Variation in folate pathway genes and distal colorectal adenoma risk: a sigmoidoscopy-based case–control study

A. Joan Levine · Won Lee · Jane C. Figueiredo · David V. Conti · David J. Vandenberg · Brian D. Davis · Christopher K. Edlund · Susanne M. Henning · David Heber · Mariana C. Stern · Robert W. Haile

Received: 4 March 2010 / Accepted: 6 January 2011 / Published online: 28 January 2011 © The Author(s) 2011. This article is published with open access at Springerlink.com

Abstract

Background Folate-associated one-carbon metabolism (FOCM) is an important pathway in colorectal neoplasia risk but data on genetic variation in this pathway are largely limited to studies of single SNPs in selected genes.

Methods We used a comprehensive tagSNP approach to study the association between genetic variation in 11 genes in the FOCM pathway and risk of incident distal colorectal adenomas in a sigmoidoscopy-based case–control study. We included 655 cases (one or more adenomas) and 695 controls (no adenomas) recruited from one of two Kaiser Permanente clinics between 1991 and 1995. We assessed a total of 159 tagSNPs selected using Haploview Tagger as well as selected non-synonymous SNPs. We used unconditional logistic regression to model the association between SNPs and risk of distal adenomas, assuming a log-additive model.

Results Five SNPs in the SLC19A1 (RFC1) gene: rs1051266 (G80A), rs283895, rs2236484, rs12482346, and rs2838958 were associated with adenoma risk after correction for multiple testing (all corrected p values ≤0.043). The non-synonymous SLC19A1 SNP G80A interacted significantly with the MTHFR C677T genotype (interaction p value = 0.018).

Conclusion Our data suggest that genetic variation in SLC19A1 may modify the risk of distal colorectal adenoma.

Keywords Folate · MTHFR · SLC19A1 · TagSNP · Folate · Colorectal adenoma

Introduction

Folate-associated one-carbon groups are essential for hundreds of intracellular transmethylation reactions including those involved in DNA methylation and DNA synthesis [1]. An inverse association between folate intake and carcinogenic changes in colorectal epithelium has been observed, both in vitro and in vivo, in both humans and animals [1–3]. However, clinical trials of folic acid supplementation [4] and animal experiments [5, 6] suggest that folate plays a dual role in the colorectum by both protecting against and promoting growth of neoplastic lesions depending on timing, dose, and source (diet vs. supplements).

Folate-associated one-carbon groups are a complex cycle of inter-related reactions that provide one-carbon groups needed for numerous intracellular processes. This pathway has been well characterized, and a mathematical model has been developed [7, 8]. Genetic variants in genes that play key roles in the FOCM pathway have been investigated as potential colorectal adenoma susceptibility genes, but studies have been limited to one or a few SNPs and a small number of genes, with mixed results [9, 10]. The MTHFR C677T and A1298C polymorphisms

Electronic supplementary material The online version of this article (doi:10.1007/s10552-011-9726-7) contains supplementary material, which is available to authorized users.

A. J. Levine (✉) · W. Lee · J. C. Figueiredo · D. V. Conti · D. J. Vandenberg · B. D. Davis · C. K. Edlund · M. C. Stern · R. W. Haile
Department of Preventive Medicine, Genetic Epidemiology, University of Southern California Keck School of Medicine, NRT 1450 Biggy Street Room 1509A, Los Angeles, CA 90033, USA
E-mail: ajoanlev@usc.edu

S. M. Henning · D. Heber
Center for Human Nutrition, David Geffen School of Medicine, University of California, Los Angeles, CA 90095, USA
have been associated with CRC but not colorectal adenoma risk [11, 12]. A recent study of genetic variability in FOCM-related genes and adenoma risk focused on 24 non-synonymous SNPs in 13 genes, reported little evidence of a major role for the folate pathway genes included in their analysis [13].

In this study, we conducted a comprehensive analysis of the role of genetic variation in 11 genes that play key roles in the FOCM pathway and colorectal adenoma risk. Figure 1 defines the genes included in the current analysis and their roles in FOCM. We included enzymes involved in the uptake of folate, nucleotide synthesis, and S-adenosylmethionine (SAM) synthesis. We also included the gene for cystathionine-β-synthase (CBS), an enzyme important for modulating intracellular homocysteine and that for gastric intrinsic factor (GIF), which is involved in the uptake of vitamin B₁₂, a key co-factor in FOCM.

Materials and methods

Study subjects

The subjects in this study were participants in the USC/Kaiser Permanente study of risk factors for colorectal adenomas. Characteristics of this study population have been previously described [14, 15]. Briefly, phase 1 subjects were recruited from one of two Kaiser Permanente clinics and received a sigmoidoscopy examination from 1991 to 1993, while phase 2 subjects were recruited from the same two clinics from 1993 to 1995. Identical criteria were used to recruit subjects from each phase. Eligible subjects were English-speaking, between the ages of 50–74 years, and living in the Los Angeles metropolitan area. Subjects were excluded if they had a history of invasive cancer, inflammatory bowel disease, familial polyposis, previous bowel surgery, or symptoms suggestive of gastrointestinal disease. Cases were those subjects who had at least one histologically confirmed adenoma during their sigmoidoscopy exam, and controls were those subjects with no evidence of an adenoma at sigmoidoscopy and who had no history of confirmed adenomas. Cases were individually matched to controls on age (within 5 years), gender, sigmoidoscopy date (within 3 months), and Kaiser Permanente clinic. All subjects signed an approved informed consent, donated a blood sample, and filled in two questionnaires.

