Author/s:
Huang, Y;Hua, X;Labadie, JD;Harrison, TA;Dai, JY;Lindstrom, S;Lin, Y;Berndt, SI;Buchanan, DD;Campbell, PT;Casey, G;Gallinger, SJ;Gunter, MJ;Hoffmeister, M;Jenkins, MA;Sakoda, LC;Schoen, RE;Diergaarde, B;Slattery, ML;White, E;Giles, G;Brenner, H;Chang-Claude, J;Joshi, A;Ma, W;Pai, RK;Chan, AT;Peters, U;Newcomb, PA

Title:
Genetic variants associated with circulating C-reactive protein levels and colorectal cancer survival: Sex-specific and lifestyle factors specific associations

Date:
2022-05-01

Citation:
Huang, Y., Hua, X., Labadie, J. D., Harrison, T. A., Dai, J. Y., Lindstrom, S., Lin, Y., Berndt, S. I., Buchanan, D. D., Campbell, P. T., Casey, G., Gallinger, S. J., Gunter, M. J., Hoffmeister, M., Jenkins, M. A., Sakoda, L. C., Schoen, R. E., Diergaarde, B., Slattery, M. L. ,... Newcomb, P. A. (2022). Genetic variants associated with circulating C-reactive protein levels and colorectal cancer survival: Sex-specific and lifestyle factors specific associations. [Abstract]. International Journal of Cancer, 150 (9), pp.1447-1454. https://doi.org/10.1002/ijc.33897.

Hua Xinwei (Orcid ID: 0000-0003-2114-2975)
Dai James (Orcid ID: 0000-0001-6479-9286)
Buchanan Daniel D (Orcid ID: 0000-0003-2225-6675)
Hoffmeister Michael (Orcid ID: 0000-0002-8307-3197)
Slattery Martha (Marty) L. (Orcid ID: 0000-0002-1655-6543)
Newcomb Polly A (Orcid ID: 0000-0001-8786-0043)

Genetic variants associated with circulating C-reactive protein levels and colorectal cancer survival: Sex- and lifestyle factors- specific associations

Yuhan Huang[1,2], Xinwei Hua[1,2,3*], Julia D. Labadie[1,2], Tabitha A. Harrison[1], James Y. Dai[1], Sara Lindstrom[1,2], Yi Lin[1], Sonja I. Berndt[4], Daniel D. Buchanan[5,6,7], Peter T. Campbell[8], Graham Casey[9], Steven J. Gallinger[10], Marc J. Gunter[11], Michael Hoffmeister[12], Mark A. Jenkins[13], Lori C. Sakoda[1,14], Robert E. Schoen[15], Brenda Diergaarde[16,17], Martha L. Slattery[18], Emily White[1,2], Graham Giles[13,19,20], Hermann Brenner[12,21,22], Jenny Chang-Claude[23,24], Amit Joshi[25], Wenjie Ma[25], Rish K. Pai[26], Andrew T. Chan[27,28,29], Ulrike Peters[1,2], Polly A. Newcomb[1,2*]

[1]Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA.
[2]Department of Epidemiology, School of Public Health, University of Washington, Seattle, Washington, USA.
[3]Clinical and Translational Epidemiology Unit and Department of Gastroenterology, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts, USA.
[4]Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, Maryland, USA.
[5]Colorectal Oncogenomics Group, Department of Clinical Pathology, University of Melbourne, Parkville, Victoria, Australia.
[6]University of Melbourne Centre for Cancer Research, Victorian Comprehensive Cancer Centre, Parkville, Victoria, Australia.
[7]Genetic Medicine and Family Cancer Clinic, The Royal Melbourne Hospital, Parkville, Victoria, Australia.
[8]Department of Population Science, American Cancer Society, Atlanta, Georgia, USA.
[9]Center for Public Health Genomics, University of Virginia, Charlottesville, Virginia, USA.
[10]Lunenfeld Tanenbaum Research Institute, Mount Sinai Hospital, University of Toronto, Toronto, Ontario, Canada.
[11]Nutrition and Metabolism Section, International Agency for Research on Cancer, World Health Organization, Lyon, France.
[12]Division of Clinical Epidemiology and Aging Research, German Cancer Research Center (DKFZ), Heidelberg, Germany.
[13]Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, University of Melbourne, Melbourne, Victoria, Australia.
[14]Division of Research, Kaiser Permanente Northern California, Oakland, California, USA.
[15]Department of Medicine and Epidemiology, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania, USA.
[16]Department of Human Genetics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, Pennsylvania, USA.
[17]UPMC Hillman Cancer Center, Pittsburgh, Pennsylvania, USA
[18]Department of Internal Medicine, University of Utah, Salt Lake City, Utah, USA.
[19]Cancer Epidemiology Division, Cancer Council Victoria, Melbourne, Victoria, Australia.
[20]Precision Medicine, School of Clinical Sciences at Monash Health, Monash University, Clayton, Australia.
[21]Division of Preventive Oncology, German Cancer Research Center (DKFZ) and National Center for Tumor Diseases (NCT), Heidelberg, Germany.

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/ijc.33897

[22]German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Heidelberg, Germany.
[23]University Medical Centre Hamburg-Eppendorf, University Cancer Centre Hamburg (UCCH), Hamburg, Germany.
[24]Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany.
[25]Clinical and Translational Epidemiology Unit, Massachusetts General Hospital, Boston, Massachusetts, USA.
[26]Department of Laboratory Medicine and Pathology, Mayo Clinic Arizona, Scottsdale, Arizona, USA
[27]Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital, Boston, Massachusetts, USA.
[28]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Harvard University, Boston, Massachusetts, USA,
[29]Department of Medicine, Massachusetts General Hospital, Boston, Massachusetts, USA.

