Research Article

Cancer Epidemiology, Biomarkers & Prevention

# Genetic Variation in the Inflammation and Innate Immunity Pathways and Colorectal Cancer Risk

Hansong Wang[1], Darin Taverna[2], Daniel O. Stram[3], Barbara K. Fortini[3], Iona Cheng[4], Lynne R. Wilkens[1], Terrilea Burnett[1], Karen W. Makar[5], Noralane M. Lindor[6], John L. Hopper[7], Steve Gallinger[8], John A. Baron[9], Robert Haile[10], Laurence N. Kolonel[1], Brian E. Henderson[3], Polly A. Newcomb[5], Graham Casey[3], David Duggan[2], Cornelia M. Ulrich[5,11], and Loïc Le Marchand[1]

## Abstract

**Background:** It is widely accepted that chronic inflammation plays a role in the etiology of colorectal cancer. Using a two-stage design, we examined the associations between colorectal cancer and common variation in 37 key genes in the inflammation and innate immunity pathways.

**Methods:** In the discovery stage, 2,322 discordant sibships (2,535 cases, 3,915 sibling controls) from the Colorectal Cancer Family Registry were genotyped for more than 600 tagSNPs and 99 single-nucleotide polymorphisms (SNP) were selected for further examination based on strength of association. In the second stage, 351 SNPs tagging gene regions covered by the 99 SNPs were tested in 4,783 Multiethnic Cohort subjects (2,153 cases, 2,630 controls).

**Results:** The association between rs9858822 in the *PPARG* gene and colorectal cancer was statistically significant at the end of the second stage (OR per allele = 1.36, Bonferroni-adjusted $P = 0.045$), based on the "effective" number of markers in stage II ($n = 306$). The risk allele C was common (frequency 0.3) in African Americans but rare (frequency < 0.03) in whites, Japanese Americans, Latinos, and Native Hawaiians. No statistically significant heterogeneity of effects across race/ethnicity, body mass index (BMI) levels, regular aspirin use, or pack-years of smoking was detected for this SNP. Suggestive associations were also observed for several SNPs in close vicinity to rs9858822.

**Conclusions:** Our results provide new evidence of association between *PPARG* variants and colorectal cancer risk.

**Impact:** Further replication in independent samples is warranted. *Cancer Epidemiol Biomarkers Prev; 22(11); 2094–101. ©2013 AACR.*

## Introduction

Chronic inflammation and innate immunity have been strongly implicated in cancer development since Virchow first proposed a connection (1). The elimination of early neoplastic cells through the cytotoxic activity of tumor-infiltrating T cells or intra-epithelial lymphocytes is now a well-established host-defense mechanism against cancer (2, 3). Normal inflammation is an integral part of the immune response to foreign antigens. It is self-limiting and is initiated by pro-inflammatory cytokines and subsequently resolved by anti-inflammatory cytokines (4). Dysregulation of the two types of molecules may lead to chronic inflammation, a consequence of the inability of tissues to remove the irritant (including the initial tumor cells; refs. 2–6). Chronic inflammation is characterized by the infiltration of damaged tissue by immune cells together with tissue destruction and constant attempts to repair (7). The cytokines and chemokines produced in chronic inflammation and chronic immune activation may affect tumor formation at various stages via induction of reactive oxygen species capable of oxidizing DNA, promoting tumor cell proliferation and survival, stimulating angiogenesis, providing the tools and path for tumor cells (as well as for leucocytes) migration, etc. (1).

Colorectal cancer is a particularly relevant disease to study the links between inflammation and cancer because

**Authors' Affiliations:** [1]University of Hawaii Cancer Center, Honolulu, Hawaii; [2]Translational Genomics Research Institute, Phoenix, Arizona; [3]Department of Preventive Medicine, Keck School of Medicine and Norris Comprehensive Cancer Center, University of Southern California, Los Angeles; [4]Cancer Prevention Institute of California, Fremont, California; [5]Fred Hutchinson Cancer Research Center, Seattle, Washington; [6]Mayo Clinic, Rochester, Minnesota; [7]University of Melbourne, Carlton, Victoria, Australia; [8]Cancer Care Ontario, Toronto, Ontario, Canada; [9]Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, North Carolina; [10]Stanford Cancer Institute, Stanford, California; and [11]National Center for Tumor Diseases and German Cancer Research Center, Heidelberg, Germany

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

**Corresponding Author:** Loïc Le Marchand, Epidemiology Program, University of Hawaii Cancer Center, 701 Ilalo Street, Honolulu, HI 96813. Phone: 808-586-2988; Fax: 808-586-2982; E-mail: loic@cc.hawaii.edu

**doi:** 10.1158/1055-9965.EPI-13-0694

of the importance of the immune system in the maintenance of homeostasis in the normal gut and because of its strong associations with inflammatory bowel disease and nonsteroidal anti-inflammatory drug use (8). In this study, we tested the hypothesis that common variants in 37 key genes involved in the inflammation and innate immunity pathways may result in predisposition to colorectal cancer, by examining associations between disease status and >600 tagSNPs in these genes (as well as 79 additional genetic risk variants with published evidence for association) in two large observational studies in two stages. We conducted both single-variant and pathway-based analysis and took into account single-nucleotide polymorphisms (SNP) correlation and gene correlation. Our approach is more comprehensive than past studies that mostly focused on single genes or single polymorphisms (9–15). The advantage of our approach over genome-wide association studies is that by utilizing prior biologic knowledge, the burden of multiple comparison adjustment is reduced so that the power of identifying genuine but modest effects is enhanced.

