UCSF

UC San Francisco Previously Published Works

Title

Genome-Wide Interaction Analysis of Genetic Variants With Menopausal Hormone Therapy for Colorectal Cancer Risk.

Permalink

https://escholarship.org/uc/item/3vq0n596

Journal

Journal of the National Cancer Institute, 114(8)

ISSN

0027-8874

Authors

Tian, Yu
Kim, Andre E
Bien, Stephanie A
et al.

Publication Date

2022-08-08

DOI

10.1093/jnci/djac094

Peer reviewed

Genome-Wide Interaction Analysis of Genetic Variants With Menopausal Hormone Therapy for Colorectal Cancer Risk

Yu Tian, PhD,[1,2,†] Andre E. Kim, PhD,[3,†] Stephanie A. Bien, PhD,[4] Yi Lin, MS,[4] Conghui Qu, MS,[4] Tabitha A. Harrison, MPH,[4] Robert Carreras-Torres, PhD,[5,6,7] Virginia Díez-Obrero, PhD,[5,6,7] Niki Dimou, PhD,[8] David A. Drew, PhD,[9,10] Akihisa Hidaka, MD, PhD,[4] Jeroen R. Huyghe, PhD,[4] Kristina M. Jordahl, PhD,[4,11] John Morrison, MS,[3] Neil Murphy, PhD,[8] Mireia Obón-Santacana, PhD,[5,6,7] Cornelia M. Ulrich, PhD,[12,13] Jennifer Ose, PhD,[12,13] Anita R. Peoples, PhD, MPH,[12,13] Edward A. Ruiz-Narváez, ScD,[14] Anna Shcherbina, PhD,[15] Mariana C. Stern, PhD,[16] Yu-Ru Su, PhD,[4,17] Franzel J.B. van Duijnhoven, PhD,[18] Volker Arndt, MD, MPH,[19] James W. Baurley, PhD,[20,21] Sonja I. Berndt, PharmD, PhD,[22] D. Timothy Bishop, PhD,[23] Hermann Brenner, MD, MPH,[19,24,25] Daniel D. Buchanan, PhD,[26,27,28] Andrew T. Chan, MD, MPH,[9,10,29,30,31,32] Jane C. Figueiredo, PhD,[33,34] Steven Gallinger, MD,[35] Stephen B. Gruber, MD, MPH, PhD,[16] Sophia Harlid, PhD,[36] Michael Hoffmeister, PhD,[19] Mark A. Jenkins, PhD,[37] Amit D. Joshi, PhD,[31] Temitope O. Keku, PhD, MSPH,[38] Susanna C. Larsson, PhD,[39] Loic Le Marchand, MD, PhD,[40] Li Li, MD, PhD,[41] Graham G. Giles, PhD,[37,42,43] Roger L. Milne, PhD,[37,42,43] Hongmei Nan, MD, PhD,[44,45] Rami Nassir, PhD,[46] Shuji Ogino, MD, PhD,[30,31,47,48] Arif Budiarto, MS,[20] Elizabeth A. Platz, ScD, MPH,[49] John D. Potter, MBBS, PhD,[4,50] Ross L. Prentice, PhD,[4,51] Gad Rennert, MD, PhD,[52,53,54] Lori C. Sakoda, PhD,[4,55] Robert E. Schoen, MD, MPH,[56] Martha L. Slattery, PhD, MPH,[57] Stephen N. Thibodeau, PhD,[58] Bethany Van Guelpen, MD, PhD, MSSc,[36,59] Kala Visvanathan, MD, MHS,[49] Emily White, PhD,[4,11] Alicja Wolk, PhD,[39] Michael O. Woods, PhD,[60] Anna H. Wu, PhD,[34] Peter T. Campbell, PhD,[61] Graham Casey, PhD,[62] David V. Conti, PhD,[16] Marc J. Gunter, PhD,[8] Anshul Kundaje, PhD,[63,64] Juan Pablo Lewinger, PhD,[3] Victor Moreno, PhD,[5,6,7,65] Polly A. Newcomb, PhD,[4,11] Bens Pardamean, PhD,[20] Duncan C. Thomas, PhD,[16] Konstantinos K. Tsilidis, PhD,[66,67] Ulrike Peters, PhD,[4,11,‡] W. James Gauderman, PhD,[16,‡] Li Hsu, PhD,[4,51,‡] and Jenny Chang-Claude, PhD[1,68,*,‡]

[1]Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany; [2]School of Public Health, Capital Medical University, Beijing, China; [3]Division of Biostatistics, Department of Population and Public Health Sciences, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA; [4]Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA; [5]Colorectal Cancer Group, ONCOBELL Program, Bellvitge Biomedical Research Institute (IDIBELL), L'Hospitalet de Llobregat, Barcelona, Spain; [6]Oncology Data Analytics Program, Catalan Institute of Oncology, L'Hospitalet de Llobregat, Barcelona, Spain; [7]Consortium for Biomedical Research in Epidemiology and Public Health (CIBERESP), Madrid, Spain; [8]Nutrition and Metabolism Branch, International Agency for Research on Cancer, Lyon, France; [9]Clinical and Translational Epidemiology Unit, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA; [10]Division of Gastroenterology, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA; [11]Department of Epidemiology, University of Washington School of Public Health, Seattle, WA, USA; [12]Huntsman Cancer Institute, Salt Lake City, UT, USA; [13]Department of Population Health Sciences, University of Utah, Salt Lake City, UT, USA; [14]Department of Nutritional Sciences, University of Michigan School of Public Health, Ann Arbor, MI, USA; [15]Biomedical Informatics Program, Department of Biomedical Data Sciences, Stanford University, Stanford, CA, USA; [16]Division of Biostatistics, Department of Population and Public Health Sciences & USC Norris Comprehensive Cancer Center, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA; [17]Kaiser Permanente Washington Health Research Institute, Seattle, WA, USA; [18]Division of Human Nutrition and Health, Wageningen University & Research, Wageningen, The Netherlands; [19]Division of Clinical Epidemiology and Aging Research, German Cancer Research Center (DKFZ), Heidelberg, Germany; [20]Bioinformatics and Data Science Research Center, Bina Nusantara University, Jakarta, Indonesia; [21]BioRealm LLC, Walnut, CA, USA; [22]Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA; [23]Leeds Institute of Cancer and Pathology, University of Leeds, Leeds, UK; [24]Division of Preventive Oncology, German Cancer Research Center (DKFZ) and National Center for Tumor Diseases (NCT), Heidelberg, Germany; [25]German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Heidelberg, Germany; [26]Colorectal Oncogenomics Group, Department of Clinical Pathology, The University of Melbourne, Parkville, Victoria, Australia; [27]University of Melbourne Centre for Cancer Research, Victorian Comprehensive Cancer Centre, Parkville, Victoria, Australia; [28]Genomic Medicine and

