Published in final edited form as:

Cancer Epidemiol Biomarkers Prev. 2015 March ; 24(3): 512–519. doi:10.1158/1055-9965.EPI-14-1161.

Association of the Colorectal CpG Island Methylator Phenotype with molecular features, risk factors and family history

Daniel J. Weisenberger[#1,2], **A. Joan Levine**[#3,#], **Tiffany I. Long**[4], **Daniel D. Buchanan**[5,6], **Rhiannon Walters**[5], **Mark Clendenning**[5], **Christophe Rosty**[7], **Amit D. Joshi**[3,&], **Mariana C. Stern**[3], **Loic LeMarchand**[8], **Noralane M. Lindor**[9], **Darshana Daftary**[10], **Steven Gallinger**[10], **Teresa Selander**[10], **Bharati Bapat**[10], **Polly A. Newcomb**[11], **Peter T. Campbell**[12], **Graham Casey**[3], **Dennis J. Ahnen**[13], **John A. Baron**[14], **Robert W. Haile**[3,#], **John L. Hopper**[6], **Joanne P. Young**[7,15], **Peter W. Laird**[1,2,4,%], **Kimberly D. Siegmund**[3,*], and **the Colon Cancer Family Registry**

[1]USC Epigenome Center, University of Southern California, Los Angeles, CA

[2]Department of Biochemistry and Molecular Biology,University of Southern California, Los Angeles, CA, USA

[3]Department of Preventive Medicine, University of Southern California, Los Angeles, CA, USA

[4]Department of Surgery, University of Southern California, Los Angeles, CA, USA

[5]Oncogenomics Group, Genetic Epidemiology Laboratory, Department of Pathology

[6]Centre for Epidemiology and Biostatistics, The University of Melbourne, Parkville, VIC, Australia

[7]Queensland Institute of Medical Research, Herston, QLD 4029, Australia

[8]Cancer Research Center of Hawaii, University of Hawaii, Honolulu, HI

[9]Department of Health Science Research, Mayo Clinic, Scottsdale, AZ, USA

[10]Department of Pathology and Laboratory Medicine, Samuel Lunenfield Research Institute, Mt. Sinai Hospital, Toronto, Ontario, Canada

[11]Epidemiology Department, University of Washington and Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA

[12]Epidemiology Research Program, American Cancer Society, Atlanta, GA

[13]Division of Gastroenterology, University of Colorado School of Medicine, Denver, Colorado

[14]Department of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

[15]The Queen Elizabeth Hospital, Woodville, SA 5011, Australia

[*] To whom correspondence should be addressed. kims@usc.edu; Phone: (323) 442-1310, FAX: (323) 442-2993..

[#]Present Address: Stanford Cancer Institute, Stanford University, Palo Alto, CA

[&]Present Address: Program in Genetic Epidemiology and Statistical Genetics, Department of Epidemiology, Harvard School of Public Health, Boston, MA

[%]Present Address: Laboratory of Cancer Epigenetics, Van Andel Research Institute, Grand Rapids, MI

Conflicts of Interest

D.J.W is a consultant for Zymo Research, Inc., which has a commercial interest in DNA methylation products. Zymo Research did not support this work, nor has an interest in the outcome of this research.

[#] These authors contributed equally to this work.

Abstract

Background—The CpG Island Methylator Phenotype (CIMP) represents a subset of colorectal cancers (CRCs) 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 CRC tumors from the multinational Colon Cancer Family Registry. Etiologic heterogeneity was assessed by a case-case study comparing risk factor frequency of CRC 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% 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 multi-vitamin or calcium supplements did not. Only for female CRCs was CIMP status associated with increased pack-years of smoking (trend p < 0.001) and body mass index (BMI) (trend p = 0.03).

Conclusions—The frequency of several CRC 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 CRC with important lifestyle and environmental exposures increase understanding of the molecular pathologic epidemiology of this heavily methylated subset of CRCs.

Keywords

CpG island methylator phenotype; CIMP; Colon Cancer Family Registry; DNA methylation; colorectal cancer; KRAS; BRAF; MSI

Introduction

Human colorectal cancer (CRC) is a worldwide heath 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 accounts for 2-5% of all CRCs (reviewed in (2)). Most CRCs are thought to result from the accumulation of somatic genetic (3-5) and epigenetic alterations (reviewed in (6,7)) often associated with gender, age, diet, lifestyle habits, and environmental exposures (8-15). The majority of non Lynch syndrome CRCs are located in the distal (descending left) colon and rectum and are enriched for *KRAS* mutations. In contrast, approximately 15% of CRCs are predominantly located in the proximal (ascending, right colon) of older age females with enrichment for $BRAF^{V600E}$ mutations, high levels of microsatellite instability