## Materials and Methods

### Stage I SNP selection, subjects, and genotypes

The 37 genes are related to tumor infiltrating and cytotoxic activity, pattern recognition receptors, pro- and anti-inflammatory cytokines and chemokines, NF-κB pathway and prostaglandin synthesis (Supplementary Table S1). Due to linkage disequilibrium (LD) structure and proximity between genes, tagSNPs were selected for 34 gene regions (including 5 kb upstream and 10 kb downstream) with $r^2 > 0.8$ based on HapMap or SeattleSNPs data (all in Europeans). TagSNPs were filtered on the basis of minor allele frequency (MAF > 0.05) and Illumina design score (>0.4), except that nonsynonymous/missense variants were forced into the selection. Seventy-nine SNPs in relevant pathways reported to be associated with colorectal cancer were also included, bringing the number of selected SNPs to 768.

The first stage was conducted in the Colorectal Cancer Family Registry (CCFR) using a case-unaffected-sibling-control design. The CCFR is a consortium of six family registries in the United States, Australia, and Canada that consist of patients and their families who represent the continuum of risk for colorectal cancer. Cases, ascertained from population-based cancer registries and family clinics, were probands with a diagnosis of invasive adenocarcinoma of the colon or rectum and controls were siblings with no history of colorectal cancer. Data on important covariates were available from risk factor questionnaires of similar format across all sites. Further information can be found elsewhere (16).

A total of 724 SNPs of the 768 SNPs (94%) were successfully genotyped for 7,086 CCFR subjects (all available sibships in CCFR with at least one colorectal cancer-discordant pair with available DNA) at the Translational Genomics Research Institute (TGen) using the Illumina GoldenGate technology. The average concordance rate was 99.8% among replicates. We used strict quality controls (QC) similar to those in genome-wide association studies (details in Supplementary Data) and included 6,450 subjects (2,355 cases, 4,095 controls) on 616 SNPs for analysis. Five hundred thirty-seven SNPs can be grouped into 34 gene regions to account for long-range LD structure (see Supplementary Table S1 for gene names and locations), excluding the 79 SNPs with previous evidence of association.

### Stage I analysis

Conditional logistic regression was used to test for associations between single SNPs (count of minor alleles) and colorectal cancers, adjusting for age and gender. Known colorectal cancer risk factors (listed in Supplementary Data) did not substantially change OR estimates (<20%) and were not included as covariates. We screened for interactions between SNPs and body mass index (BMI; quartiles), gender, aspirin use (ever/never taken aspirin $\geq$ twice per week for > 1 month) and pack-years of smoking (0, $\leq$20, >20), and for interactions between SNPs from different gene regions (>1 Mb apart) using cross-product terms. Imputation of untyped SNPs was conducted with BEAGLE 3.2 (17) and the Europeans in the 1000 Genomes Project, treating siblings as independent as no imputation program has been designed for sib-ship data. Only SNPs ($n = 2{,}305$) with $R^2 > 0.9$ were analyzed for main effects using expected allelic dosages. We used arbitrary and liberal $P$ value thresholds (see Results) for promoting markers with any main or interaction effects for further testing in stage II. An SNP-set approach based on Kolmogorov–Smirnov-like running sums of ranked single SNP $\chi^2$ statistics (18) was also implemented to examine the collective effect of all SNPs in a gene (see Supplementary Data). Stage I analysis was mainly for discovery; the expectation was that true genetic effects would appear as extreme statistics in some of the performed tests and would be followed up in the next stage. *SAS* (9.2) and *R* 2.10 were used in data analysis.

### Stage II SNP selection, subjects, and genotypes

The second stage was conducted in the Multiethnic Cohort (MEC; ref. 19). The MEC includes 215,000 men and women aged 45 to 75 years at recruitment, primarily from 5 racial/ethnic groups (African Americans, Japanese Americans, Latinos, Native Hawaiians, and whites) in Hawaii and California (ref. 19; more details in Supplementary Data). Identification of incident colorectal cancer cases is through linkage with the Hawaii Tumor Registry, the Los Angeles County Cancer Surveillance Program, and the State of California Cancer Registry. Controls were randomly selected from the pool of MEC cancer-free subjects who provided a blood sample and were frequency-matched to cases on age at cohort entry within each ethnic group. A total of 2,237 colorectal cancer cases and 2,697 controls from the 5 groups with available DNA were genotyped.

For the genetic regions covered by the promising variants from stage I, plus 5 kb up- and 10 kb downstream, tagSNPs were selected to capture common variation (MAF > 0.05) with $r^2$ > 0.9 in the HapMap CEU (Europeans), ASW (African Americans), CHB (Chinese), JPT (Japanese), MKK (Africans), and MEX (Mexicans). A sequential algorithm was used so that tagSNPs selected for a previous population were forced into the subsequent round(s) of selection. We note that genetic variation in Native Hawaiians, which are admixed with Asian, European, and Polynesian ancestries, may be less well covered as no Polynesian population is included in any public datasets. A total of 384 tagSNPs were selected and 364 were successfully genotyped by TGen on the Illumina GoldenGate platform. The average concordance rate was 99.5% among replicates pairs. After QC (details in Supplementary Data), genotypes for 351 SNPs and 4,783 subjects (2,630 controls, 2,153 cases) were used in analysis. Principal components (PC) based on 93 ancestry informative markers (AIM) that separate major continental ancestries (20) were derived, except for 144 subjects for whom AIMs genotyping was considered to have failed.