* Co-corresponding authors

Address correspondence to:
Polly A. Newcomb, PhD, MPH
Professor, Cancer Prevention Program
Public Health Sciences Division, Fred Hutch
Address: 1100 Fairview Ave. N., M4-B402, Seattle, WA 98109-1024
Email: pnewcomb@fredhutch.org

Xinwei Hua
Clinical and Translational Epidemiology Unit and Department of Gastroenterology
Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts
Twitter Handle: @XinweiHua
Address: 100 Cambridge Street, Boston, MA 02114
Email: xhua1@mgh.harvard.edu

Keywords: C-reactive protein, colorectal cancer, genetic variants, survival

Abbreviations: BMI=Body Mass Index, CCFR= Colon Cancer Family Registry, CI=Confidence Interval, CPS II= Cancer Prevention Study II, CRC=Colorectal Cancer, CRP=C-Reactive Protein, DACHS= Darmkrebs: Chancen der Verhütung durch Screening, DALS= Diet, Activity, and Lifestyle Study, EDRN= Early Detection Research Network, EPIC=European Prospective Investigation into Cancer, GWAS= Genome-Wide Association Studies, HPFS=Health Professionals Follow-up Study, HR=Hazard Ratio, ICD= International Classification of Diseases, ISACC= International Survival Analysis in Colorectal Cancer Consortium, LD=Linkage Disequilibrium, MCCS= Melbourne Collaborative Cohort Study, NHGRI-EBI=National Human Genome Research Institute-European Bioinformatics Institute, NHS= Nurses' Health Study, PCA=Principal Components Analysis, PHS= Physicians' Health Study, PLCO= Prostate, Lung, Colorectal, and Ovarian Study, QC=Quality Control, SNP= Single Nucleotide Polymorphism, UKB= UK Biobank, VITAL= VITamins And Lifestyle Study, WHI= Women's Health Initiative.

Novelty and Impact: This is the first study evaluating the gene-environment interactions of C-reactive protein level-related genetic variants with sex and lifestyle factors in relation to colorectal cancer-specific mortality. We found evidence of a statistically significant interaction

between variant rs1933736 at *FRK* gene and sex. Women with C (vs. T) allele of variant rs1933736 were associated with increased colorectal cancer-specific mortality, while an inverse association was observed for men.

Article category: Cancer Genetics and Epigenetics

Abstract

Elevated blood levels of C-reactive protein (CRP) have been linked to colorectal cancer (CRC) survival. We evaluated genetic variants associated with CRP levels and their interactions with sex and lifestyle factors in association with CRC-specific mortality.

This study included 16,142 CRC cases from the International Survival Analysis in Colorectal Cancer Consortium. We identified 618 common single nucleotide polymorphisms (SNPs) associated with CRP levels from the NHGRI-EBI GWAS Catalog. Cox proportional hazards regression was used to estimate hazard ratios (HRs) and 95% confidence intervals (CIs) for associations between SNPs and CRC-specific mortality adjusting for age, sex, genotyping platform/study, and principal components. We investigated their interactions with sex and lifestyle factors using likelihood ratio tests.

Of 5,472 (33.9%) deaths accrued over up to 10 years of follow-up, 3,547 (64.8%) were due to CRC. No variants were associated with CRC-specific mortality after multiple comparison correction. We observed strong evidence of interaction between variant rs1933736 at *FRK* gene and sex in relation to CRC-specific mortality (corrected $P_{interaction}$ = 0.0004); women had higher CRC-specific mortality associated with the minor allele (HR = 1.11, 95% CI = 1.04 to 1.19) whereas an inverse association was observed for men (HR = 0.88, 95% CI = 0.82 to 0.94). There was no evidence of interactions between CRP-associated SNPs and alcohol, obesity or smoking.

Our study observed a significant interaction between sex and a CRP-associated variant in relation to CRC-specific mortality. Future replication of this association and functional annotation of the variant are needed.

Introduction

Chronic inflammation is strongly associated with colorectal cancer (CRC) through several mechanisms [1]. C-reactive protein (CRP) is a sensitive marker of chronic low-grade inflammation. Increasing evidence has shown associations of elevated CRP levels with higher CRC risk and worse prognosis [2-6]. Although the mechanisms underlying these relationships are incompletely understood, germline genetic variants are thought to play an important role.

Genome-wide association studies (GWAS) have been successful in identifying genetic variants associated with CRP levels [7, 8], and have facilitated investigations of the role of CRP-related variants in CRC incidence [9, 10]. Slattery et al. and Nimptsch et al. reported several genetic variants in *CRP* gene associated with the risk of colon or rectal cancer. However, few studies evaluated the association between CRP-related variants and CRC prognosis.

Additionally, CRP concentrations are positively associated with several prognostic factors of CRC such as sex, obesity, smoking, as well as excessive alcohol consumption [11-20]. CRC patients who were smokers, obese, heavy drinkers, or males were shown to have worse survival outcomes [16-18, 20]. To further evaluate the interactions between CRP-related genetic variants and these environmental factors in relation to CRC mortality can help to improve our understanding of the relationship between CRP-associated genetic variants and CRC mortality, and to identify populations at differential risk of CRC death and provide insights into underlying mechanisms of progression and prognosis.

In this consortium study, we examined CRP-related variants in relation to CRC survival and whether these associations differed according to sex and lifestyle factors among patients diagnosed with CRC.

Method

Study population

This study included individuals diagnosed with incident, invasive CRC from the International Survival Analysis in Colorectal Cancer Consortium (ISACC), a consortium of case-control studies, cohort studies, and clinical trials from around the world. The following 14 ISACC studies were included: the Colon Cancer Family Registry (CCFR) [21], Cancer Prevention Study II (CPSII) [22], Darmkrebs: Chancen der Verhütung durch Screening (DACHS) [23], Diet, Activity, and Lifestyle Study (DALS) [24], Early Detection Research Network (EDRN) [25], European Prospective Investigation into Cancer (EPIC) [26], Health Professionals Follow-up Study (HPFS) [27], Melbourne Collaborative Cohort Study (MCCS) [28], Nurses' Health Study (NHS) [29, 30], Physicians' Health Study (PHS) [31], Prostate, Lung, Colorectal, and Ovarian Study (PLCO) [32], UK Biobank (UKB) [33], VITamins And Lifestyle Study (VITAL) [34], Women's Health Initiative (WHI) [35]. Study-specific details are described in the Supplementary Table 1.

