and 104 (3.2%) for missing tumor site data or sampling weights (described in Statistical methods section). The final analysis included 3,119 primary colorectal cancer. The CIMP results for the 116 tumors from affected relatives were used to study the concordance for CIMP in tumors from affected relatives.

KRAS and B-raf mutation testing

The somatic T>A mutation at nucleotide 1,799 causing the V600E mutation in B-raf was determined using a fluorescent allele-specific PCR assay that amplified a 97-bp product for the mutant allele (A1799) and a 94-bp product for the wild-type allele (T1799), as previously described (33). Positive controls were run in each experiment and 10% of samples were replicated with 100% concordance. KRAS mutation analysis of codons 12 and 13 was performed using direct Sanger sequencing of a 169-bp PCR amplified product as previously described (34). The larger amplification size for KRAS analysis compared with B-raf*V600E contributed to a slightly higher proportion of the FFPE tumor DNA samples failing to amplify for the KRAS assay compared with B-raf*V600E assay.

MSI testing

MSI was tested using DNA from tumor and matched normal tissue as described (35) using 10 microsatellite loci (BAT25, BAT26, BAT40, BAT34C4, D5S346, D17S250, ACTC, D18S55, D10S197, and MYCL). Samples were classified as MSI-H if >30% showed instability, MSS if no markers showed instability, and MSI-L otherwise. Tumor classification was based on ≥4 interpretable markers.

Statistical methods

Contingency tables present the frequency of patient and tumor characteristics by tumor CIMP status. All analyses were weighted based on the (inverse) sampling probability that the case proband was recruited into the registry to ensure the numbers represent the entire population of colorectal cancer cases at each study site. Subjects were included from all sites except Hawaii, because their sampling design precluded this type of weighted analysis. Frequencies are based on the weighted number of tumors in each category.

We tested for differences in distributions of individual risk factors by CIMP status using a case–case analysis. Case–case odds ratios (ccOR) and 95% confidence intervals (CI) were estimated using standard logistic regression, with weights to correct for sampling bias. These ccORs represent the relative odds for the risk factor in CIMP colorectal cancer compared with that in non-CIMP colorectal cancer and cannot be interpreted in terms of the magnitude of the risk for either tumor phenotype (36). The case–case analysis was the most powerful for testing etiologic heterogeneity of tumor subtype, as it was not affected by heterogeneity due to the recruitment and use of different control types (related or unrelated) by different C-CFR centers. Models were stratified by sex and adjusted for age and tumor site. Analyses of proximal tumors only yielded similar results, as the low numbers of CIMP in distal and rectal tumors precluded our ability to estimate separate ccORs by tumor site. We tested linear trend by modeling the levels of the ordered categorical variable as continuous. Interaction P values were obtained by including interaction terms (e.g., sex × pack-year category) in the model and using a multiple degree of freedom test. Statistical significance was defined as a Wald test P ≤ 0.05 in a 2-sided test. All statistical analyses were performed using SAS 9.3 software (SAS Institute Inc.).

Results

Characteristics of study population

After weighting, the 3,119 patients with colorectal cancer in this study represented an estimated 6,253 colorectal cancer cases. The estimated frequency of CIMP colorectal cancer was 12.6%, with frequencies ranging from 7% to 18% depending on the C-CFR study represented an estimated 6,253 colorectal cancer cases. The study represented an estimated 6,253 colorectal cancer cases. The estimated frequency of CIMP colorectal cancer was 12.6%, with frequencies ranging from 7% to 18% depending on the C-CFR study population (Supplementary Table S1). CIMP colorectal cancer was associated with increased patient age (P < 0.0001) and Australia and USC, the study populations with the lowest frequencies of CIMP colorectal cancer also had the lowest averages for age of colorectal cancer diagnosis (data not shown). CIMP colorectal cancer frequency varied by sex (16.8% in females versus 9.3% in males, P = 0.0001) and was statistically significantly associated with location in the proximal colon in both males and females (Table 1). In addition, we observed variation in CIMP prevalence by race (Supplementary Table S1). In African Americans, the CIMP prevalence was 4.5% and in Asians it was 4.0%, compared with 13.4% in non-Hispanic Whites and 12.3% in
Patient age
c<50 y 18 (5.5) 465 (14.7) 1.0
51–69 y 185 (56.7) 1,842 (58.1) 2.76 (1.67–4.58) 0.0001
>70 y 124 (37.9) 865 (27.3) 3.35 (1.99–5.63) 0.0001
Trend 1.52 (1.25–1.84) 0.0001