Family Cancer Clinic, The Royal Melbourne Hospital, Parkville, Victoria, Australia; [29]Channing Division of Network Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA; [30]Broad Institute of Harvard and MIT, Cambridge, MA, USA; [31]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Harvard University, Boston, MA, USA; [32]Department of Immunology and Infectious Diseases, Harvard T.H. Chan School of Public Health, Harvard University, Boston, MA, USA; [33]Department of Medicine, Samuel Oschin Comprehensive Cancer Institute, Cedars-Sinai Medical Center, Los Angeles, CA, USA; [34]Department of Population and Public Health Sciences, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA; [35]Lunenfeld Tanenbaum Research Institute, Mount Sinai Hospital, University of Toronto, Toronto, Ontario, Canada; [36]Department of Radiation Sciences, Oncology Unit, Umeå University, Umeå, Sweden; [37]Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, Victoria, Australia; [38]Center for Gastrointestinal Biology and Disease, University of North Carolina, Chapel Hill, NC, USA; [39]Institute of Environmental Medicine, Karolinska Institute, Stockholm, Sweden; [40]University of Hawaii Cancer Center, Honolulu, HI, USA; [41]Department of Family Medicine, University of Virginia, Charlottesville, VA, USA; [42]Cancer Epidemiology Division, Cancer Council Victoria, Melbourne, Victoria, Australia; [43]Precision Medicine, School of Clinical Sciences at Monash Health, Monash University, Clayton, Victoria, Australia; [44]Department of Global Health, Richard M. Fairbanks School of Public Health, Indianapolis, IN, USA; [45]Department of Epidemiology, Richard M. Fairbanks School of Public Health, Indianapolis, IN, USA; [46]Department of Pathology, School of Medicine, Umm Al-Qura'a University, Mecca, Saudi Arabia; [47]Program in MPE Molecular Pathological Epidemiology, Department of Pathology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA; [48]Department of Oncologic Pathology, Dana-Farber Cancer Institute, Boston, MA, USA; [49]Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA; [50]Research Centre for Hauora and Health, Massey University, Wellington, New Zealand; [51]Department of Biostatistics, University of Washington, Seattle, WA, USA; [52]Department of Community Medicine and Epidemiology, Lady Davis Carmel Medical Center, Haifa, Israel; [53]Ruth and Bruce Rappaport Faculty of Medicine, Technion-Israel Institute of Technology, Haifa, Israel; [54]Clalit National Cancer Control Center, Haifa, Israel; [55]Division of Research, Kaiser Permanente Northern California, Oakland, CA, USA; [56]Department of Medicine and Epidemiology, University of Pittsburgh Medical Center, Pittsburgh, PA, USA; [57]Department of Internal Medicine, University of Utah, Salt Lake City, UT, USA; [58]Division of Laboratory Genetics, Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN, USA; [59]Wallenberg Centre for Molecular Medicine, Umeå University, Umeå, Sweden; [60]Memorial University of Newfoundland, Discipline of Genetics, St. John's, NL,Canada; [61]Department of Population Science, American Cancer Society, Atlanta, GA, USA; [62]Center for Public Health Genomics, University of Virginia, Charlottesville, VA, USA; [63]Department of Genetics, Stanford University, Stanford, CA, USA; [64]Department of Computer Science, Stanford University, Stanford, CA, USA; [65]Department of Clinical Sciences, Faculty of Medicine, University of Barcelona, Barcelona, Spain; [66]Department of Epidemiology and Biostatistics, Imperial College London, School of Public Health, London, UK; [67]Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece; and [68]University Cancer Centre Hamburg (UCCH), University Medical Centre Hamburg-Eppendorf, Hamburg, Germany

[†]These authors contributed equally to this work.

[‡]These authors jointly supervised this work.

[*]**Correspondence to:** Jenny Chang-Claude, PhD, Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 581, 69120 Heidelberg, Germany (e-mail: j.chang-claude@dkfz.de)

Abstract

Background: The use of menopausal hormone therapy (MHT) may interact with genetic variants to influence colorectal cancer (CRC) risk. **Methods:** We conducted a genome-wide, gene-environment interaction between single nucleotide polymorphisms and the use of any MHT, estrogen only, and combined estrogen-progestogen therapy with CRC risk, among 28 486 postmenopausal women (11 519 CRC patients and 16 967 participants without CRC) from 38 studies, using logistic regression, 2-step method, and 2– or 3–degree-of-freedom joint test. A set-based score test was applied for rare genetic variants. **Results:** The use of any MHT, estrogen only and estrogen-progestogen were associated with a reduced CRC risk (odds ratio [OR] = 0.71, 95% confidence interval [CI] = 0.64 to 0.78; OR = 0.65, 95% CI = 0.53 to 0.79; and OR = 0.73, 95% CI = 0.59 to 0.90, respectively). The 2-step method identified a statistically significant interaction between a *GRIN2B* variant rs117868593 and MHT use, whereby MHT-associated CRC risk was statistically significantly reduced in women with the GG genotype (OR = 0.68, 95% CI = 0.64 to 0.72) but not within strata of GC or CC genotypes. A statistically significant interaction between a *DCBLD1* intronic variant at 6q22.1 (rs10782186) and MHT use was identified by the 2–degree-of-freedom joint test. The MHT-associated CRC risk was reduced with increasing number of rs10782186-C alleles, showing odds ratios of 0.78 (95% CI = 0.70 to 0.87) for TT, 0.68 (95% CI = 0.63 to 0.73) for TC, and 0.66 (95% CI = 0.60 to 0.74) for CC genotypes. In addition, 5 genes in rare variant analysis showed suggestive interactions with MHT (2-sided $P < 1.2 \times 10^{-4}$). **Conclusion:** Genetic variants that modify the association between MHT and CRC risk were identified, offering new insights into pathways of CRC carcinogenesis and potential mechanisms involved.

The use of menopausal hormone therapy (MHT) has been identified to be associated with a reduced risk of colorectal cancer (CRC) (1-4). In a meta-analysis including 20 studies, ever use of estrogen-only MHT (relative risk [RR] = 0.79, 95% confidence interval [CI] = 0.69 to 0.91) and ever use of combined estrogen-progestogen MHT (RR = 0.74, 95% CI = 0.68 to 0.81) were associated with a reduced CRC risk (1).

Previous gene-environment (GxE) interaction studies that investigated the association of MHT use with CRC risk according to genetic variants (5-10) have reported a few potential genetic modifiers of CRC risk associated with the use of MHT; however, these studies were based on limited candidate genes and/or pathways or limited sample size. We conducted a comprehensive genome-wide GxE analysis of common and rare genetic variants, using the largest known study sample to date, on one hand, to identify novel genetic variants that may modify the beneficial influence of MHT on CRC risk to obtain insight into

potential mechanisms behind the association between MHT and CRC risk. On the other hand, the analysis can yield novel genetic susceptibility alleles for CRC risk, which may not be identified without accounting for the GxE component.

Methods

Study Participants

We included 38 studies from North America, Australia, and Europe participating in the multicentered Colon Cancer Family Registry, the Colorectal Transdisciplinary Study, the Genetics and Epidemiology of Colorectal Cancer Consortium, and the United Kingdom Biobank, which were included in genome-wide association studies (GWAS) as described previously (11-13). Study details and descriptions can be found in the Supplementary Methods (available online). All studies were

approved by their respective institutional review boards, and study participants provided informed consent.

Exposure Assessment

Information on demographics and environmental risk factors were collected by interviews and/or structured questionnaires. We carried out a multistep data-harmonization procedure at the Genetics and Epidemiology of Colorectal Cancer Consortium coordinating center (Fred Hutchinson Cancer Research Center) as described previously (10,14,15).

Postmenopausal status was defined by using 1) menopausal status derived from studies, if available; 2) self-reported menopausal status, if study derived was not available; or 3) aged older than 55 years, if neither study derived nor self-report were available (Supplementary Table 1, available online). MHT use was considered as any MHT use or estrogen-only use or estrogen-progestogen use at or up to the reference time. Nonusers of any MHT at or up to reference time were used as the reference group.

Genotyping, Quality Control, and Imputation

Details on genotyping, imputation, and quality control have been reported previously (16). In brief, genotyped single nucleotide polymorphisms (SNPs) were excluded on the basis of call rate (<98%), evidence of departure from Hardy-Weinberg equilibrium in controls ($P < 1 \times 10^{-4}$). All autosomal SNPs in all studies were imputed to the Haplotype Reference Consortium r1.1 (2016) reference panel via the Michigan Imputation Server (17) and converted into a binary format for data management and analyses using R package BinaryDosage (18). Imputed common SNPs were restricted based on a pooled minor allele frequency (MAF) of at least 1% and imputation accuracy ($R^2 > 0.8$). After imputation and quality control analyses, more than 7.2 million common SNPs were included. All analyses were restricted to samples clustering with the Utah residents of Northern and Western European ancestry (the CEU population) in principal component analysis.

Statistical Methods

Statistical analyses of all data were conducted centrally on individual-level data. All tests of statistical significance were 2-sided. Unless otherwise indicated, we adjusted for age at the reference time, study center, and the first 3 principal components (Plink2) to account for potential population substructure. SNPs were treated as continuous variables (ie, log-additive effects). To evaluate MHT main effects, each study was analyzed separately using logistic regression models, and study-specific results were combined using fixed- and random-effects meta-analysis methods to obtain summary odds ratios (ORs) and 95% confidence intervals across studies. We calculated the heterogeneity P values using Cochran Q statistics (19). Quantile-quantile plots were used to assess whether the distribution of the P values was consistent with the null distribution (except for the extreme tail).

Genome-wide interaction scans of common markers were conducted using R package GxEScanR (20), which implements several interaction testing methods. To test for multiplicative statistical interactions between each SNP and environmental risk factors (MHT, estrogen only, estrogen-progestogen), we primarily used conventional case-control logistic regression

analysis and 2-step methods (21-23) to test the GxE interaction term. Additionally, we also used a 2–degree-of-freedom (2-df) joint test (24) and 3-df joint test (25) to test GxE interaction in the context of simultaneously testing for the association between SNPs and CRC, and the association between SNPs and environmental risk factors (G|E) (MHT, estrogen only, estrogen-progestogen) associations. For the 2- and 3-df test, we do not report on known loci (16). For all novel findings, we examined the odds ratios of MHT, estrogen only, and estrogen-progestogen stratified by genotypes of statistically significant SNPs. More details in these testing methods can be found in the Supplementary Methods (available online).