Risk factor data

The risk factor questionnaire queried demographic information, family cancer history, smoking history, history of...
using selected over the counter and prescription drugs (e.g., aspirin and other non-steroidal anti-inflammatory agents, laxatives), specified dietary supplements (multivitamins, calcium), a physical activity history, a history of usual sun exposure, a brief description of their usual method for cooking red meats, and, for women, a brief reproductive history. A separate food frequency questionnaire (FFQ) was used to estimate nutrient intake. For phase 1 subjects, we used a modified form of the Block food frequency questionnaire [16]. For phase 2 subjects, we used a diet questionnaire developed and validated at the University of Hawaii [17]. The two questionnaires were similar with respect to the number and type of foods queried and the method for estimating nutrient intakes, in grams, milligrams, or micrograms per day, as appropriate for the nutrient. Nutrient intakes were based on expected nutrient levels for the stated portion size and estimated consumption frequency per week for specific foods and supplements using the nutritionist food and supplement databases developed by the designers of the FFQs. Dietary folate intake was categorized into quartiles using study phase-specific cutpoints and intakes from the controls. For alcohol, we used predefined cutpoints of 0, 0.1–10, 11–20, and >20 g/day.

Plasma folate

Blood samples were taken from fasting subjects into EDTA-coated tubes in the morning and immediately put on ice until processing to separate the plasma. In both study phases, samples were stored at −70°C until processing. For phase 1 subjects, plasma folate was determined for the first 370 male subjects (52.3%) and the first 316 samples female subjects (86.8%) using the Quantaphase Radioassay as described [18]. For phase 2 subjects, plasma folate was estimated for 335 men (97.7%) and 199 women (97.5%) using the Quantaphase Radioassay II.

Plasma total homocysteine (tHcy)

Plasma tHcy was not estimated for phase 1 subjects. For phase 2 subjects plasma tHcy (μmol/L) was estimated for 335 men (97.7%) and 199 women (97.5%). Blood samples were taken from fasting subjects into EDTA-coated tubes in the morning and immediately put on ice until processing to separate the plasma. In both study phases, samples were stored at −70°C until processing. For determination of plasma tHcy, we used the reversed-phase HPLC method of Kuo et al. [19]. All chemicals were purchased from Sigma. Plasma thiols were measured using fluorescence detection after derivatization with ammonium 7-fluorobenzo-2-oxa-1,3-diazole-4-sulfonate (SBDF). HPLC separation was carried out using gradient analysis with fluorescence detection (Varian Assoc., Sugar Land Texas) and the Agilent Technology HPLC 1100 system (Agilent Technology, Wilmington DE). Thiols were separated using a Bakerbond C18 4.6 × 250 mm column (Mallinkrodt Baker, Phillipsburg, NJ). The mobile phase consisted of 8% methanol to 40% at time 12 min. For quality control purposes, pooled plasma samples were analyzed with each batch of samples. The intra-assay coefficient of variance was 5.7% and inter-assay CV was 4.4% (8.6 μM).

SNP selection

We used a comprehensive tagSNP approach to assess gene-level risk in a pre-folic acid-fortification population of colorectal adenoma patients and controls. Supplemental Table 1 describes the SNPs included in this analysis by race/ethnic group. SNPs were selected using HaplovTagger [20] and based on the CEPH data using the following criteria: MAF ≥ 5%, pairwise r² ≥ 0.95, and distance from closest SNP greater than 60 base pairs on the Illumina platform. The linkage disequilibrium blocks were determined using data from HapMap data release #16c.1, June 2005, on NCBI B34 assembly, dbSNP b124. For each gene, we extended the 5' and 3' UTR regions to include the 5’- and 3’-most SNP within the LD block (approximately 10 kb upstream and 5 kb downstream). In regions of no- or low-LD, SNPs with an MAF ≥ 5% at a density of approximately 1 per kb were selected from either HapMap or dbSNP. Finally, non-synonymous and expert-curated SNPs regardless of MAF were included. The SNPs included in this study are listed in Supplementary Table 1.

SNP genotyping

Except for the MTHFR C677T polymorphism (described later), SNPs were genotyped on the Illumina GoldenGate platform [21]. We implemented a series of quality control checks based on Illumina metrics, and SNPs were excluded from analysis based on the following criteria: GenTrain score < 0.4, 10% GC score < 0.25, AB T Dev > 0.1239, call rate < 0.95, more than 2 Mendelian errors > 2 or discordance with HapMap > 3. Inter- and intraplate replicates were included, and SNPs were excluded from the analysis if there were greater than 2 errors on replicate genotypes. In addition, genotype data from 30 CEPH trios (Coriell Cell Repository, Camden, NJ) were used to confirm reliability and reproducibility of the genotyping. SNPs were excluded from the analysis if more than 3 discordant genotypes were discovered in comparison with genotypes from the International HapMap Project [22]. After deleting SNPs with a minor allele frequency < 0.05 in the total study sample and those with a p value for Hardy–Weinberg equilibrium (HWE) < 0.0003 (the Bonferroni-corrected p value), we calculated effect estimates for 159 SNPs.
**MTFHR C677T genotype**

We have previously reported on associations between folate and the MTHFR C677T genotype on adenoma risk in the phase 1 subjects [23]. For phase 1 subjects, the MTHFR C677T genotype (rs1801133) was determined by the PCR–RFLP method of Frostell et al. [24] using their published primer pairs. MTHFR 677 genotyping was not available for 14 phase 1 subjects and 485 phase 2 subjects. For these subjects, the 677 genotype was determined by imputation using the MACH software [25]. Accuracy was 97.4% for non-Hispanic whites, 99.8% for Hispanics, 93.4% for African Americans, and 94.5% for Asians.

**Statistical analysis**

Minor allele frequency was estimated from genotype data from cases and controls combined. Hardy–Weinberg equilibrium was assessed using a χ² test. Pairwise linkage disequilibrium between SNPs was estimated using Haploview [20]. We used the square of the correlation coefficient (r²) between markers to define linkage using the data from the study population. We have previously established that using unconditional logistic regression adjusting for the matching factors led to the same results as conditional logistic regression while allowing use of the entire study population [23]. Therefore, we used unconditional logistic regression controlling for the matching factors (age, sex, clinic and examination date), study phase and, in non-ethnic group-specific analyses, race/ethnic group to estimate main effects and stratum-specific odds ratios. Additional control for dietary folate (mcg/ml), total dietary fiber, multivitamin use (yes/no), and alcohol intake (0, >0–10, 11–20 and >20 g/day) did not change the results, and we present only the matching factor, race/ethnicity (where appropriate), and phase-adjusted models here.