Tumor site
dProximal 241 (73.8) 908 (28.7) 8.48 (5.97–12.10) 0.0001
Distal 48 (14.7) 1,037 (32.7) 1.49 (0.97–2.31) 0.07
Rectal 38 (11.5) 1,226 (38.7) 1.0

BRAF (V600E)
eMutated 148 (45.4) 55 (1.8) 35.7 (23.4–52.5) 0.0001
Not mutated 178 (54.6) 3,084 (98.3) 1.0

KRASf
fMutated 95 (34.3) 814 (32.1) 0.76 (0.58–1.0) 0.06
Not mutated 182 (65.7) 1,226 (67.9) 1.0

MSI status
gMSI-H 122 (37.5) 275 (7.9) 3.86 (2.86–5.20) <0.0001
MSI-L 46 (14.3) 481 (15.8) 1.28 (0.90–1.82) 0.17
MSS 158 (48.4) 2,394 (78.9) 1.0

Missing 0 33

Interaction P (df)

As defined in the text, the sampling weights are the inverse of the sampling fraction that corrected for the biased sampling of case probands by age, race, and family history.

Defined as a PMR ≥ 10 for at least 3 of 5 genes: CACNA1G, IGF2, NEUROGI, RUNX3, and SOCS1.

Logistic regression model using proband weights and controlling for tumor site.

Logistic regression model using proband weights and controlling for age (<50, 51–69, >70 years).

Logistic regression model using proband weights and controlling for age (<50, 51–69, >70 years) and tumor site.

Hispanics. CIMP prevalence was significantly lower in African Americans (P = 0.0098) and Asians (P = 0.0182).

Association of CIMP status with BRAF mutation, KRAS mutation, and MSI

In screening for known KRAS and BRAF mutations in the sample cohort, we found a high frequency of the BRAFV600E mutation for CIMP proband tumors (63.8%) but not non-CIMP proband tumors (2.1% Supplementary Table S1). KRAS mutations were more prevalent for non-CIMP compared with CIMP colorectal cancer (33.3% vs. 21%; P = 0.0001). These associations remained significant after controlling for age, sex, and tumor site (adjusted cOR = 59.8; 95% CI, 45.8–78.0 for BRAF and adjusted cOR = 0.44; 95% CI, 0.35–0.54 for KRAS). There was a strong mutual exclusivity of BRAF and KRAS mutations in the tumor cohort, with only 2 CIMP colorectal cancer and one non-CIMP colorectal cancer displaying mutations in both genes. This could be explained for CIMP colorectal cancer by variations in BRAF and KRAS mutation frequency by age. BRAFV600E mutation frequencies for CIMP colorectal cancer were 36%, 59%, and 75% in patients diagnosed at <50, 50–69, and >70 years. KRAS mutation frequencies for the same subgroups were 26%, 29%, and 10%. For CIMP colorectal cancer, 58.7% were MSI-H, 11.2% MSI-L, and 30.2% MSS. For non-CIMP tumors, these figures were 10.7% MSI-H, 17.8% MSI-L, and 71.5% MSS.

The associations between CIMP status and BRAF, KRAS, and MSI-H were stronger for females than males (Table 1, all interaction P = 0.0012). The BRAFV600E mutation occurred in 77.3% of CIMP colorectal cancer for females and 45.4% of CIMP colorectal cancer for males; KRAS mutation appeared in only 11.3% of CIMP colorectal cancer for females versus 34.3% of the same for males and MSI-H occurred in 68.4% of CIMP colorectal cancer for females versus 43.2% of the same for males. KRAS mutation data was missing for 16.7% of CIMP colorectal cancer and 20% non-CIMP colorectal cancer (P = 0.0017; Supplementary Table S1).

