Supplementary Online Content

Nan H, Hutter CM, Lin Y, et al; CCFR and GECCO. Association of aspirin and NSAID use with risk of colorectal cancer according to genetic variants. *JAMA.* doi:10.1001/jama.2015.1815

This supplementary material has been provided by the authors to give readers additional information about their work.

eAppendix. Description of Study Populations

This study is based on the Colon Cancer Family Registry (CCFR) and nine cohorts from the Genetics and Epidemiology of Colorectal Cancer Consortium (GECCO), which include nested case-control studies within five prospective US cohorts: Health Professionals Follow-up Study (HPFS); Nurses' Health Study (NHS); Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial (PLCO); VITamins And Lifestyle (VITAL); and Women's Health Initiative (WHI); and four case-control studies: Darmkrebs: Chancen der Verhütung durch Screening (DACHS) study; Diet, Activity and Lifestyle Study (DALS); Ontario Familial Colorectal Cancer Registry (OFCCR); and Postmenopausal Hormone study-Colon Cancer Family Registry (PMH-CCFR). In the following we describe each study population used in the genome-wide gene by environment (G X E) interaction analysis.

Colon Cancer Family Registry (CCFR)

The CCFR is an NCI-supported consortium consisting of six centers dedicated to the establishment of a comprehensive collaborative infrastructure for interdisciplinary studies in the genetic epidemiology of colorectal cancer.[1] The CCFR includes data from approximately 30,500 total subjects (10,500 probands, and 20,000 unaffected and affected relatives and unrelated controls). Cases and controls were recruited at the six participating centers beginning in 1998. CCFR implemented a standardized questionnaire that is administered to all participants, and includes established and suspected risk factors for colorectal cancer, which includes questions on medical history and medication use, reproductive history (for female participants), family history, physical activity, demographics, alcohol and tobacco use, and dietary factors. For genome-wide interaction analysis we only included the CCFR Set 1 scan, which has been described previously,[2] includes population-based cases and age-matched controls from the three population-based centers: Seattle, Toronto and Australia. Cases were genetically enriched by over-sampling those with a young age at onset or positive family history. Controls were matched to cases on age and sex. All cases and controls were self-reported as White, which was confirmed with genotype data. The CCFR Set 2 scan was not included as

controls were same generation family members and the statistical methods used are not easily applicable to this design.

Darmkrebs: Chancen der Verhütung durch Screening (DACHS) [3,4]

This German study was initiated as a large population-based case-control study in 2003 in the Rhine-Neckar-Odenwald region (southwest region of Germany) to assess the potential of endoscopic screening for reduction of colorectal cancer risk and to investigate etiologic determinants of disease, particularly lifestyle/environmental factors and genetic factors. Cases with a first diagnosis of invasive colorectal cancer (ICO-10 codes C18-C20) who were at least 30 years of age (no upper age limit), German speaking, a resident in the study region, and mentally and physically able to participate in a one-hour interview, were recruited by their treating physicians either in the hospital a few days after surgery, or by mail after discharge from the hospital. Cases were confirmed based on histologic reports and hospital discharge letters following diagnosis of colorectal cancer. All hospitals treating colorectal cancer patients in the study region participated. Based on estimates from population-based cancer registries, more than 50% of all potentially eligible patients with incident colorectal cancer in the study region were included. Community-based controls were randomly selected from population registries, employing frequency matching with respect to age (5-year groups), sex, and county of residence. Controls with a history of colorectal cancer were excluded. Controls were contacted by mail and follow-up calls. The participation rate was 51%. During an in-person interview, data were collected on demographics, medical history, family history of colorectal cancer, and various life-style factors, as were blood and mouthwash samples. The Set 1 scan consisted of a subset of participants recruited up to 2007, and samples were frequency matched on age and gender. The Set 2 scan consisted of additional subjects that were recruited up to 2010 as part of this ongoing study.

Diet, Activity and Lifestyle Study (DALS) [5]

DALS is a population-based case-control study of colon cancer. Participants were recruited between 1991 and 1994 from three locations: the Kaiser Permanente Medical Care Program (KPMCP) of Northern California, an eight-county area in Utah, and the

metropolitan Twin Cities area of Minnesota. Eligibility criteria for cases included age at diagnosis between 30 and 79 years, diagnosis with first primary colon cancer (ICD-O-2 codes 18.0 and 18.2-18.9) between October 1st 1991 and September 30th 1994, English speaking, and competency to complete the interview. Individuals with cancer of the rectosigmoid junction or rectum were excluded, as were those with a pathology report noting familial adenomatous polyposis, Crohn's disease, or ulcerative colitis. A rapid-reporting system was used to identify all incident cases of colon cancer resulting in the majority of cases being interviewed within four months of diagnosis. Controls from KPMCP were randomly selected from membership lists. In Utah, controls under 65 years of age were randomly selected through random-digit dialing and driver license lists. Controls, 65 years of age and older, were randomly selected from Health Care Financing Administration lists. In Minnesota, controls were identified from Minnesota driver's license or state ID lists. Controls were matched to cases by 5-year age groups and sex. The Set 1 scan consisted of a subset of the study designed above, from Utah, Minnesota, and KPMCP, and was restricted to subjects who self-reported as White non-Hispanic. The Set 2 scan consisted of subjects from Utah and Minnesota that were not genotyped in Set 1. Set 2 was restricted to subjects who self-reported as White non-Hispanic and those that had appropriate consent to post data to dbGaP.

Health Professionals Follow-up Study (HPFS) [6]

The HPFS is a parallel prospective study to the Nurses' Health Study (NHS). The HPFS cohort comprises 51,529 men who, in 1986, responded to a mailed questionnaire. The participants are U.S. male dentists, optometrists, osteopaths, podiatrists, pharmacists, and veterinarians born between 1910 and 1946. Participants have provided information on health related exposures, including: current and past smoking history, age, weight, height, diet, physical activity, aspirin and/or NSAID use, and family history of colorectal cancer. Colorectal cancer and other outcomes were reported by participants or next-of-kin and followed up through review of the medical and pathology record by physicians. Overall, more than 97% of self-reported colorectal cancers were confirmed by medical record review. Information was abstracted on histology and primary location. Follow-up has been excellent, with 94% of the men responding to date. Colorectal cancer cases were

ascertained through January 1, 2008. In 1993-95, 18,825 men in HPFS mailed in blood samples by overnight courier which were aliquoted into buffy coat and stored in liquid nitrogen. In 2001-04, 13,956 men in HPFS who had not previously provided a blood sample mailed in a "swish-and-spit" sample of buccal cells. Incident cases are defined as those occurring after the subject provided a blood or buccal sample. Prevalent cases are defined as those occurring after enrollment in the study in 1986, but prior to the subject providing either a blood or buccal sample. After excluding participants with histories of cancer (except non-melanoma skin cancer), ulcerative colitis, or familial polyposis, two case-control sets were constructed from which DNA was isolated from either buffy coat or buccal cells for genotyping: 1) a case-control set with cases of colorectal cancer matched to randomly selected controls who provided a blood sample and were free of colorectal cancer at the same time the colorectal cancer was diagnosed in the cases; 2) a case-control set with cases of colorectal cancer matched to randomly selected controls who provided a buccal sample and were free of colorectal cancer at the same time the colorectal cancer was diagnosed in the cases. For both case-control sets, matching criteria included year of birth (within 1 year) and month/year of blood or buccal cell sampling (within six months). Cases were pair matched 1:1, 1:2, or 1:3 with a control participant(s).