### Stage II analysis

Logistic regression with adjustment for age, gender, and racial/ethnic groups was conducted to estimate ORs and 95% confidence intervals (CI) for each increase in allelic count. Further adjustment for up to 8 PCs in the subset of data with PCs did not change ORs importantly (change <4%), indicating that adjustment for self-reported race was sufficient to control for potential population stratification; ORs in this subset were also similar to those from the full data set. Adjusting additionally for colorectal cancer risk factors did not change ORs substantially (change <13%). Thus, these covariates were not considered further in main-effect analyses. Effect modification across race/ethnicity, BMI, regular aspirin use (yes/no), and pack-years of smoking (0, ≤20, >20) was tested with a likelihood ratio test (LRT) comparing models with and without the cross-product interaction terms, with further adjustment for BMI, physical activity (average MET hours), total intake of red meat, pack-years of smoking, aspirin use, and daily intake of calcium and folate, where appropriate. Risk factors were grouped on the basis of tertiles unless otherwise specified. Heterogeneity across anatomic subsites (left colon, right colon, and rectum) was tested using multinomial logistic regression in *SAS*. Bonferroni correction was conducted to adjust for multiple comparisons at significance level 0.05 (2-sided) based on the number of independent/effective markers estimated by the "Keffective" program (21): For each chromosome, estimates were obtained for each ethnic group and then the maximum among all groups was taken; the sum of the maximums across all chromosomes was 306, leading to a single-test significance threshold $1.63 \times 10^{-4}$. Imputation was performed with BEAGLE (3.3) and ethnic specific reference panels from the 1000 Genomes Project: ASN for Japanese Americans, EUR for whites, EUR and AFR for African Americans, ASN and EUR for Native Hawaiians, EUR and AMR for Latinos. Only markers ($n = 1{,}541$) with $R^2 > 0.8$ in all 5 groups were kept for analysis. QC and data analysis was performed with PLINK (22) and *R* 2.13 except otherwise noted. All presented *P* values are before Bonferroni correction, unless otherwise specified.

### Combining the two stages

When a SNP was available (i.e., was genotyped or passed imputation quality measures) in both stages, a fixed-effect model was applied to combine ORs. $I^2$ was calculated, where $I^2 > 25\%$ indicates high level of heterogeneity. We used the combined ORs as a supplemental index of the noteworthiness of a risk variant.

### Summary effect of nonsignificant genes in stage II

To explore whether genes with no statistically significant single-variant association contributed to colorectal cancer risk globally, we first selected the most important SNPs (conditional $P < 0.05$) in these genes with stepwise regression, adjusting for sex and ethnicity (missing values were replaced with mean dosage within sex and ethnic groups to preserve sample size). The sum of the selected "risk" alleles (with OR > 1) was then tested for association with disease. The significance of this summary effect was assessed with permutation: disease status was randomly shuffled within ethnicity and sex groups and the above procedures were repeated for each permuted dataset; *P* value was the proportion among 4,000 permutated datasets for which the LRT statistics for the sum of the selected "risk" alleles were greater than that observed in the original data.

## Results

### Stage I

The 2,322 discordant sibships from CCFR ($n = 6{,}450$) were of sizes 2 (54%), 3 (26%), 4 (12%), 5 (5%), and 6 to 13 (3%) and most (83%) were of European descent. The proportion of men was 51.0% in cases and 44.6% in controls. Cases (mean = 55.1, SD = 11.3) and controls (mean = 53.8, SD = 11.8) were similar in age distribution. One thousand five hundred thirty-five sibships included only colon cancer cases, 720 only rectal cancer and 67 both colon and rectal cancers. From available parental information, 2 pairs of sibships (9 subjects) were related as half-sibs, and 44 parents and their siblings (total 124 subjects) were also present. Sensitivity analysis showed that excluding the 133 subjects that complicated the sibship structure did not materially alter the results. Model convergence was questionable for 15 SNPs among 616 analyzed SNPs due to sparse data so results were available for 601 markers. We used the following criteria to promote 99 markers for further examination: (i) $P < 0.1$ for genotyped markers ($n = 49$) or <0.05 for imputed markers ($n = 31$) in the main-effect analysis; (ii) $P < 0.001$ for interactions between SNPs and risk factors ($n = 5$); and (iii) $P < 0.0001$ for interaction tests between SNP pairs from

different gene regions ($n = 14$). These SNPs were distributed on 15 chromosomes (see Supplementary Table S3 for main-effect association results). The SNP-set based analysis did not reveal more genes/regions of interest, as only 2 gene-specific $P$ values were <0.05 ($P = 0.0024$ and 0.01, respectively, for IL10 and PTGIS) and both had been included for follow-up based on single-variant analysis (Supplementary Table S3).