Association of CIMP with known risk factors of colorectal cancer

Using the available clinical history and lifestyle information, we next determined whether CIMP correlated with known colorectal cancer risk factors, including smoking history, alcohol use, physical activity, BMI, and family colorectal cancer history (Table 2). A CIMP colorectal cancer was negatively correlated with family history for both men and women (both P < 0.001), occurring more often in cases without a family history of colorectal cancer. However, only 2 of the 94 colorectal cancer affected relatives (2%) were concordant for CIMP colorectal cancer and 16 were discordant (17%; Supplementary Table S1). Limited in power and not statistically significant, this reflected a 2-fold higher frequency of CIMP colorectal cancer for affected relatives of a proband with CIMP colorectal cancer compared with a proband with non-CIMP colorectal cancer (25% vs. 12%).

We found associations of CIMP status with smoking and BMI only for female cases (interaction P = 0.0002 and 0.0001, respectively). We observed a significant trend of increased frequency of smoking in women with CIMP colorectal cancer compared with those with non-CIMP colorectal cancer (P trend = 0.0001); no such association was observed for men (P trend = 0.18). With respect to BMI, CIMP was inversely associated with overweight status for men (BMI, 25–29.9 kg/m²), but there was not a significant trend across BMI groups (P trend = 0.13). For female cases, both the overweight (P = 0.03) and obese (P = 0.0001) groups showed an increased frequency of having CIMP colorectal cancer and the trend was significant (P trend = 0.0001). Alcohol use did not show heterogeneity by CIMP subgroup in men with CR colorectal cancer, but alcohol use in women presented lower frequencies.
Table 2. Associations between CIMP and selected CRC risk factors by gender

<table>
<thead>
<tr>
<th>Risk Factor</th>
<th>Male-weighted (OR)</th>
<th>Female-weighted (OR)</th>
<th>Interaction P (df)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family CRC history</td>
<td>CIMP (%)</td>
<td>OR (95% CI)</td>
<td>P</td>
</tr>
<tr>
<td>None</td>
<td>280 (86.4)</td>
<td>2.534 (79.9)</td>
<td>1.0</td>
</tr>
<tr>
<td>≥1</td>
<td>44 (13.6)</td>
<td>0.55 (0.39-0.78)</td>
<td>0.0007</td>
</tr>
<tr>
<td>Smoking (pack-years)*</td>
<td>0</td>
<td>90 (28.0)</td>
<td>894 (29.0)</td>
</tr>
<tr>
<td>1–10</td>
<td>24 (7.5)</td>
<td>456 (14.7)</td>
<td>0.42 (0.26-0.68)</td>
</tr>
<tr>
<td>11–20</td>
<td>53 (16.6)</td>
<td>466 (15.1)</td>
<td>1.32 (0.90-1.92)</td>
</tr>
<tr>
<td>21–40</td>
<td>83 (25.8)</td>
<td>621 (20.1)</td>
<td>1.26 (0.90-1.76)</td>
</tr>
<tr>
<td>≥40</td>
<td>72 (22.2)</td>
<td>654 (21.2)</td>
<td>0.94 (0.66-1.33)</td>
</tr>
<tr>
<td>Trend</td>
<td>1.06 (0.98–1.14)</td>
<td>0.18</td>
<td></td>
</tr>
<tr>
<td>Alcohol (drinks/wk)</td>
<td>0</td>
<td>103 (33.0)</td>
<td>1.067 (35.0)</td>
</tr>
<tr>
<td>1</td>
<td>170 (54.3)</td>
<td>1.070 (55.9)</td>
<td>0.94 (0.72-1.23)</td>
</tr>
<tr>
<td>≥1</td>
<td>40 (12.6)</td>
<td>281 (9.1)</td>
<td>1.39 (0.92-2.09)</td>
</tr>
<tr>
<td>Trend</td>
<td>1.10 (0.90-1.34)</td>
<td>0.34</td>
<td></td>
</tr>
<tr>
<td>BMI, kg/m²</td>
<td>18–24.9</td>
<td>95 (29.1)</td>
<td>737 (23.2)</td>
</tr>
<tr>
<td>25–29.9</td>
<td>147 (45.0)</td>
<td>1.732 (54.6)</td>
<td>0.51 (0.38-0.69)</td>
</tr>
<tr>
<td>≥30</td>
<td>84 (25.8)</td>
<td>703 (22.6)</td>
<td>0.78 (0.56-1.08)</td>
</tr>
<tr>
<td>Trend</td>
<td>0.87 (0.73–1.04)</td>
<td>0.13</td>
<td></td>
</tr>
<tr>
<td>Physical activity</td>
<td>0–5.7</td>
<td>86 (28.1)</td>
<td>688 (22.6)</td>
</tr>
<tr>
<td>5.8–14.5</td>
<td>59 (19.4)</td>
<td>616 (20.3)</td>
<td>0.76 (0.53–1.10)</td>
</tr>
<tr>
<td>14.6–30.8</td>
<td>68 (22.3)</td>
<td>766 (25.5)</td>
<td>0.77 (0.54–1.09)</td>
</tr>
<tr>
<td>≥30.8</td>
<td>92 (30.2)</td>
<td>963 (31.6)</td>
<td>0.87 (0.63–1.20)</td>
</tr>
<tr>
<td>Trend</td>
<td>0.96 (0.86–1.07)</td>
<td>0.46</td>
<td></td>
</tr>
</tbody>
</table>