Nurses' Health Study (NHS) [7]

The NHS cohort began in 1976 when 121,700 married female registered nurses aged 30 to 55 years returned the initial questionnaire that ascertained a variety of important health-related exposures. Since 1976, follow-up questionnaires have been mailed every two years. Colorectal cancer and other outcomes were reported by participants or next-of-kin and followed up through review of the medical and pathology record by physicians. Overall, more than 97% of self-reported colorectal cancers were confirmed by medical-record review. Information was abstracted on histology and primary location. Follow-up has been high: as a proportion of the total possible follow-up time, follow-up has been over 92%. Colorectal cancer cases were ascertained through June 1, 2008. In 1989-90, 32,826 women in NHS mailed in blood samples by overnight courier which were aliquoted into buffy coat and stored in liquid nitrogen. In 2001-04, 29,684 women in

NHS who did not previously provide a blood sample mailed in a "swish-and-spit" sample of buccal cells. Incident cases are defined as those occurring after the subject provided a blood or buccal sample. Prevalent cases are defined as those occurring after enrollment in the study in 1976, but prior to the subject providing either a blood or buccal sample. After excluding participants with histories of cancer (except non-melanoma skin cancer), ulcerative colitis, or familial polyposis, we constructed two case-control sets from which DNA was isolated from either buffy coat or buccal cells for genotyping: 1) a case-control set with cases of colorectal cancer matched to randomly selected controls who provided a blood sample and were free of colorectal cancer at the same time the colorectal cancer was diagnosed in the cases; 2) a case-control set with cases of colorectal cancer matched to randomly selected controls who provided a buccal sample and were free of colorectal cancer at the same time the colorectal cancer was diagnosed in the cases. For both case-control sets, matching criteria included year of birth (within one year) and month/year of blood or buccal cell sampling (within six months). Cases were pair matched 1:1, 1:2, or 1:3 with a control participant(s).

Ontario Familial Colorectal Cancer Registry (OFCCR)

A subset of the Assessment of Risk in Colorectal Tumours in Canada (ARCTIC) from the Ontario Registry for Studies of Familial Colorectal Cancer (OFCCR) was used. Both the case-control study[8] and the OFCCR[9] have been described in detail previously, as have genome-wide association study (GWAS) results.[10] In brief, cases were confirmed incident colorectal cancer cases ages 20 to 74 years, residents of Ontario identified through comprehensive registry and diagnosed between July 1998 and June 2003. Population-based controls were randomly selected among Ontario residents (random-digit-dialing and listing of all Ontario residents), and matched by sex and 5-year age groups. A total of 1,236 colorectal cancer cases and 1,223 controls were successfully genotyped on at least one of the Illumina 1536 GoldenGate assay, the Affymetrix GeneChip® Human Mapping 100K and 500K Array Set, and a 10K non-synonymous SNP chip. Analysis was based on a set of unrelated subjects who were non-Hispanic, White by self-report or by investigation of genetic ancestry. We further excluded subjects if there was a sample mix-up, if they were missing epidemiologic questionnaire data, if

they were appendix cases, or if they were overlapped with the Colon Cancer Family Registry GWAS. Additionally, only samples genotyped on the Affymetrix GeneChip® 500K Array were utilized in order to avoid coverage issues in imputation.

Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial (PLCO)

PLCO enrolled 154,934 participants (men and women, aged between 55 and 74 years) at ten centers into a large, randomized, two-arm trial to determine the effectiveness of screening to reduce cancer mortality. Sequential blood samples were collected from participants assigned to the screening arm. Participation was 93% at the baseline blood draw. In the observational (control) arm, buccal cells were collected via mail using the "swish-and-spit" protocol and participation rate was 65%. Details of this study have been previously described[11,12] and are available online (http://dcp.cancer.gov/plco). The Set 2 GWAS data used in this study included a subset of 485 colorectal cancer cases from both arms of the trial. Samples were excluded if participants did not sign appropriate consents, if DNA was unavailable, if baseline questionnaire data with follow-up were unavailable, if they had a history of colon cancer prior to the trial, if they were a rare cancer, and if they were already in colon GWAS, or if they were a control in the prostate or lung populations. Controls were frequency matched 1:1 to cases without replacement, and cases were not eligible to be controls. Matching criteria were age at enrollment (two year blocks), enrollment date (two year blocks), sex, race/ethnicity, trial arm, and study year of diagnosis (i.e., controls must be cancer free into the case's year of diagnosis).

Postmenopausal Hormone study-Colon Cancer Family Registry (PMH-CCFR) [13]

Eligible case patients included all female residents, ages 50 to 74 years, residing in the 13 counties in Washington State reporting to the Cancer Surveillance SEER program, who were newly diagnosed with invasive colorectal adenocarcinoma (ICD-O C18.0, C18.2-.9, C19.9, C20.0-.9) between October 1998 and February 2002. Eligibility for all individuals was limited to those who were English-speaking with available telephone numbers, in which they could be contacted. On average, cases were identified within four months of diagnosis. The overall response proportion of eligible cases identified was 73%. Community-based controls were randomly selected according to age distribution (in 5-

year age intervals) of the eligible cases by using lists of licensed drivers from the Washington State Department of Licensing for individuals, ages 50 to 64 years, and rosters from the Health Care Financing Administration (now the Centers for Medicare and Medicaid) for individuals older than 64 years. The overall response proportion of eligible controls was 66%. In GECCO, samples with sufficient DNA extracted from blood were genotyped. Only participants that were not part of the CCFR Seattle site were included in the sample set.

VITamins And Lifestyle (VITAL)

The VITamins And Lifestyle (VITAL) cohort comprises of 77,721 Washington State men and women aged 50 to 76 years, recruited from 2000 to 2002 to investigate the association of supplement use and lifestyle factors with cancer risk. Subjects were recruited by mail, from October 2000 to December 2002, using names purchased from a commercial mailing list. All subjects completed a 24 page questionnaire and buccal cell specimens for DNA were self-collected by 70% of the participants. Subjects are followed for cancer by linkage to the western Washington Surveillance, Epidemiology, and End Results (SEER) cancer registry and are censored when they move out of the area covered by the registry or at time of death. Details of this study have been previously described.[14] In GECCO, a nested case-control set was genotyped. Samples included, colorectal cancer cases with DNA, excluding subjects with colorectal cancer before baseline, in situ cases, (large cell) neuroendocrine carcinoma, squamous cell carcinoma, carcinoid tumor, Goblet cell carcinoid, any type of lymphoma, including non-Hodgkin, Mantle cell, large B-cell, or follicular lymphoma. Controls were matched on age at enrollment (within one year), enrollment date (within one year), sex, and race/ethnicity. One control was randomly selected per case among all controls that matched on the four factors above and where the control follow-up time was greater than follow-up time of the case until diagnosis.

Women's Health Initiative (WHI)

WHI is a long-term health study of 161,808 post-menopausal women aged 50 to 79 years at 40 clinical centers throughout the US. WHI comprises a Clinical Trial (CT) arm, an Observational Study (OS) arm, and several extension studies. The details of WHI have

been previously described[15,16] and are available online (https://cleo.whi.org/SitePages/Home.aspx). In GECCO, Set 1 cases were selected from the September 12, 2005 database and were comprised of centrally adjudicated colon cancer cases from the Observational Study (OS) who self-reported as White. Controls were first selected among controls previously genotyped as part of a Hip Fracture GWAS conducted within the WHI-OS and matched to cases on age (within three years), enrollment date (within 365 days), hysterectomy status, and prevalent conditions at baseline. For 37 cases, there was not a control match in the Hip Fracture GWAS. For these participants, we identified a matched control in the WHI-OS based on same criteria. In the Set 2 scan, cases were selected from the August 2009 database and were comprised of centrally adjudicated colorectal cancer cases from the OS and CT who were not genotyped in Set 1. In addition, case and control participants were subject to the following exclusion criteria: a prior history of colorectal cancer at baseline, IRB approval not available for data submission into dbGaP, and not sufficient DNA available. Matching criteria included age (within years), race/ethnicity, WHI date (within three years), WHI Calcium and Vitamin D study date (within three years), and randomization arms (OS flag, hormone therapy assignments, dietary modification assignments, calcium/vitamin D assignments). In addition, they were matched on the four regions of randomization centers. Each case was matched with one control (1:1) that exactly met the matching criteria. Control selection was done in a time-forward manner, selecting one control for each case first from the risk set at the time of the case's event. The matching algorithm was allowed to select the closest match based on a criterion to minimize an overall distance measure.[17] Each matching factor was given the same weight. Additional available controls that were genotyped as part of the Hip Fracture GWAS were included to improve power.