Demographic and epidemiologic factors were collected using self- or interviewer-administered questionnaires at enrollment according to study-specific protocols. A multistep data harmonization process was conducted centrally using an iterative process to reconcile the differences of the protocols and data collection instruments across studies, as described previously [36]. Briefly, within each study, all exposure information, including age at diagnosis, sex, body mass index (BMI), ever smoking, alcohol consumption, was collected by in-person or phone interviews, self-administered structured questionnaires, or both with the reference time for cohort studies as the time of enrollment. For this study, we restricted the population to CRC cases who all had available data on genotyping, epidemiologic factors, and CRC survival outcomes, and were of European genetic ancestry. The final study population for the analysis of CRP-levels related gene variants and CRC mortality consisted of 16,142 CRC cases after

excluding cases with non-European ancestry or extreme low BMI (BMI<18.5 kg/m^2). For the analyses of gene x environment interactions, we only included individuals with complete data of the specific lifestyle factor (i.e., BMI, ever smoking, and alcohol consumption).

Genotype data

Details of genotyping and quality control (QC) methods have been reported previously [37]. Briefly, genomic DNA was extracted from blood or buccal samples using conventional methods. Genotyping was performed using several platforms (Supplementary Table 1). All genotype data underwent standardized QC procedures, such as exclusion of samples and variants with low call rates (<98%), variants in regions with chromosomal anomalies, samples with discrepant self-reported vs. genetic sex, and variants departing from Hardy–Weinberg Equilibrium ($P < 10^{-4}$). To investigate population structure, PLINK (v1.9) was used for principal components analysis (PCA) [38]. We restricted analyses to participants of European genetic ancestry based on PCA. The first three eigenvectors explained 46% of the genetic variation and thus were used as covariates in analysis. Variants were phased using SHAPEIT2 and imputed to the Haplotype Reference Consortium panel using the University of Michigan Imputation Server [39-41]. Genotype probabilities were converted to allelic dosages after imputation.

A total of 781 single nucleotide polymorphisms (SNPs) associated with CRP levels were identified from the National Human Genome Research Institute-European Bioinformatics Institute (NHGRI-EBI) GWAS catalog (October 30, 2020), of which 651 have minor allele frequency >5% in European ancestry. Among those common CRP-related SNPs, 618 were directly genotyped or imputed in our dataset and, therefore, included in analysis (Supplementary Table 2).

Definitions of covariates

Age at diagnosis was ascertained via cancer registries and/or medical records. BMI defined as weight/height^2 was categorized as normal weight (18.5-24.9 kg/m^2), overweight (25-29.9 kg/m^2), and obese ($\geq$ 30 kg/m^2). Ever smoking was defined as a binary variable (yes for smokers and former smokers; no for never smokers). Alcohol consumption was categorized as nondrinker (≤1 g/day), moderate drinker (1.01-28 g/day), and heavy drinker (> 28 g/day). A joint variable of study and genotyping platform was created.

Ascertainment of survival outcomes

All study participants were followed up for vital status. For CCFR, CPSII, DACHS, DALS, EDRN, EPIC, MCCS, UKB, and VITAL, date and cause of death were ascertained through linkages to the National Death Index or cancer registries, which link to death certificates. For HPFS, NHS, PHS, PLCO, and WHI, date and cause of death were obtained via active follow-up and verified by death certificates and/or medical records. We computed person-time of follow-up for each participant from the diagnosis of CRC to the date of death, last date of contact, or the end of follow-up. Participants who died from causes other than CRC were censored at the time of death. We used the International Classification of Diseases-9 (ICD-9) or ICD-10 (depending on year of linkage) to define CRC-specific deaths (ICD-9: 153.0-153.4, 153.6-153.9, or 154.0-154.1; ICD-10: C18.0-20.0 or C26.0).

Statistical methods

Statistical analyses were conducted using individual-level data. To evaluate the association of each CRP-related SNP with CRC mortality, we used Cox proportional hazards regression to estimate hazard ratios (HRs) and 95% confidence intervals (CIs). Variants were analyzed as continuous variables assuming log-additive effects. We adjusted for age at diagnosis, sex, a joint variable of study and genotyping platform, and the first three principal components.

Violations of the proportional hazards assumption were evaluated by Schoenfeld residuals [42]. Age at diagnosis (continuous), sex, and the joint variable of study/genotyping platform violated the proportional hazards assumption. To resolve this, models were stratified by age categories (<50, 50-59, 60-69, 70-79, and 80+ years), sex and study/genotyping platform. We used likelihood ratio tests to assess multiplicative gene-environment interactions between CRP-related SNPs and sex, obesity, smoking, or alcohol consumption. Sub-analyses stratifying by these factors were also conducted for SNPs with statistically significant interactions.

To account for multiple comparisons, Pearson correlation coefficients were computed to determine the correlations between every pair of SNPs within the same chromosome, and principal components analysis was performed to obtain the effect number of independent tests (M_{eff_G}). M_{eff_G} of 556 was used for type I error control in Bonferroni correction in single-SNP survival analysis, with $P_{corrected}$ <0.05 considered statistically significant [43]. All statistical tests and *P*-values were two-sided. All analyses were conducted using R 3.6.0.

Results

Participant characteristics are presented in Table 1. Study-specific characteristics are presented in Supplementary Table 1. The mean (standard deviation) age at diagnosis across studies was 66.4 (9.8) years, and 50.3% of participants were male. After a median of 4.7 years (interquartile range = 2.3 to 7.7 years) of follow-up since diagnosis, 5,472 (33.9%) deaths accrued, 3,547 (64.8%) of which were due to CRC.

The main effects of 618 SNPs were not significantly associated with CRC mortality after multiple comparison correction (Supplementary Table 2). One variant, rs4148191 at *ABCG5* gene, was borderline significantly associated with higher CRC-specific mortality (HR = 1.18, 95% CI = 1.08 to 1.29; $P_{corrected}$ = 0.08).