### Stage II

The main characteristics of the 4,783 MEC subjects (2,630 controls, 2,153 cases) are shown in Table 1. There were 480 rectal, 910 right colon, and 565 left colon cancer cases (counts are mutually exclusive). We found one genotyped SNP, rs9858822 in the second intron of *PPARG* on Chr 3, to be statistically significantly associated with colorectal cancer (Bonferroni-adjusted $P = 0.045$, raw $P = 1.48 \times 10^{-4}$; Table 2). The OR for each copy of allele C was 1.36 and the genotypic ORs for AC and CC, versus AA, were 1.24 (95% CI, 0.99–1.56) and 2.03 (95% CI, 1.42–2.89), respectively. A log-additive genetic model was preferable based on the Bayesian Information Criterion, compared to dominant, recessive or co-dominant models. The C allele was common (MAF = 0.35) in African Americans but was relatively rare (MAF $\leq$ 0.03) in the other 4 ethnic/racial groups, consistent with the HapMap data. As a result, the observed association was mainly driven by the African Americans where the OR per C allele was 1.36 ($P = 4.9 \times 10^{-4}$), whereas in the other 4 groups combined, the OR was 1.19 ($P = 0.53$). This effect was not confounded by known colorectal cancer risk factors (further adjustment resulted in <5% change in OR) or differed across ethnicity/race ($P_{\text{interaction}} = 0.99$, perhaps owing to the sparse data in some ethnic groups). Similarly, there did not seem to be important effect modification by pack-years of smoking, BMI, or regular aspirin use (raw $P_{\text{interaction}} \geq 0.08$). For example, the OR (95% CI) for allele C was 1.28 (0.97–1.70) among regular aspirin users ($n = 1{,}544$) and 1.38 (1.08–1.76) among non-users ($n = 2{,}517$), with adjustment of other risk factors. Because rs9858822 was not genotyped or successfully imputed in stage I (as expected, given its low frequency in Europeans), no OR for the combined studies was available.

The *PPARG* gene maps to 12,329,348 to 12,475,854 at 3p25.2 and SNPs in our data were located from 12,330,730 to 12,479,552. LD structure for *PPARG* is complex and LD blocks are difficult to define in the 1000 Genomes data (Supplementary Fig. S1–S3 for LD in Europeans, Africans, and East Asians in the 1000 Genomes Project). A few other SNPs located nearby in the *PPARG* region showed suggestive associations in stage II (Fig. 1), even though none passed the stringent Bonferroni adjustment. For example, the raw $P$ value for rs4135304 (12,394,601 bp) and rs6778740 (12,398,636 bp) was $2.9 \times 10^{-4}$ and $2.5 \times 10^{-4}$, respectively, and these SNPs were in high LD ($r^2 = 0.78$ and 0.93) with rs9858822 in African Americans. Logistic regression of all variants in *PPARG* conditional on rs9858822 did not reveal better candidate markers.

Rs9858822 lies in close proximity to a H3K4me1 ChIP-seq peak in the colon cancer cell line HCT-116 (ENCODE, Stanford/Yale/USC/Harvard; ref. 23). This histone modification is a mark of a putative enhancer element. A few SNPs in high LD ($r^2 \geq 0.69$, $|D'| = 1$) with rs9858822 in the HapMap ASW (African Americans) are located

**Table 1.** Main characteristics of 4,783 subjects by ethnicity in the MEC

| | Ethnicity/race | | | | |
|---|---|---|---|---|---|
| | AA | NH | JA | LA | EA |
| *N* (% male) | | | | | |
| Cases | 443 (45.2) | 114 (58.8) | 725 (58.9) | 489 (59.5) | 382 (51.1) |
| Controls | 824 (72.7) | 177 (40.7) | 799 (46.6) | 414 (60.6) | 416 (52.4) |
| Mean age (SD) | | | | | |
| Cases | 63.5 (8.1) | 58.8 (8.2) | 62.6 (8.4) | 62.3 (7.2) | 62.6 (8.2) |
| Controls | 62.9 (7.6) | 56.6 (6.9) | 60.3 (8.3) | 60.5 (6.6) | 60.0 (8.0) |
| Mean BMI (SD) | | | | | |
| Cases | 28.7 (5.0) | 30.0 (5.2) | 25.2 (4.0) | 27.9 (4.5) | 26.8 (4.6) |
| Controls | 27.8 (4.8) | 28.9 (5.5) | 24.6 (3.5) | 27.6 (4.3) | 26.2 (4.9) |
| Pack-years of smoking[a] | | | | | |
| Cases | 3.9 (14.2) | 7.1 (27.5) | 2.9 (19.8) | 1.3 (12.0) | 8.3 (31.8) |
| Controls | 6.2 (19.8) | 6.2 (19.8) | 0 (12.0) | 1.3 (10.3) | 2.0 (19.8) |
| % Aspirin use[b] | | | | | |
| Cases | 44.6 | 44.6 | 27.8 | 43.9 | 48.1 |
| Controls | 47.4 | 27.5 | 25.8 | 41.2 | 49.6 |

Abbreviations: AA, African Americans; EA, European Americans; JA, Japanese Americans; LA, Latinos; NH, Native Hawaiians.

[a]Data are medians (upper quartiles). Lower quartiles are all zero.

[b]Ever used aspirin at least twice per week consecutively for more than a month.