For interaction analysis of rare genetic risk variants (MAF < 1%) and MHT, we conducted the Mixed effects Score Tests for interaction (MiSTi) (26), a set-based statistical framework providing mixed effects score tests for GxE interaction and addressing issues of power and low effect sizes, to discover genes that interact with MHT in relation to CRC risk (see the Supplementary Methods, available online). Because more than 20 000 genes were tested (22 476 genes for any MHT use, 20 609 for estrogen only, and 20 360 for estrogen-progestogen), interactions with a P value less than 2.5×10^{-6} were considered statistically significant, whereas those with a P value less than 1.2×10^{-4} were considered as suggestive.

Functional Annotation

We performed bioinformatic follow-up for genome-wide interaction study (GWIS) variants that were deemed statistically significant for downstream analysis (for more details, see the Supplementary Methods, available online). Relevant regional plots were generated using the command line version (Standalone) of LocusZoom v1.3 (27). Measures of linkage disequilibrium (LD) were estimated using study population controls.

Results

Detailed descriptive characteristics of the participants are shown in Table 1. MHT use was associated with reduced CRC risk both in cohort studies and case-control studies (Figures 1-3).

Genome-Wide MHT-Interaction Scans for CRC Risk

Statistical interaction results for genetic variants are summarized in Table 2. Although conventional case-control logistic regression models with a Bonferroni correction for multiple testing did not identify any statistically significant interactions between the use of any MHT, estrogen only, or estrogen-progestogen, and genetic variants (data not shown), we identified 2 interactions with common genetic variants reaching statistical significance for the 2-step method and 2-df joint test. The 2-step method (with G|E in step 1) identified a statistically significant interaction for any MHT use with SNP rs117868593 located 20 kb downstream of GRIN2B (Glutamate Ionotropic Receptor N-methyl D-aspartate Type Subunit 2B) variant at 12p13.1 ($P_{observed} = .003$, $P_{threshold} = .005$; Supplementary Figures 1 and 2, available online). The 2-df joint test identified a further statistically significant interaction for any MHT use with a DCBLD1 (Discoidin, CUB [Complement C1r/C1s, Uegf, Bmp1] And LCCL [Limulus factor C, Coch-5b2 and Lgl1] Domain Containing 1) intronic variant at 6q22.1 (rs10782186; joint $P_{observed} = 4.23 \times 10^{-8}$, $P_{threshold} = 5 \times 10^{-8}$; Supplementary Figures 3 and 4,

ARTICLE

Table 1. Descriptive characteristics of study participants included in the genome-wide interaction analysis between common variants and menopausal hormone therapy for risk of colorectal cancer[a]

Study	CRC patients						Participants without CRC					
	Total No.	No use of MHT No. (%)	Any MHT No. (%)	E-only No.	E+P No.	Age at diagnosis Mean (SD), y	Total No.	No use of MHT No. (%)	Any MHT No. (%)	E-only No.	E+P No.	Age at enrollment Mean (SD), y
CCFR Set 1	259	183 (70.7)	76 (29.3)	35	34	58.5 (9.9)	372	206 (55.4)	166 (44.6)	93	65	61.7 (7.9)
CCFR Set 3	427	292 (68.4)	135 (31.6)	68	42	61.8 (7.6)	250	155 (62.0)	95 (38.0)	54	27	62.7 (7.4)
CCFR Set 4	383	259 (67.6)	124 (32.4)	78	41	62.1 (9.3)	118	77 (65.3)	41 (34.7)	16	18	61.7 (9.3)
CLUEII	114	98 (86)	16 (14.0)	—	—	74.9 (9.5)	108	97 (89.8)	11 (10.2)	—	—	65.0 (9.4)
Colo 2&3	37	18 (48.6)	19 (51.4)	—	—	66.5 (11.4)	44	17 (38.6)	27 (61.4)	—	—	67.3 (9.2)
CPSII_1	263	176 (66.9)	87 (33.1)	50	37	74.8 (5.9)	255	142 (55.7)	113 (44.3)	83	30	74.3 (5.8)
CPSII_2	172	116 (67.4)	56 (32.6)	35	21	79.4 (6.1)	177	101 (57.1)	76 (42.9)	46	30	79.0 (6.0)
CRCGEN	274	266 (97.1)	8 (2.9)	—	—	68.8 (9.9)	394	377 (95.7)	17 (4.3)	—	—	65.9 (9.2)
DACHS_1	630	416 (66.0)	214 (34.0)	—	—	71.1 (9.5)	630	294 (46.7)	336 (53.3)	—	—	70.4 (8.7)
DACHS_2	229	161 (70.3)	68 (29.7)	—	—	72.2 (9.8)	162	88 (54.3)	74 (45.7)	—	—	72.3 (9.0)
DACHS_3	420	297 (70.7)	123 (29.3)	—	—	71.4 (9.5)	195	113 (57.9)	82 (42.1)	—	—	70.3 (10.1)
DALS_1	267	204 (76.4)	63 (23.6)	—	—	68.0 (7.7)	270	189 (70.0)	81 (30.0)	—	—	67.7 (7.9)
DALS_2	159	127 (79.9)	32 (20.1)	—	—	67.5 (7.5)	194	137 (70.6)	57 (29.4)	—	—	67.5 (8.2)
EPIC	771	544 (70.6)	227 (29.4)	—	—	67.2 (6.6)	865	619 (71.6)	246 (28.4)	—	—	72.5 (5.9)
ESTHER_VERDI	70	52 (74.3)	18 (25.7)	—	—	68.4 (6.8)	70	49 (70.0)	21 (30.0)	—	—	65.8 (6.7)
Kentucky	397	184 (46.3)	213 (53.7)	100	56	64.4 (8.9)	525	150 (28.6)	375 (71.4)	166	86	66.7 (6.6)
LCCS	116	90 (77.6)	26 (22.4)	—	—	66.1 (6.9)	108	88 (81.5)	20 (18.5)	—	—	65.7 (5.5)
MCCS_1	211	159 (75.4)	52 (24.6)	—	—	72.0 (7.1)	184	132 (71.7)	52 (28.3)	—	—	71.1 (7.2)
MCCS_2	85	65 (76.5)	20 (23.5)	—	—	74.3 (8.4)	86	65 (75.6)	21 (24.4)	—	—	73.8 (8.0)
MEC_1	99	55 (55.6)	44 (44.4)	27	—	70.3 (7.9)	115	42 (36.5)	73 (63.5)	37	—	70.3 (7.6)
MEC_2	15	2 (13.3)	13 (86.7)	5	—	80.1 (6.2)	30	4 (13.3)	26 (86.7)	12	—	74.6 (6.1)
MECC_3	309	260 (84.1)	49 (15.9)	—	—	69.5 (10.3)	367	290 (79.0)	77 (21.0)	—	—	73.0 (10.0)
NCCCSII	219	128 (58.4)	91 (41.6)	—	—	63.8 (9.8)	221	89 (40.3)	132 (59.7)	—	—	65.4 (9.4)
NFCCR_2	60	51 (85.0)	9 (15.0)	—	—	61.1 (7.9)	130	104 (80.0)	26 (20.0)	—	—	60.2 (7.2)
NHS_1_2	328	174 (53.0)	154 (47.0)	23	7	68.0 (7.4)	673	321 (47.7)	352 (52.3)	42	7	68.5 (6.9)
NHS_3_AD	410	187 (45.6)	223 (54.4)	21	10	68.1 (6.7)	335	133 (39.7)	202 (60.3)	15	7	67.9 (6.7)
PLCO_1_Rematch	216	125 (57.9)	91 (42.1)	—	—	68.8 (6.0)	123	61 (49.6)	62 (50.4)	—	—	67.5 (6.2)
PLCO_2	196	110 (56.1)	86 (43.9)	—	—	70.6 (6.6)	163	90 (55.2)	73 (44.8)	—	—	70.6 (6.3)
PLCO_3	295	157 (53.2)	138 (46.8)	—	—	67.0 (7.2)	1964	900 (45.8)	1064 (54.2)	—	—	62.1 (5.3)
PLCO_4_AD	434	241 (55.5)	193 (44.5)	—	—	64.0 (5.9)	587	274 (46.7)	313 (53.3)	—	—	61.9 (5.3)
REACH_AD	9	7 (77.8)	2 (22.2)	—	—	62.9 (4.0)	75	47 (62.7)	28 (37.3)	—	—	62.3 (5.6)
SMC_COSM	179	90 (50.3)	89 (49.7)	—	—	69.7 (9.7)	330	145 (43.9)	185 (56.1)	—	—	64.6 (7.8)
UKB_1	1073	996 (92.8)	77 (7.2)	—	—	65.4 (5.3)	4254	3928 (92.3)	326 (7.7)	—	—	65.4 (5.3)