*As defined in the text, the sampling weights are the inverse of the sampling fraction which corrected for the oversampling of case probands by age, race, and family history.

*Defined as a PMR ≥ 10 for at least 3 of 5 genes: CACNA1G, IF2G, NEUROGI, RUNX3, and SOCSI.

*95% confidence limits estimated using logistic regression and controlling for age (≤50, 51–69, ≥70 years) and tumor site.

*The subject reported a history of CRC in one or more first-degree relatives (parents, siblings, or children).

*Number of reported cigarettes per day multiplied by years of smoking.

*Met-hours for 10 different physical activities were summed across up to three age groups (<30/31–49/50) based on subjects' age at the time of the questionnaire and the mean of the total MET-hours per week.

Association of CIMP with prediction diagnosis use of vitamin supplements, NSAIDs, and hormone therapies

We also evaluated use of multivitamins, calcium supplements, and NSAIDs before colorectal cancer diagnosis in CIMP and non-CIMP cancers for men and women separately and use of HRT by CIMP in women with colorectal cancer (Table 3). Multivitamins or calcium supplement were not associated with CIMP subtype for men or women; however, men and women who used NSAIDs before diagnosis showed an increased frequency of CIMP colorectal cancer (\( P = 0.0001 \) for men; \( P = 0.02 \) for women; Table 3) The association between CIMP status and NSAID use varied between men and women (\( P_{\text{interaction}} = 0.0008 \)). The small increase in the frequency of CIMP colorectal cancer with HRT use for women was not statistically significant (\( P = 0.17 \)).

Discussion

The global health concern regarding colorectal cancer necessitates an understanding of the contributions of family history and modifiable risk factors to the onset of disease. Because colorectal cancer can be classified into different molecular groups, we were specifically interested in whether CIMP colorectal cancer, as defined by DNA methylation analyses, is differentially associated with lifestyle, obesity status, and/or family history compared with CIMP-negative tumors. We took advantage of the extensive sample collection of the C-CFR, together with patient information, to determine how CIMP status correlates with known colorectal cancer risk factors in a large population-based setting. In this case–case analysis, associations between risk factor and CIMP status indicate etiologic heterogeneity between CIMP and non-CIMP tumors and does not inform us on direction of risk relative to nondiseased individuals. Furthermore, lack of association suggests no evidence of etiologic heterogeneity between the cancer subtypes.

Our data are in general agreement with previous reports that CIMP was more common in women with colorectal cancer, patients with later age of diagnosis, and colorectal cancer located in the proximal colon (22, 23, 25, 26). Furthermore, our data showed that CIMP CRC occurred more often for patients without a family history of colorectal cancer and that modifiable risk factors may contribute differently to CIMP tumor development. Several risk factors showed different distributions in CIMP and non-CIMP tumors, with some of the associations modified by gender. For instance, smoking was associated with frequency of CIMP-positive tumors for women but not men. There was also a significant gender difference for the association between CIMP status and NSAID use. Finally, while CIMP was nonsignificantly inversely correlated with BMI for men, both overweight and obese statuses were positively correlated with CIMP status for women with a significant trend as BMI increases. These differences may not be due to female...
hormones in this population given that a history of HRT was not significantly associated with CIMP status.