**Table 2.** Association between colorectal cancer and rs9858822 in the MEC

| SNP | Chr | BP (HG19) | Minor/ major | Group | Cases/ controls | MAF (controls) | MAF (cases) | OR[a] (95% CI) | $P^b$ |
|---|---|---|---|---|---|---|---|---|---|
| rs9858822 | 3 | 12411238 | C/A | All | 2152/2630 | | | 1.36 (1.16–1.60) | $1.5 \times 10^{-4}$ |
| | | | | LA | 489/414 | 0.027 | 0.034 | 1.26 (0.72–2.20) | 0.42 |
| | | | | JA | 725/799 | 0 | 0 | na | na |
| | | | | AA | 443/824 | 0.31 | 0.39 | 1.36 (1.15–1.62) | 0.0005 |
| | | | | EA | 381/416 | 0.0036 | 0.0013 | 0.31 (0.03–3.05) | 0.32 |
| | | | | NH | 114/177 | 0.0028 | 0.0044 | 2.65 (0.16–44.7) | 0.50 |

Abbreviations: AA, African Americans; EA, European Americans; JA, Japanese Americans; LA, Latinos; NH, Native Hawaiians.
[a]OR for the minor allele.
[b]*P* values were from log-additive models, adjusted for age, gender, and racial/ethnic groups. No heterogeneity across ethnic groups was detected ($P_{het} > 0.3$).

within the putative enhancer elements, such as rs9812856 (12,404,491), rs4135343 (12,447,964), and rs4135346 (12,448,498). Notably, rs4135343 is found at the 3′ end of a shorter *PPARG* splice variant, in both H3K4me1 and H3K27ac ChIP-seq peaks, and directly downstream of many transcription factor ChIP-seq signals (ENCODE, HAIB). These SNPs were imputed with $R^2 > 0.91$ only in African Americans in stage II and were not analyzed in other ethnic groups or in stage I. The *P* values of association were 0.0002, 0.0014, and 0.0015, respectively, for the three variants in African Americans.

Rs1801282 (Pro12Ala, C/G) in *PPARG* has been inconsistently associated with colorectal cancer in past studies. In our data, it was genotyped in stage I and was imputed in stage II with $R^2 \geq 0.8$ in all ethnic groups. However, the SNP was not associated with colorectal cancer in either stage, or when the 2 studies were combined (all $P > 0.4$, Supplementary Table S5). In stage II, no important association was observed in subsite-specific analyses (colon or rectum), or when adjusted for BMI and diabetes status, or when stratified by ethnicity (all $P > 0.05$, Supplementary Table S5).

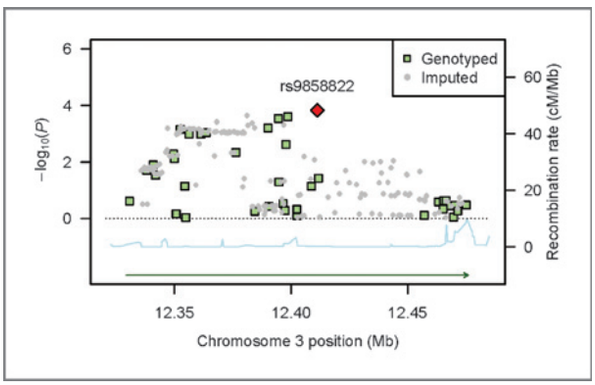


Figure 1. *P* values in the MEC for genotyped and imputed markers in the *PPARG* region as a function of genetic position (HG19), along with recombination rates (combined rates from HapMap) that reflect local LD structure. The top hit SNP is shown as a red diamond. SNP density is not ideal at some regions because only SNPs with high imputation accuracy ($R^2 > 0.8$) in ALL 5 ethnic groups were analyzed and $R^2$ can vary a lot due to differences in allele frequencies and LD patterns across populations.

Overall, we did not detect important interaction effects between SNPs and ethnicity or known risk factors (BMI, pack-years of smoking, or aspirin use) after multiple comparison adjustment ($P_{interaction} \geq 0.0003$). Similarly, no strong heterogeneity by anatomical location of the tumor was observed ($P_{heterogeneity} > 0.005$). All 31 SNPs with $P < 0.05$ in stage II main-effect analysis are shown in Supplementary Materials. Results from imputed SNPs were similar to those from genotyped ones.

### Summary effect of nonsignificant genes in stage II

Among the 313 SNPs in genes other than *PPARG* where no variant was statistically significantly associated with colorectal cancer, stepwise regression selected 12 SNPs with conditional $P < 0.05$. The sum of the selected "risk" alleles (with OR > 1) seemed associated with colorectal cancer ($\chi^2 = 63.3$, $P = 1.8 \times 10^{-15}$). However, this effect was statistically nonsignificant from permutation analysis ($P = 0.26$), when the so-called "winner's curse" was corrected for.

## Discussion

Using a discovery and a validation step, both conducted in large and well-annotated studies, we examined associations between colorectal cancer and common variation in 37 key genes involved in the inflammation and immunity pathways. At the end of the validation step, only one SNP in the second intron of *PPARG*, rs9858822 (12,411,238 bp on Chr 3), showed statistically significant association after stringent Bonferroni-type multiple comparison adjustment. This SNP is common (MAF = 0.3) in African Americans, a population at high risk for colorectal cancer, but rare in Europeans, Hispanics, East Asians, and Native Hawaiians (MAF < 0.03).