(continued)

Table 1. (continued)

Study	CRC patients						Participants without CRC					
	Total No.	No use of MHT No. (%)	Any MHT No. (%)	E-only No.	E + P No.	Age at diagnosis Mean (SD), y	Total No.	No use of MHT No. (%)	Any MHT No. (%)	E-only No.	E + P No.	Age at enrollment Mean (SD), y
USC_HRT_CRC	296	127 (42.9)	169 (57.1)	75	67	66.3 (5.5)	400	150 (37.5)	250 (62.5)	116	82	65.0 (6.8)
VITAL	114	61 (53.5)	53 (46.5)	—	—	70.5 (6.4)	126	60 (47.6)	66 (52.4)	—	—	71.5 (6.6)
WHI_1	450	297 (66.0)	153 (34.0)	95	58	71.0 (7.1)	519	282 (54.3)	237 (45.7)	137	100	71.2 (7.0)
WHI_2	977	576 (59.0)	401 (41.0)	202	199	72.2 (7.4)	990	512 (51.7)	478 (48.3)	260	217	72.0 (7.2)
WHI_3	556	313 (56.3)	243 (43.7)	117	126	78.6 (6.9)	558	267 (47.8)	291 (52.2)	148	142	78.5 (6.9)
Total	11519	7664 (66.5)	3855 (33.5)	931	698	—	16967	10795 (63.6)	6172 (36.4)	1225	811	—

ᵃ CCFR = Colon Cancer Family Registry; CLUEII = Campaign against Cancer and Heart Disease II; Colo 2&3 = Hawaii Colorectal Cancer Studies 2 & 3; CPSII = Cancer Prevention Study-II; CRCGEN = Colorectal Cancer Genetics & Genomics; DACHS = Darmkrebs: Chancen der Verhütung durch Screening; DALS = Diet, Activity, and Lifestyle Study; E + P = combined estrogen-progestogen; E-only = estrogen only; EPIC = European Prospective Investigation into Cancer; ESTHER_VERDI = Epidemiologische Studie zu Chancen der Verhütung = Früherkennung und optimierten THerapie chronischer ERkrankungen in der älteren Bevölkerung; Kentucky = Kentucky Case-Control Study; LCCS = Leeds Colorectal Cancer Study; MCCS = Melbourne Collaborative Cohort Study; MEC = Multiethnic Cohort Study; MECC = Molecular Epidemiology of Colorectal Cancer Study; MHT = menopausal hormone therapy; NCCCSII = The North Carolina Colon Cancer Study II; NHS = Nurses' Health Study; NFCCR = Newfoundland Case-Control Study; PLCO = Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial; REACH = Colon Cancer Pathways: Hyperplastic Polyps and Adenomas; SMC_COSM = Swedish Mammography Cohort and Swedish Men Cohort; UKB = UK Biobank; USC_HRT_CRC = University of Southern California Hormone Replacement Therapy Colorectal Cancer Study; VITAL = Cancer Screening Trial VITamins And Lifestyle cohort; WHI = Women's Health Initiative. The sign "—" = Not available

ARTICLE
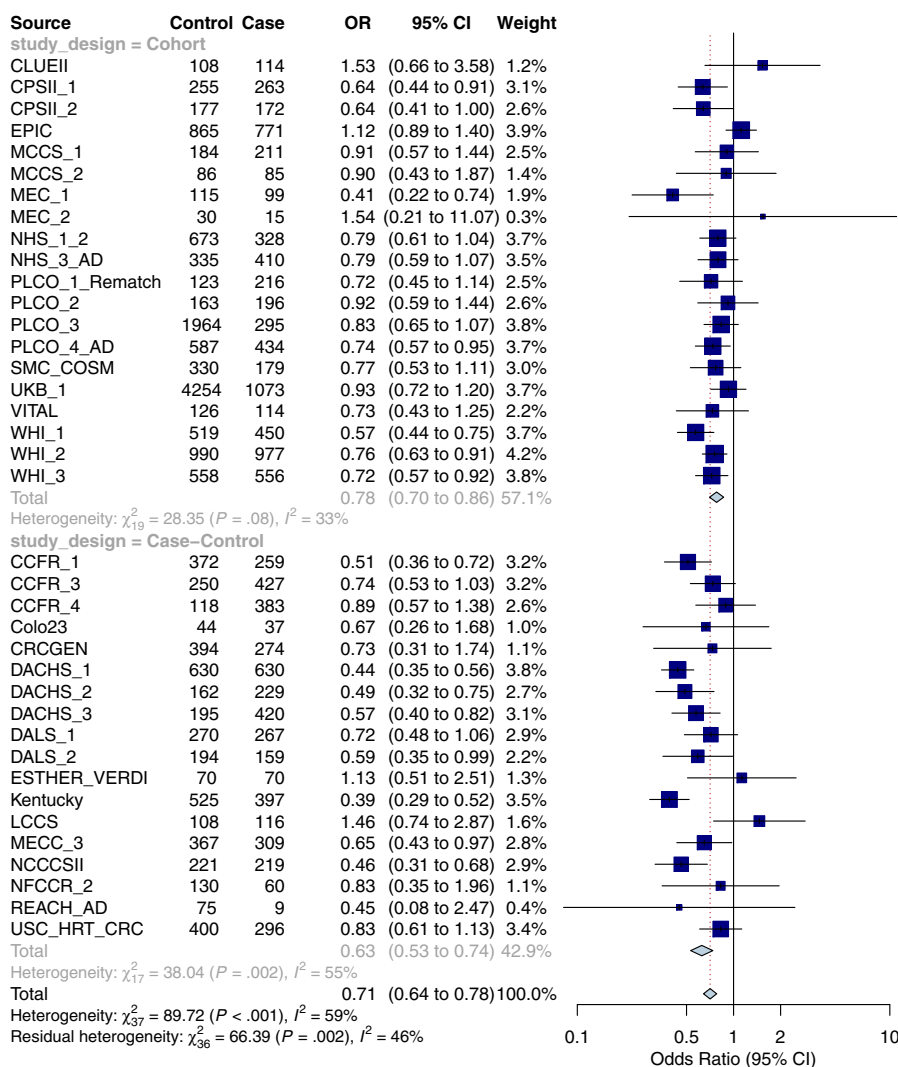
Figure 1. Association of any menopausal hormone therapy use with the risk of colorectal cancer. CI = confidence interval; OR = odds ratio.

available online). Several *DCBLD1* intronic variants at 6q22.1 (rs4945586, rs9320604, rs4946260), which were in LD with rs10782186, also yielded low P values using the 2-*df* joint test although not genome-wide significant (5.28×10^{-8}, 5.60×10^{-8}, and 5.70×10^{-8}; Supplementary Figures 3 and 4, available online). We did not identify any genome-wide statistically significant interactions between estrogen-only use or estrogen-progestogen use and common genetic variants for CRC risk. Common variants that reached the suggestive interaction level ($P < 5 \times 10^{-6}$) with MHT use for CRC risk are shown in Supplementary Tables 2-4 (available online), which included 87 SNPs with any MHT use, 80 with estrogen-only use, and 137 with estrogen-progestogen use. We also performed GWIS stratified by colon and rectal cancer, but the common variant analysis did not yield any statistically significant interactions for the MHT variables, respectively (data not shown).

Table 3 presents associations of MHT use with CRC risk by the genotype of the 2 SNPs that were found to be statistically significant. For rs117868593, there was a statistically significant protective effect of any MHT use only among women with the *GG* homozygotes (OR = 0.68, 95% CI = 0.64 to 0.72; $P = 4.3 \times 10^{-37}$) but not in women with the *GC* genotype (OR = 0.91, 95% CI = 0.77 to 1.09; P = .31) or with the *CC* genotype (OR = 0.64, 95% CI = 0.22 to

1.85; P = .41). When stratified by MTH use, there was a statistically significant per-minor allele association with CRC risk in users of any MHT (OR = 1.20, 95% CI = 1.05 to 1.37) but not in nonusers (OR = 0.93, 95% CI = 0.83 to 1.03). For rs10782186, the protective effect of any MHT use compared with women not using any MHT was increasingly stronger for women with an increasing number of C alleles: TT (OR = 0.78, 95% CI = 0.70 to 0.87; $P = 4.3 \times 10^{-6}$), TC (OR = 0.68, 95% CI = 0.63 to 0.73; $P = 1.4 \times 10^{-22}$), and CC (OR = 0.66, 95% CI = 0.60 to 0.74; $P = 5.7 \times 10^{-14}$). When rs10782186 was investigated in relation to CRC risk among strata of MTH use, the per-minor allele odds ratio for CRC risk was attenuated in users of any MHT (OR = 1.05, 95% CI = 0.99 to 1.11) compared with nonusers (OR = 1.14, 95% CI = 1.09 to 1.19).