Except for the MTHFR C677T (rs1801133) and MTHFR A1298C (rs1801131) polymorphisms, for which there are data supporting a recessive model, we assumed a log-additive model to assess genotype/adenoma associations. To minimize population stratification, we conducted all analyses separately by race/ethnic group. Since prior data suggest specific effects for MTHFR C677T and A1298C, p values for these two SNPs were not corrected for multiple testing in any analysis. Additionally, tests based on prior data from the literature (e.g., the interaction between SLC19A1 G80A genotype and MTHFR C677T genotype) were not corrected for multiple testing. For the log-additive model and within each gene, p values for all other SNPs were adjusted for multiple testing taking into account correlated tagSNPs using a modified test of Conneely and Boehnke (p_{act}) [26]. The p_{act} also adjusts for the correlation structure of each gene within each ethnic group, using the race/ethnic group-specific data to compute the underlying correlation structure specific to that group. For all stratified analyses, we included interaction terms in the regression models to get the interaction p values. p Values for trend were estimated by entering the quartile value (1–4) into the regression model. The likelihood ratio test was used to test for heterogeneity between strata. For these multiple degrees of freedom likelihood ratio tests, the Bonferroni method was used to reset the significance level to 0.0003, based on a total of 159 SNPs.

We assessed the association between genotype and plasma tHcy using a multiple regression where plasma tHcy was modeled as a continuous variable, and the change in plasma tHcy per each variant allele (e.g., 0, 1, or 2) was assessed using a 1-df likelihood ratio test, corrected for multiple testing using the method of Conneely and Boehnke [26]. Since plasma tHcy is not normally distributed, it was first transformed to its natural logarithm.

Comparisons between cases and controls for selected baseline characteristics and comparisons between subjects with and without genotyping data were made using multiple regression controlling for the matching factors, race/ethnicity, and study phase. For dietary factors, we also controlled for calories. All statistical analyses were conducted using the R programming language and SAS v9.1.

**Results**

Genotyping results were available for a maximum of 1,354 of 1,621 subjects with questionnaire data (83.5%). Phase 2 subjects did not differ from phase 1 subjects in age, gender, ethnicity, and smoking patterns but had higher estimated dietary folate and alcohol intakes. The subjects with genotyping data did not differ significantly from those with no genotyping data on selected adenoma risk factors (data not shown). Since data on plasma folate was not available for all subjects, we compared those with data on plasma folate to those with no data on plasma folate. Those with plasma folate data differed significantly from those with no plasma folate data on a number of adenoma risk factors (e.g., dietary folate, total dietary fiber, and red meat intakes), and thus plasma folate was not used in any of the analyses.

The characteristics of the study population are presented in Table 1. Cases were more likely to be current smokers and to drink more alcohol, had less dietary folate and fiber intakes, more saturated fat intake, and ate more red meat. After adjustment for the matching factors, race/ethnic group, study phase and calories, smoking, dietary folate, and fiber intakes remained significantly different between cases and controls.
FOCM genes and adenoma risk

The per allele associations with adenoma risk (except for MTHFR C677T and A1298C as noted in the “Methods” section) for all 159 SNPs analyzed, by race/ethnicity, are shown in Supplemental Table 2. When considering all subjects combined, we observed 5 SNPs in the SLC19A1 gene that were statistically associated with adenoma risk.
after correcting for multiple testing (Table 2). Three of these 5 SNPs, rs12482346 (OR = 1.24; 95% CI = 1.07–1.44; $p_{act} = 0.032$), rs2838958 (OR = 0.79; 95% CI = 0.68–0.92; $p_{act} = 0.02$) and rs1051266 (OR = 1.25; 95% CI = 1.07–1.45; $p_{act} = 0.028$) were in moderate linkage disequilibrium (LD) in the total study population (rs12482346–rs2838958, $r^2 = 0.63$; rs12482346–rs1051266, $r^2 = 0.79$; rs2838958–rs1051266, $r^2 = 0.69$). The remaining 2 tagSNPs, rs2838951 (OR = 0.78; 95% CI = 0.67–0.91; $p_{act} = 0.015$) and rs2236484 (OR = 1.25; 95% CI = 1.07–1.46; $p_{act} = 0.027$), were not in linkage disequilibrium with these SNPs or each other. Heterogeneity across race/ethnicity in the total study population was statistically significant for one SNP (MTHFD1 rs11627525; $p = 0.0000886$). MTHFD1 rs11627525 was not significantly associated with risk (Table 2). Five SNPs, rs12482346 (OR = 1.46; 95% CI = 1.07–1.97; $p = 0.015$) and rs2236484 (OR = 1.25; 95% CI = 1.07–1.46; $p_{act} = 0.027$), were not in linkage disequilibrium with these SNPs or each other. Heterogeneity across race/ethnicity in the total study population was statistically significant for one SNP (MTHFD1 rs11627525; $p = 0.0000886$). MTHFD1 rs11627525 was not significantly associated with risk (Table 2).

Among the SNPs significant in non-Hispanic whites, there were no significant associations between MTR D919G (rs1801133) nor the MTHFR A1298C CC genotype (rs1801131) and adenoma risk; OR = 0.93, 95% CI = 0.66–1.32 and OR = 0.97, 95% CI = 0.66–1.43 for the 677 TT and 1298 CC genotypes, respectively. Similarly, there were no associations between MTHFD1 R653Q (rs2236225), SHMT1 L474F (rs1979277), or MTHFR A1793Q (rs2274976) and adenoma risk overall or in non-Hispanic whites (Supplementary Table 2).