We observed strong evidence of interaction between variant rs1933736 (MAF = 0.42) at *FRK* gene and sex in relation to CRC-specific mortality (corrected $P_{interaction}$ = 0.0004); women had higher CRC-specific mortality associated with the minor allele (HR = 1.11, 95% CI = 1.04 to 1.19) whereas an inverse association was observed for men (HR = 0.88, 95% CI = 0.82 to 0.94) (Table 2 and Supplementary Table 3). The overall association of this variant with CRC-specific mortality was not significant (HR = 0.99, 95% CI = 0.94 to 1.04; $P_{corrected}$ >0.99). There was no evidence of statistically significant interactions between CRP-associated SNPs and alcohol, obesity or smoking in association with CRC-specific mortality (Supplementary Table 4, 5, and 6).

Discussion

To our knowledge, this is the first study evaluating the gene-environment interactions of CRP level-related genetic variants identified by GWAS catalog with sex and lifestyle factors in relation to CRC-specific mortality. We found evidence of a statistically significant interaction between rs1933736 at *FRK* gene and sex. Women with C (vs. T) allele of variant rs1933736 were associated with increased CRC-specific mortality, while an inverse association was observed for men.

Few studies have reported associations of CRP-related variants and CRC survival. Slattery et al. investigated four CRP-related variants, rs1205, rs1417938, rs1800947, and rs3093075, selected based on linkage disequilibrium (LD) (r^2 >0.90) and minor allele frequency > 4%. They did not observe association of these variants with colon or rectal cancer mortality [10]. The variant rs1800947 was the only one that was also investigated by our study, and our results were

consistent (Supplement Table 2). There is limited evidence for the influence of CRP-related variants on CRC incidence. Slattery et al. reported association of rs1205 with colon cancer incidence and association of rs3093075 with rectal cancer incidence, while no associations of rs1417938 and rs1800947 were observed [10]. Nimptsch et al. analyzed five tagging SNPs to cover variations in CRP gene common to European population, and reported that two CRP-related variants, rs1205 and rs11308864, were associated CRC incidence, while rs1800947, rs3093077 and rs2808630 were not associated with CRC risk [9].

Although no associations were observed for CRP-related SNPs and CRC-specific mortality, we observed an interaction between sex and the variant rs1933736. This genetic variant locates at the intronic region of the gene *FRK,* which is known to negatively regulates cell proliferation and positively regulates PTEN protein stability and may function as a tumor suppressor. Han et al reported that C allele of variant rs1933736 was associated with decreased CRP level [44]. The mechanism of the interaction is unclear. Although we cannot rule out that the interaction may be due to chance, one possible explanation could be the difference in sex hormones-binding globulin levels between women and men. CRP concentrations are higher in women than men and it is possible that CRP level-related variants have stronger impact on CRP concentrations in women as compared to men [19, 45]. Cross-sectional studies also suggested an inverse association between sex hormone-binding globulin and CRP in men and women [46-49]. In addition, Ruth et al. reported a variant rs1890426 at gene *FRK*, which is in high LD with the variant rs1933736 (r^2=1), was positively associated with sex hormone-binding globulin levels in men while no association was observed in women of European ancestry [50], which further supports this hypothesis.

Our study has several strengths. First, our large sample size facilitated examination of gene-environment interactions. Second, we standardized and harmonized environmental data, and thus bias attributable to heterogeneity in the definitions of environmental variables is likely to be minimal. Moreover, we had a relatively long follow-up period and cause of death was verified, reducing information bias.

We acknowledge some limitations. First, we only analyzed 618 out of 651 CRP-related SNPs identified by GWAS catalog due to lack of data. Thus, some influential SNPs may have been omitted. Second, we had some missing data on smoking and alcohol consumption slightly reducing our ability to exam those factors. However, the proportions of missing data (2.9% for smoking and alcohol consumption) were small and unlikely to bias our results. In addition, our findings have not been replicated due to lack of a replication cohort. Lastly, although our sample size was large for our main analysis, some stratified analyses for rare variants may be underpowered. Although we used conservative approach to correct for multiple comparisons in our GxE analysis, it is possible that our finding on the interaction between SNP rs1933736 and sex may be due to chance.

In summary, we found that 618 CRP-associated variants were not associated with CRC-specific mortality after correcting for multiple comparisons. However, we observed an interaction between the variant rs1933736 and sex. Although further validation is needed, these findings may provide novel insights for strategies of improving CRC prognosis in population subgroups. Validation in additional populations and investigation of the biological function of this SNP in relation to CRC are needed.

Conflict of Interest: The authors declare no potential conflicts of interest.

Data Availability Statement
The summary statistics of the 618 common CRP-related SNPs analyzed in this study are listed in Supplementary Table 2. Further data that support the findings of this study are available from the corresponding author upon reasonable request.

Ethics Statement
All participants provided written or oral informed consent, and studies were reviewed and approved by their respective institutional review boards or ethics committees.

Funding:
ISACC: National Cancer Institute, National Institutes of Health, U.S. Department of Health and Human Services (R01 CA176272). This research was funded in part through the NIH/NCI Cancer Center Support Grant P30 CA015704. Scientific Computing Infrastructure at Fred Hutch funded by ORIP grant S10OD028685.