The GxE interactions between rs117868593 or rs10782186 and any MHT were not heterogeneous across studies overall (P = .98, P = .56, respectively) or stratified by study regions (North America, Australia, and Europe). The corresponding forest plots are shown in Supplementary Figures 5 and 6 (available online).

Rare Variants for CRC Risk

The rare variant analysis did not yield any statistically significant interactions ($P < 2.5 \times 10^{-6}$) for the MHT variables.
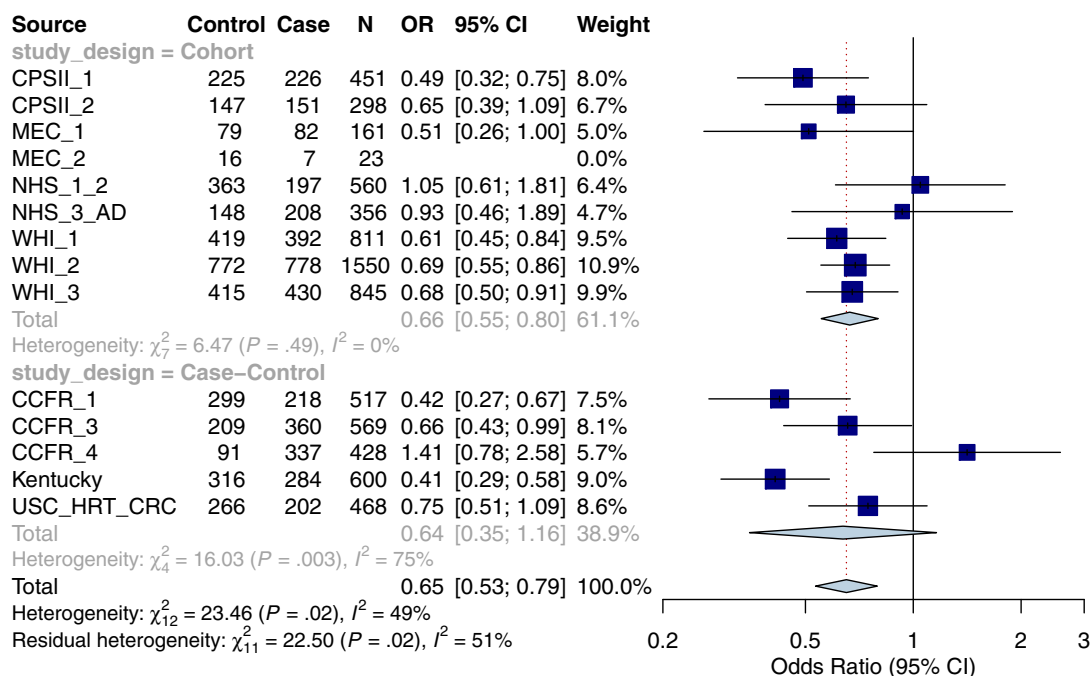
Figure 2. Association of use of estrogen only with the risk of colorectal cancer. CI = confidence interval; OR = odds ratio.
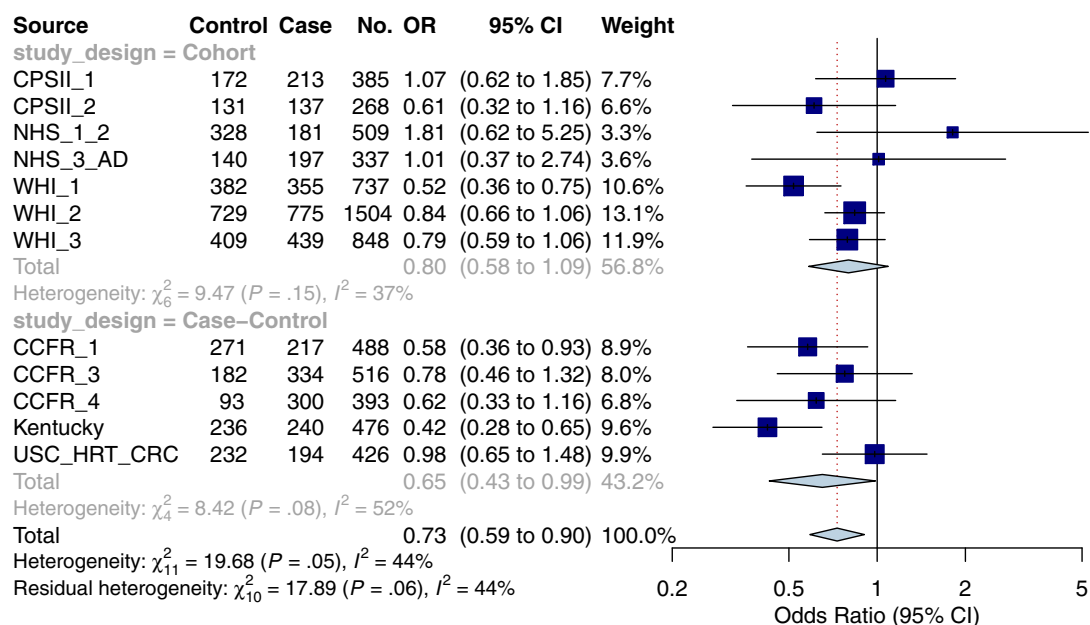
Figure 3. Association of use of combined estrogen-progestogen with the risk of colorectal cancer. CI = confidence interval; OR = odds ratio.

However, several genes were found to reach the suggestive level for interaction ($P < 1.2 \times 10^{-4}$) for CRC risk: *PREX1* (Phosphatidylinositol-3,4,5-Trisphosphate Dependent Rac Exchange Factor 1) with any MHT use ($P = 5.02 \times 10^{-5}$), *SOS2* (SOS Ras/Rho Guanine Nucleotide Exchange Factor 2) with estrogen-only therapy ($P = 9.23 \times 10^{-5}$), as well as *TMEM189-UBE2V1* (Transmembrane protein 189 - Ubiquitin Conjugating Enzyme E2 V1) ($P = 2.46 \times 10^{-5}$), *FAM149A* (Family With Sequence Similarity 149 Member A) ($P = 9.67 \times 10^{-5}$), and *RPS13* (Ribosomal Protein S13) ($P = 1.02 \times 10^{-5}$) with estrogen-

progestogen therapy (Table 4; quantile-quantile plots shown in Supplementary Figures 7-9, available online).

Functional Annotations of Genetic Loci

We performed bioinformatic analysis of the 2 loci showing statistically significant interactions with MHT use (rs117868593 located 20 kb downstream of *GRIN2B* variant at 12p13.1 and a *DCBLD1* intronic variant rs10782186 at 6q22.1). Annotation was performed for all variants tagged by the most statistically

ARTICLE

Table 2. Results of genome-wide interaction analyses with menopausal hormone therapy for colorectal cancer risk among postmenopausal women[a]

MHT type	SNP	Chr	BP position	Locus	Gene	Count allele	Count allele frequency	Statistical method used to detect the GxMHT interaction	$P_{threshold}$ for GxMHT interaction	P observed for GxMHT interaction	$P_{heterogeneity}$	No. of studies included
Any MHT	rs117868593	12	13670508	12p13.1	GRIN2B	C	0.05	2-step method (by G|E in step 1)	5×10^{-3}	.003	.98	38
Any MHT	rs10782186	6	117823508	6q22.1	DCBLD1	C	0.50	2-df joint test	5×10^{-8}	4.23×10^{-8}	.56	38

[a]Directly genotyped SNPs were coded as 0, 1, or 2 copies of the count allele. Imputed SNPs were coded as expected gene dosage. Multiplicative interaction terms were modeled as the product of MHT and each SNP of interest. All statistical tests were 2-sided. 2-df = 2-degree-of-freedom; Chr = chromosome; BP position = base pair position based on NCBI Build37; MHT = menopausal hormone therapy; SNP = single nucleotide polymorphism; GxMHT = interaction between SNPs and MHT; G|E = associations between SNPs and environmental risk factors in the combined case-control population.