Dietary folate, alcohol, and adenoma risk

We assessed possible modifications of the genotype/adenoma associations by dietary folate intake and alcohol use for all 159 tagSNPs in the total study population and non-Hispanic whites. There was no interaction between any SLC19A1 or MTHFD1 SNP and dietary folate or alcohol in either the total study population or non-Hispanic whites (data not shown). The interactions between MTHFR C677T and A1298C and folate availability and alcohol use, assuming a recessive model, are shown in Table 3. Sample size was not sufficient to assess genotype effects for an index of folate and alcohol combined (there were 3 cells with fewer than 5 subjects, including cells with 0 subjects, in MTHFR A1298C and 1 such cell for MTHFR C677T). For the 677 TT genotype, there was no evidence of a linear trend across quartiles of dietary folate intake ($p$ value for trend = 0.793). When we stratified on alcohol intake, the OR for the 677 TT genotype and high alcohol consumption ($>$20 g/day) was non-significantly greater than 1.0 (OR = 1.46 (0.60–3.55) and about 1.0 for those drinking 0 g/day (OR = 0.96 (0.55–1.68), but, again, there was no linear trend ($p = 0.68$). For A1298C (rs1801131), there was significant heterogeneity across strata of dietary folate ($p = 0.004$) due to a significant decrease in risk for those in the second dietary folate quartile (OR = 0.22 (0.07–0.65) and an increase in risk for those in the highest dietary folate quartile (OR = 2.07 (0.92–4.64)). In the first and third quartiles, the ORs were 1.27 (0.57–2.85) and 1.15 (0.57–2.33), respectively ($p$ value for linear trend = 0.12). There was no modification by alcohol consumption for A1298C (heterogeneity $p$ value = 0.98; trend $p$ value = 0.94).

There were no significant interactions between any of the remaining tagSNPs and dietary folate or alcohol (data not shown). No SNP significantly interacted with dietary B12. However, there was a nominally significant interaction between the vitamin B12 transport protein GTF rs519221 and dietary vitamin B12 (heterogeneity $p$ value = 0.02; trend $p$ value = 0.003). ORs were less than 1.0 in the lowest B12 quartiles and non-significantly greater than 1.0 in the highest B12 quartile (ORs = 0.63 (0.43–0.92); 0.66 (0.46–0.95); 0.96 (0.67–1.39) and 1.33 (0.89–1.96) for the lowest intake to the highest intake quartiles, respectively).

Interactions between FOCM genes and sex

There were no significant interactions between any SNPs and sex at the pre-defined level of significance except for MTHFR A1298C. MTHFR A1298C interacted with sex at a nominal level of significance (interaction $p$ value 0.007). Assuming a recessive model, the association between homozygosity for the C allele, compared to those with at least one A allele, and adenoma risk was significantly decreased in women (OR = 0.50 (0.26–0.96) and non-significantly increased in men (OR = 1.52 (0.92–2.51)).

FOCM genes, plasma tHcy, and adenoma risk

The association between SNPs and plasma tHcy was assessed for phase 2 subjects (data not shown). One SNP in MTHFR was associated with plasma tHcy at a nominal level of significance (rs9651118; $\beta = -1.09$, $\pm 0.36$, $p_{act} = 0.03$). There were no other associations with plasma tHcy overall. After stratifying on dietary folate quartiles, there was no longer an association between rs9651118 and
| Gene/SNP   | Totala | Non-Hispanic White | Black | Hispanic | Asian | \( p \) Heterogeneity |
|-----------|--------|--------------------|-------|----------|-------|-----------------------|
| **SLC19A1** |        |                    |       |          |       |                       |
| rs2838951 |        |                    |       |          |       |                       |
| \( n \)   | 1,337  | 719                | 221   | 143      | 193   | 0.08                  |
| MAF\(^b\) | 0.37   | 0.41               | 0.28  | 0.35     | 0.38  |                       |
| OR (95% CI) | 0.78 (0.67, 0.92) | 0.70 (0.57, 0.88) | 1.28 (0.85, 1.92) | 0.97 (0.58, 1.62) | 0.71 (0.46, 1.09) |
| \( p \) \(^{act}\) | 0.02   | 0.01               | 0.97  | 0.99     | 0.64  |                       |
| rs2838958 |        |                    |       |          |       |                       |
| \( n \)   | 1,348  | 722                | 224   | 147      | 195   | 0.13                  |
| MAF\(^b\) | 0.46   | 0.47               | 0.27  | 0.42     | 0.46  |                       |
| OR (95% CI) | 0.79 (0.68, 0.93) | 0.73 (0.5, 0.90)  | 1.30 (0.85, 1.98) | 0.85 (0.51, 1.42) | 0.75 (0.49, 1.15) |
| \( p \) \(^{act}\) | 0.02   | 0.02               | 0.85  | 0.99     | 0.91  |                       |
| rs2236484 |        |                    |       |          |       |                       |
| \( n \)   | 1,325  | 712                | 223   | 142      | 188   | 0.14                  |
| MAF\(^b\) | 0.46   | 0.42               | 0.47  | 0.47     | 0.47  |                       |
| OR (95% CI) | 1.26 (1.08, 1.47) | 1.44 (1.16, 1.79) | 0.90 (0.63, 1.28) | 0.92 (0.57, 1.49) | 1.39 (0.91, 2.13) |
| \( p \) \(^{act}\) | 0.03   | 0.01               | 0.99  | 0.99     | 0.90  |                       |
| rs1051266 |        |                    |       |          |       |                       |
| \( n \)   | 1,344  | 721                | 224   | 146      | 193   | 0.68                  |
| MAF\(^b\) | 0.48   | 0.45               | 0.44  | 0.48     | 0.47  |                       |
| OR (95% CI) | 1.24 (1.06, 1.44) | 1.34 (1.08, 1.65) | 0.98 (0.68, 1.41) | 1.14 (0.71, 1.82) | 1.19 (0.79, 1.77) |
| \( p \) \(^{act}\) | 0.04   | 0.03               | 0.99  | 0.99     | 0.98  |                       |
| rs12482346|        |                    |       |          |       |                       |
| \( n \)   | 1,340  | 721                | 222   | 143      | 194   | 0.23                  |
| MAF\(^b\) | 0.47   | 0.45               | 0.49  | 0.50     | 0.49  |                       |
| OR (95% CI) | 1.24 (1.06, 1.44) | 1.42 (1.15, 1.76) | 0.88 (0.61, 1.26) | 1.06 (0.66, 1.71) | 1.19 (0.79, 1.80) |
| \( p \) \(^{act}\) | 0.04   | 0.01               | 0.99  | 0.96     | 0.99  |                       |
| rs7499    |        |                    |       |          |       |                       |
| \( n \)   | 1,334  | 714                | 223   | 144      | 192   | 0.35                  |
| MAF\(^b\) | 0.43   | 0.41               | 0.41  | 0.46     | 0.46  |                       |
| OR (95% CI) | 1.22 (1.04, 1.42) | 1.38 (1.11, 1.71) | 0.92 (0.64, 1.32) | 1.00 (0.62, 1.60) | 1.21 (0.79, 1.86) |
| \( p \) \(^{act}\) | 0.07   | 0.01               | 0.99  | 1.0      | 1.0   |                       |
| **MTHFD1** |        |                    |       |          |       |                       |
| rs11627387|        |                    |       |          |       |                       |
| \( n \)   | 1,346  | 719                | 225   | 147      | 195   | 0.24                  |
| MAF\(^b\) | 0.29   | 0.31               | 0.24  | 0.23     | 0.29  |                       |
| OR (95% CI) | 0.78 (0.66, 0.92) | 0.68 (0.54, 0.85) | 0.70 (0.45, 1.08) | 1.01 (0.57, 1.81) | 1.10 (0.71, 1.71) |
| \( p \) \(^{act}\) | 0.07   | 0.01               | 0.96  | 0.99     | 0.96  |                       |
| rs8016556 |        |                    |       |          |       |                       |
| \( n \)   | 1,348  | 722                | 225   | 145      | 195   | 0.03                  |
| MAF\(^b\) | 0.33   | 0.35               | 0.26  | 0.25     | 0.39  |                       |
| OR (95% CI) | 0.87 (0.74, 1.02) | 0.69 (0.55, 0.86) | 1.32 (0.85, 2.04) | 0.84 (0.47, 1.49) | 1.20 (0.81, 1.77) |
| \( p \) \(^{act}\) | 0.67   | 0.01               | 0.88  | 0.99     | 0.99  |                       |