*PPARG* encodes a nuclear transcription factor, *PPARγ*, that regulates target gene expression upon ligand activation and through interaction with transcription co-factors. Natural ligands for *PPARγ* include fatty acids and their

derivatives. *PPARγ* has a predominant role in lipid metabolism and adipocyte differentiation, as well as in maintaining glucose homeostasis and insulin sensitivity, and in inflammation. PPARγ proteins are expressed in a broad range of tissues including heart, skeletal muscle, small and large intestines, kidney, etc., depending on the isoforms (24–26). Some links have been demonstrated between PPARγ and colorectal tumor growth or differentiation, although the exact mechanism remains unclear (27–31). PPARγ and its ligands have been shown to block pro-inflammatory genes in activated macrophages, monocytes, and colon cancer cell lines by inhibiting the NF-κB (a master regulator of inflammatory processes), AP-1 and STAT pathways (32–34). There is growing evidence that obesity/type 2 diabetes, inflammation, and cancer (in particular colorectal cancers) are etiologically related (35, 36). Being at the crossroads of all these processes, *PPARγ* may be a key element to understanding the pathophysiology of colorectal cancer.

The association of rs9858822 with colorectal cancer has not been reported before, probably owing to its low frequency in non-African descent populations. Because of the extended and complex LD structure in the *PPARG* region, we view this signal as an indication that the neighboring *PPARG* region may contain causal risk variants for colorectal cancer, possibly representing an ethnic-specific susceptibility locus in African Americans only. This signal is unlikely a chance finding given the supporting evidence (small *P* values) from neighboring SNPs (Fig. 1).

The Pro12Ala polymorphism (rs1801282), the most studied *PPARG* variant to date, has been consistently associated with type 2 diabetes (37, 38). It has also been linked to colorectal cancer risk, but inconsistently so. For example, a reduced risk was reported in a Spanish (12), a Greek (39), and a Singapore-Chinese study (11); however, no effect was observed in several other studies in the United States or among West Asians (13, 15, 40–42). Owing to the various designs and study populations, it is hard to reconcile these past studies. As discussed before, technical artifact during genotyping and the presence of multiple genetic ancestries within individual studies (population stratification; ref. 43) may not have been carefully controlled in earlier candidate gene studies. We did not observe any statistically significant association between Pro12Ala and colorectal cancer in our data, consistent with a recent systematic review of colorectal cancer and genetic polymorphisms (44) that examined 15 SNPs in the inflammation pathways, including 2 in *PPARG*, rs3856806 (C1431T) and rs1801282 (Pro12Ala). With more than 15,000 cases and 15,000 controls, rs1801282 was not associated with colorectal cancer ($P > 0.35$ for both heterozygous and homozygous genotypes; ref. 44); neither was rs3856806.

The strengths of our study include a comprehensive SNP tagging approach to capture common genetic variation in key genes in the inflammation and immunity pathways, rigorous genotype quality control procedures, large sample sizes in both stages, a replication in a multiethnic population where ethnic-specific risk can be examined, the ability to control for a variety of environmental risk factors and to examine gene–environment interactions, and a strict multiple comparison adjustment approach. The value of conducting association studies in multiethnic populations is evident from our result for rs9858822. Given its low frequency in European-descent populations, the association with colorectal cancer (if confirmed) could not have been easily detected if not tested in African Americans. The use of Bonferroni correction based on the number of independent/effective markers estimated from Keffective (21) has been shown to be conservative (45). This was especially so in our analysis because we used the maximum estimate across 5 ethnic groups as the number of independent/effective markers for each chromosome.

The limitations of our study include treating sibs in stage I (in CCFR) as independent subjects during imputation. To reduce the possibility of bias caused by this unmet assumption, we restricted analysis to those markers imputed with a high $R^2$ (>0.9), which eliminated many imputed markers from stage I and prohibited us from combining the results from both studies for these SNPs. However, this issue is not relevant to our findings with rs9858822, because this variant is rare in whites and could not have been reliably imputed even if the discovery phase had been conducted in unrelated subjects. Other limitations included that some interactions may have been missed due to insufficient power and/or model misspecification.

In summary, our findings from 2 large studies with different designs and populations suggest that a new region in *PPARG* tagged by rs9858822 may constitute a susceptibility locus for colorectal cancer. Replication in independent samples, especially in African Americans, is warranted. In addition, a summary effect for the other studied genes (excluding *PPARG*) in the inflammation and innate immunity pathways was not important, indicating little effects of common variants in these genes with respect to colorectal cancer risk.

## Disclosure of Potential Conflicts of Interest

C.M. Ulrich is a consultant/advisory board member of Bayer. No potential conflicts of interest were disclosed by the other authors.

## Disclaimer

The sponsoring agency played no role in study design and in collection, analysis, and interpretation of data.

## Authors' Contributions

**Conception and design:** H. Wang, D. Taverna, D.O. Stram, L.R. Wilkens, J.L. Hopper, C.M. Ulrich, L. Le Marchand

**Development of methodology:** H. Wang, D. O. Stram, J.L. Hopper, D. Duggan, C.M. Ulrich, L. Le Marchand

**Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.):** T. Burnett, K.W. Makar, N.M. Lindor, J.L. Hopper, S. Gallinger, R.W. Haile, L.N. Kolonel, P.A. Newcomb, G. Casey, D. Duggan, L. Le Marchand

**Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis):** H. Wang, D. Taverna, B. K. Fortini, I. Cheng, N.M. Lindor, J.L. Hopper, C.M. Ulrich, L. Le Marchand