Harmonization of environmental data:

All exposure information within each study, including regular use of aspirin and/or nonsteroidal anti-inflammatory drug (NSAID) and other colorectal cancer-related factors, was collected by in-person interviews and/or structured questionnaires, as detailed previously.[1,3,5,11,16,18-20] We carried out a multi-step data harmonization procedure,

reconciling each study's unique protocols and data-collection instruments at the GECCO coordinating center (Fred Hutchinson Cancer Research Center). First, we defined common data elements (CDEs). We examined the questionnaires and data dictionaries for each study to identify study-specific data elements that could be mapped to the CDEs. Through an iterative process, we communicated with each data contributor to obtain relevant data and coding information. The data elements were combined into a single dataset with common definitions, standardized permissible values and coding. The mapping and resulting data were reviewed for quality assurance, and range and logic checks were performed to assess data distributions within and between studies. Outlying samples were truncated to the minimum or maximum value of a pre-defined range for each variable. The reference time for cohort studies was time of enrollment (WHI, PLCO, and VITAL) or blood draw (HPFS and NHS). Dichotomous variables for regular use of either aspirin and/or NSAIDs (yes or no) or aspirin-only (yes or [no, regardless of use of other NSAIDs]) at the reference time were used for data analyses. The exact definition of regular use of aspirin and/or NSAIDs (including use of aspirin only, NSAIDs only, or both aspirin and NSAIDs), which was determined individually by each study cohort, is provided in **Table 1**. Non-regular users were considered as the reference. Data harmonization was performed using SAS and T-SQL.

Genotyping, quality assurance/quality control and imputation:

All analyses were based on genotyped data generated from genome-wide association scans and imputation to HapMap II. We note that genotyping for some cohorts was conducted at two different time points (i.e., sets 1 and 2) based on the availability of funds and samples. We accounted for this accordingly in the statistical analysis by analyzing each set separately before meta-analyzing data. Also, we have genotyped the cases and their matched controls together at the same time to avoid bias. CCFR genotyping was based on Illumina Human1M.[2] Phase one genotyping of DALS Set 1 and WHI Set 1 was done using Illumina HumanHap 550K/610K and Illumina 550Kduo/610K, respectively, and has been described previously.[21] OFCCR was genotyped using Affymetrix platforms.[10] DACHS Set 1, DALS Set 2, PMH-CCFR, PLCO Set 2, VITAL,

and WHI Set 2 were genotyped using Illumina HumanCytoSNP. HPFS, NHS, and DACHS Set 2 were genotyped using Illumina HumanOmniExpress.

DNA was extracted from blood samples or, for a subset of DACHS, HPFS, NHS, and PLCO samples, and for all VITAL samples, from buccal cells, using conventional methods. All studies included 1 to 6% blinded duplicates to monitor quality of the genotyping. All individual-level genotype data were managed, and underwent quality assurance and quality control (QA/QC) at University of Southern California (CCFR), the Ontario Institute for Cancer Research (OFCCR), the University of Washington Genetics Coordinating Center (HPFS, NHS, and DACHS Set 2), or the GECCO Coordinating Center at the Fred Hutchinson Cancer Research Center (all other studies). Details on the QA/QC can be found in **Supplementary Table 1**. In brief, samples were excluded based on call rate, heterozygosity, unexpected duplicates, gender discrepancy, and unexpectedly high identity-by-descent or unexpected genotype concordance (> 65%) with another individual. All analyses were restricted to samples clustering with the Utah residents with Northern and Western European ancestry from the CEPH collection (CEU) population in principal component analysis,[22] including the HapMap II populations as reference. Single nucleotide polymorphisms (SNPs) were excluded if they were triallelic, not assigned a rs number, or were reported or observed as not performing consistently across platforms. Additionally, genotyped SNPs were excluded based on call rate (< 98%), lack of Hardy-Weinberg Equilibrium in controls (HWE, $P < 1x10^{-4}$), and minor allele frequency (MAF < 5% for WHI Set 1, DALS Set 1, and OFCCR; MAF < 5 / # of samples for each other study). As imputation of genotypes is established as standard practice in the genetic association analysis, all autosomal SNPs of each study were imputed to the CEU population in HapMap II release 24, with the exception of OFCCR, which was imputed to HapMap II release 22. CCFR was imputed using IMPUTE,[10] OFCCR was imputed using BEAGLE,[23] and all other studies were imputed using MACH.[24] Imputed data were merged with genotype data such that genotype data were used if a SNP had both types of data, unless there was a difference in terms of reference allele frequency (> 0.1) or position (> 100 base pairs), in which case imputed data were used. Given the high agreement of imputation accuracy among MACH, IMPUTE, and BEAGLE,[25] the

common practice of using different imputation programs is unlikely to cause heterogeneity[26] and the results can be combined without any further correction. We calculated R^2 as a measurement of imputation accuracy. SNPs were restricted based on per study MAF > 5 / # of samples and per study imputation accuracy ($R^2 > 0.3$). After imputation and quality control (QC) analyses, a total of about 2.7 million SNPs were used in the analysis. In the statistical analyses, both genotyped and imputed SNPs were examined as continuous variables (i.e., assuming log-additive effects). Briefly, under the log-additive model, the statistical effect of a homozygous variant genotype is assumed to be twice the statistical effect of a heterozygous genotype on a logit-scale. This is equivalent to considering genotype according to dosage or number of variant alleles (0, 1 and 2) and evaluating its contribution to the model as a continuous covariate. For imputed genotypes, we obtained the posterior probabilities for heterozygous and homozygous variant genotypes from the MACH imputation program to calculate the expected dosage as 2Pr(Genotype=AA) + Pr(Genotype=Aa). Because the posterior probabilities are constrained between 0 and 1, the expected dosage will be between 0 and 2. We have previously shown that the expected dosage provides a valid inference of the actual number of variant alleles.[27] To evaluate overall performance, we calculated the genomic inflation factor (λ) to measure the over-dispersion of the test-statistics from the marginal association tests by dividing the median of the squared Z statistics by 0.455, the median of a chi-squared distribution with 1 degree of freedom. The inflation factor λ was between 0.999 and 1.044 for individual studies based on all SNPs including both directly genotyped and imputed, indicating there is little evidence of residual population substructure, cryptic relatedness, or differential genotyping between cases and controls. This result was consistent with the visual inspection of the study-specific quantile-quantile (Q-Q) plots.

Statistical models for interaction analyses:

For the conventional logistic regression analysis, we modeled G X E interaction using the cross-product of number of copies of the variant allele for the SNP and the regular use of aspirin and/or NSAIDs while simultaneously adjusting for the main associations of the SNP and use of aspirin and/or NSAIDs with colorectal cancer risk. For conventional

logistic regression analysis, we fitted the log-additive model: Logit(Pr(D=1)) = b0 + b1*(NSAID=1) + b2*E(G) + b3*(NSAID=1)*E(G), where E(G) is expected dosage for imputed SNPs and dosage for genotyped SNPs. For case-only interaction analysis, we also fitted conventional logistic regression but in colorectal cancer cases only. The models are: log(prob(G=1|D=1)/prob(G=0|D=1)) = b01 + b3*(NSAID=1); and log(prob(G=2|D=1)/prob(G=0|D=1)) = b02 + 2b3*(NSAID=1); note that b3 in the case-only logistic regression model is the same parameter as the interaction statistical effect b3 in the case-control logistic regression model. The G and E association in case-only analysis is equivalent to G X E interaction analysis when G and E are independent in the population and the disease is rare, because in this case the correlation of G and E is approximately 0 in the controls. The case-only test improves statistical power considerably compared with the conventional case-control interaction test under some scenarios, as the analysis does not need to account for the variation in the control population when the G and E are independent in the population.