Table 3. Associations with colorectal cancer risk stratified by use of any menopausal hormone therapy and genotypes of SNPs of interest[a]

		Genotype of SNP									Per minor allele within strata of MHT use	
		Homozygous noncarriers			Heterozygous			Homozygous carries of the minor allele				
SNP	MHT use	N Ca/Co	OR (95% CI)	P	N Ca/Co	OR (95% CI)	P	N Ca/Co	OR (95% CI)	P	OR (95% CI)	P
rs117868593		GG			GC			CC			Per C allele within strata of MHT use	
	No	6991.4/9745.2	1.00 (Referent)	—	652.6/1026.1	0.91 (0.81 to 1.03)	.13	20/23.7	1.09 (0.53 to 2.22)	.81	0.93 (0.83 to 1.03)	0.17
	Yes	3390.5/5537.1	0.68 (0.64 to 0.72)	4.3×10^{-37}	450.7/614.9	0.83 (0.72 to 0.96)	.011	13.8/20.1	0.69 (0.31 to 1.54)	.37	1.20 (1.05 to 1.37)	.008
	OR (95% CI)	—	0.68 (0.64 to 0.72)	4.3×10^{-37}	—	0.91 (0.77 to 1.09)	.31	—	0.64 (0.22 to 1.85)	.41		
rs10782186		TT			TC			CC			Per C allele within strata of MHT use	
	No	1861.2/2936.4	1.00 (Referent)	—	3806.9/5361.8	1.15 (1.07 to 1.24)	3.5×10^{-4}	1995.9/2496.9	1.29 (1.18 to 1.41)	1.9×10^{-8}	1.14 (1.09 to 1.19)	1.8×10^{-8}
	Yes	993/1624.8	0.78 (0.70 to 0.87)	4.3×10^{-6}	1861.9/3068.3	0.78 (0.71 to 0.85)	3.8×10^{-8}	1000/1478.9	0.85 (0.77 to 0.95)	.004	1.05 (0.99 to 1.11)	.14
	OR (95% CI)	—	0.78 (0.70 to 0.87)	4.3×10^{-6}	—	0.68 (0.63 to 0.73)	1.4×10^{-22}	—	0.66 (0.60 to 0.74)	5.7×10^{-14}		

[a]Case/control counts were calculated by imputed genotype probabilities. Ca/Co = case/control; CI = confidence interval; MHT = menopausal hormone therapy; OR = odds ratio; P = probability value; SNP = single nucleotide polymorphism.

Table 4. Suggestive association ($P < 1.2 \times 10^{-4}$) of genes from rare variants analyses of G×E with menopausal hormone therapy for colorectal cancer risk among postmenopausal women[a]

MHT type	Gene	Gene name	Chr	No. of SNPs	P
Any MHT	ENSG00000124126	PREX1	20	45	5.02×10^{-5}
E-only	ENSG00000100485	SOS2	14	15	9.23×10^{-5}
E + P	ENSG00000124208	TMEM189-UBE2V1	20	57	2.46×10^{-5}
E + P	ENSG00000109794	FAM149A	4	8	9.67×10^{-5}
E + P	ENSG00000110700	RPS13	11	5	1.02×10^{-4}

[a]Chr = chromosome; E + P = combined estrogen-progestogen; E-only = estrogen only; MHT = menopausal hormone therapy; P = Fisher P value by the set-based score (MiSTi) test; SNP = single nucleotide polymorphism.

significant SNPs ($r^2 > 0.5$) using our novel functional annotation analyses. The *GRIN2B* rs117868593 locus is in LD with rs17822202 ($D' = 0.93$ and $r^2 = 0.85$ in 1000 Genomes Project CEU), which is downstream of the GRIN2B gene. We noted that this SNP was associated with more pronounced enhancer activity in colon tumor and cancer cell lines than in normal colon tissues (Supplementary Figure 10, available online). The *DCBLD1* rs10782186 is in high LD with rs9320604 ($D' = 0.99$ and $r^2 = 0.98$ in 1000 Genomes Project CEU); a SNP overlapping histone methylation patterns with enhancer activity in normal colon tissues, colon tumor, and cancer cell lines, and associated with strong DNase hypersensitivity in tumor tissues (Supplementary Figure 11, available online).

Based on BarcUVa-Seq expression quantitative trait loci (eQTL) analysis (Supplementary Methods, available online), we identified 4 genes—*EMP1*(Epithelial membrane protein 1), *RPL13A20* (Ribosomal Protein L13a Pseudogene 20), *FAM234B* (Family With Sequence Similarity 234 Member B), and *CDKN1B* (Cyclin Dependent Kinase Inhibitor 1B)—whose expression in normal colon tissue was statistically significantly associated with the SNP rs117868593 or the SNPs in LD ($R^2 > 0.5$) ($P < .05$) (Supplementary Table 5 and Figure 12, available online), as well as 2 genes—*ROS1* (ROS Proto-Oncogene 1, Receptor Tyrosine Kinase) and *GOPC* (Golgi Associated PDZ And Coiled-Coil Motif Containing)—with the SNP rs10782186 or the SNPs in LD ($P < .05$) (Supplementary Table 6 and Figure 13, available online). These eQTL effects persisted when restricting the sample to postmenopausal women although statistically significant for rs10782186_ROS1, rs117868593_RPL13AP20, and rs1806217_FAM234B.

Discussion

We identified novel GxE interactions between the use of any MHT and common variants at 2 loci for CRC risk among postmenopausal women. The putative target genes underlying these interactions include *EMP1*, *RPL13AP20*, *FAM234B*, *CDKN1B*, *ROS1*, and *GOPC*. In addition, we found suggestive interactions between the use of MHT and rare variants in *PREX1*, *SOS2*, *TMEM189-UBE2V1*, *FAM149A*, and *RPS13*. Using independent samples in the current study, the previously found SNPs for GxE interactions (Supplementary Table 7, available online) (7,10) did not show statistically significant interaction with MHT with respect to CRC risk. These earlier studies used a candidate gene approach, different covariable adjustment, or different exposure and nonexposure definitions compared with our GWAS study. Additionally, power could be further reduced by variations in the underlying distribution of MHT as new studies were introduced to the larger cohort.

Currently, the underlying etiologic mechanisms by which MHT affects CRC are not yet well understood. It is likely that protective cellular effects of estrogen and progesterone in the development of CRC are mediated through estrogen receptor α, estrogen receptor β (ESR2), and progesterone receptor (28-30). Estrogen and progestin may play a role in the pathway leading to DNA hypermethylation (31,32), which regulates gene expression including that of tumor suppressor genes and thereby play a crucial role in tumorigenesis of CRC. Estrogen has also been found to have an impact on a large number of serum proteome, which plays a role in mucosal protection and repair in the gastrointestinal tract (33) as well as colon transcriptome (34). In addition, estrogen may contribute to maintaining the genomic stability in colonic epithelial cells by upregulation of mismatch repair genes (35). MHT use has also been reported to have growth-inhibiting effects on colon cancer cells through upregulating cell cycle regulators (eg, TP53) (36). Consortium efforts that are powered to explore the relationships of MHT with specific subtypes of CRC may yield further insights to GxE interactions with respect to hormonal contributions to the pathogenesis of CRC (37).

The SNP rs117868593 located about 20 kb downstream from *GRIN2B* was not found to be associated with expression of the nearest gene *GRIN2B* but with *EMP1*, *RPL13AP20*, *FAM234B*, and *CDKN1B*. Expression of *EMP1* has been found to be lower in human CRC than normal adjacent colorectal tissues (38), and overexpression of *EMP1* was observed to reduce proliferation and induce apoptosis of CRC cells (39), which are consistent with our findings, that is, lower expression of *EMP1* and higher risk of CRC associated with G allele of rs117868593. We found the MHT users with GG have a stronger statistically significant reduction of CRC risk, suggesting that EMP1 may function as an oncogene in hormone-dependent epithelium, which has been observed for EMP2, a paralog of EMP1 (40). Downregulation of CDKN1B, which mainly results from increased ubiquitin-mediated proteasomal degradation, has been associated with tumor progression in CRC (41), and CDKN1B could be induced through ESR2-mediated repression of the F-box protein p45 (SKP2), which has been identified as the substrate recognition component that targets and binds CDKN1B for ubiquitination and subsequent degradation (41-43). The link between CDKN1B and ESR2 might explain the observed interaction of *CDKN1B* with MHT. Potential mechanisms through which *RPL13AP20* and *FAM234B* act in modifying MHT-associated CRC risk are unknown.

