Research Article

Cancer Epidemiology, Biomarkers & Prevention

Associations between Genes in the One-Carbon Metabolism Pathway and Advanced Colorectal Adenoma Risk in Individuals with Low Folate Intake

Summer S. Han[1], Laura Y. Sue[1], Sonja I. Berndt[1], Jacob Selhub[2], Laurie A. Burdette[3], Philip S. Rosenberg[1], and Regina G. Ziegler[1]

Abstract

Background: Folate is essential for one-carbon metabolism, a pathway required by DNA synthesis, methylation, and repair. Low dietary and circulating folate and polymorphic variation in this pathway are associated with increased risk of colorectal adenoma and cancer.

Methods: We genotyped 882 single nucleotide polymorphisms (SNP) in 82 one-carbon metabolism genes for 1,331 cases of advanced colorectal adenoma, identified by sigmoidoscopy at baseline, and 1,501 controls from the Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial (PLCO). We evaluated associations between one-carbon genes and adenoma risk in all subjects and stratified by folate intake. We applied the Adaptive Rank Truncated Product (ARTP) method to assess statistical significance at the gene and pathway levels.

Results: Folate intake was inversely associated with advanced colorectal adenoma risk [odds ratio (OR) by quartile = 0.85, $P = 1.9 \times 10^{-5}$]. We found no statistically significant associations between one-carbon genes and adenoma risk in all subjects. As hypothesized, we observed a statistically significant pathway-level association ($P = 0.038$) in the lowest quartile of folate; no significant associations were found in higher quartiles. Several genes including adenosine deaminase (*ADA*) and cysteine dioxygenase (*CDO1*) contributed to this signal (gene-level $P = 0.001$ and 0.0073, respectively). The most statistically significant SNP was rs244072 in *ADA* ($P = 2.37 \times 10^{-5}$).

Conclusions and Impact: Stratification by dietary folate and application of the ARTP method revealed statistically significant pathway- and gene-level associations between one-carbon metabolism genes and risk of advanced colorectal adenoma, which were not apparent in analysis of the entire population. Folate intake may interact with associations between common variants in one-carbon metabolism genes and colorectal adenoma risk. *Cancer Epidemiol Biomarkers Prev; 21(3); 417–27. ©2012 AACR.*

Introduction

Rates of colorectal cancer incidence and mortality have decreased in recent years; however, colorectal cancer remains the third most common cancer and third leading cause of cancer-related death in both men and women in the United States (1). Although most colorectal adenomatous polyps will not develop into cancer, these are commonly accepted as precursor lesions of colorectal cancer (2, 3), and screening by colonoscopy or sigmoidoscopy and polyp excision is standard of care for adults ages 50 and older (4, 5).

Folate is essential for one-carbon metabolism, a pathway required by DNA synthesis, methylation, and repair (6). Folate is a water-soluble B-vitamin found in leafy green vegetables, beans and peas, many fruits, and fortified grains and cereals. In epidemiologic studies, low folate intake and low circulating folate levels have consistently been associated with increased risk of colorectal adenoma and colorectal cancer (7–11).

Genes in the folate-mediated one-carbon metabolism pathway have been of interest as candidate genes for influencing colorectal adenoma risk (10, 12–15). Studies of these folate-related genes and folate intake may reveal gene–environment interactions and shed light on complex patterns of risk. There is expanding discussion of the possible importance of gene–environment interaction in

Authors' Affiliations: [1]Epidemiology and Biostatistics Program, Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH, Bethesda, Maryland; [2]Jean Mayer USDA Human Nutrition Research Center on Aging, Tufts University, Boston, Massachusetts; and [3]Core Genotyping Facility at the Advanced Technology Center, Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH, Gaithersburg, Maryland

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

S.S. Han and L.Y. Sue contributed equally to this work.

Corresponding Author: Summer S. Han, Biostatistics Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, 6120 Executive Blvd., EPS/8089, Bethesda, MD 20892. Phone: 301-435-3994; Fax 301-402-0081; E-mail: summer.han@nih.gov

doi: 10.1158/1055-9965.EPI-11-0782

chronic disease etiology, and how to assess efficiently whether interaction exists (12, 13, 15–18). Notably, several studies have found associations between colorectal adenoma risk and common polymorphisms in methylenetetrahydrofolate reductase (*MTHFR*; 12, 15) and thymidylate synthase (*TYMS*; 13, 14), along with evidence of gene–environment interaction; specifically some of these studies reported elevated risk for polymorphisms in these genes among individuals with low folate intake and colorectal adenoma (14, 15) as well as for colorectal cancer (19, 20). Recently, there have been several genome-wide association studies (GWAS) for colorectal cancer (21–23), but so far no GWAS have been reported for colorectal adenoma.