(MSI-H), *MLH1* epigenetic silencing and the CpG Island Methylator Phenotype (CIMP) (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 five-gene diagnostic panel to identify CIMP tumors: *CACNA1G*, *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 $BRAF^{V600E}$ 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 CRC with environmental exposures are well documented. The risk of CRC 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 CRC risk factors differ in CIMP and non-CIMP tumors, including family CRC history, physical activity, smoking history, history of alcohol use, use of non-steroidal anti-inflammatory drugs (NSAID's) 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 CRCs. Accompanying these samples are a rich data resource of family history, and the level of use/intake of the known CRC risk factors. We evaluated etiological heterogeneity of these risk factors using a case-case study, directly comparing the distribution of known CRC 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 CRC cases, family members and population-based controls, which utilized 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 six centers: the University of Southern California (USC) Consortium (Arizona, Colorado, New Hampshire, Minnesota, North Carolina, and Los Angeles, California), University of Hawaii (Honolulu), 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 CRC, or CRC case probands diagnosed under age 50 to target families with increased CRC risk. First-degree and some second-degree relatives with CRC

were also recruited from families with multiple CRC cases. In this study, we included only CRC cases recruited from 1997- 2002 (30), who signed a written informed consent and completed the risk factor questionnaire (RFQ) within 5 years of their CRC 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 three-category variable: ≤ 50, 51-69 and ≥ 70 years. Family history of CRC was self-reported and was considered positive if the case reported CRC 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 three groups based on WHO criteria for overweight and obesity: 18-24.9, 25-29.9 and ≥30 kg/m^2.

The average weekly hours of physical activity was derived for each of 10 common activities within three 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 (31) and summed within age categories to derive total MET-hours per week during each age category. To reflect lifetime average physical activity, we calculated the mean average MET-hours of all relevant age categories. Adulthood average met-hours per week was grouped by quartiles: 0-5.7, 5.8-14.5, 14.6-30.8, >30.8.

Alcohol use was queried for the same three age groups as physical activity, with total drinks classified as 0, 1 or > 1 per week.

Supplement intake was a three-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 one year prior to 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 one year prior to 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 six 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 CRC-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 two unstained 5-micron tissue sections embedded in paraffin from each tumor on positively charged “plus” glass slides without coverslips.

Slides were randomized to avoid batch effects attributed to source site and reagents. We deparaffinized each slide, microdissected tumor tissues and extracted genomic DNA as previously described (32). Proteinase K was inactivated by heating at 100°C for 10 min. An aliquot was then removed for bisulfite conversion using the Zymo EZ-96 DNA methylation kit (Zymo Research, Irvine, CA) as specified by the manufacturer. CIMP status in each sample was determined using a five-gene MethyLight-based signature (*CACNA1G*, *IGF2*, *NEUROG1*, *RUNX3* and *SOCS1)* 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 two 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 CRCs with CIMP results: 3,544 primary CRCs from case probands and 116 CRCs from affected relatives. Associations between tumor CIMP status and demographic, molecular, and environmental risk factors were performed using the population-based CRC samples from case probands. Of these primary CRCs, 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 CRCs. 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 BRAF Mutation Testing

The somatic T>A mutation at nucleotide 1799 causing the V600E mutation in *BRAF* was determined using a fluorescent allele-specific PCR assay that amplified a 97bp product for the mutant allele (A1799) and a 94bp product for the wildtype 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 169bp PCR amplified product as previously described (34). The larger amplicon size for *KRAS* analysis compared with $BRAF^{V600E}$ contributed to a slightly higher proportion of the FFPE tumor DNA samples failing to amplify for the *KRAS* assay compared with $BRAF^{V600E}$ assay.

MSI testing

MSI was tested using DNA from tumor and matched normal tissue as described in (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 CRC 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 ccOR's represent the relative odds for the risk factor in CIMP CRC compared to that in non-CIMP CRC 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 since 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-value < 0.05 in a two-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 CRC patients in this study represented an estimated 6,253 colorectal cancer cases. The estimated frequency of CIMP CRC was 12.6%, with frequencies ranging from 7% to 18% depending on the CCFR study population (Supplemental Table 1). CIMP CRC was associated with increased patient age ($p<0.0001$) and Australia and USC, the study populations with the lowest frequencies of CIMP CRC also had the lowest averages for age of CRC diagnosis (data not shown). CIMP CRC 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 (Supplemental Table 1). In African Americans the CIMP prevalence was 4.5% and in

Asians it was 4.0%, compared to 13.4% in non-Hispanic Whites and 12.3 % in 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 microsatellite instability

In screening for known *KRAS* and *BRAF* mutations in the sample cohort, we found a high frequency of the $BRAF^{V600E}$ mutation for CIMP proband tumors (63.8%), but not non-CIMP proband tumors (2.1%) (Supplemental Table 1). *KRAS* mutations were more prevalent for non-CIMP compared to CIMP CRC (33.3%, versus 21%) ($p < 0.0001$). These associations remained significant after controlling for age, sex and tumor site (adjusted ccOR = 59.8, 95% CI=45.8-78.0 for *BRAF* and adjusted ccOR = 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 two CIMP CRC and one non-CIMP CRC displaying mutations in both genes. This could be explained for CIMP CRC by variations in *BRAF* and *KRAS* mutation frequency by age. $BRAF^{V600E}$ mutation frequencies for CIMP CRC 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 CRC 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 CRC for females and 45.4% of CIMP CRC for males; *KRAS* mutation appeared in only 11.3% of CIMP CRC for females versus 34.3% of the same for males; and MSI-H occurred in 68.4% of CIMP CRC for females versus 43.2% of the same for males. *KRAS* mutation data was missing for 16.7% of CIMP CRC and 20% non-CIMP CRC (p=0.0017) (Supplemental Table 1).