Stratified analysis:

We performed stratified analysis for the SNPs showing gene-environment (G X E) interaction with aspirin and/or NSAID use using conventional logistic regression. We estimated the association of aspirin and/or NSAID use with colorectal cancer risk stratified by SNP genotypes, as well as the associations in strata defined by SNP and use of aspirin and/or NSAID with one common reference group. We pooled the studies for the stratified analyses to minimize strata with small sample sizes. Briefly, to evaluate the associations between aspirin and/or NSAID use and colorectal cancer stratified by genotypes accounting for imputation, we fit the following model: logit(Pr[D=1]) = $b_0 + b_1e + c_1p_1 + c_2p_2 + \beta_1p_1e + \beta_2p_2e$ + covariates, where p_1 and p_2 are the imputation posterior probabilities for genotypes A/B and B/B. The stratified effects of aspirin and/or NSAID use were estimated by $\hat{b}_1$, $\hat{b}_1+\hat{\beta}_1$, $\hat{b}_1+\hat{\beta}_2$ for genotype A/A, A/B, and B/B, respectively with standard errors obtained by using the standard formula for linear combination of two parameters based on the covariance matrix of these parameter estimators.

Calculation of absolute risk:

We calculated absolute risks for each genotype of the SNPs showing G X E interaction. Briefly, based upon the Surveillance, Epidemiology, and End Results (SEER) age-adjusted colorectal cancer incidence rate (denoted by "I") between 2007-2011 among the White population of 42.9 per 100,000 men and women per year, we estimated the reference incidence rate of colorectal cancer (denoted by "I_{reference}") using the following formula: I_{reference} = I/(P(AA, non-E) + OR_{Aa/aa, non-E} P(Aa/aa, non-E) + OR_{AA, E} P(AA, E) + OR_{Aa/aa, E} P(Aa/aa, E)), where P(genotype, E (or non-E)) is the prevalence of aspirin and/or NSAID use (or non-use) in each corresponding genotype category among controls (non-cases). Based on this reference incidence rate of colorectal cancer (i.e., I_{reference}), we further calculated absolute colorectal cancer incidence rates within each subgroup defined by genotype of the SNPs according to aspirin and/or NSAID use or non-use by multiplying the I_{reference} with each corresponding OR.

Calculation of D′ and r^2:

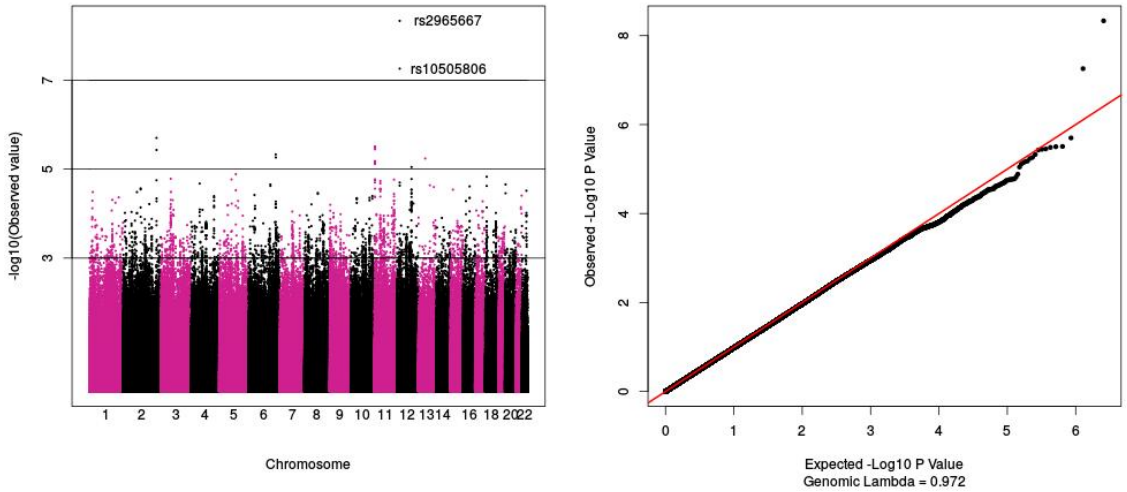
To examine whether the two SNPs identified from conventional logistic regression analysis are correlated, we obtained D′ and r^2 using HapMap CEU population data. Briefly, the deviation of the observed frequency of two loci from the expected is a quantity called the linkage disequilibrium (LD) and is commonly denoted by D. r^2 is the squared correlation, where *r* scales D by the standard deviations of the allele frequencies at two loci. D′ scales D by dividing it by the theoretical maximum for the observed allele frequencies. A value of 0 for D′ indicates that the examined loci are in fact independent of one another, while a value of 1 demonstrates complete dependency (i.e., two SNPs are highly correlated).

eReferences

1. Newcomb PA, Baron J, Cotterchio M, et al. Colon Cancer Family Registry: an international resource for studies of the genetic epidemiology of colon cancer. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology.* Nov 2007;16(11):2331-2343.
2. Figueiredo JC, Lewinger JP, Song C, et al. Genotype-environment interactions in microsatellite stable/microsatellite instability-low colorectal cancer: results from a genome-wide association study. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology.* May 2011;20(5):758-766.
3. 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.* Jan 4 2011;154(1):22-30.
4. Lilla C, Verla-Tebit E, Risch A, et al. Effect of NAT1 and NAT2 genetic polymorphisms on colorectal cancer risk associated with exposure to tobacco smoke and meat consumption. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology.* Jan 2006;15(1):99-107.
5. Slattery ML, Potter J, Caan B, et al. Energy balance and colon cancer--beyond physical activity. *Cancer Res.* Jan 1 1997;57(1):75-80.
6. 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.* Nov 1990;1(6):466-473.
7. Belanger CF, Hennekens CH, Rosner B, Speizer FE. The nurses' health study. *Am J Nurs.* Jun 1978;78(6):1039-1040.
8. Cotterchio M, Manno M, Klar N, McLaughlin J, Gallinger S. Colorectal screening is associated with reduced colorectal cancer risk: a case-control study within the population-based Ontario Familial Colorectal Cancer Registry. *Cancer Causes Control.* Sep 2005;16(7):865-875.
9. Cotterchio M, McKeown-Eyssen G, Sutherland H, et al. Ontario familial colon cancer registry: methods and first-year response rates. *Chronic Dis Can.* 2000;21(2):81-86.
10. Zanke BW, Greenwood CM, Rangrej J, et al. Genome-wide association scan identifies a colorectal cancer susceptibility locus on chromosome 8q24. *Nature genetics.* Aug 2007;39(8):989-994.
11. Prorok PC, Andriole GL, Bresalier RS, et al. Design of the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial. *Controlled clinical trials.* Dec 2000;21(6 Suppl):273S-309S.
12. Gohagan JK, Prorok PC, Hayes RB, Kramer BS. The Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial of the National Cancer Institute: history, organization, and status. *Controlled clinical trials.* Dec 2000;21(6 Suppl):251S-272S.