\(^{a}\) Includes 61 individuals who characterized their race/ethnicity as ‘other’

\(^{b}\) Minor allele frequency in cases and controls

\(^{c}\) Odds ratio and 95% confidence interval estimated by logistic regression assuming a log-additive (per allele) model and controlling for the matching factors (age, sex, clinic, and examination date) and study phase. The estimates for the total study population are also adjusted for race/ethnicity

\(^{d}\) Adjusted for multiple testing taking into account correlated tagSNPs using a modified test of Conneely and Boehnke [26]
plasma tHcy in any quartile and no interaction between dietary folate and genotype (interaction \( p = 0.67 \)) on plasma tHcy levels. Neither \( MTHFR \) C677T (interaction \( p = 0.97 \), trend \( p = 0.68 \)) nor \( MTHFR \) A1298C (interaction \( p = 0.98 \), trend \( p = 0.87 \)) was associated with plasma tHcy in any dietary folate quartile. Mean plasma tHcy was not significantly higher for those with the \( MTHFR \) 677 TT or 1298 CC genotypes and low dietary folate compared to those with the highest dietary folate (data not shown).

### Gene–gene interactions

Gene–gene interactions were assessed for genes for which there was prior data suggesting interactions with the \( MTHFR \) C677T genotype (Supplemental Table 3). This included SNPs in \( CBS, MTR, TYMS, \) and \( SLC19A1 \). In the total study population, adenoma risk per A allele at \( SLC19A1 \) rs1051266 was significantly increased for those with at least one \( MTHFR \) 677 C allele (OR = 1.33 (1.21–1.56)) but risk per A allele was decreased for those with the 677 TT genotype (OR = 0.74 (0.47–1.17)). The interaction \( p \) value was 0.018. Nominally significant interactions were also observed for 5 other \( SLC19A1 \) SNPs (rs17004785, rs2297291, rs3788205, rs3939250, and rs3827266), none of which were in LD with rs1051266. There was no interaction between \( MTR \) A2756G (rs1805087) or any \( CBS \) SNP and the \( MTHFR \) C677T genotype. There were nominally significant interactions between the \( MTHFR \) C677T genotype and 9 SNPs in \( MTR \) (rs12759827, rs1252252, rs10737812, rs12070633, rs3788205, rs10802569, and rs10737812) and one SNP in \( TYMS \) (rs2298582). Except for rs12759827 and rs4659730, the remaining seven \( MTR \) SNPs were in LD (\( r^2 > 0.70 \)) with each other in non-Hispanic whites. Only one SNP, rs4659730, was in LD with rs1805087 and that only in non-Hispanic whites.

### Table 3

| Genotype  | 528–2,191 mcg/day\(^a\)   | 360–527 mcg/day | 262–359 mcg/day | 42–261 mcg/day | \( p \) Trend | \( p \) Heterogeneity |
|-----------|---------------------------|----------------|---------------|---------------|--------------|------------------------|
| Dietary folate quartiles  
677TT/677CC + CT  
CC + CT: cases/controls | 124/145 | 127/149 | 122/152 | 187/144 | 0.70 | 0.12 |
| TT: cases/controls | 14/20 | 13/24 | 24/15 | 17/18 | 0.95 | 0.12 |
| OR (95% CI)\(^d\) | 0.91 (0.44, 1.9) | 0.66 (0.32, 1.36) | 2.01 (1.0, 4.05) | 0.78 (0.39, 1.6) | 0.12 | 0.12 |
| 1298CC/1298AA + AC  
AA + AC: cases/controls | 125/158 | 127/157 | 152/157 | 195/162 | 0.12 | 0.004 |
| CC: cases/controls | 14/20 | 13/24 | 24/15 | 17/18 | 0.95 | 0.12 |
| OR (95% CI) | 2.07 (0.92, 4.6) | 1.15 (0.57, 2.33) | 0.22 (0.07, 0.65) | 1.27 (0.57, 2.85) | 0.12 | 0.12 |

| Alcohol use  
677TT/677CC + CT  
CC + CT: cases/controls | 265/277 | 153/196 | 60/50 | 80/67 | 0.52 | 0.68 |
| TT: cases/controls | 26/30 | 18/29 | 8/9 | 16/9 | 0.95 | 0.12 |
| OR (95% CI)\(^d\) | 0.96 (0.55, 1.68) | 0.83 (0.44, 1.56) | 0.69 (0.24, 1.93) | 1.46 (0.60, 3.55) | 0.94 | 0.98 |
| 1298CC/1298AA + AC  
AA + AC: cases/controls | 278/297 | 163/210 | 63/53 | 94/74 | 0.94 | 0.98 |
| CC: cases/controls | 24/28 | 18/23 | 8/6 | 4/4 | 0.95 | 0.12 |
| OR (95% CI) | 0.92 (0.52, 1.64) | 1.11 (0.57, 2.13) | 1.04 (0.33, 3.25) | 0.88 (0.21, 3.67) | 0.94 | 0.98 |