The Colon Cancer Family Registry (CCFR, www.coloncfr.org) is supported in part by funding from the National Cancer Institute (NCI), National Institutes of Health (NIH) (award U01 CA167551). Support for case ascertainment was provided in part from the Surveillance, Epidemiology, and End Results (SEER) Program and the following U.S. state cancer registries: AZ, CO, MN, NC, NH; and by the Victoria Cancer Registry (Australia) and Ontario Cancer Registry (Canada). The CCFR Set-1 (Illumina 1M/1M-Duo) and Set-2 (Illumina Omni1-Quad) scans were supported by NIH awards U01 CA122839 and R01 CA143247 (to GC). The CCFR Set-3 (Affymetrix Axiom CORECT Set array) was supported by NIH award U19 CA148107 and R01 CA81488 (to SBG). The CCFR Set-4 (Illumina OncoArray 600K SNP array) was supported by NIH award U19 CA148107 (to SBG) and by the Center for Inherited Disease Research (CIDR), which is funded by the NIH to the Johns Hopkins University, contract number HHSN268201200008I. Additional funding for the OFCCR/ARCTIC was through award GL201-043 from the Ontario Research Fund (to BWZ), award 112746 from the Canadian Institutes of Health Research (to TJH), through a Cancer Risk Evaluation (CaRE) Program grant from the Canadian Cancer Society (to SG), and through generous support from the Ontario Ministry of Research and Innovation. The SFCCR Illumina HumanCytoSNP array was supported in part through NCI/NIH awards U01/U24 CA074794 and R01 CA076366 (to PAN). The content of this manuscript does not necessarily reflect the views or policies of the NCI, NIH 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, any cancer registry, or the CCFR.

CPS-II: The American Cancer Society funds the creation, maintenance, and updating of the Cancer Prevention Study-II (CPS-II) cohort. This study was conducted with Institutional Review Board approval.

DACHS: This work was supported by the German Research Council (BR 1704/6-1, BR 1704/6-3, BR 1704/6-4, CH 117/1-1, HO 5117/2-1, HE 5998/2-1, KL 2354/3-1, RO 2270/8-1 and BR 1704/17-1), the Interdisciplinary Research Program of the National Center for Tumor Diseases (NCT), Germany, and the German Federal Ministry of Education and Research (01KH0404, 01ER0814, 01ER0815, 01ER1505A and 01ER1505B).

DALS: National Institutes of Health (R01 CA048998 to M. L. Slattery).

EDRN: This work is funded and supported by the NCI, EDRN Grant (U01 CA 084968-06).

EPIC: The coordination of EPIC is financially supported by International Agency for Research on Cancer (IARC) and also by the Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London which has additional infrastructure support provided by the NIHR Imperial Biomedical Research Centre (BRC). The national cohorts are supported by:

Danish Cancer Society (Denmark); Ligue Contre le Cancer, Institut Gustave Roussy, Mutuelle Générale de l'Education Nationale, Institut National de la Santé et de la Recherche Médicale (INSERM) (France); German Cancer Aid, German Cancer Research Center (DKFZ), German Institute of Human Nutrition Potsdam- Rehbruecke (DIfE), Federal Ministry of Education and Research (BMBF) (Germany); Associazione Italiana per la Ricerca sul Cancro-AIRC-Italy, Compagnia di SanPaolo and National Research Council (Italy); Dutch Ministry of Public Health, Welfare and Sports (VWS), Netherlands Cancer Registry (NKR), LK Research Funds, Dutch Prevention Funds, Dutch ZON (Zorg Onderzoek Nederland), World Cancer Research Fund (WCRF), Statistics Netherlands (The Netherlands); Health Research Fund (FIS) - Instituto de Salud Carlos III (ISCIII), Regional Governments of Andalucía, Asturias, Basque Country, Murcia and Navarra, and the Catalan Institute of Oncology - ICO (Spain); Swedish Cancer Society, Swedish Research Council and County Councils of Skåne and Västerbotten (Sweden); Cancer Research UK (14136 to EPIC-Norfolk; C8221/A29017 to EPIC-Oxford), Medical Research Council (1000143 to EPIC-Norfolk; MR/M012190/1 to EPIC-Oxford). (United Kingdom).

Harvard cohorts (HPFS, NHS, PHS): HPFS is supported by the National Institutes of Health (P01 CA055075, UM1 CA167552, U01 CA167552, R01 CA137178, R01 CA151993, and R35 CA197735), NHS by the National Institutes of Health (R01 CA137178, P01 CA087969, UM1 CA186107, R01 CA151993, and R35 CA197735) and PHS by the National Institutes of Health (R01 CA042182).

MCCS cohort recruitment was funded by VicHealth and Cancer Council Victoria. The MCCS was further supported by Australian NHMRC grants 509348, 209057, 251553 and 504711 and by infrastructure provided by Cancer Council Victoria. Cases and their vital status were ascertained through the Victorian Cancer Registry (VCR) and the Australian Institute of Health and Welfare (AIHW), including the National Death Index and the Australian Cancer Database.

PLCO: Intramural Research Program of the Division of Cancer Epidemiology and Genetics and supported by contracts from the Division of Cancer Prevention, National Cancer Institute, NIH, DHHS. Funding was provided by National Institutes of Health (NIH), Genes, Environment and Health Initiative (GEI) Z01 CP 010200, NIH U01 HG004446, and NIH GEI U01 HG 004438.

UK Biobank: This research has been conducted using the UK Biobank Resource under Application Number 8614

VITAL: National Institutes of Health (K05 CA154337).

WHI: The WHI program is funded by the National Heart, Lung, and Blood Institute, National Institutes of Health, U.S. Department of Health and Human Services through contracts HHSN268201100046C, HHSN268201100001C, HHSN268201100002C, HHSN268201100003C, HHSN268201100004C, and HHSN271201100004C.

Acknowledgements:

CCFR: The Colon CFR graciously thanks the generous contributions of their 42,505 study participants, dedication of study staff, and the financial support from the U.S. National Cancer Institute, without which this important registry would not exist. The content of this manuscript does not necessarily reflect the views or policies of the NIH or any of the collaborating centers in the CCFR, nor does mention of trade names, commercial products, or organizations imply endorsement by the US Government, any cancer registry, or the CCFR.

CPS-II: The authors thank the CPS-II participants and Study Management Group for their invaluable contributions to this research. The authors would also like to acknowledge the contribution to this study from central cancer registries supported through the Centers for Disease Control and Prevention National Program of Cancer Registries, and cancer registries supported by the National Cancer Institute Surveillance Epidemiology and End Results program.

DACHS: We thank all participants and cooperating clinicians, and everyone who provided for excellent technical assistance.