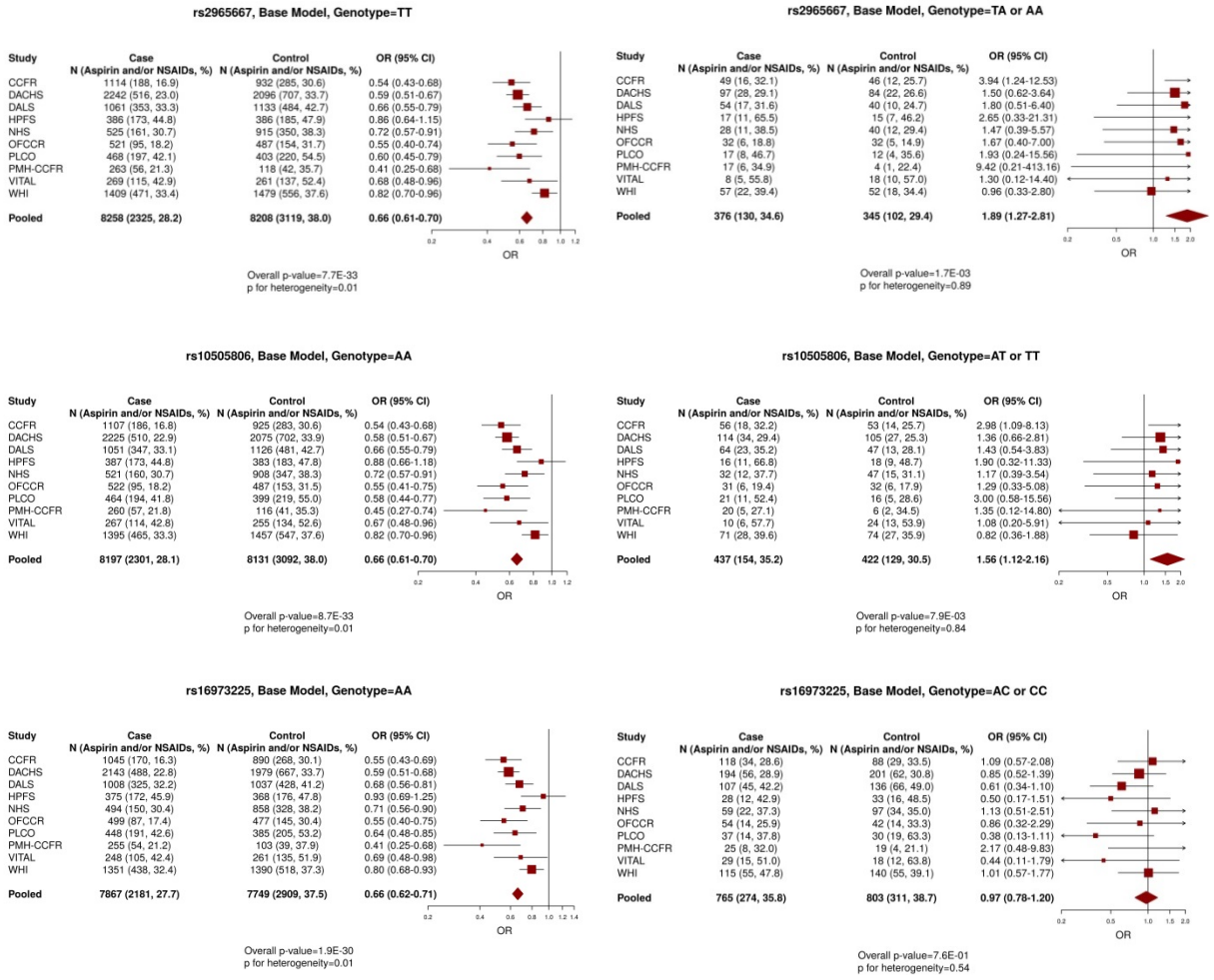
13. Newcomb PA, Zheng Y, Chia VM, et al. Estrogen plus progestin use, microsatellite instability, and the risk of colorectal cancer in women. *Cancer Res.* Aug 1 2007;67(15):7534-7539.
14. White E, Patterson RE, Kristal AR, et al. VITamins And Lifestyle cohort study: study design and characteristics of supplement users. *American journal of epidemiology.* Jan 1 2004;159(1):83-93.
15. Hays J, Hunt JR, Hubbell FA, et al. The Women's Health Initiative recruitment methods and results. *Ann Epidemiol.* Oct 2003;13(9 Suppl):S18-77.
16. Design of the Women's Health Initiative clinical trial and observational study. The Women's Health Initiative Study Group. *Controlled clinical trials.* Feb 1998;19(1):61-109.
17. Pruim RJ, Welch RP, Sanna S, et al. LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics.* Sep 15 2010;26(18):2336-2337.
18. Kury S, Buecher B, Robiou-du-Pont S, et al. Combinations of cytochrome P450 gene polymorphisms enhancing the risk for sporadic colorectal cancer related to red meat consumption. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology.* Jul 2007;16(7):1460-1467.
19. Colditz GA, Hankinson SE. The Nurses' Health Study: lifestyle and health among women. *Nat Rev Cancer.* May 2005;5(5):388-396.
20. Giovannucci E, Rimm EB, Stampfer MJ, Colditz GA, Ascherio A, Willett WC. Aspirin use and the risk for colorectal cancer and adenoma in male health professionals. *Ann Intern Med.* Aug 15 1994;121(4):241-246.
21. Sever ML, Salo PM, Haynes AK, Zeldin DC. Inner-city environments and mitigation of cockroach allergen. *Am J Prev Med.* Aug 2011;41(2 Suppl 1):S55-56.
22. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet.* Aug 2006;38(8):904-909.
23. Browning SR, Browning BL. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *American journal of human genetics.* Nov 2007;81(5):1084-1097.
24. Li Y, Willer CJ, Ding J, Scheet P, Abecasis GR. MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genetic epidemiology.* Dec 2010;34(8):816-834.
25. Nothnagel M, Ellinghaus D, Schreiber S, Krawczak M, Franke A. A comprehensive evaluation of SNP genotype imputation. *Human genetics.* Mar 2009;125(2):163-171.
26. Gogele M, Minelli C, Thakkinstian A, et al. Methods for meta-analyses of genome-wide association studies: critical assessment of empirical evidence. *American journal of epidemiology.* Apr 15 2012;175(8):739-749.
27. Jiao S, Hsu L, Hutter CM, Peters U. The use of imputed values in the meta-analysis of genome-wide association studies. *Genetic epidemiology.* Nov 2011;35(7):597-605.

eFigure 1. Manhattan Plot and Q-Q Plot for the Interaction Results With Aspirin and/or NSAIDs (Meta-analysis) From Conventional Logistic Regression Analysis



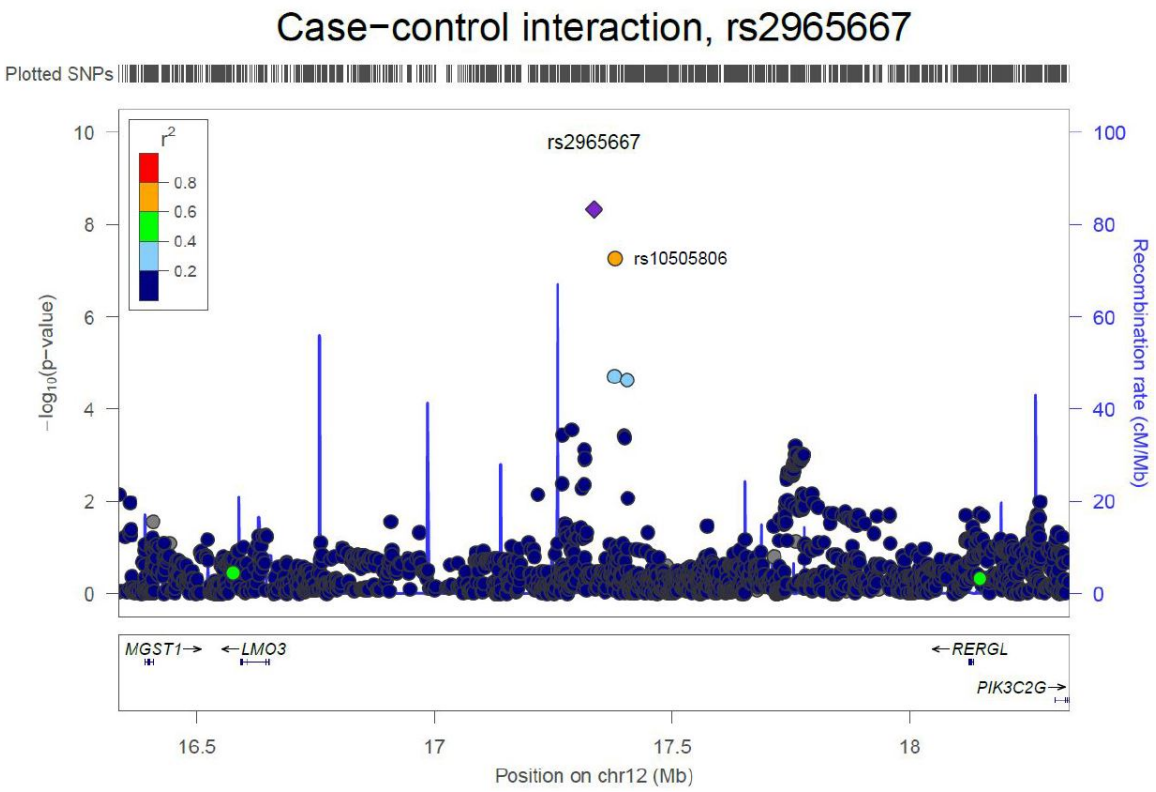
"Aspirin and/or NSAIDs" includes the regular use of aspirin only, NSAIDs only, or both aspirin and NSAIDs.