\( a \) Quartile cutpoints were based on the distribution of the controls and were study phase specific  
\( b \) Dietary folate values reflect reported average intake in the year before the sigmoidoscopy  
\( c \) Numbers do not add up to total due to missing values for genotype for some whose race/ethnic group classification was ‘other’ (27 cases/29 controls for C677T and 2 cases/1 control for A1298C) or alcohol (1 case)  
\( d \) All ORs and 95% confidence intervals were estimated by unconditional logistic regression assuming a recessive model and adjusting for the matching factors (age, sex, clinic, examination date), race/ethnic group and study phase  
\( e \) Alcohol cutpoints were pre-determined and not based on the population distribution of alcohol consumption
Discussion

In this study, we conducted a comprehensive analysis of 159 tagSNPs in 11 genes involved in folate-associated one-carbon metabolism and the risk of distal colorectal adenomas. Our results suggest that genetic variability in SLC19A1 (also known as RFC1) may be associated with distal colorectal adenoma risk. There was no significant modification by dietary folate intake or alcohol for any SNP except MTHFR A1298C and dietary folate. The data were consistent with a possible interaction between the MTHFR 1298 CC genotype and sex. There was a significant interaction between SLC19A1 rs1051266 (G80A, H27R) and the MTHFR C677T genotype. Seven SNPs in the MTR gene that were within a single LD block among non-Hispanic whites, also interacted with the MTHFR C677T genotype.

To our knowledge, this is the most comprehensive study of genetic variation in the folate pathway and risk of colorectal adenoma conducted to date. Hazra et al. [13] assessed the association between 24 non-synonymous SNPs in 13 folate pathway genes and colorectal adenoma risk in the Nurses’ Health Study. They reported a significant association only for the transcobalamin II (TCN2) P259R polymorphism, assuming a dominant model. As in our study, no associations were seen for non-synonymous SNPs in MTHFR, MTHFD1, SHMT1, or MTR.

SLC19A1 codes for a ubiquitously expressed transmembrane protein necessary for the uptake of reduced folates such as 5-MTHF, the main circulating folate. The SLC19A1 protein transports 5-MTHF from the lumen of the gut into the cells of the small intestine and from the blood stream into epithelial cells. 5-MTHF is the enzymatically active form of folate and is required for the synthesis of methionine from homocysteine, the regeneration of tetrahydrofolate, and nucleotide synthesis [27]. Thus, functional change in the activity of SLC19A1 could modulate folate metabolism by increasing or decreasing intracellular folate availability. A plausible role for SLC19A1 function and colorectal neoplasia risk is suggested by animal studies showing that mice with an inactive gene have an increased susceptibility to chemically induced CRC [28].

To our knowledge, this is the first study of the relationship between rs1051266 and colorectal adenomas. However, two previous studies, both also conducted in non-folic acid-supplemented populations, have reported on the association between this SNP and colorectal cancer [29, 30]. Contrary to our results (OR = 1.25 (95% CI 1.07–1.45) per A allele), there was no association with CRC risk in either study population but, unlike our data, both found a possible decrease in risk associated with the A allele in those with low folate intake. Further studies of this SNP in other colorectal adenoma patients will be necessary before we can draw any conclusions.

Our data are consistent with previous studies of the association between the MTHFR C677T and A1298C polymorphisms and colorectal adenoma in suggesting that neither of these polymorphisms are associated with adenoma risk overall [11, 12]. Our data provide limited support for the hypothesis that the 677 TT genotype may increase colorectal adenoma risk when folate is low and decrease risk when folate is high as has been seen in some studies including our own analysis of the phase 1 subjects [23, 31–34]. However, our results are consistent with other studies not reporting such modification [35–38].

The data were consistent with a possible U-shaped relationship between the MTHFR A1298C genotype and dietary folate intake. A recent European study also observed an increased adenoma risk for the 1298 CC genotype and high folate intake [37]. However, our data are inconsistent with those of most other studies assessing this interaction for adenoma [36, 39] or colorectal cancer risk [40–43].

We also observed a significant interaction between the MTHFR 1298 genotype and sex with a decreased risk for women and an increased risk for men with the CC genotype, relative to those with at least one A allele. These findings are consistent with those of two other adenoma studies [37, 39] but opposite to those of a recent CRC study [44] and inconsistent with studies that did not observe any modification by sex for CRC [45, 46] or rectal cancer [47]. Given these mixed results, future studies should assess such modification so that some consensus can be achieved.

Data from several studies have suggested a gene–gene interaction between the SLC19A1 G80A and MTHFR C677T genotypes [29, 48, 49]. In all studies, including our own, the data suggest that the A allele might be associated with lower risk in those with the 677 TT genotype. Although the MTR A2756G polymorphism did not interact significantly with MTHFR C677T, as suggested in some previous studies [50–54], we did observe nominally significant interactions for a set of seven linked MTR SNPs that may indicate some effect modification for this linkage block, at least in non-Hispanic whites. To our knowledge, this is the first study to suggest such an interaction, and it remains for future studies to further define this linkage block and replicate the interaction before any conclusions can be drawn.