EDRN: We acknowledge all the following contributors to the development of the resource: University of Pittsburgh School of Medicine, Department of Gastroenterology, Hepatology and Nutrition: Lynda Dzubinski; University of Pittsburgh School of Medicine, Department of Pathology: Michelle Bisceglia; and University of Pittsburgh School of Medicine, Department of Biomedical Informatics.

EPIC: Where authors are identified as personnel of the International Agency for Research on Cancer/World Health Organization, the authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy or views of the International Agency for Research on Cancer/World Health Organization.

Harvard cohorts (HPFS, NHS, PHS): The study protocol was approved by the institutional review boards of the Brigham and Women's Hospital and Harvard T.H. Chan School of Public Health, and those of participating registries as required. We would like to thank the participants and staff of the HPFS, NHS and PHS for their valuable contributions as well as the following state cancer registries for their help: AL, AZ, AR, CA, CO, CT, DE, FL, GA, ID, IL, IN, IA, KY, LA, ME, MD, MA, MI, NE, NH, NJ, NY, NC, ND, OH, OK, OR, PA, RI, SC, TN, TX, VA, WA, WY. The authors assume full responsibility for analyses and interpretation of these data.

PLCO: The authors thank the PLCO Cancer Screening Trial screening center investigators and the staff from Information Management Services Inc and Westat Inc. Most importantly, we thank the study participants for their contributions that made this study possible. Cancer incidence data have been provided by the District of Columbia Cancer Registry, Georgia Cancer Registry, Hawaii Cancer Registry, Minnesota Cancer Surveillance System, Missouri Cancer Registry, Nevada Central Cancer Registry, Pennsylvania Cancer Registry, Texas Cancer Registry, Virginia Cancer Registry, and Wisconsin Cancer Reporting System. All are supported in part by funds from the Center for Disease Control and Prevention, National Program for Central Registries, local states or by the National Cancer Institute, Surveillance, Epidemiology, and End Results program. The results reported here and the conclusions derived are the sole responsibility of the authors.

SFCCR: The authors would like to thank the study participants and staff of the Seattle Colon Cancer Family Registry and the Hormones and Colon Cancer study (CORE Studies).

WHI: The authors thank the WHI investigators and staff for their dedication, and the study participants for making the program possible. A full listing of WHI investigators can be found at: http://www.whi.org/researchers/Documents%20%20Write%20a%20Paper/WHI%20Investigator%20Short%20List.pdf

Reference

1. Lasry A, Zinger A, Ben-Neriah Y. Inflammatory networks underlying colorectal cancer. *Nat Immunol* 2016;**17**: 230-40.

2. Erlinger TP, Platz EA, Rifai N, Helzlsouer KJ. C-reactive protein and the risk of incident colorectal cancer. *JAMA* 2004;**291**: 585-90.

3. Gunter MJ, Stolzenberg-Solomon R, Cross AJ, Leitzmann MF, Weinstein S, Wood RJ, Virtamo J, Taylor PR, Albanes D, Sinha R. A prospective study of serum C-reactive protein and colorectal cancer risk in men. *Cancer Res* 2006;**66**: 2483-7.

4. Koike Y, Miki C, Okugawa Y, Yokoe T, Toiyama Y, Tanaka K, Inoue Y, Kusunoki M. Preoperative C-reactive protein as a prognostic and therapeutic marker for colorectal cancer. *J Surg Oncol* 2008;**98**: 540-4.

5. Otani T, Iwasaki M, Sasazuki S, Inoue M, Tsugane S, Japan Public Health Center-Based Prospective Study G. Plasma C-reactive protein and risk of colorectal cancer in a nested case-control study: Japan Public Health Center-based prospective study. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* 2006;**15**: 690-5.

6. Woo HD, Kim K, Kim J. Association between preoperative C-reactive protein level and colorectal cancer survival: a meta-analysis. *Cancer causes & control : CCC* 2015;**26**: 1661-70.

7. Dehghan A, Dupuis J, Barbalic M, Bis JC, Eiriksdottir G, Lu C, Pellikka N, Wallaschofski H, Kettunen J, Henneman P, Baumert J, Strachan DP, et al. Meta-analysis of genome-wide association studies in >80 000 subjects identifies multiple loci for C-reactive protein levels. *Circulation* 2011;**123**: 731-8.

8. Ligthart S, Vaez A, Vosa U, Stathopoulou MG, de Vries PS, Prins BP, Van der Most PJ, Tanaka T, Naderi E, Rose LM, Wu Y, Karlsson R, et al. Genome Analyses of >200,000 Individuals Identify 58 Loci for Chronic Inflammation and Highlight Pathways that Link Inflammation and Complex Disorders. *Am J Hum Genet* 2018;**103**: 691-706.

9. Nimptsch K, Aleksandrova K, Boeing H, Janke J, Lee YA, Jenab M, Bueno-de-Mesquita HB, Jansen EH, Tsilidis KK, Trichopoulou A, Weiderpass E, Wu C, et al. Association of CRP genetic variants with blood concentrations of C-reactive protein and colorectal cancer risk. *Int J Cancer* 2015;**136**: 1181-92.

10. Slattery ML, Curtin K, Poole EM, Duggan DJ, Samowitz WS, Peters U, Caan BJ, Potter JD, Ulrich CM. Genetic variation in C-reactive protein in relation to colon and rectal cancer risk and survival. *Int J Cancer* 2011;**128**: 2726-34.

11. Choi J, Joseph L, Pilote L. Obesity and C-reactive protein in various populations: a systematic review and meta-analysis. *Obes Rev* 2013;**14**: 232-44.

12. Shiels MS, Katki HA, Freedman ND, Purdue MP, Wentzensen N, Trabert B, Kitahara CM, Furr M, Li Y, Kemp TJ, Goedert JJ, Chang CM, et al. Cigarette smoking and variations in systemic immune and inflammation markers. *J Natl Cancer Inst* 2014;**106**.

13. Imhof A, Froehlich M, Brenner H, Boeing H, Pepys MB, Koenig W. Effect of alcohol consumption on systemic markers of inflammation. *Lancet* 2001;**357**: 763-7.