eFigure 2. Risk for Colorectal Cancer According to Regular Use of Aspirin and/or NSAIDs, Stratified by the Genotypes of rs2965667, rs10505806, and rs16973225



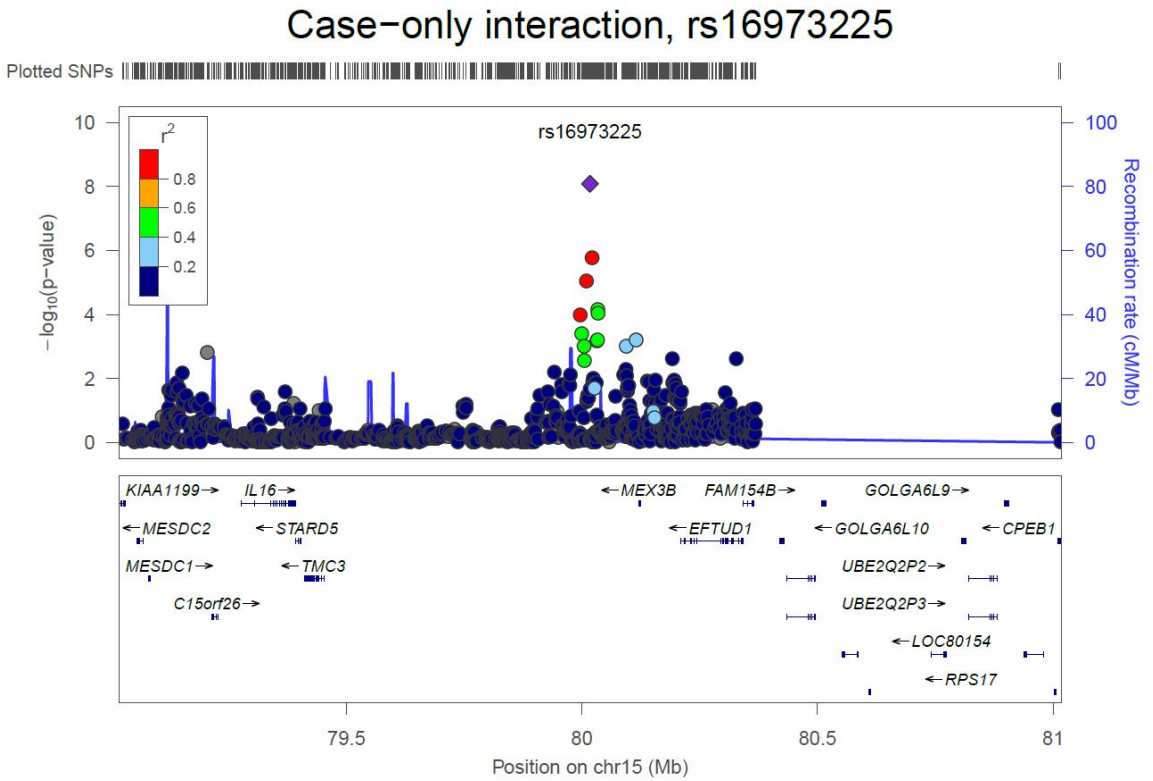
"Aspirin and/or NSAIDs" includes the regular use of aspirin only, NSAIDs only, or both aspirin and NSAIDs. The size of the data markers is proportional to the precision of the estimate, which is the inverse of the variance.

eFigure 3. Regional Association Plot of 1000 kb for the Interaction Between Regular Use of Aspirin and/or NSAIDs and rs2965667, as Well as Surrounding SNPs



"Aspirin and/or NSAIDs" includes the regular use of aspirin only, NSAIDs only, or both aspirin and NSAIDs. The top half of the figure has physical position along the x-axis, and the -log10 of the meta-analysis *p*-value on the y-axis. Each dot on the plot represents the *p*-value of the interaction for one SNP in relation to colorectal cancer conducted across all studies. The most significant SNP in the region (index SNP) is marked as a purple diamond. The color scheme represents the pairwise correlation (r^2) for the SNPs across the region with the index SNP. Interaction was calculated using the HapMap CEU data. The bottom half of the figure shows the position of the genes across the region. The genomic coordinate is in NCBI36.1/hg18.

eFigure 4. Regional Association Plot of 1000 kb for the Interaction Between Regular Use of Aspirin and/or NSAIDs and rs16973225, as Well as Surrounding SNPs



"Aspirin and/or NSAIDs" includes the regular use of aspirin only, NSAIDs only, or both aspirin and NSAIDs. The top half of the figure has physical position along the x-axis, and the -log10 of the meta-analysis *p*-value on the y-axis. Each dot on the plot represents the *p*-value of the interaction for one SNP in relation to colorectal cancer conducted across all studies. The most significant SNP in the region (index SNP) is marked as a purple diamond. The color scheme represents the pairwise correlation (r^2) for the SNPs across the region with the index SNP. Interaction was calculated using the HapMap CEU data. The bottom half of the figure shows the position of the genes across the region. The genomic coordinate is in NCBI36.1/hg18.

eTable 1. Details on Genotyping Platform and Quality Assurance and Quality Control (QA/QC Measurements)[a]

Study	Genotyping Platform[b]	Duplicate Concordance	Sample Call Rate	SNP Exclusions[c]	SNPs Passing QC	SNP Call Rate	No. of Imputed SNPs by R^2		
		(%)	(Mean)	(#)	(#)	(Mean)	< 0.3	0.3-0.8	> 0.8
DACHS Set 1	300K	99.9%	99.93%	33,588	255,208	99.90%	70,989	434,295	1,869,458
DACHS Set 2	730K	100%	99.84%	32,159	609,115	99.85%	18,551	154,813	1,865,294
DALS Set 1	550K, 610K	>97%[d]	99.69%	34,644	516,631	99.82%	20,173	180,322	1,912,832
DALS Set 2	300K	100%	99.94%	32,885	250,320	99.94%	69,289	438,282	1,867,371
HPFS Set 1	730K	99.9%	99.93%	32,953	612,091	99.93%	18,257	150,880	1,857,252
HPFS Set 2	730K	99.9%	99.83%	51,725	590,132	99.84%	20,040	160,464	1,861,553
NHS Set 1	730K	100%	99.93%	47,295	628,541	99.93%	17,142	147,723	1,855,814
NHS Set 2	730K	100%	99.81%	53,328	594,015	99.81%	19,434	160,804	1,875,767
PLCO Set 2	300K	99.9%	99.80%	38,655	253,702	99.90%	68,059	434,769	1,870,311
PMH-CCFR	300K	99.9%	99.89%	39,275	256,743	99.92%	67,818	429,887	1,875,260
VITAL	300K	99.9%	99.81%	36,805	243,625	99.89%	73,966	461,036	1,845,318
WHI Set 1	550Kduo, 610K	>97%[d]	99.60%	40,276	511,251	99.77%	21,655	184,833	1,914,909
WHI Set 2	300K	100%	99.96%	27,392	251,707	99.96%	72,272	442,111	1,864,141

We note that genotyping for some cohorts was conducted at two different time points (i.e., sets 1 and 2) based on the availability of funds and samples. We accounted for this accordingly in the statistical analysis by analyzing each set separately before meta-analyzing data. Also, we have genotyped the cases and their matched controls together at the same time to avoid bias.