Our data show significant variation in smoking by CIMP status in women, but not in men, with colorectal cancer. This agrees with the results of other studies (9, 21). In a women-only study, age-related methylation of CpG islands in normal mucosa was confined to the proximal colon in the presence of smoking (29). However, Worthley and colleagues reported no difference in methylation status of a panel of CIMP markers in the normal colon between smokers and nonsmokers when looking in men and women combined (37). Given our results, the analysis of the smoking/tumor phenotype association separately by gender is indicated.

Aspirin and other NSAIDs are protective against colorectal neoplasia (38). In a recent study of normal colorectal tissues from women, the use of aspirin and HRT resulted in suppressed rates of DNA methylation gain at sites commonly hypermethylated in CRC (29). In a study of advanced serrated polyps, the acknowledged precursors of CIMP colorectal cancer, aspirin was associated with a decreased risk of developing these lesions in the proximal colon (39). In our population, NSAID use was significantly more frequent in CIMP colorectal cancer than non-CIMP colorectal cancer for both men and women, suggesting that NSAID use is either not as protective against CIMP colorectal cancer as it is for non-CIMP colorectal cancer CRC or it increases risk of CIMP colorectal cancer. Slattery and colleagues reported significant protective effects for NSAIDs that were similar for both CIMP-low (0 or 1 marker methylated) and CIMP-high tumors (>2 of 5 markers methylated; ref. 40). The CIMP markers used in that study were substantially different from ours, and there were notable differences between these 2 marker sets in a study comparing them directly (23). In our study, NSAID use was missing for more study participants with CIMP colorectal cancer than with non-CIMP colorectal cancer (6% vs. 2%), which if not missing at random, could introduce some bias in our reported frequencies. Whether NSAID use affects colorectal cancer risk differently for the CIMP subset of tumors, and a possible interaction with gender, needs to be assessed in more study populations before any conclusions can be drawn.

Subsequent to this study, CIMP has been subcategorized into 2 groups, CIMP-high (CIMP-H) and CIMP-low (CIMP-L). In addition, the CIMP2 subgroup was also identified, which has similarities to CIMP-L tumors (41). CIMP-H is representative of classic CIMP, with MSI-H, the BRAF\textsuperscript{V600E} mutation and extensive DNA hypermethylation of a subset of CpG islands (25, 26). Alternatively, CIMP-L tumors, first described by Ogino and colleagues (42), display attenuated DNA methylation of CIMP-defining loci, but these tumors are enriched for KRAS mutations and are generally chromosome stable. Recently, The Cancer Genome Atlas (TCGA) reported CIMP-H in about 15% of colorectal tumors, the majority of which also showed elevated mutation rates (hypermutated) and few somatic copy number alterations (25). The MethyLight panel used here is analogous to the CIMP-H subtype. While our study did not characterize CIMP-L status, previous findings demonstrating the nonassociation of CIMP-L with smoking in colorectal tumors are intriguing and may suggest that there are molecular features altered between CIMP-H and CIMP-L tumors that may help to explain these different relationships. Although BRAF\textsuperscript{V600E} mutation and MLH1 DNA hypermethylation are both highly associated with CIMP, only 64% of CIMP tumors harbored the BRAF\textsuperscript{V600E} mutation and about 50% were MSI-H. Small differences from frequencies in other studies might be