The region in which *DCBLD1* is located, chromosome 6q22.1, has been reported as one of the suggestive susceptibility regions ($P = 3.20 \times 10^{-6}$) in a GWAS meta-analysis on CRC risk (12). Association estimates for the index SNP rs10782186 and correlated SNPs (rs4945586, rs9320604, and rs4946260) reported in the above-mentioned GWAS paper. The significance ($P = 4.23 \times 10^{-8}$) of the interaction in our GWIS using the 2-*df* joint test was mainly driven by the genetic association

ARTICLE

$(P = 6.79 \times 10^{-8})$ and was further strengthened by the GxE product term $(P = .03)$. Thus, incorporating the GxE component helped uncover genetic susceptibility variants for CRC risk, which did not reach genome-wide significance level in GWAS. Analyses of associated gene expression indicated the involvement of *ROS1* and *GOPC*. ROS1 is a transmembrane receptor tyrosine kinase that often shows genetic rearrangements in colorectal tumor tissue, such as intrachromosomal fusion with GOPC because of microdeletions at 6q22.1, which is highly prevalent in CRC (44,45). GOPC-ROS fusion proteins have been shown to activate the downstream signaling pathway, signal transducers, and activators of transcription-3 that play a important role in progression of CRC (45,46). The transcription factor of signal transducers and activators of transcription-3 in epithelial cells is activated by interleukin–6, promoting CRC tumorigenesis (29,47), whereas ESR2 mediates the downregulation of the inflammatory cytokine interleukin–6 network (48), which may explain the observed interaction with MHT.

There are still considerable challenges in investigating GxE interaction of rare genetic variants because of the scarcity of subjects with data on both these variants and the relevant environmental and lifestyle exposures. Therefore, the role that rare predisposition alleles play in modifying the association between environmental factors and CRC risk remains poorly understood. Our study used MiSTi to tackle the challenge for GxE interaction analysis of rare variants, which strengthened statistical power to robustly uncover potential rare variant GxE association signals. Through this method, we found suggestive interaction for MHT use with rare variants in 5 genes for CRC risk. Despite their as yet unknown mechanisms in modifying CRC risk associated with MHT use, our application of GxE interaction analysis for CRC risk to rare variants alongside common variants represents a novel and rigorous approach. GxE interaction studies of rare genetic variants that incorporate functional genomic information ideally accounting for MHT effects and studies with larger sample sizes and hence with greater statistical power may contribute to understanding any missing heritability of cancer that remains unexplained by common variants.

Our study has several strengths. First, our large sample size, including more than 28 000 participants, facilitated the most powerful scan for gene-MHT interaction to date. Second, we used recently developed statistical approaches that can provide greater statistical power than conventional case-control logistic regression (49). Because no single approach provides the best power across all possible patterns of GxE interaction, we used a combination of approaches to maximize the chance of identifying novel loci in this discovery analysis. MiSTi, used for rare variant analysis, helped identify suggestive associations with CRC risk through interaction with MHT for 5 genes that warrant further follow-up. Third, we carefully harmonized environmental data on MHT use and other covariates across studies to minimize between-study heterogeneity bias as previously described (11). We acknowledge, however, that our analysis was limited to populations of European ancestry; thus the results might not be generalizable to other race and ethnicity groups. Measurement error of the primarily self-reported exposure assessment might also have contributed to reduced power; however, previous studies have found the high validity for self-reported MHT use when compared with population-based prescription databases as references (50) and a high concordance between self-reported MHT use and that of physicians' reports (51). Despite our sizable sample size and use of advanced statistical methods, we acknowledge that statistical power remains limited to detect small to modest-sized interaction effects in a genome-wide scan setting.

This might explain the relatively small number of novel findings. To overcome these issues, it will be critical to expand sample sizes of well-characterized studies as well as incorporated functional genomic data relevant to CRC and MHT use, such as multi-omics data of normal and tumor colon tissue exposed and unexposed to MHT.

From a comprehensive genome-wide GxE interaction investigation, we identified 2 common loci, which were statistically significantly associated with CRC risk in conjunction with MHT use, as well as 5 genes, which showed suggestive evidence of GxE interaction through rare variant set analysis. The putative target genes of the 2 identified loci (*EMP1*, *RPL13AP20*, *FAM234B*, *CDKN1B*, *ROS1*, and *GOPC*) may explain the GxE interactions with MHT and offer new insights into CRC etiological mechanisms and pathways of CRC carcinogenesis. Further downstream, follow-up studies for exploring potential genetic functions are warranted to confirm the involvement of these genetic variants or genes in CRC risk associated with MHT use.

Funding

Notes

Role of the funders: The funders had no role in the design of the study; the collection, analysis, and interpretation of the data; the writing of the manuscript; and the decision to submit the manuscript for publication.

Disclosures: All authors have no conflicts of interest to disclose. KV, who is a *JNCI* Associate Editor and co-author on this paper, was not involved in the editorial review or decision to publish the manuscript.

Author contributions: Writing—original draft: YT, AEK. Analysis—statistical methods: YT, AEK, YL, CQ, VDO. Supervising: UP, WJG, LH, JCC. Writing- reviewing & editing: YT, AEK, SAB, YL, CQ, TH, RCT, VDO, ND, DAD, AH, JRH, KMJ, JM, NM, MOS, CMU, JO, ARP, EAR, AS, MS, YS, FJD, VA, JB, SIB, DTB, HB, DDB, ATC, JCF, SG, SBG, SH, MH, MAJ, ADJ, TOK, SCL, LLM, LL, GGG, RLM, HN, RN, SO, AB, EAP, JDP, RLP, GR, LCS, RES, MLS, SNT, BVG, KV, EW, AW, MOW, AHW, PTC, GC, DVC, MJG, AK, JPL, VM, PAN, BP, DCT, KKT, UP, WJG, LH, JCC.

Data Availability

The data underlying this article are available in dbGaP at https://www.ncbi.nlm.nih.gov/gap/, and can be accessed with accession numbers phs001415.v1.p1, phs001315.v1.p1, phs001078.v1.p1, phs001499.v1.p1, phs001903.v1.p1, phs001856.v1.p1, phs001045.v1.p1, and phs001499.v1.p1.