[a] CCFR and OFCCR had QA/QC performed separately by CCCR and OFCCR investigators as documented in Zanke et al. 2007 and Figueiredo et al. 2011. All QA/QC numbers are based on the total number of subjects with GWAS data per study.

[b] All platforms were Illumina assays, except for OFCCR, which was genotyped using Affymetrix products.

[c] Directly genotyped SNPs were excluded for a call rate < 98%, *P*-value for Hardy Weinberg Equilibrium (HWE) < 1×10^{-4}, and low minor allele frequency (MAF < 5% for WHI Set 1 and DALS Set 1; MAF< 5 / # of samples for each other study; this MAF reflects exclusions going into imputation step, not exclusions for marginal association analysis), and if SNPs reportedly did not perform consistently across platforms.

[d] Blinded duplicates were assessed across DALS set 1 and WHI Set 1; exact concordance was not recorded, but all 98 pairs were identified as having concordance > 97%.

eTable 2. Interaction Between Regular Use of Aspirin Only and rs2965667 on the Risk of Colorectal Cancer

		rs2965667 genotype			**OR (95% CI) for genotype within strata of aspirin**
	TT		TA/AA		
	N Cases/Controls	OR (95% CI)	N Cases/Controls	OR (95% CI)	
Non-regular aspirin users	5,603/5,207	1.00	238/237	0.92 (0.73-1.15)	0.91 (0.72-1.15)
				P= 0.46	*P*= 0.43
Regular aspirin users	1,714/2,353	0.68 (0.63-0.74)	101/81	1.58 (1.09-2.29)	2.27 (1.54-3.35)
		P= $1x10^{-21}$		*P*= 0.016	*P*= $3.4x10^{-5}$
OR (95% CI) for aspirin within strata of genotype		0.68 (0.63-0.74)		1.72 (1.12-2.65)	
		P= $1x10^{-21}$		*P*= 0.014	

ORs are calculated after adjusting for age at the reference time, sex, center, and the first three principal components from EIGENSTRAT.

eTable 3. Imputation Quality for Three SNPs (rs2965667, rs10505806 and rs16973225) Identified in this Study

rs2965667	**Study**	**Imputed/Genotyped**	**Allele 'A' frequency (%)**	**Imputation R^2**
	CCFR	Imputed	2.4	0.703
	OFCCR	Imputed	3.2	0.977
	DACHS Set 1	Imputed	1.9	0.625
	DACHS Set 2	Imputed	2.1	0.634
	DALS Set 1	Imputed	1.6	0.689
	DALS Set 2	Imputed	2.0	0.697
	HPFS Set 1	Imputed	2.0	0.669
	HPFS Set 2	Imputed	1.9	0.620
	NHS Set 1	Imputed	2.1	0.683
	NHS Set 2	Imputed	2.3	0.601
	PLCO Set 2	Imputed	1.5	0.587
	PMH-CCFR	Imputed	1.7	0.749
	VITAL	Imputed	3.2	0.627
	WHI Set 1	Imputed	2.1	0.649
	WHI Set 2	Imputed	1.8	0.606
rs10505806	**Study**	**Imputed/Genotyped**	**Allele 'T' frequency**	**Imputation R^2**
	CCFR	Imputed	2.8	0.787
	OFCCR	Imputed	3.2	1.000
	DACHS Set 1	Imputed	1.8	0.790
	DACHS Set 2	Imputed	2.4	0.797
	DALS Set 1	Imputed	2.4	0.779
	DALS Set 2	Imputed	2.4	0.831
	HPFS Set 1	Imputed	2.2	0.787
	HPFS Set 2	Imputed	2.2	0.739
	NHS Set 1	Imputed	2.3	0.794
	NHS Set 2	Imputed	3.2	0.771
	PLCO Set 2	Imputed	2.0	0.793
	PMH-CCFR	Imputed	2.3	0.842
	VITAL	Imputed	4.4	0.827
	WHI Set 1	Imputed	2.6	0.726
	WHI Set 2	Imputed	2.7	0.807
rs16973225	**Study**	**Imputed/Genotyped**	**Allele 'C' frequency**	**Imputation R^2**
	CCFR	Imputed	4.6	0.955
	OFCCR	Imputed	4.1	0.991
	DACHS Set 1	Genotyped	4.8	NA
	DACHS Set 2	Genotyped	4.6	NA
	DALS Set 1	Imputed	6.0	0.930
	DALS Set 2	Genotyped	5.8	NA
	HPFS Set 1	Genotyped	4.1	NA
	HPFS Set 2	Genotyped	4.4	NA
	NHS Set 1	Genotyped	5.7	NA
	NHS Set 2	Imputed	2.5	1.000
	PLCO Set 2	Genotyped	3.7	NA
	PMH-CCFR	Genotyped	7.8	NA
	VITAL	Imputed	3.4	0.805
	WHI Set 1	Imputed	4.9	0.928
	WHI Set 2	Genotyped	4.6	NA

eTable 4. Additional Descriptive Characteristics of Study Populations

Study	Female No. (%)		Mean Age (range, yrs)		Smoking[a] No. (%)		BMI (kg/cm^2) Mean (SD)		Alcohol (g/day) Mean (SD)		Red meat (serving/day) Mean (SD)	
	Cases	Controls	Cases	Controls	Cases	Controls	Cases	Controls	Cases	Controls	Cases	Controls
CCFR	558 (48)	509 (52)	51.1 (17-81)	58 (21-76)	553 (47.5)	549 (56.1)	28.2 (7.4)	26.9 (6)	-	-	0.7 (0.6)	0.6 (0.5)
DACHS	952 (40.7)	849 (38.9)	68.5 (33-94)	69 (34-99)	1389 (59.4)	1216 (55.8)	27 (4.1)	26.3 (3.7)	15.9 (21.4)	14.5 (18.9)	0.8 (0.4)	0.7 (0.3)
DALS	497 (44.6)	530 (45.2)	63.7 (30-78)	64 (28-79)	636 (57)	597 (50.9)	27.7 (5.3)	26.4 (4.5)	11 (23.3)	9.2 (18.7)	1.1 (0.8)	1 (0.8)
HPFS	-	-	65.2 (48-82)	65.2 (48-83)	218 (54.1)	208 (51.9)	26.3 (3.2)	25.4 (3.3)	14.3 (17.4)	12.3 (15)	0.9 (0.7)	0.7 (0.6)
NHS	553 (100)	955 (100)	59.5 (44-69)	59.9 (44-69)	326 (59)	529 (55.4)	25.4 (4.5)	25.5 (4.3)	5.9 (10.1)	5.8 (10.6)	0.7 (0.6)	0.7 (0.5)
OFCCR	352 (63.7)	225 (43.4)	61.6 (33-77)	62.7 (29-77)	309 (55.9)	305 (58.8)	26.2 (4.3)	26.3 (4.5)	-	-	0.6 (0.6)	0.6 (0.5)
PMH-CCFR	280 (100)	122 (100)	63.3 (48-73)	61.6 (48-73)	38 (13.6)	15 (12.3)	27.8 (6.1)	25.5 (4.8)	-	-	0.4 (0.3)	0.4 (0.4)
PLCO	207 (42.7)	175 (42.2)	63.7 (55-75)	63.6 (55-75)	270 (55.7)	212 (51.1)	27.5 (4.4)	27.3 (4.3)	13.2 (26)	11.8 (21.7)	1.2 (1)	1.2 (1)
VITAL	133 (48)	135 (48.4)	66.4 (51-76)	66.6 (50-76)	176 (63.5)	153 (54.8)	28.1 (5.7)	26.9 (4.6)	12.5 (21.2)	7.5 (13.9)	0.7 (0.5)	0.6 (0.5)
WHI	1466 (100)	1531 (100)	66.3 (50-79)	66.4 (50-79)	769 (52.5)	724 (47.3)	28.3 (5.6)	27.6 (5.5)	5.4 (10.7)	5.2 (9.8)	0.7 (0.6)	0.7 (0.6)