### Table 3. Associations between prediagnosis supplement use and CIMP status by gender

<table>
<thead>
<tr>
<th>CIMP (%)</th>
<th>Male-weighted n*</th>
<th>Female-weighted n</th>
<th>P</th>
<th>Interaction P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CIMP (%)</td>
<td>Non-CIMP (%)</td>
<td>OR (95% CI)*</td>
<td>P</td>
</tr>
<tr>
<td>Prediagnosis multivitamins(\textsuperscript{a})</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-user</td>
<td>180 (56.3)</td>
<td>1,734 (55.1)</td>
<td>1.0</td>
<td>206 (45.0)</td>
</tr>
<tr>
<td>Former user</td>
<td>41 (12.9)</td>
<td>638 (20.4)</td>
<td>0.63 (0.44–0.91)</td>
<td>0.01</td>
</tr>
<tr>
<td>User</td>
<td>99 (30.9)</td>
<td>764 (24.4)</td>
<td>1.18 (0.90–1.55)</td>
<td>0.24</td>
</tr>
<tr>
<td>Missing</td>
<td>7</td>
<td>46</td>
<td></td>
<td>7</td>
</tr>
<tr>
<td>Prediagnosis calcium(\textsuperscript{a})</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-user</td>
<td>291 (91.4)</td>
<td>2,826 (89.9)</td>
<td>1.0</td>
<td>242 (53.3)</td>
</tr>
<tr>
<td>Former user</td>
<td>10 (3.1)</td>
<td>152 (4.8)</td>
<td>0.75 (0.38–1.47)</td>
<td>0.40</td>
</tr>
<tr>
<td>User</td>
<td>17 (5.5)</td>
<td>165 (5.2)</td>
<td>0.91 (0.53–1.54)</td>
<td>0.71</td>
</tr>
<tr>
<td>Missing</td>
<td>8</td>
<td>30</td>
<td></td>
<td>9</td>
</tr>
<tr>
<td>Prediagnosis NSAIDs(\textsuperscript{a})</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-user</td>
<td>105 (34.3)</td>
<td>1,748 (56.2)</td>
<td>1.0</td>
<td>227 (52.0)</td>
</tr>
<tr>
<td>Former user</td>
<td>74 (24.3)</td>
<td>607 (19.5)</td>
<td>1.75 (1.27–2.42)</td>
<td>0.0007</td>
</tr>
<tr>
<td>User</td>
<td>127 (41.6)</td>
<td>753 (24.2)</td>
<td>2.30 (1.72–3.07)</td>
<td>0.0001</td>
</tr>
<tr>
<td>Missing</td>
<td>21</td>
<td>63</td>
<td></td>
<td>28</td>
</tr>
</tbody>
</table>

*As defined in the text, the sampling weights are the inverse of the sampling fraction which corrected for the oversampling of case probands by age, race, and family history.

†Defined as a PMR \textsuperscript{d} > 10 for at least 3 of 5 genes: CACNA1G, IG2F, NEUROG1, RUNX3, and SOCS1.

\(\textsuperscript{a}\)ORs and 95% confidence limits estimated using logistic regression and controlling for age (<50, 51–69, >70 years) and tumor site.

\(\textsuperscript{b}\)Users were those that answered “yes” to the question “have you ever used [supplement] at least 2 times a week for more than a month” and indicated that they were taking that supplement 1 year before CRC diagnosis. Former users include an unknown number of subjects who began using the supplement after colorectal cancer diagnosis.

\(\textsuperscript{c}\)Users were defined as NSAID users if they had used either aspirin or ibuprofen and nonusers if they had not used either at least 1 year before colorectal cancer diagnosis. Former users include an unknown number of subjects who began using the supplement after colorectal cancer diagnosis.

\(\textsuperscript{d}\)Answered “yes” to the question “have you ever used a pill or patch form of hormone replacement therapy” for any hormone replacement preparation (estrogen only or estrogen + progesterone) for 6 months or longer.

\(\textsuperscript{e}\)Refers to the proximal colon in the presence of smoking (29).

\(\textsuperscript{f}\)Refers to the PMR (995.5–10,500).
explained by a different average age of diagnosis (43–48). This suggests that the use of MSI-H status, MLHI DNA methylation, or BRAF mutation status as a surrogate for CIMP will result in misclassification of CIMP status. Also, several CIMP marker panels have been developed, as the initial Toyota report in 1999 (22), and although the 5-gene CIMP panel used in our study was chosen as a definitive panel, reports using other panels have been published (28, 49). Sensitivities and specificities may differ between panels, contributing to varying CIMP calls. In addition, these findings have some implications for understanding which types of serrated polyps give rise to CIMP colorectal cancer. Although the canonical serrated neoplasia pathway has its founding in the BRAF-mutated sessile serrated adenoma/polyp, other pathways to malignant transformation are needed to explain the diversity of CIMP colorectal cancer subtypes in this study, including KRAS-mutated CIMP colorectal cancer, which has been previously thought to be rare (23, 50). Some of the non-BRAF–mutated CIMP colorectal cancer may harbor mutations in PIK3CA (51).

The strengths of our study include its large size and population-based sample and the use of a set of well-characterized markers to define CIMP status, thereby minimizing misclassification. The risk factor data were standardized across the different tumor collection sites using validated questions. However, we did not characterize associations with respect to CIMP-L status. To the extent that risk factors for CIMP-L cases are similar to those for CIMP-H we will have underestimated associations by including exposed CIMP-L cases in the non-CIMP group. However, we cannot predict the direction of bias in cases where risk factors for CIMP-L are significantly different from those for CIMP-H. Future studies should evaluate associations of risk factors with CIMP-L once validated marker panels are developed.