In this study, we investigated associations between advanced colorectal adenomas and 82 one-carbon metabolism genes, selected on the basis of the known biology of this pathway and an extensive literature search. We studied 1,331 cases and 1,501 controls from the Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial (PLCO), a large, multicenter randomized screening trial coordinated by the National Cancer Institute (NCI). We systematically looked for associations with each of 882 tag single nucleotide polymorphisms (SNP) in these genes in the entire population, and then examined associations by level of folate intake. To further investigate gene- and pathway-level associations both in the entire population and by the folate levels, we applied the Adaptive Rank Truncated Product (ARTP) method that combines association signals over a set of SNPs and genes while accounting for multiple testing.

Methods

Study population

The PLCO is a randomized screening trial in which men and women, ages 55 to 74 years, were enrolled from 10 different centers in the United States between 1993 and 2001 (24). All subjects included in this study had been assigned to the screening arm of the trial, completed a baseline risk factor questionnaire, provided a blood specimen at baseline, and consented to participate in etiologic studies of cancer and related diseases. In addition, subjects had completed at baseline a food frequency questionnaire, called the PLCO Dietary Questionnaire (DQx).

Cases were all those participants with a histologically confirmed advanced distal colorectal adenoma diagnosed at baseline sigmoidoscopy. Advanced colorectal adenomas were defined as adenomas more than 1 cm in size, with villous characteristics, and/or with high-grade dysplasia, and carcinomas *in situ*. Controls were randomly selected from those participants without distal colorectal polyps at baseline sigmoidoscopy, and frequency matched to cases on age at baseline, sex, and race. A total of 1,331 cases (850 men and 481 women) and 1,501 controls (950 men and 551 women) were included in this analysis.

A subset ($n = 500$ cases and 500 controls) was chosen for folate measurements in baseline plasma. During the PLCO randomization period (1993–2001), the United States had required that enriched breads, cereals, flours, corn meals, pastas, rice, and other grain products be fortified with folic acid, specifically by January 1, 1998 (25). Therefore, this subset of cases was randomly selected from those advanced adenoma cases with baseline blood draws either before December, 1997 or after January, 1998 so that plasma folate values would represent circulating levels either prior or subsequent to the mandatory folate fortification in 1998. The subset of controls was matched to the cases on date of baseline blood draw.

Dietary assessment

Folate intake and alcohol consumption were derived from the DQx, which asked about the typical frequency of intake during the past year of 137 food items, including beer, wine, and hard liquor, and typical portion sizes for 77 of the food items (24).

We calculated 7 measures of folate intake, in micrograms per day, all of which were based on the single DQx administered to each participant at the time of study entry. Two of the dietary folate estimates were based on USDA (U.S. Department of Agriculture) food composition databases: the first for the U.S. food supply prior to folate fortification in 1998 and the second for the U.S. food supply after folate fortification. The USDA databases relied on microbiologic folate assays to estimate the bioavailability of the folate naturally occurring in foods and the synthetic folic acid added in fortification. A third measure, of total folate, combined vitamin supplement information from the DQx with dietary folate estimates based on the postfortification USDA database.

Four University of Minnesota food composition databases from the Nutrition Data Systems for Research software (NDS-R) were also used; all were for the U.S. food supply after folate fortification in 1998. The first estimated natural folate in foods and the second estimated synthetic folate added to foods. The third and fourth combined natural and synthetic folate occurring in foods by bioavailability based on microbiologic assays and by weighting synthetic folate as 1.67 times more bioavailable than natural folate, respectively.

Unless otherwise specified, for all statistical analyses, we used dietary folate estimated with the USDA prefortification food composition database. Quartile cut points were based on the frequency distribution in the study controls. The quartiles of dietary folate intake were 0 to 262 μg/d (Q1), 263 to 359 μg/d (Q2), 360 to 466 μg/d (Q3), and more than 466 μg/d (Q4).