Association of CIMP with known risk factors of colorectal cancer

Using the available clinical history and lifestyle information, we next determined if CIMP correlated with known CRC risk factors, including smoking history, alcohol use, physical activity, BMI and family CRC history (Table 2). A CIMP CRC 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 CRC. However, only two of the 94 CRC affected relative pairs (2%) were concordant for CIMP CRC and 16 were discordant (17%) (Supplemental Table 1). Limited in power and not statistically significant, this reflected a 2-fold higher frequency of CIMP CRC for affected relatives of a proband with CIMP CRC compared to a proband with non-CIMP CRC (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 CRCs compared to those with non-CIMP CRCs ($P_{trend} = 0.0001$); 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), 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 CRCs and the trend was significant ($P_{trend} = 0.0001$). Alcohol use did not show heterogeneity by CIMP subgroup in men with CRC, but alcohol use in women presented lower frequencies of CIMP CRCs ($P_{trend} = 0.01$). In general, men and women who engaged in higher levels of physical activity showed lower frequencies of CIMP CRCs (Heterogeneity P = 0.01; Supplemental Table 1).

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

We also evaluated use of multivitamins, calcium supplements and NSAIDs prior to CRC diagnosis in CIMP and non-CIMP cancers for men and women separately, and use of hormone replacement therapies by CIMP in women with CRCs (Table 3). Multivitamins or calcium supplement were not associated with CIMP subtype for men or women, however, men and women who used NSAIDs prior to diagnosis showed an increased frequency of CIMP CRC (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 ($P_{interaction} = 0.0008$). The small increase in the frequency of CIMP CRC with HRT use for women was not statistically significant (P= 0.17).

Discussion

The global health concern regarding CRC necessitates an understanding of the contributions of family history and modifiable risk factors to the onset of disease. Since CRC can be classified into different molecular groups, we were specifically interested in whether CIMP CRC, as defined by DNA methylation analyses, is differentially associated with lifestyle, obesity status and/or family history compared to 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 CRC 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 non-diseased 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 CRC, patients with later age of diagnosis, and CRC 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 CRC, and that modifiable risk factors may contribute differentially 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 non-significantly 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 hormone replacement therapy was not significantly associated with CIMP status.

Our data show significant variation in smoking by CIMP status in women, but not in men, with CRC. 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 et al. reported no difference in methylation status of a panel of CIMP markers in the normal colon between smokers and non-smokers 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 nonsteroidal anti-inflammatory drugs (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 CRC, 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 CRC than non-CIMP CRC for both men and women, suggesting that NSAID use is either not as protective against CIMP CRCs as it is for non-CIMP CRC or it increases risk of CIMP CRC. Slattery et al. 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) (40). The CIMP markers used in that study were substantially different from ours, and there were notable differences between these two marker sets in a study comparing them directly (23). In our study, NSAID use was missing for more study participants with CIMP CRCs than with non-CIMP CRCs (6% vs 2%), which if not missing at random, could introduce some bias in our reported frequencies. Whether NSAID use affects CRC 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 two 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 $BRAF^{V600E}$ 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 ~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 non-association 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* mutation and *MLH1* DNA hypermethylation are both highly associated with CIMP, only 64% of CIMP tumors harbored the $BRAF^{V600E}$ mutation and about 50% were MSI-H. Small differences from frequencies in other studies might be explained by a different average age of diagnosis (43-48). This suggests that the use of MSI-H status, *MLH1* 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 since the initial Toyota report in 1999 (22), and although the five-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 CRC. Though the canonical serrated neoplasia pathway has its foundation in the *BRAF*-mutated sessile serrated adenoma/polyp, other pathways to malignant transformation are needed to explain the diversity of CIMP CRC subtypes in this study, including *KRAS*-mutated CIMP CRC which has been previously thought to be rare (23,50). Some of the non-*BRAF*-mutated CIMP CRC 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 utilized the large, population-based inventory of primary colorectal tumors from the C-CFR to analyze the associations between common CRC 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.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We 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). 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 (U01/U24 CA097735) (to J.L. Hopper), USC Consortium Colorectal Cancer

Family Registry (U01/U24 CA074799) (to R.W. Haile), Mayo Clinic Cooperative Family Registry for Colon Cancer Studies (U01/U24 CA074800) (to N.L. Lindor), Ontario Registry for Studies of Familial Colorectal Cancer (U01/U24 CA074783) (to S. Gallinger), Seattle Colorectal Cancer Family Registry (U01/U24 CA074794) (to J.D. Potter and P.A. Newcomb), University of Hawaii Colorectal Cancer Family Registry (U01/U24 CA074806) (to L. Le Marchand). The Jeremy Jass Memorial Pathology Bank provided CCFR paraffin embedded tissue and pathology-related variables for this study.