Our study has several strengths including the large sample size and the comprehensive approach to defining genetic variation in a large number of folate pathway genes. In addition, we had data on adenoma risk factors and folate pathway intermediates including plasma tHcy. Our large sample size allowed us to conduct several stratified analyses with reasonable statistical power, including analyses by race/
ethnic group, folate intake, and alcohol use. Weaknesses of our study include the fact that we did not have data on all the genes relevant to the folate pathway. Also, because this was a sigmoidoscopy-based study, some controls may have had undetected adenomas in the proximal colon, and so our findings are limited to distal adenomas. Additionally, this study was conducted prior to fortification of the US food supply with folic acid and folate intakes were low compared to those of supplemented populations. Thus, our findings may be specific to un-supplemented populations. Finally, the use of a Bonferroni correction for multiple testing in all stratified analyses may have been too conservative, and so there may be some false-negative results. Alternatively, given the large number of total comparisons made, false-positive findings are also possible.

In conclusion, in this comprehensive tagSNP analysis of 11 folate pathway genes, we observed that variation in SLC19A1 may play a role in colorectal adenoma risk in non-Hispanic whites. Our data suggest a possible interaction between the MTHFR A1298C genotype and folate intake as well as an interaction with sex. The SLC19A1 G80A polymorphism and seven SNPs defining a linkage disequilibrium block in the MTR gene in non-Hispanic whites may interact with MTHFR C677T genotype in determining distal colorectal adenoma risk.

Acknowledgments We would like to thank all the subjects who contributed their time to this study. We would also like to thank Terry Kolb and Jessie Lin for their assistance with data collection, cleaning, and management and Anh Diep and for her assistance with biospecimen management. This work was supported by the National Cancer Institute, National Institutes of Health under 1R01 CA096830 and 5P01 CA42710.

Conflicts of interest D. Conti is a consultant for Pfizer Inc.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

1. Kim YI (2003) Role of folate in colon cancer development and progression. J Nutr 133:3731S–3739S
2. Giovannucci E (2002) Epidemiologic studies of folate and colorectal neoplasia: a review. J Nutr 132:2350S–2355S
3. Mason JB, Choi SW (2000) The mechanisms by which folate depletion enhances colorectal carcinogenesis: a unified scheme. Nestle Nutr Workshop Ser Clin Perform Programme 4:87–99; discussion 99–101
4. Cole BF, Baron JA, Sandler RS, Haile RW, Ahnen DJ, Bresalier RS, McKeown-Eyssen G, Summers RW, Rothstein RI, Burke CA et al (2007) Folic acid for the prevention of colorectal adenomas: a randomized clinical trial. JAMA 297:2351–2359
5. Song J, Medline A, Mason JB, Gallinger S, Kim YI (2000) Effects of dietary folate on intestinal tumorigenesis in the apcMin mouse. Cancer Res 60:5434–5440
6. Song J, Sohn KJ, Medline A, Ash C, Gallinger S, Kim YI (2000) Chemopreventive effects of dietary folate on intestinal polyps in Apc ± Msh2−/− mice. Cancer Res 60:3191–3199
7. Nijhout HF, Reed MC, Ulrich CM (2008) Mathematical models of folate-mediated one-carbon metabolism. Vitam Horm 79: 45–82
8. Ulrich CM, Neuhouser M, Liu AY, Boynton A, Gregory JF III, Shane B, James SJ, Reed MC, Nijhout HF (2008) Mathematical modeling of folate metabolism: predicted effects of genetic polymorphisms on mechanisms and biomarkers relevant to carcinogenesis. Cancer Epidemiol Biomarkers Prev 17:1822–1831
9. Ulrich CM (2005) Nutrigenetics in cancer research–folate metabolism and colorectal cancer. J Nutr 135:2698–2702
10. Xu X, Liu AY, Ulrich CM, Chen J (2009) Folate and cancer: epidemiological perspective. Taylor and Francis, New York
11. Huang Y, Han S, Li Y, Mao Y, Xie Y (2007) Different roles of MTHFR C677T and A1298C polymorphisms in colorectal adenoma and colorectal cancer: a meta-analysis. J Hum Genet 52:73–85
12. Kono S, Chen K (2005) Genetic polymorphisms of methylene-tetrahydrofolate reductase and colorectal cancer and adenaoma. Cancer Sci 96:553–542
13. Hazra A, Wu K, Craft P, Fuchs CS, Giovannucci EL, Hunter DJ (2007) Twenty-four non-synonymous polymorphisms in the one-carbon metabolic pathway and risk of colorectal adenoma in the Nurses’ Health Study. Carcinogenesis 28:1510–1519
14. Stern MC, Siegumd KD, Conti DV, Corral R, Haile RW (2006) XRC1C, XRC3C, and XPD polymorphisms as modifiers of the effect of smoking and alcohol on colorectal adenoma risk. Cancer Epidemiol Biomarkers Prev 15:2384–2390
15. Sterne MC, Siegumd KD, Corral R, Haile RW (2005) XRCC1 and XRCC3 polymorphisms and their role as effect modifiers of unsaturated fatty acids and antioxidant intake on colorectal ade- nomas risk. Cancer Epidemiol Biomarkers Prev 14:609–615
16. Rimm EB, Giovannucci EL, Stampfer MJ, Colditz GA, Litin LB, Willett WC (1992) Reproducibility and validity of an expanded self-administered semiquantitative food frequency questionnaire among male health professionals. Am J Epidemiol 135:1114–1126; discussion 1127–1136
17. Strom D, Hankin JH, Wilken LS, Pike MC, Monroe KR, Park S, Henderson BE, Nomura AM, Earle ME, Nagamine FS et al (2000) Calibration of the dietary questionnaire for a multiethnic cohort in Hawaii and Los Angeles. Am J Epidemiol 151:358–370
18. Bird CL, Swendsen ME, Witte JS, Shikany JM, Hunt IF, Frankl HD, Lee ER, Longnecker MP, Haile RW (1995) Red cell and plasma folate, folate consumption, and the risk of colorectal adenomatous polyps. Cancer Epidemiol Biomarkers Prev 4:709–714
19. Kuo K, Still R, Cale S, McDowell I (1997) Standardization (external and internal) of HPLC assay for plasma homocysteine. Clin Chem 43:1653–1655
20. Barrett JC, Fry B, Maller J, Daly MJ (2005) Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21:263–265
21. Shen R, Fan JB, Campbell D, Chang W, Chen J, Doucet D, Yeakley J, Bihioko V, Wickham Garcia E, McBride C et al (2005) High-throughput SNP genotyping on universal bead arrays. Mutat Res 573:70–82
22. Consortium TIIH (2005) A haplotype map of the human genome. Nature 437:1299–1320
23. Levine AJ, Siegumd KD, Ervin CM, Diep A, Lee ER, Frankl HD, Haile RW (2000) The methylenetetrahydrofolate reductase
26. Conneely KN, Boehnke M (2007) So many correlated tests, so little time! Rapid adjustment of p values for multiple correlated tests. Am J Hum Genet 81:1158–1168