Laboratory assays

For the subset of 500 cases and 500 controls, heparin plasma was assayed for folate, using a *Lactobacillus casei* microbial assay, at the USDA Human Nutrition Research Center on Aging at Tufts University (26). Plasma collections from 4 subjects were used for quality control. Four anonymized quality control samples were randomly inserted into each of the 28 assay batches, with 2 of the

quality control samples in each batch from the same individual. The total laboratory coefficient of variation (CV) for plasma folate, based on all 28 batches, averaged 14.5%, with individual CVs of 4.4%, 10.4%, 15.0%, and 18.0% for the 4 subjects. Quartiles of plasma folate were calculated based on the plasma folate distribution within the controls.

Gene selection and genotyping

A total of 82 genes related to one-carbon metabolism and regulation were genotyped at the NCI Core Genotyping Facility as part of a custom iSelect bead chip (Illumina Custom Infinium) that included over 29,000 SNPs with coverage of more than 1,300 genes postulated to be related to cancer (27). The one-carbon metabolism pathway genes included on the platform were selected through an extensive literature search and based on the known biology of this pathway.

Tag SNPs were selected for the 82 genes inclusive of the region 20 kb upstream and 10 kb downstream of each gene. We used the HapMap CEU, JPT, CHB, and YRI populations and the Carlson method, as implemented in Tagzilla with a r^2 threshold of 0.8 and minor allele frequency (MAF) of 5% or more. Coverage of tag SNPs for each gene was calculated with Tagzilla; the median percentage coverage of loci needed to be tagged at a threshold of $r^2 = 0.8$ was 79.62%. SNPs with known or putative functional significance (i.e., nonsynonymous, promoter, intron–exon splice sites) were also included whenever possible (27), for an initial total of 1,082 SNPs selected in one-carbon metabolism genes.

Of the tag SNPs originally selected, we excluded SNPs with a call rate less than 90%, in violation of Hardy–Weinberg equilibrium ($P < 0.05/1,082$), and/or with an MAF less than 5%. We also excluded SNPs on the X chromosome with a large number of haploid heterozygotes (>10% among men). After these exclusions, 882 SNPs remained for the 82 one-carbon metabolism genes and were used for this analysis.

Statistical analyses

Association between dietary folate and colorectal adenoma risk. The association between dietary folate and advanced colorectal adenoma was assessed with multivariate logistic regression adjusting for age, gender, race, and study center. We used quartiles of folate as a continuous variable and a trend test with one degree of freedom.

SNP-level genetic association tests. Genetic association tests for individual SNPs were conducted with multivariate logistic regression, adjusting for the same set of covariates described above and additionally including dietary folate using all subjects in the data. In addition, analyses stratified by folate intake were conducted by applying the association test to each SNP among subjects within each quartile of dietary folate. An additive genetic model and the corresponding one degree of freedom likelihood ratio test were used.

To correct for multiple comparisons involved in these genetic association tests, we used a permutation method that takes into account both the multiple SNPs and multiple strata used in the analyses. We simulated 5,000 null data sets by permuting case/control status; for each data set, we conducted 4 sets of association tests across quartiles of dietary folate. The minimum P value from each data set across the 882 SNPs and across the 4 strata was recorded, and a distribution of minimum P values was used as a reference distribution for assessing the statistical significance of our findings for individual SNPs.

Gene- and pathway-level genetic association tests. To test our prior hypothesis that gene- and pathway-level associations existed within the lowest quartile of dietary folate, we applied the ARTP method that combines association signals over a set of SNPs while accounting for multiple testing (28). An underlying assumption of this method is that for each gene (or pathway) to which the test is applied, there may be more than one SNP (or gene) that contributes to disease risk. Thus, instead of using only the smallest P value, a summary statistic based on the top K smallest P values is used ($K = 1, 2, 3, 4,$ and 5). Specifically, the product of the top K P values is used as the test statistic and its significance is assessed with permutations. The ARTP method is sensitive and powerful when there are a number of small association signals. Additional information is provided elsewhere (29, 30). We applied the ARTP method to each gene and obtained 82 gene-level P values; we then applied the ARTP method again to the 82 gene-level P values and obtained a single summary pathway-level P value. Gene- and pathway-level P values were calculated independently for each quartile of dietary folate, and do not take into account testing across the multiple strata.