14. Raum E, Gebhardt K, Buchner M, Schiltenwolf M, Brenner H. Long-term and short-term alcohol consumption and levels of C-reactive protein. *Int J Cardiol* 2007;**121**: 224-6.

15. Rexrode KM, Pradhan A, Manson JE, Buring JE, Ridker PM. Relationship of total and abdominal adiposity with CRP and IL-6 in women. *Ann Epidemiol* 2003;**13**: 674-82.

16. Walter V, Jansen L, Ulrich A, Roth W, Blaker H, Chang-Claude J, Hoffmeister M, Brenner H. Alcohol consumption and survival of colorectal cancer patients: a population-based study from Germany. *Am J Clin Nutr* 2016;**103**: 1497-506.

17. Daniel CR, Shu X, Ye Y, Gu J, Raju GS, Kopetz S, Wu X. Severe obesity prior to diagnosis limits survival in colorectal cancer patients evaluated at a large cancer centre. *Br J Cancer* 2016;**114**: 103-9.

18. Phipps AI, Baron J, Newcomb PA. Prediagnostic smoking history, alcohol consumption, and colorectal cancer survival: the Seattle Colon Cancer Family Registry. *Cancer* 2011;**117**: 4948-57.

19. Cartier A, Cote M, Lemieux I, Perusse L, Tremblay A, Bouchard C, Despres JP. Sex differences in inflammatory markers: what is the contribution of visceral adiposity? *Am J Clin Nutr* 2009;**89**: 1307-14.

20. White A, Ironmonger L, Steele RJC, Ormiston-Smith N, Crawford C, Seims A. A review of sex-related differences in colorectal cancer incidence, screening uptake, routes to diagnosis, cancer stage and survival in the UK. *BMC Cancer* 2018;**18**: 906.

21. Newcomb PA, Baron J, Cotterchio M, Gallinger S, Grove J, Haile R, Hall D, Hopper JL, Jass J, Le Marchand L, Limburg P, Lindor N, et al. Colon Cancer Family Registry: an international resource for studies of the genetic epidemiology of colon cancer. *Cancer Epidemiol Biomarkers Prev* 2007;**16**: 2331-43.

22. Calle EE, Rodriguez C, Jacobs EJ, Almon ML, Chao A, McCullough ML, Feigelson HS, Thun MJ. The American Cancer Society Cancer Prevention Study II Nutrition Cohort: rationale, study design, and baseline characteristics. *Cancer* 2002;**94**: 500-11.

23. Brenner H, Chang-Claude J, Seiler CM, Rickert A, Hoffmeister M. Protection from colorectal cancer after colonoscopy: a population-based, case-control study. *Ann Intern Med* 2011;**154**: 22-30.

24. Slattery ML, Potter J, Caan B, Edwards S, Coates A, Ma KN, Berry TD. Energy balance and colon cancer--beyond physical activity. *Cancer Res* 1997;**57**: 75-80.

25. Amin W, Singh H, Dzubinski LA, Schoen RE, Parwani AV. Design and utilization of the colorectal and pancreatic neoplasm virtual biorepository: An early detection research network initiative. *J Pathol Inform* 2010;**1**: 22.

26. Riboli E, Kaaks R. The EPIC Project: rationale and study design. European Prospective Investigation into Cancer and Nutrition. *Int J Epidemiol* 1997;**26 Suppl 1**: S6-14.

27. Rimm EB, Stampfer MJ, Colditz GA, Chute CG, Litin LB, Willett WC. Validity of self-reported waist and hip circumferences in men and women. *Epidemiology* 1990;**1**: 466-73.

28. Giles GG, English DR. The Melbourne Collaborative Cohort Study. *IARC Sci Publ* 2002;**156**: 69-70.

29. Belanger CF, Hennekens CH, Rosner B, Speizer FE. The Nurses' Health Study 1978;**78**: 1039-40.

30. Colditz GA, Manson JE, Hankinson SE. The Nurses' Health Study: 20-year contribution to the understanding of health among women. *J Womens Health* 1997;**6**: 49-62.

31. Christen WG, Gaziano JM, Hennekens CH. Design of Physicians' Health Study II--a randomized trial of beta-carotene, vitamins E and C, and multivitamins, in prevention of cancer, cardiovascular disease, and eye disease, and review of results of completed trials. *Ann Epidemiol* 2000;**10**: 125-34.

32. Gohagan JK, Prorok PC, Hayes RB, Kramer BS, Prostate LC, Ovarian Cancer Screening Trial Project T. The Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial of the National Cancer Institute: history, organization, and status. *Control Clin Trials* 2000;**21**: 251S-72S.

33. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, Downey P, Elliott P, Green J, Landray M, Liu B, Matthews P, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med* 2015;**12**: e1001779.

34. White E, Patterson RE, Kristal AR, Thornquist M, King I, Shattuck AL, Evans I, Satia-Abouta J, Littman AJ, Potter JD. VITamins And Lifestyle cohort study: study design and characteristics of supplement users. *Am J Epidemiol* 2004;**159**: 83-93.

35. Hays J, Hunt JR, Hubbell FA, Anderson GL, Limacher M, Allen C, Rossouw JE. The Women's Health Initiative recruitment methods and results. *Ann Epidemiol* 2003;**13**: S18-77.

36. Hutter CM, Chang-Claude J, Slattery ML, Pflugeisen BM, Lin Y, Duggan D, Nan H, Lemire M, Rangrej J, Figueiredo JC, Jiao S, Harrison TA, et al. Characterization of gene-environment interactions for colorectal cancer susceptibility loci. *Cancer Res* 2012;**72**: 2036-44.

37. Huyghe JR, Bien SA, Harrison TA, Kang HM, Chen S, Schmit SL, Conti DV, Qu C, Jeon J, Edlund CK, Greenside P, Wainberg M, et al. Discovery of common and rare genetic risk variants for colorectal cancer. *Nat Genet* 2019;**51**: 76-87.

38. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007;**81**: 559-75.