The content of this manuscript 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 US Government or the CCFR.

References

1. Siegel R, Desantis C, Jemal A. Colorectal cancer statistics, 2014. CA Cancer J Clin. Mar-Apr;2014 64:104–117. [PubMed: 24639052]
2. Lynch HT, Smyrk T. Hereditary nonpolyposis colorectal cancer (Lynch syndrome). An updated review. Cancer. Sep 15.1996 78:1149–1167.
3. Greenman C, Stephens P, Smith R, Dalgliesh GL, Hunter C, Bignell G, et al. Patterns of somatic mutation in human cancer genomes. Nature. Mar 8.2007 446:153–158. [PubMed: 17344846]
4. Vogelstein B, Kinzler KW. Cancer genes and the pathways they control. Nat Med. Aug.2004 10:789–799. [PubMed: 15286780]
5. Wood LD, Parsons DW, Jones S, Lin J, Sjoblom T, Leary RJ, et al. The genomic landscapes of human breast and colorectal cancers. Science. Nov 16.2007 318:1108–1113. [PubMed: 17932254]
6. Jones PA, Baylin SB. The fundamental role of epigenetic events in cancer. Nat Rev Genet. Jun.2002 3:415–428. [PubMed: 12042769]
7. Shen H, Laird PW. Interplay between the cancer genome and epigenome. Cell. Mar 28.2013 153:38–55. [PubMed: 23540689]
8. Chan AT, Giovannucci EL. Primary prevention of colorectal cancer. Gastroenterology. Jun.2010 138:2029–2043. e2010. [PubMed: 20420944]
9. Limsui D, Vierkant RA, Tillmans LS, Wang AH, Weisenberger DJ, Laird PW, et al. Cigarette smoking and colorectal cancer risk by molecularly defined subtypes. Journal of the National Cancer Institute. Jul 21.2010 102:1012–1022. [PubMed: 20587792]
10. Rasool S, Kadla SA, Rasool V, Ganai BA. A comparative overview of general risk factors associated with the incidence of colorectal cancer. Tumour Biol. Oct.2013 34:2469–2476. [PubMed: 23832537]
11. Slattery ML, Curtin K, Wolff RK, Herrick JS, Caan BJ, Samowitz W. Diet, physical activity, and body size associations with rectal tumor mutations and epigenetic changes. Cancer Causes Control. 2010; 21:1237–1245. [PubMed: 20383576]
12. Steinmetz KA, Potter JD. Vegetables, fruit, and cancer prevention: a review. J Am Diet Assoc. Oct.1996 96:1027–1039. [PubMed: 8841165]
13. Michels KB, Edward G, Joshipura KJ, Rosner BA, Stampfer MJ, Fuchs CS, et al. Prospective study of fruit and vegetable consumption and incidence of colon and rectal cancers. Journal of the National Cancer Institute. Nov 1.2000 92:1740–1752. [PubMed: 11058617]
14. Boyle P, Langman JS. ABC of colorectal cancer. BMJ. 2001
15. Gay LJ, Arends MJ, Mitrou PN, Bowman R, Ibrahim AE, Happerfield L, et al. MLH1 promoter methylation, diet, and lifestyle factors in mismatch repair deficient colorectal cancer patients from EPIC-Norfolk. Nutr Cancer. 2011; 63:1000–1010. [PubMed: 21875327]
16. Gonsalves WI, Mahoney MR, Sargent DJ, Nelson GD, Alberts SR, Sinicrope FA, et al. Patient and Tumor Characteristics and BRAF and KRAS Mutations in Colon Cancer, NCCTG/Alliance N0147. Journal of the National Cancer Institute. Jul.2014 :106. [PubMed: 25174031]
17. Kambara T, Simms LA, Whitehall VL, Spring KJ, Wynter CV, Walsh MD, et al. BRAF mutation is associated with DNA methylation in serrated polyps and cancers of the colorectum. Gut. Aug. 2004 53:1137–1144. [PubMed: 15247181]