In conclusion, we have used the large, population-based inventory of primary colorectal tumors from the C-CFR to analyze the associations between common colorectal cancer risk factors and tumor CIMP status to assess etiologic heterogeneity in cancer subtypes. The findings in this study show differential lifestyle and risk factor contributions to a subset of colorectal tumors with unique molecular characteristics.

Disclosure of Potential Conflicts of Interest

D.J. Weisenberger has ownership interest (including patents) in and is a consultant/advisory board member for Zymo Research, Inc. No potential conflicts of interest were disclosed by the other authors.

Disclaimer

The content of this article does not necessarily reflect the views or policies of the National Cancer Institute or any of the collaborating centers in the Colon Cancer Family Registry (CCFR), nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. government or the CCFR.

Authors’ Contributions

Conception and design: D.J. Weisenberger, A.J. Levine, D.J. Ahnen, R.W. Haile, J.L. Hopper, J.P. Young, K.D. Siegmund, P.W. Laird


Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): D.J. Weisenberger, T.I. Long, D.D. Buchanan, R. Walters, M. Clendenning, L. Le Marchand, N.M. Lindor, S. Gallinger, T. Selander, B. Bapat, P.A. Newcomb, P.T. Campbell, G. Casey, D.J. Ahnen, R.W. Haile, J.L. Hopper, J.P. Young, P.W. Laird


Writing, review, and/or revision of the manuscript: D.J. Weisenberger, A.J. Levine, D.D. Buchanan, R. Walters, C. Rosty, A.D. Joshi, M.C. Stern, L. Le Marchand, N.M. Lindor, B. Bapat, P.A. Newcomb, P.T. Campbell, G. Casey, D.J. Ahnen, A. Baron, R.W. Haile, J.L. Hopper, J.P. Young, P.W. Laird, K.D. Siegmund

Administrative, technical, or material support (i.e., reporting or organizing data, constructing data tables): T.I. Long, M. Clendenning, L. Le Marchand, P.A. Newcomb, J.L. Hopper, J.P. Young

Study supervision: D.J. Weisenberger, J.L. Hopper, J.P. Young, P.W. Laird

Acknowledgments

The authors thank the Colon Cancer Family Registry for their contributions and dedicated work on this project and all the subjects who provided their time and effort in providing the data.

Grant Support

This work was supported by NIH/NCI grant R01 CA118699 (to P.W. Laird) and R01 HG006705 (to K.D. Siegmund). This work was also supported by grant UM1 CA167551 (to R.W. Haile; M.A. Jenkins, N.M. Lindor) from the National Cancer Institute and through the cooperative agreements with the following CCFR centers: Australasian Colorectal Cancer Family Registry (U01 CA074778; to J.R. Jass) and (U10/IU24 CA097735; to J.L. Hopper), USC Consortium Colorectal Cancer Family Registry (U10/IU24 CA074799; to R.W. Haile), Mayo Clinic Cooperative Family Registry for Colon Cancer Studies (U10/IU24 CA074800; to N.L. Lindor), Ontario Registry for Studies of Familial Colorectal Cancer (U10/IU24 CA074860; to L. Le Marchand). The Jeremy Jass Memorial Pathology Bank provided CCFR paraffin-embedded tissue and pathology-related variables for this study.

Received October 13, 2014; revised December 29, 2014; accepted January 6, 2015; published OnlineFirst January 13, 2015.


Association of the Colorectal CpG Island Methylator Phenotype with Molecular Features, Risk Factors, and Family History


Cancer Epidemiol Biomarkers Prev  Published OnlineFirst January 13, 2015.

Updated version  Access the most recent version of this article at:
doi:10.1158/1055-9965.EPI-14-1161

Supplementary Material  Access the most recent supplemental material at:
http://cebp.aacrjournals.org/content/suppl/2015/01/14/1055-9965.EPI-14-1161.DC1

E-mail alerts  Sign up to receive free email-alerts related to this article or journal.
Reprints and Subscriptions  To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.
Permissions  To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.