24. Frosst P, Blom HJ, Milos R, Goyette P, Sheppard CA, Matthews RG, Boers GJ, den Heijer M, Kluitjimans LA, van den Heuvel LP et al (1995) A candidate genetic risk factor for vascular disease: a common mutation in methylenetetrahydrofolate reductase. Nat Genet 10:111–113

25. Li Y, Abecasis GR (2006) Mach 1.0: rapid haplotype reconstruction and missing genotype inference. Am J Hum Genet 79:2290

23. Conneely KN, Boehnke M (2007) So many correlated tests, so little time! Rapid adjustment of p values for multiple correlated tests. Am J Hum Genet 81:1158–1168

22. Lucock M (2000) Folic acid: nutritional biochemistry, molecular biology, and role in disease processes. Mol Genet Metab 71:121–138

21. Ma DW, Finnell RH, Davidson LA, Callaway ES, Spiegelstein O, Martin C, Sandler R (2002) 5,10-Methylenetetrahydrofolate reductase codon 677 and A1298C polymorphisms: diet, estrogen, and risk of colon cancer. Cancer Epidemiol Biomarkers Prev 13:285–292

20. Keku T, Millikan R, Welsey K, Winkle S, Eaton A, Biscocho L, Martin C, Sandler R (2002) 5,10-Methylenetetrahydrofolate reductase codon 677 and 1298 polymorphisms and colon cancer in African Americans and whites. Cancer Epidemiol Biomarkers Prev 11:1611–1621

19. Matsuo K, Ito H, Wakai K, Hirose K, Saito T, Suzuki T, Kato T, Hirai T, Kanemitsu Y, Hamajima H et al (2005) One-carbon metabolism related gene polymorphisms interact with alcohol drinking to influence the risk of colorectal cancer in Japan. Carcinogenesis 26:2164–2171

18. Otani T, Iwasaki M, Hanaoka T, Kobayashi M, Ishihara J, Natsukawa S, Shaura K, Kozumi Y, Kasuga Y, Yoshimura K et al (2005) Folate, vitamin B6, vitamin B12, and vitamin B2 intake, genetic polymorphisms of related enzymes, and risk of colorectal cancer in a hospital-based case–control study in Japan. Nutr Cancer 53:42–50

17. Koushik A, Kraft P, Fuchs CS, Hankinson SE, Willett WC, Giovannucci EL, Hunter DJ (2006) Nonsynonymous polymorphisms in genes in the one-carbon metabolism pathway and associations with colorectal cancer. Cancer Epidemiol Biomarkers Prev 15:2408–2417

16. Curtin K, Slattery ML, Ulrich CM, Bigler J, Levin TR, Wolff RK, Albertsens H, Potter JD, Samowitz WS (2007) Genetic polymorphisms in one-carbon metabolism: associations with Cpg island methylator phenotype (CIMP) in colon cancer and the modifying effects of diet. Carcinogenesis 28:1672–1679

15. Yin G, Kono S, Toyomura K, Hagiwara T, Nagano J, Mizoue T, Miyazaki R, Tanaka M, Kakeji Y, Maehara Y et al (2004) Methyltetrahydrofolate reductase C677T and A1298C polymorphisms and colorectal cancer: the Fukuoka colorectal cancer study. Cancer Sci 95:908–913

14. Murtha MA, Curtin K, Sweeney C, Wolff RK, Holubkov R, Caan BJ, Slattery ML (2007) Dietary intake of folate and co-factors in folate metabolism, MTHFR polymorphisms, and reduced rectal cancer. Cancer Causes Control 18:153–163

13. Chango A, Emery-Fillon N, de Courcy GP, Lambert D, Pfister M, Rosenblatt DS, Nicolas JP (2000) A polymorphism (80G→A) in the reduced folate carrier gene and its associations with folate status and homocysteinaemia. Mol Genet Metab 70:310–315

12. Devlin AM, Clarke R, Birks J, Evans JG, Halsted CH (2006) Interactions among polymorphisms in folate-metabolizing genes and serum total homocysteine concentrations in a healthy elderly population. Am J Clin Nutr 83:708–713

11. Feix A, Fritsche-Polanz R, Kletzmayr J, Vychytil A, Horl WH, Sunder-Plassmann G, Fodinger M (2001) Increased prevalence of combined MTR and MTHFR genotypes among individuals with severely elevated total homocysteine plasma levels. Am J Kidney Dis 38:956–964
51. Gross M, Friedman G, Eliashar R, Koren-Morag N, Goldschmidt N, Atta IA, Ben-Yehuda A (2006) Impact of methionine synthase gene and methylenetetrahydrofolate reductase gene polymorphisms on the risk of sudden sensorineural hearing loss. Audiol Neurootol 11:287–293

52. Le Marchand L, Donlon T, Hankin JH, Kolonel LN, Wilkens LR, Seifried A (2002) B-vitamin intake, metabolic genes, and colorectal cancer risk (United States). Cancer Causes Control 13:239–248

53. Mostowska A, Hozas KK, Wojcicki P, Dziegelewska M, Jagodzinski PP (2009) Associations of folate and choline metabolism gene polymorphisms with orofacial clefts. J Med Genet 47:809–815

54. Terruzzi I, Senesi P, Fermo I, Lattuada G, Luzi L (2007) Are genetic variants of the methyl group metabolism enzymes risk factors predisposing to obesity? J Endocrinol Invest 30:747–753