18. McGivern A, Wynter CV, Whitehall VL, Kambara T, Spring KJ, Walsh MD, et al. Promoter hypermethylation frequency and BRAF mutations distinguish hereditary non-polyposis colon cancer from sporadic MSI-H colon cancer. Fam Cancer. 2004; 3:101–107. [PubMed: 15340260]
19. Nagasaka T, Sasamoto H, Notohara K, Cullings HM, Takeda M, Kimura K, et al. Colorectal cancer with mutation in BRAF, KRAS, and wild-type with respect to both oncogenes showing different patterns of DNA methylation. J Clin Oncol. Nov 15.2004 22:4584–4594. [PubMed: 15542810]
20. Samowitz WS, Albertsen H, Herrick J, Levin TR, Sweeney C, Murtaugh MA, et al. Evaluation of a large, population-based sample supports a CpG island methylator phenotype in colon cancer. Gastroenterology. Sep.2005 129:837–845. [PubMed: 16143123]
21. Samowitz WS, Albertsen H, Sweeney C, Herrick J, Caan BJ, Anderson KE, et al. Association of smoking, CpG island methylator phenotype, and V600E BRAF mutations in colon cancer. Journal of the National Cancer Institute. Dec 6.2006 98:1731–1738. [PubMed: 17148775]
22. Toyota M, Ahuja N, Ohe-Toyota M, Herman JG, Baylin SB, Issa JP. CpG island methylator phenotype in colorectal cancer. Proc Natl Acad Sci U S A. Jul 20.1999 96:8681–8686. [PubMed: 10411935]
23. Weisenberger DJ, Siegmund KD, Campan M, Young J, Long TI, Faasse MA, et al. CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. Nature genetics. 2006; 38:787–793. [PubMed: 16804544]
24. Yang S, Farraye FA, Mack C, Posnik O, O'Brien MJ. BRAF and KRAS Mutations in hyperplastic polyps and serrated adenomas of the colorectum: relationship to histology and CpG island methylation status. Am J Surg Pathol. Nov.2004 28:1452–1459. [PubMed: 15489648]
25. Cancer Genome Atlas Research Network T. Comprehensive molecular characterization of human colon and rectal cancer. Nature. 2012; 487:330–337. [PubMed: 22810696]
26. Hinoue T, Weisenberger DJ, Lange CP, Shen H, Byun HM, Van Den Berg D, et al. Genome-scale analysis of aberrant DNA methylation in colorectal cancer. Genome Res. Feb.2012 22:271–282. [PubMed: 21659424]
27. Hinoue T, Weisenberger DJ, Pan F, Campan M, Kim M, Young J, et al. Analysis of the association between CIMP and BRAF in colorectal cancer by DNA methylation profiling. PloS one. 2009; 4:e8357. [PubMed: 20027224]
28. Yagi K, Akagi K, Hayashi H, Nagae G, Tsuji S, Isagawa T, et al. Three DNA methylation epigenotypes in human colorectal cancer. Clin Cancer Res. Jan 1.2010 16:21–33. [PubMed: 20028768]
29. Noreen F, Roosli M, Gaj P, Pietrzak J, Weis S, Urfer P, et al. Modulation of age- and cancer-associated DNA methylation change in the healthy colon by aspirin and lifestyle. Journal of the National Cancer Institute. Jul.2014 :106. [PubMed: 25174031]
30. Newcomb PA, Baron J, Cotterchio M, Gallinger S, Grove J, Haile R, et al. Colon Cancer Family Registry: an international resource for studies of the genetic epidemiology of colon cancer. Cancer Epidemiol Biomarkers Prev. Nov.2007 16:2331–2343. [PubMed: 17982118]
31. Ainsworth BE, Haskell WL, Whitt MC, Irwin ML, Swartz AM, Strath SJ, et al. Compendium of physical activities: an update of activity codes and MET intensities. Med Sci Sports Exerc. Sep. 2000 32:S498–504. [PubMed: 10993420]
32. Wang X, Kang GH, Campan M, Weisenberger DJ, Long TI, Cozen W, et al. Epigenetic subgroups of esophageal and gastric adenocarcinoma with differential GATA5 DNA methylation associated with clinical and lifestyle factors. PloS one. 2011; 6:e25985. [PubMed: 22028801]
33. Buchanan DD, Sweet K, Drini M, Jenkins MA, Win AK, English DR, et al. Risk factors for colorectal cancer in patients with multiple serrated polyps: a cross-sectional case series from genetics clinics. PloS one. 2010; 5:e11636. [PubMed: 20661287]
34. Stewart CJ, Leung Y, Walsh MD, Walters RJ, Young JP, Buchanan DD. KRAS mutations in ovarian low-grade endometrioid adenocarcinoma: association with concurrent endometriosis. Hum Pathol. Aug.2012 43:1177–1183. [PubMed: 22305241]
35. Bapat B, Lindor NM, Baron J, Siegmund K, Li L, Zheng Y, et al. The association of tumor microsatellite instability phenotype with family history of colorectal cancer. Cancer Epidemiol Biomarkers Prev. Mar.2009 18:967–975. [PubMed: 19258475]