39. Das S, Forer L, Schonherr S, Sidore C, Locke AE, Kwong A, Vrieze SI, Chew EY, Levy S, McGue M, Schlessinger D, Stambolian D, et al. Next-generation genotype imputation service and methods. *Nat Genet* 2016;**48**: 1284-7.

40. Delaneau O, Marchini J, Genomes Project C, Genomes Project C. Integrating sequence and array data to create an improved 1000 Genomes Project haplotype reference panel. *Nat Commun* 2014;**5**: 3934.

41. McCarthy S, Das S, Kretzschmar W, Delaneau O, Wood AR, Teumer A, Kang HM, Fuchsberger C, Danecek P, Sharp K, Luo Y, Sidore C, et al. A reference panel of 64,976 haplotypes for genotype imputation. *Nat Genet* 2016;**48**: 1279-83.

42. GRAMBSCH PM, THERNEAU TM. Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika* 1994;**81**: 515-26.

43. Gao X, Starmer J, Martin ER. A multiple testing correction method for genetic association studies using correlated single nucleotide polymorphisms. *Genet Epidemiol* 2008;**32**: 361-9.

44. Han X, Ong JS, An J, Hewitt AW, Gharahkhani P, MacGregor S. Using Mendelian randomization to evaluate the causal relationship between serum C-reactive protein levels and age-related macular degeneration. *Eur J Epidemiol* 2020;**35**: 139-46.

45. Shanahan L, Copeland WE, Worthman CM, Erkanli A, Angold A, Costello EJ. Sex-differentiated changes in C-reactive protein from ages 9 to 21: the contributions of BMI and physical/sexual maturation. *Psychoneuroendocrinology* 2013;**38**: 2209-17.

46. Joffe HV, Ridker PM, Manson JE, Cook NR, Buring JE, Rexrode KM. Sex hormone-binding globulin and serum testosterone are inversely associated with C-reactive protein levels in postmenopausal women at high risk for cardiovascular disease. *Ann Epidemiol* 2006;**16**: 105-12.

47. Kupelian V, Chiu GR, Araujo AB, Williams RE, Clark RV, McKinlay JB. Association of sex hormones and C-reactive protein levels in men. *Clin Endocrinol (Oxf)* 2010;**72**: 527-33.

48. Maggio M, Ceda GP, Lauretani F, Bandinelli S, Corsi AM, Giallauria F, Guralnik JM, Zuliani G, Cattabiani C, Parrino S, Ablondi F, Dall'aglio E, et al. SHBG, sex hormones, and inflammatory markers in older women. *J Clin Endocrinol Metab* 2011;**96**: 1053-9.

49. Liao CH, Li HY, Yu HJ, Chiang HS, Lin MS, Hua CH, Ma WY. Low serum sex hormone-binding globulin: marker of inflammation? *Clin Chim Acta* 2012;**413**: 803-7.

50. Ruth KS, Day FR, Tyrrell J, Thompson DJ, Wood AR, Mahajan A, Beaumont RN, Wittemans L, Martin S, Busch AS, Erzurumluoglu AM, Hollis B, et al. Using human genetics to understand the disease impacts of testosterone in men and women. *Nat Med* 2020;**26**: 252-8.

Novelty and Impact:

Chronic inflammation is strongly associated with colorectal cancer (CRC). C-reactive protein (CRP) is a sensitive marker of this type of inflammation, and may be associated with poor prognosis. In this large genetic study, the authors found that none of over 600 CRP-associated SNPs were associated with CRC-specific mortality. However, they did identify a genetic variant of the FRK gene that may affect levels of CRP. This variant was associated with increased CRC mortality in women, while it showed the opposite effect in men. These findings may provide novel insights for improving CRC prognosis.

Table 1. Baseline characteristics of Study Participants[a]

Variables		Number of colorectal cancer cases, N (%) (n=16142)	Number of colorectal cancer deaths, N (%) (n=3547)
Age at diagnosis, years, mean (SD)		66.4 (9.8)	66.6 (10.0)
Age at diagnosis, years			
	<50	890 (5.5)	197 (5.6)
	50-59	2598 (16.1)	591 (16.7)
	60-69	6267 (38.8)	1296 (36.5)
	70-79	5146 (31.9)	1147 (32.3)
	80+	1241 (7.7)	316 (8.9)
Male		8125 (50.3)	1798 (50.7)
Body Mass Index			
	Normal	5189 (32.1)	1137 (32.1)
	Overweight	7041 (43.6)	1494 (42.1)
	Obese	3912 (24.2)	916 (25.8)
Ever smoker			
	No	6940 (43.0)	1527 (43.1)
	Yes	8732 (54.1)	1897 (53.5)
	Missing	470 (2.9)	123 (3.5)
Alcohol consumption			
	Nondrinker (0-1 g/day)	5580 (34.6)	1350 (38.1)
	Moderate drinker (1.01-28 g/day)	7643 (47.3)	1561 (44.0)
	Heavy drinker (>28 g/day)	2449 (15.2)	553 (15.6)
	Missing	470 (2.9)	83 (2.3)

SD = Standard Deviation

[a] Age at diagnosis, sex, body mass index, ever smoking, and alcohol consumption, were measured by in-person or phone interviews, structured questionnaires, or both with the reference time for cohort studies as the time of enrollment.

Table 2. HR and 95% CI for the association between SNP rs1933736 and CRC-specific mortality, stratified by sex

SNP	Gene	CA/RA	CAF	HR (95% CI)[a] Female	HR (95% CI)[a] Male	$P_{interaction}$	Corrected $P_{interaction}$[b]
rs1933736	*FRK*	C/T	0.42	1.11 (1.04, 1.19)	0.88 (0.82, 0.94)	8.00×10^{-7}	0.0004

SNP = Single-Nucleotide Polymorphisms; CA = Coded Allele; RA = Reference Allele; CAF = Coded Allele Frequency; HR = Hazard Ratio; CI = Confidence Interval; CRC = Colorectal Cancer.

[a] Adjusted for age at diagnosis categories, genotyping platform/study, and the first three principal components.

[b] *P* values were adjusted using Bonferroni method using M_{eff_G}.