**Writing, review, and/or revision of the manuscript:** H. Wang, D. Taverna, D. O. Stram, I. Cheng, L.R. Wilkens, K.W. Makar, N.M. Lindor, J.L. Hopper, J.A. Baron, R.W. Haile, L.N. Kolonel, B.E. Henderson, P.A. Newcomb, G. Casey, D. Duggan, C.M. Ulrich, L. Le Marchand
**Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases):** H. Wang, J.L. Hopper, S. Gallinger, L.N. Kolonel, L. Le Marchand
**Study supervision:** T. Burnett, L. Le Marchand

## Grant Support

This study was funded by the NIH grant R01 CA129063. The MEC study is funded by the NIH grants UM1 CA164973, R37 CA54281, P01 CA33619, R01 CA63464, U01 CA98758, and U01 CA136792. The CCFR study is funded by the NIH grant U01CA097735 (Australasian Colorectal Cancer Family Registry), U01 CA074799 (Familial Colorectal Neoplasia Collaborative Group [USC]), U01 CA074800 (Mayo Clinic Cooperative Family Registry for Colon Cancer Studies), U01 CA074783 (Ontario Registry for Studies of Familial Colorectal Cancer), U01 CA074794 (Seattle Colorectal Cancer Family Registry), and U01 CA074806 (University of Hawaii Colorectal Cancer Family Registry).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received July 10, 2013; revised September 3, 2013; accepted September 4, 2013; published OnlineFirst September 17, 2013.

## References

1. Balkwill F, Mantovani A. Inflammation and cancer: back to Virchow? Lancet 2001;357:539–45.
2. Imai K, Matsuyama S, Miyake S, Suga K, Nakachi K. Natural cytotoxic activity of peripheral-blood lymphocytes and cancer incidence: an 11-year follow-up study of a general population. Lancet 2000;356:1795–9.
3. Shankaran V, Ikeda H, Bruce AT, White JM, Swanson PE, Old LJ, et al. IFNgamma and lymphocytes prevent primary tumour development and shape tumour immunogenicity. Nature 2001;410:1107–11.
4. Dvorak HF. Tumors: wounds that do not heal. Similarities between tumor stroma generation and wound healing. N Engl J Med 1986; 315:1650–9.
5. Burnet FM. The concept of immunological surveillance. Prog Exp Tumor Res 1970;13:1–27.
6. Herberman RB, Ortaldo JR. Natural killer cells: their roles in defenses against disease. Science 1981;214:24–30.
7. Macarthur M, Hold GL, El-Omar EM. Inflammation and Cancer II. Role of chronic inflammation and cytokine gene polymorphisms in the pathogenesis of gastrointestinal malignancy. Am J Physiol Gastrointest Liver Physiol 2004;286:G515–20.
8. Ulrich CM, Bigler J, Potter JD. Non-steroidal anti-inflammatory drugs for cancer prevention: promise, perils and pharmacogenetics. Nat Rev Cancer 2006;6:130–40.
9. Cacev T, Radosevic S, Krizanac S, Kapitanovic S. Influence of interleukin-8 and interleukin-10 on sporadic colon cancer development and progression. Carcinogenesis 2008;29:1572–80.
10. Castro FA, Forsti A, Buch S, Kalthoff H, Krauss C, Bauer M, et al. TLR-3 polymorphism is an independent prognostic marker for stage II colorectal cancer. Eur J Cancer 2011;47:1203–10.
11. Koh WP, Yuan JM, Van Den Berg D, Ingles SA, Yu MC. Peroxisome proliferator-activated receptor (PPAR) gamma gene polymorphisms and colorectal cancer risk among Chinese in Singapore. Carcinogenesis 2006;27:1797–802.
12. Landi S, Moreno V, Gioia-Patricola L, Guino E, Navarro M, de Oca J, et al. Association of common polymorphisms in inflammatory genes interleukin (IL)6, IL8, tumor necrosis factor alpha, NFKB1, and peroxisome proliferator-activated receptor gamma with colorectal cancer. Cancer Res 2003;63:3560–6.
13. Slattery ML, Murtaugh MA, Sweeney C, Ma KN, Potter JD, Caan BJ, et al. PPARgamma, energy balance, and associations with colon and rectal cancer. Nutr Cancer 2005;51:155–61.
14. Suchy J, Klujszo-Grabowska E, Kladny J, Cybulski C, Wokolorczyk D, Szymanska-Pasternak J, et al. Inflammatory response gene polymorphisms and their relationship with colorectal cancer risk. BMC Cancer 2008;8:112.
15. Tsilidis KK, Helzlsouer KJ, Smith MW, Grinberg V, Hoffman-Bolton J, Clipp SL, et al. Association of common polymorphisms in IL10, and in other genes related to inflammatory response and obesity with colorectal cancer. Cancer Causes Control 2009;20:1739–51.
16. Newcomb PA, Baron J, Cotterchio M, Gallinger S, Grove J, Haile R, et al. Colon Cancer Family Registry: an international resource for studies of the genetic epidemiology of colon cancer. Cancer Epidemiol Biomarkers Prev 2007;16:2331–43.
17. Browning SR, Browning BL. Rapid and accurate haplotype phasing and missing data inference for whole genome association studies using localized haplotype clustering. Am J Hum Genet 2007;81:1084–97.
18. Wang K, Li M, Bucan M. Pathway-based approaches for analysis of genomewide association studies. Am J Hum Genet 2007;81: 1278–83.
19. Kolonel LN, Henderson BE, Hankin JH, Nomura AM, Wilkens LR, Pike MC, et al. A multiethnic cohort in Hawaii and Los Angeles: baseline characteristics. Am J Epidemiol 2000;151:346–57.
20. Kosoy R, Nassir R, Tian C, White PA, Butler LM, Silva G, et al. Ancestry informative marker sets for determining continental origin and admixture proportions in common populations in America. Hum Mutat 2009;30:69–78.
21. Moskvina V, Schmidt KM. On multiple-testing correction in genome-wide association studies. Genet Epidemiol 2008;32:567–73.
22. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81:559–75.
23. The ENCODE Consortium Project. A user's guide to the encyclopedia of DNA elements (ENCODE). PLoS Biol 2011;9:e1001046.
24. Berger J, Moller DE. The mechanisms of action of PPARs. Ann Rev Med 2002;53:409–35.
25. Fajas L, Auboeuf D, Raspe E, Schoonjans K, Lefebvre A-M, Saladin R, et al. The organization, promoter analysis, and expression of the human PPARgamma gene. J Biol Chem 1997;272:18779–89.
26. Fajas L, Fruchart J-C, Auwerx J. PPARgamma3 mRNA: a distinct PPARgamma mRNA subtype transcribed from an independent promoter. FEBS Lett 1998;438:55–60.
27. DuBois RN, Gupta R, Brockman J, Reddy BS, Krakow SL, Lazar MA. The nuclear eicosanoid receptor, PPARgamma, is aberrantly expressed in colonic cancers. Carcinogenesis 1998;19:49–53.
28. Sarraf P, Mueller E, Jones D, King FJ, DeAngelo DJ, Partridge JB, et al. Differentiation and reversal of malignant changes in colon cancer through PPARgamma. Nat Med 1998;4:1046–52.
29. Sarraf P, Mueller E, Smith WM, Wright HM, Kum JB, Aaltonen LA, et al. Loss-of-function mutations in PPAR gamma associated with human colon cancer. Mol Cell 1999;3:799–804.
30. Saez E, Tontonoz P, Nelson MC, Alvarez JG, Ming UT, Baird SM, et al. Activators of the nuclear receptor PPARgamma enhance colon polyp formation. Nat Med 1998;4:1058–61.
31. Lefebvre AM, Chen I, Desreumaux P, Najib J, Fruchart JC, Geboes K, et al. Activation of the peroxisome proliferator-activated receptor gamma promotes the development of colon tumors in C57BL/6J-APCMin/+mice. Nat Med 1998;4:1053–7.
32. Jiang C, Ting AT, Seed B. PPAR-[gamma] agonists inhibit production of monocyte inflammatory cytokines. Nature 1998;391:82–6.
33. Ricote M, Li AC, Willson TM, Kelly CJ, Glass CK. The peroxisome proliferator-activated receptor-gamma is a negative regulator of macrophage activation. Nature 1998;391:79–82.
34. Su CG, Wen X, Bailey ST, Jiang W, Rangwala SM, Keilbaugh SA, et al. A novel therapy for colitis utilizing PPAR-gamma ligands to inhibit the epithelial inflammatory response. J Clin Invest 1999;104:383–9.