36. Martinez ME, Cruz GI, Brewster AM, Bondy ML, Thompson PA. What can we learn about disease etiology from case-case analyses? Lessons from breast cancer. Cancer Epidemiol Biomarkers Prev. Nov.2010 19:2710–2714. [PubMed: 20870734]
37. Worthley DL, Whitehall VL, Buttenshaw RL, Irahara N, Greco SA, Ramsnes I, et al. DNA methylation within the normal colorectal mucosa is associated with pathway-specific predisposition to cancer. Oncogene. Mar 18.2010 29:1653–1662. [PubMed: 19966864]
38. Garcia-Albeniz X, Chan AT. Aspirin for the prevention of colorectal cancer. Best Pract Res Clin Gastroenterol. Aug.2011 25:461–472. [PubMed: 22122763]
39. Wallace K, Grau MV, Ahnen D, Snover DC, Robertson DJ, Mahnke D, et al. The association of lifestyle and dietary factors with the risk for serrated polyps of the colorectum. Cancer Epidemiol Biomarkers Prev. Aug.2009 18:2310–2317. [PubMed: 19661090]
40. Slattery ML, Curtin K, Sweeney C, Levin TR, Potter J, Wolff RK, et al. Diet and lifestyle factor associations with CpG island methylator phenotype and BRAF mutations in colon cancer. Int J Cancer. Feb 1.2007 120:656–663. [PubMed: 17096326]
41. Shen L, Toyota M, Kondo Y, Lin E, Zhang L, Guo Y, et al. Integrated genetic and epigenetic analysis identifies three different subclasses of colon cancer. Proc Natl Acad Sci U S A. Nov 20.2007 104:18654–18659. [PubMed: 18003927]
42. Ogino S, Kawasaki T, Kirkner GJ, Loda M, Fuchs CS. CpG island methylator phenotype-low (CIMP-low) in colorectal cancer: possible associations with male sex and KRAS mutations. J Mol Diagn. Nov.2006 8:582–588. [PubMed: 17065427]
43. Kang GH. Four molecular subtypes of colorectal cancer and their precursor lesions. Arch Pathol Lab Med. Jun.2011 135:698–703. [PubMed: 21631262]
44. Kim JH, Shin SH, Kwon HJ, Cho NY, Kang GH. Prognostic implications of CpG island hypermethylator phenotype in colorectal cancers. Virchows Arch. Dec.2009 455:485–494. [PubMed: 19911194]
45. Hutchins G, Southward K, Handley K, Magill L, Beaumont C, Stahlschmidt J, et al. Value of mismatch repair, KRAS, and BRAF mutations in predicting recurrence and benefits from chemotherapy in colorectal cancer. J Clin Oncol. Apr 1.2011 29:1261–1270. [PubMed: 21383284]
46. Saridaki Z, Papadatos-Pastos D, Tzardi M, Mavroudis D, Bairaktari E, Arvanity H, et al. BRAF mutations, microsatellite instability status and cyclin D1 expression predict metastatic colorectal patients' outcome. Br J Cancer. Jun 8.2010 102:1762–1768. [PubMed: 20485284]
47. Lochhead P, Kuchiba A, Imamura Y, Liao X, Yamauchi M, Nishihara R, et al. Microsatellite instability and BRAF mutation testing in colorectal cancer prognostication. Journal of the National Cancer Institute. Aug 7.2013 105:1151–1156. [PubMed: 23878352]
48. Kim JH, Rhee YY, Bae JM, Kwon HJ, Cho NY, Kim MJ, et al. Subsets of microsatellite-unstable colorectal cancers exhibit discordance between the CpG island methylator phenotype and MLH1 methylation status. Mod Pathol. Jul.2013 26:1013–1022. [PubMed: 23370766]
49. Ogino S, Kawasaki T, Kirkner GJ, Kraft P, Loda M, Fuchs CS. Evaluation of markers for CpG island methylator phenotype (CIMP) in colorectal cancer by a large population-based sample. J Mol Diagn. Jul.2007 9:305–314. [PubMed: 17591929]
50. English DR, Young JP, Simpson JA, Jenkins MA, Southey MC, Walsh MD, et al. Ethnicity and risk for colorectal cancers showing somatic BRAF V600E mutation or CpG island methylator phenotype. Cancer Epidemiol Biomarkers Prev. Jul.2008 17:1774–1780. [PubMed: 18628431]
51. Rosty C, Young JP, Walsh MD, Clendenning M, Sanderson K, Walters RJ, et al. PIK3CA activating mutation in colorectal carcinoma: associations with molecular features and survival. PloS one. 2013; 8:e65479. [PubMed: 23785428]

Table 1

Distribution of CIMP status by center, age and tumor location stratified by gender (unweighted N=1,628 Males/1,491 Females)