[a] Sample size of ever smokers in each study, i.e., including both former and current smokers.

eTable 5. Risk for Colorectal Cancer According to Regular Use of Aspirin and/or NSAIDs, Stratified by the Genotypes of rs2965667, rs10505806, and rs16973225

rs2965667[a]	**Non-regular aspirin and/or NSAID users**		**Regular aspirin and/or NSAID users**		***P*-value**
TT					
Cases/Controls	5,933/5,088		2,325/3,119		
Base Model (OR)[c]	1.00		0.66 (0.61-0.70)		$1.1x10^{-32}$
Multivariable-Adjusted Model (OR)[d]	1.00		0.63 (0.59-0.68)		$3.1x10^{-35}$
TA					
Cases/Controls	243/240		126/101		
Base Model (OR)[c]	1.00		1.74 (1.16-2.61)		0.01
Multivariable-Adjusted Model (OR)[d]	1.00		1.62 (1.06-2.48)		0.03
AA					
Cases/Controls	3/4		4/1		
Base Model (OR)[c]	1.00		-		-
Multivariable-Adjusted Model (OR)[d]	1.00		-		-
P for interaction[e]	$4.6x10^{-9}$				
rs10505806[a]	**Non-regular aspirin and/or NSAID users**		**Regular aspirin and/or NSAID users**		***P*-value**
AA					
Cases/Controls	5,896/5,039		2,301/3,092		
Base Model (OR)[c]	1.00		0.66 (0.61-0.70)		$1.0x10^{-32}$
Multivariable-Adjusted Model (OR)[d]	1.00		0.63 (0.59-0.68)		$4.7x10^{-35}$
AT					
Cases/Controls	279/287		150/128		
Base Model (OR)[c]	1.00		1.47 (1.05-2.05)		0.02
Multivariable-Adjusted Model (OR)[d]	1.00		1.34 (0.94-1.90)		0.10
TT					
Cases/Controls	4/6		4/1		
Base Model (OR)[c]	1.00		-		-
Multivariable-Adjusted Model (OR)[d]	1.00		-		-
P for interaction[e]	$5.5x10^{-8}$				

rs16973225[b]	Non-regular aspirin and/or NSAID users		Regular aspirin and/or NSAID users		*P*-value
AA					
Cases/Controls	5,686/4,840		2,181/2,909		
Base Model (OR)[c]	1.00		0.66 (0.62-0.71)		1.9×10^{-30}
Multivariable-Adjusted Model (OR)[d]	1.00		0.63 (0.59-0.68)		3.6×10^{-33}
AC					
Cases/Controls	475/483		266/305		
Base Model (OR)[c]	1.00		0.97 (0.78-1.20)		0.80
Multivariable-Adjusted Model (OR)[d]	1.00		0.94 (0.75-1.18)		0.58
CC					
Cases/Controls	16/9		8/6		
Base Model (OR)[c]	1.00		0.85 (0.21-3.37)		0.81
Multivariable-Adjusted Model (OR)[d]	1.00		0.81 (0.20-3.30)		0.77
P for interaction[e]	8.2×10^{-9}				

The numbers of cases and controls were from the Base Model.
We note that because the stratified analyses were based on the three genotypes, the *p*-values corresponding to the wild-genotype are slightly different from that in Table 2 where the homozygous variant genotype was grouped with the heterozygous genotype due to the low count of homozygous variant genotype.
"Aspirin and/or NSAIDs" includes the regular use of aspirin only, NSAIDs only, or both aspirin and NSAIDs.
" - ": ORs (95% CIs) and *p*-values cannot be estimated due to small sample size in this group.
[a] SNPs rs2965667 and rs10505806 were identified from conventional logistic regression analysis.
[b] SNP rs16973225 was identified from case-only interaction analysis.
[c] ORs in Base Models are adjusted for age at the reference time, sex, center, and the first three principal components from EIGENSTRAT.
[d] ORs in Multivariable-Adjusted Models are adjusted for age at the reference time, sex, center, the first three principal components, smoking status (never, former, or current smoker), BMI, alcohol consumption, and red meat consumption.
[e] *P*-values for interactions were calculated after adjusting for age at the reference time, sex, center, and the first three principal components from EIGENSTRAT.

eTable 6. Interactions Between Regular Use of Aspirin and/or NSAIDs and Genotypes of rs2965667, rs10505806, and rs16973225 on the Risk of Colorectal Cancer

		rs2965667 genotype			**OR (95% CI) for genotype within strata of aspirin and/or NSAIDs**
	TT		TA/AA		
	N Cases/Controls	OR (95% CI)	N Cases/Controls	OR (95% CI)	
Non-regular aspirin and/or NSAID users	5,933/5,088	1.00	246/244	0.81 (0.64-1.01)	0.80 (0.63-1.00)
				P= 0.06	*P*= 0.05
Regular aspirin and/or NSAID users	2,325/3,119	0.66 (0.61-0.70)	130/102	1.52 (1.09-2.12)	2.36 (1.67-3.34)
		P= 7.7×10^{-33}		*P*= 0.014	*P*= 1.1×10^{-6}
OR (95% CI) for aspirin and/or NSAIDs within strata of genotype		0.66 (0.61-0.70)		1.89 (1.27-2.81)	
		P= 7.7×10^{-33}		*P*= 0.002	
		rs10505806 genotype			**OR (95% CI) for genotype within strata of aspirin and/or NSAIDs**
	AA		AT/TT		
	N Cases/Controls	OR (95% CI)	N Cases/Controls	OR (95% CI)	
Non-regular aspirin and/or NSAID users	5,896/5,039	1.00	283/293	0.78 (0.64-0.94)	0.78 (0.64-0.94)
				P= 0.011	*P*= 0.10
Regular aspirin and/or NSAID users	2,301/3,092	0.66 (0.61-0.70)	154/129	1.21 (0.93-1.59)	1.88 (1.42-2.49)
		P= 8.7×10^{-33}		*P*= 0.16	*P*= 1.2×10^{-5}
OR (95% CI) for aspirin and/or NSAIDs within strata of genotype		0.66 (0.61-0.70)		1.56 (1.12-2.16)	
		P= 8.7×10^{-33}		*P*= 0.008	
		rs16973225 genotype			**OR (95% CI) for genotype within strata of aspirin and/or NSAIDs**
	AA		AC/CC		
	N Cases/Controls	OR (95% CI)	N Cases/Controls	OR (95% CI)	
Non-regular aspirin and/or NSAID users	5,686/4,840	1.00	491/492	0.83 (0.72-0.95)	0.82 (0.72-0.94)
				P= 0.006	*P*= 0.005

Regular aspirin and/or NSAID users	2,181/2,909	0.66 (0.62-0.71)	274/311	0.80 (0.67-0.95)	1.23 (1.03-1.47)
		$P = 1.9 \times 10^{-30}$		$P = 0.012$	$P = 0.025$
OR (95% CI) for aspirin and/or NSAIDs within strata of genotype		0.66 (0.62-0.71)		0.97 (0.78-1.20)	
		$P = 1.9 \times 10^{-30}$		$P = 0.76$	

ORs are calculated after adjusting for age at the reference time, sex, center, and the first three principal components from EIGENSTRAT.
"Aspirin and/or NSAIDs" includes the regular use of aspirin only, NSAIDs only, or both aspirin and NSAIDs.