35. Brower V. Illuminating the diabetes-cancer link. J Natl Cancer Inst 2012;104:1048–50.
36. Wang D, Dubois RN. Associations between obesity and cancer: the role of fatty acid synthase. J Natl Cancer Inst 2012;104:343–5.
37. Saxena R, Voight BF, Lyssenko V, Burtt NP, de Bakker PI, Chen H, et al. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. Science 2007;316:1331–6.
38. Altshuler D, Hirschhorn JN, Klannemark M, Lindgren CM, Vohl M-C, Nemesh J, et al. The common PPAR[gamma] Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes. Nat Genet 2000;26:76–80.
39. Theodoropoulos G, Papaconstantinou I, Felekouras E, Nikiteas N, Karakitsos P, Panoussopoulos D, et al. Relation between common polymorphisms in genes related to inflammatory response and colorectal cancer. World J Gastroenterol 2006;12:5037–43.
40. Abuli A, Fernandez-Rozadilla C, Alonso-Espinaco V, Munoz J, Gonzalo V, Bessa X, et al. Case-control study for colorectal cancer genetic susceptibility in EPICOLON: previously identified variants and mucins. BMC Cancer 2011;11:339.
41. Jiang J, Gajalakshmi V, Wang J, Kuriki K, Suzuki S, Nakamura S, et al. Influence of the C161T but not Pro12Ala polymorphism in the peroxisome proliferator-activated receptor-gamma on colorectal cancer in an Indian population. Cancer Sci 2005;96:507–12.
42. Murtaugh MA, Ma KN, Caan BJ, Sweeney C, Wolff R, Samowitz WS, et al. Interactions of peroxisome proliferator-activated receptor {gamma} and diet in etiology of colorectal cancer. Cancer Epidemiol Biomarkers Prev 2005;14:1224–9.
43. Hirschhorn JN, Gajdos ZK. Genome-wide association studies: results from the first few years and potential implications for clinical medicine. Annu Rev Med 2011;62:11–24.
44. Theodoratou E, Montazeri Z, Hawken S, Allum GC, Gong J, Tait V, et al. Systematic meta-analyses and field synopsis of genetic association studies in colorectal cancer. J Natl Cancer Inst 2012;104:1433–57.
45. Gao X, Becker LC, Becker DM, Starmer JD, Province MA. Avoiding the high Bonferroni penalty in genome-wide association studies. Genet Epidemiol 2010;34:100–5.