	Male Weighted[1] N				Female Weighted N				
	CIMP[2] (%)	Non-CIMP (%)	OR (95% CI)[3]	P-Value	CIMP (%)	Non-CIMP (%)	OR (95% CI)	P-Value	Interaction P-Value (df)
Patient Age[4]									
≤ 50	18 (5.5)	465 (14.7)	1.0		34 (7.3)	515 (22.5)	1.0		
51-69	185 (56.7)	1,842 (58.1)	2.76 (1.67-4.58)	**0.0001**	225 (48.5)	1,277 (55.8)	2.59 (1.76-3.83)	**0.0001**	
≥ 70	124 (37.9)	865 (27.3)	3.35 (1.99-5.63)	**0.0001**	205 (44.2)	497 (21.7)	4.61 (3.09-6.88)	**0.0001**	
Trend			1.52 (1.25-1.84)	**0.0001**			2.01 (1.69-2.38)	**0.0001**	0.065 (2)
Tumor Site[5]									
Proximal	241 (73.8)	908 (28.7)	8.48 (5.97-12.10)	**0.0001**	399 (86.0)	792 (34.6)	9.59 (6.51-14.14)	**0.0001**	**0.014** (2)
Distal	48 (14.7)	1037 (32.7)	1.49 (0.97-2.31)	0.07	35 (7.5)	859 (37.5)	0.84 (0.51-1.39)	0.50	
Rectal	38 (11.5)	1226 (38.7)	1.0		30 (6.5)	638 (27.9)	1.0		
BRAF(V600E)[6]									
Mutated	148 (45.4)	55 (1.8)	35.7 (24.3-52.5)	**0.0001**	342 (77.3)	56 (2.5)	90.6 (62.4-131.5)	**0.0001**	**0.0011** (1)
Not Mutated	178 (54.6)	3084 (98.3)	1.0		100 (22.7)	2196 (97.5)	1.0		
Missing	0	33			22	38			
KRAS[6]									
Mutated	95 (34.3)	814 (32.1)	0.76 (0.58-1.0)	0.06	43 (11.3)	643 (35.1)	0.22 (0.16-0.32)	**0.0001**	**<0.0001** (1)
Not Mutated	182 (65.7)	1721 (67.9)	1.0		339 (88.7)	1192 (65.0)	1.0		
Missing	50	637			82	455			
MSI Status[6]									
MSI-H	122 (37.5)	275 (7.9)	3.86 (2.86-5.20)	**<0.0001**	306 (66.1)	192 (8.4)	12.61 (9.56-16.6)	**<0.0001**	**<0.0001** (2)
MSI-L	46 (14.1)	481 (13.8)	1.28 (0.90-1.82)	0.17	34 (7.5)	299 (13.1)	1.34 (0.89-2.02)	0.17	
MSS	158 (48.4)	2394 (68.9)	1.0		122(26.4)	1792 (78.5)	1.0		
Missing	0	22			2	6			

[1] 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.

[2] Defined as a PMR ≥10 for at least 3 of 5 genes: *CACNA1G*, *IGF2*, *NEUROG1*, *RUNX3* and *SOCS1*.

[3] OR = Odds ratio CI = Confidence interval

[4] logistic regression model using proband weights and controlling for tumor site

[5] logistic regression model using proband weights and controlling for age (≤50, 51-69,≥70).

[6] logistic regression model using proband weights and controlling for age (≤50, 51-69,≥70) and tumor site.

Cancer Epidemiol Biomarkers Prev. Author manuscript; available in PMC 2016 March 01.

Table 2

Associations between CIMP and selected CRC risk factors by gender

	Male Weighted[1] N				Female Weighted N				
	CIMP[2] (%)	**Non-CIMP (%)**	**OR (95% CI)[3]**	**P-Value**	**CIMP (%)**	**Non-CIMP (%)**	**OR (95% CI)[3]**	**P-Value**	**Interaction P-Value (df)**
Family CRC history[4]									
None	280 (86.4)	2,534 (79.9)	1.0		356 (77.3)	1,708 (74.6)	1.0		
≥ 1	44 (13.6)	636 (20.1)	0.55 (0.39-0.78)	**0.0007**	105 (22.7)	580 (25.3)	0.54 (0.41-0.70)	**0.0001**	0.83 (1)
Smoking (pack years)[5]									
0	90 (28.0)	894 (29.0)	1.0		208 (44.9)	1,122 (50.0)	1.0		
1-10	24 (7.5)	456 (14.7)	0.42 (0.26-0.68)	**0.0004**	80 (17.4)	454 (20.2)	1.18 (0.87-1.61)	0.30	
11-20	53 (16.6)	466 (15.1)	1.32 (0.90-1.92)	0.16	53 (11.4)	234 (10.4)	1.13 (0.78-1.64)	0.53	
21-40	83 (25.8)	621 (20.1)	1.26 (0.90-1.76)	0.17	69 (14.8)	263 (11.7)	1.84 (1.31-2.60)	**0.0005**	
≥ 40	72 (22.2)	654 (21.2)	0.94 (0.66-1.33)	0.73	54 (11.6)	172 (7.7)	2.07 (1.40-3.01)	**0.0003**	
Trend			1.06 (0.98-1.14)	0.18			1.20 (1.10-1.30)	**0.0001**	**0.0002** (4)
Alcohol (drinks/week)									
0	103 (33.0)	1,067 (35.0)	1.0		280 (65.9)	1,260 (57.7)	1.0		
1	170 (54.3)	1,709 (55.9)	0.94 (0.72-1.23)	0.65	121 (28.4)	803 (36.8)	0.67 (0.52-0.86)	**0.0019**	
>1	40 (12.8)	281 (9.1)	1.39 (0.92-2.09)	0.11	24 (5.6)	122 (5.6)	0.83 (0.50-1.36)	0.45	
Trend			1.10 (0.90-1.34)	0.34			0.78 (0.64-0.94)	**0.01**	0.15 (2)
BMI kg/m[2]									
18-24.9	95 (29.1)	737 (23.2)	1.0		200 (43.1)	1,184 (51.7)	1.0		
25-29.9	147 (45.0)	1,732 (54.6)	0.51 (0.38-0.69)	**0.0001**	133 (28.8)	636 (27.8)	1.42 (1.09-1.86)	**0.03**	
≥ 30	84 (25.8)	703 (22.2)	0.78 (0.56-1.08)	0.13	131 (28.2)	470 (20.5)	1.93 (1.09-2.56)	**0.0001**	
Trend			0.87 (0.73-1.04)	0.13			1.39 (1.21-1.60)	**0.0001**	**0.0001** (2)
Physical Activity[6]									
0-5.7	86 (28.1)	688 (22.6)	1.0		146 (34.4)	580 (26.7)	1.0		
5.8-14.5	59 (19.4)	616 (20.3)	0.76 (0.53-1.10)	0.15	87 (20.5)	519 (23.9)	0.60 (0.44-0.83)	**0.002**	