References

1. Lin KJ, Cheung WY, Lai JY-C, Giovannucci EL. The effect of estrogen vs. combined estrogen-progestogen therapy on the risk of colorectal cancer. *Int J Cancer*. 2012;130(2):419-430.
2. Botteri E, Støer NC, Sakshaug S, et al. Menopausal hormone therapy and colorectal cancer: a linkage between nationwide registries in Norway. *BMJ Open*. 2017;7(11):e017639.
3. Mørch LS, Lidegaard Ø, Keiding N, Løkkegaard E, Kjær SK. The influence of hormone therapies on colon and rectal cancer. *Eur J Epidemiol*. 2016;31(5):481-489.
4. Rennert G, Rennert HS, Pinchev M, Lavie O, Gruber SB. Use of hormone replacement therapy and the risk of colorectal cancer. *J Clin Oncol*. 2009;27(27):4542-4547.
5. Lin J, Zee RYL, Liu K-Y, et al. Genetic variation in sex-steroid receptors and synthesizing enzymes and colorectal cancer risk in women. *Cancer Causes Control*. 2010;21(6):897-908.
6. Lin JH, Manson JE, Kraft P, et al. Estrogen and progesterone-related gene variants and colorectal cancer risk in women. *BMC Med Genet*. 2011;12(1):78.
7. Rudolph A, Sainz J, Hein R, et al. Modification of menopausal hormone therapy-associated colorectal cancer risk by polymorphisms in sex steroid signaling, metabolism and transport related genes. *Endocr Relat Cancer*. 2011;18(3):371-384.
8. Slattery ML, Lundgreen A, Herrick JS, et al. Variation in the CYP19A1 gene and risk of colon and rectal cancer. *Cancer Causes Control*. 2011;22(7):955-963.
9. Kantor ED, Hutter CM, Minnier J, et al. Gene-environment interaction involving recently identified colorectal cancer susceptibility loci. *Cancer Epidemiol Biomarkers Prev*. 2014;23(9):1824-1833.
10. Garcia-Albeniz X, Rudolph A, Hutter C, et al. CYP24A1 variant modifies the association between use of oestrogen plus progestogen therapy and colorectal cancer risk. *Br J Cancer*. 2016;114(2):221-229.
11. Hutter CM, Chang-Claude J, Slattery ML, et al. Characterization of gene-environment interactions for colorectal cancer susceptibility loci. *Cancer Res*. 2012;72(8):2036-2044.
12. Peters U, Jiao S, Schumacher FR, et al. Identification of genetic susceptibility loci for colorectal tumors in a genome-wide meta-analysis. *Gastroenterology*. 2013;144(4):799-807.e24.
13. Schmit SL, Edlund CK, Schumacher FR, et al. Novel common genetic susceptibility loci for colorectal cancer. *J Natl Cancer Inst*. 2019;111(2):146-157.
14. Gong J, Hutter CM, Newcomb PA, et al.; for the CCFR and GECCO. Genome-wide interaction analyses between genetic variants and alcohol consumption and smoking for risk of colorectal cancer. *PLoS Genet*. 2016;12(10):e1006296.
15. Jeon J, Du M, Schoen RE, et al.; for the Colorectal Transdisciplinary Study and Genetics and Epidemiology of Colorectal Cancer Consortium. Determining risk of colorectal cancer and starting age of screening based on lifestyle, environmental, and genetic factors. *Gastroenterology*. 2018;154(8):2152-2164.e19.
16. Huyghe JR, Bien SA, Harrison TA, et al. Discovery of common and rare genetic risk variants for colorectal cancer. *Nat Genet*. 2019;51(1):76-87.
17. Das S, Forer L, Schönherr S, et al. Next-generation genotype imputation service and methods. *Nat Genet*. 2016;48(10):1284-1287.
18. Morrison J. BinaryDosage: creates, merges, and reads binary dosage files. R package version 1.0.0; 2020. https://CRAN.R-project.org/package=BinaryDosage Accessed January 30, 2020.
19. Cochran WG. The combination of estimates from different experiments. *Biometrics*. 1954;10(1):101-129.
20. Morrison J. GxEScanR: run GWAS/GWEIS scans using binary dosage files. R package version 2.0.2; 2020. https://CRAN.R-project.org/package=GxEScanR Accessed November 6, 2020.
21. Kooperberg C, Leblanc M. Increasing the power of identifying gene x gene interactions in genome-wide association studies. *Genet Epidemiol*. 2008;32(3):255-263.
22. Murcray CE, Lewinger JP, Gauderman WJ. Gene-environment interaction in genome-wide association studies. *Am J Epidemiol*. 2009;169(2):219-226.
23. Gauderman WJ, Zhang P, Morrison JL, Lewinger JP. Finding novel genes by testing G × E interactions in a genome-wide association study. *Genet Epidemiol*. 2013;37(6):603-613.
24. Kraft P, Yen YC, Stram DO, Morrison J, Gauderman WJ. Exploiting gene-environment interaction to detect genetic associations. *Hum Hered*. 2007;63(2):111-119.
25. Gauderman WJ, Kim A, Conti DV, et al. A unified model for the analysis of gene-environment interaction. *Am J Epidemiol*. 2019;188(4):760-767.
26. Su YR, Di CZ, Hsu L; for the Genetics and Epidemiology of Colorectal Cancer Consortium. A unified powerful set-based test for sequencing data analysis of GxE interactions. *Biostatistics*. 2017;18(1):119-131.
27. Pruim RJ, Welch RP, Sanna S, Teslovich TM, et al. LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics*. 2010;26(18):2336-2337.
28. Barzi A, Lenz AM, Labonte MJ, Lenz H-J. Molecular pathways: estrogen pathway in colorectal cancer. *Clin Cancer Res*. 2013;19(21):5842-5848.
29. Caiazza F, Ryan EJ, Doherty G, Winter DC, Sheahan K. Estrogen receptors and their implications in colorectal carcinogenesis. *Front Oncol*. 2015;5:19.

ARTICLE

30. Williams C, DiLeo A, Niv Y, Gustafsson J-Å. Estrogen receptor beta as target for colorectal cancer prevention. *Cancer Lett.* 2016;372(1):48-56.

31. Newcomb PA, Zheng Y, Chia VM, et al. Estrogen plus progestin use, microsatellite instability, and the risk of colorectal cancer in women. *Cancer Res.* 2007; 67(15):7534-7539.

32. Neumeyer S, Popanda O, Butterbach K, et al. DNA methylation profiling to explore colorectal tumor differences according to menopausal hormone therapy use in women. *Epigenomics.* 2019;11(16):1765-1778.

33. Katayama H, Paczesny S, Prentice R, et al. Application of serum proteomics to the Women's Health Initiative conjugated equine estrogens trial reveals a multitude of effects relevant to clinical findings. *Genome Med.* 2009;1(4):47.

34. Hases L, Archer A, Indukuri R, et al. High-fat diet and estrogen impacts the colon and its transcriptome in a sex-dependent manner. *Sci Rep.* 2020;10(1): 16160.

35. Slattery ML, Potter JD, Curtin K, et al. Estrogens reduce and withdrawal of estrogens increase risk of microsatellite instability-positive colon cancer. *Cancer Res.* 2001;61(1):126-130.

36. Hsu HH, Cheng SF, Wu CC, et al. Apoptotic effects of over-expressed estrogen receptor-beta on LoVo colon cancer cell is mediated by p53 signalings in a ligand-dependent manner. *Chin J Physiol.* 2006;49(2):110-116.

37. Ogino S, Chan AT, Fuchs CS, Giovannucci E. Molecular pathological epidemiology of colorectal neoplasia: an emerging transdisciplinary and interdisciplinary field. *Gut.* 2011;60(3):397-411.

38. Sun GG, Wang YD, Cui DW, Cheng YJ, Hu WN. Epithelial membrane protein 1 negatively regulates cell growth and metastasis in colorectal carcinoma. *World J Gastroenterol.* 2014;20(14):4001-4010.

39. Ahmat AM, Shimizu A, Ogita H. The pivotal roles of the epithelial membrane protein family in cancer invasiveness and metastasis. *Cancers (Basel).* 2019; 11(11):1620.

40. Wang YW, Cheng HL, Ding YR, Chou LH, Chow NH. EMP1, EMP 2, and EMP3 as novel therapeutic targets in human cancer. *Biochim Biophys Acta Rev Cancer.* 2017;1868(1):199-211.

41. Ogino S, Kawasaki T, Kirkner GJ, Yamaji T, Loda M, Fuchs CS. Loss of nuclear p27 (CDKN1B/KIP1) in colorectal cancer is correlated with microsatellite instability and CIMP. *Mod Pathol.* 2007;20(1):15-22.

42. Kondakova IV, Shashova EE, Sidenko EA, Astakhova TM, Zakharova LA, Sharova NP. Estrogen receptors and ubiquitin proteasome system: mutual regulation. *Biomolecules.* 2020;10(4):500.

43. Hartman J, Edvardsson K, Lindberg K, et al. Tumor repressive functions of estrogen receptor *β* in SW480 colon cancer cells. *Cancer Res.* 2009;69(15): 6100-6106.

44. AACR project GENIE Consortium. AACR project GENIE: powering precision medicine through an international consortium. *Cancer Discov.* 2017;7(8):818-831.

45. Davies KD, Doebele RC. Molecular pathways: ROS1 fusion proteins in cancer. *Clin Cancer Res.* 2013;19(15):4040-4045.

46. Huang Y, Wang J, Cao F, et al. SHP2 associates with nuclear localization of STAT3: significance in progression and prognosis of colorectal cancer. *Sci Rep.* 2017;7(1):17597.

47. Grivennikov S, Karin E, Terzic J, et al. IL-6 and Stat3 are required for survival of intestinal epithelial cells and development of colitis-associated cancer. *Cancer Cell.* 2009;15(2):103-113.

48. Edvardsson K, Ström A, Jonsson P, Gustafsson J, Williams C. Estrogen receptor *β* induces antiinflammatory and antitumorigenic networks in colon cancer cells. *Mol Endocrinol.* 2011;25(6):969-979.

49. Gauderman WJ, Mukherjee B, Aschard H, et al. Update on the state of the science for analytical methods for gene-environment interactions. *Am J Epidemiol.* 2017;186(7):762-770.

50. Løkkegaard EL, Johnsen SP, Heitmann BL, et al. The validity of self-reported use of hormone replacement therapy among Danish nurses. *Acta Obstet Gynecol Scand.* 2004;83(5):476-481.

51. Kropp S, Terboven T, Hedicke J, et al. Good agreement between physician and self-reported hormone therapy data in a case-control study. *J Clin Epidemiol.* 2007;60(12):1280-1287.

ARTICLE