	Male Weighted[1] N				Female Weighted N				
	CIMP[2] (%)	Non-CIMP (%)	OR (95% CI)[3]	P-Value	CIMP (%)	Non-CIMP (%)	OR (95% CI)[3]	P-Value	Interaction P-Value (df)
14.6-30.8	68 (22.3)	776 (25.5)	0.77 (0.54-1.09)	0.14	89 (21.0)	509 (23.4)	0.74 (0.54-1.02)	0.070	
> 30.8	92 (30.2)	963 (31.6)	0.87 (0.63-1.20)	0.40	102 (24.1)	566 (26.0)	0.74 (0.54-1.00)	0.051	
Trend			0.96 (0.86-1.07)	0.46			0.92 (0.83-1.02)	0.101	0.71 (3)

[1] 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.

[2] Defined as a PMR ≥10 for at least 3 of 5 genes: *CACNA1G, IGF2, NEUROG1, RUNX3* and *SOCS1*.

[3] Odds ratios and 95% confidence limits estimated using logistic regression and controlling for age (≤ 50, 51-69, ≥ 70) and tumor site.

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

[5] Number of reported cigarettes per day multiplied by the number of years of smoking.

[6] 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.

Table 3

Associations between pre-diagnosis supplement use and CIMP status by gender

	Male Weighted N[1]				Female Weighted N				
	CIMP[2] (%)	Non-CIMP	OR (95% CI)[3]	P-Value	CIMP (%)	Non-CIMP (%)	OR (95% CI)	P-Value	Interaction P-Value
Pre-diagnosis Multivitamins[4]									
Non-User	180 (56.3)	1,734 (55.1)	1.0		206 (45.0)	965 (42.9)	1.0		
Former user	41 (12.9)	638 (20.4)	0.63 (0.44-0.91)	**0.01**	78 (17.0)	489 (21.7)	0.70 (0.52-0.96)	**0.03**	
User	99 (30.9)	764 (24.4)	1.18 (0.90-1.55)	0.24	174 (38.1)	795 (35.4)	0.80 (0.62-1.03)	0.08	0.11 (2)
Missing	7	46			7	41			
Pre-diagnosis Calcium[4]									
Non-user	291(91.4)	2,826 (89.9)	1.0		242 (53.3)	1,277 (57.5)	1.0		
Former user	10 (3.1)	152 (4.8)	0.75 (0.38-1.47)	0.40	64 (14.0)	381 (17.2)	1.04 (0.74-1.45)	0.83	
User	17 (5.5)	165 (5.2)	0.91 (0.53-1.54)	0.71	149 (32.8)	564 (25.4)	1.22 (0.95-1.58)	0.13	0.49 (2)
Missing	8	30			9	68			
Pre-diagnosis NSAIDs[5]									
Non-user	105 (34.3)	1,748 (56.2)	1.0		227 (52.0)	1,230 (54.9)	1.0		
Former user	74 (24.1)	607 (19.5)	1.75 (1.27-2.42)	**0.0007**	84 (19.3)	567 (25.3)	0.75 (0.56-1.01)	0.06	
User	127 (41.6)	753 (24.2)	2.30 (1.72-3.07)	**0.0001**	125 (28.7)	445 (19.8)	1.40 (1.07-1.83)	**0.02**	**0.0008** (2)
Missing	21	63			28	47			
HRT[6]									
Non-user					216 (49.5)	1,263 (55.9)	1.0		-
User					220 (50.5)	996 (44.1)	1.17 (0.93-1.47)	0.17	
					28	30			

[1] 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.

[2] Defined as a PMR ≥10 for at least 3 of 5 genes: *CACNA1G*, *IGF2*, *NEUROG1*, *RUNX3* and *SOCS1*.

[3] Odds ratios and 95% confidence limits estimated using logistic regression and controlling for age (≤ 50, 51-69, ≥ 70) and tumor site.

[4]Users were those that answered ‘yes’ to the question “have you ever used [supplement] at least two times a week for more than a month” and indicated that they were taking that supplement one year prior to CRC diagnosis. Former users include an unknown number of subjects who began using the supplement after CRC diagnosis.

[5]Users were defined as NSAID users if they had used either aspirin or Ibuprofen and non-users if they had not used either at least one year prior to CRC diagnosis. Former users include an unknown number of subjects who began using the supplement after CRC diagnosis.

[6]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 six months or longer.

Cancer Epidemiol Biomarkers Prev. Author manuscript; available in PMC 2016 March 01.