Testing for gene–environment interaction. To test formally for the statistical significance of interactions between a SNP and folate intake, we used a likelihood ratio test with 6 degrees of freedom based on the model using 4 categories of dietary folate and 3 categories for a given SNP (0, 1, or 2 copies of a minor allele). In addition, we conducted a one degree of freedom test treating both the quartiles of dietary folate and the number of copies of the minor allele as continuous variables.

Results

Characteristics of cases and controls are presented in Table 1. A moderate inverse association between dietary folate and advanced colorectal adenoma was observed [odds ratio (OR) = 0.85 per quartile, 95% confidence interval (95% CI) = 0.79–0.92, $P_{trend} = 1.92 \times 10^{-05}$]. Compared with persons in the highest quartile of dietary folate, persons in the lowest quartile were at 50% increased risk of advanced colorectal adenoma (OR = 1.46, 95% CI = 1.17–1.82).

The top findings from association tests conducted for individual SNPs in the entire population are summarized in Table 2. The strongest association was for rs11627387 in

Table 1. Characteristics of advanced colorectal adenoma cases and controls in PLCO genotyped for this study

Characteristic	Cases	%	Controls	%	p-value
No. of individuals	1,331		1,501		
Gender					0.75
Female	481	36.1	551	36.7	
Male	850	63.9	950	63.3	
Age at randomization (y)					0.11
55–59	395	29.7	509	33.9	
60–64	423	31.8	458	30.5	
65–69	330	24.8	340	22.7	
70–74	183	13.8	194	12.9	
Race					0.13
White non-Hispanic	1,234	92.7	1,368	91.1	
Other	97	7.3	133	8.9	
Family history of colorectal cancer					0.01
Yes	169	13.2	147	10.2	
No	1,331	86.8	1,298	89.8	
Alcohol consumption (g/d)					0.06
<1	486	40.1	580	41.5	
1–15	392	32.3	479	34.3	
15–30	131	10.8	158	11.3	
≥30	203	16.8	181	13.0	
Prefortification dietary folate (mcg/d)					0.0002
<400	822	67.8	850	60.8	
≥400	390	32.2	548	39.2	
Postfortification dietary folate (mcg/d)					<0.0001
<400	604	49.8	587	42.0	
≥400	608	50.2	811	58.0	
Multivitamin use					
Ever	141	16.5	156	15.1	0.42
Never	714	83.5	875	84.9	
Body mass index (kg/m^2)					0.08
<18.5	8	0.6	8	0.5	
18.5–25	348	26.3	440	29.5	
25–30	597	45.2	683	45.8	
≥30	369	27.9	359	24.1	
Strenuous physical activity (h/wk)					
0	225	18.6	182	13.1	0.0002
<1	217	18.0	231	16.6	
1	155	12.8	179	12.9	
2	184	15.2	218	15.7	
3	184	15.2	215	15.4	
≥4	243	20.1	368	26.4	
Smoking status					<0.0001
Never	510	38.4	718	47.8	
Former	196	14.7	105	7.0	
Current	624	46.9	678	45.2	

NOTE: May not sum to total number of participants due to missing covariate values.

methylenetetrahydrofolate dehydrogenase (*MTHFD1*; OR = 0.84 per allele, 95% CI = 0.75–0.94, nominal P = 0.0027, nonrisk/risk allele: *A*/*G*). However, this result, along with the other top findings in Table 2, did not remain statistically significant after adjusting for multiple testing using a permutation method (multiple-testing adjusted P = 0.34 for rs11627387). We then applied the ARTP method to all genotyped SNPs in each of the 82 one-

Table 2. Top findings from association tests in the entire population

Gene	SNP	Location	Risk allele	Nominal *P*	OR (95% CI)
MTHFD1	rs11627387	IVS26-955G>A	*G*	0.0027	0.84 (0.75–0.94)
GLS	rs12185688	IVS14-3683T>C	*T*	0.0033	0.79 (0.67–0.92)
MTHFD1	rs2230491	Ex1-1015C>T	*C*	0.0045	0.80 (0.69–0.93)
UNG2	rs231625	-4520T>C	*C*	0.0088	1.26 (1.06–1.50)
TYMS	rs2298582	IVS13+122A>C	*A*	0.0090	0.80 (0.67–0.95)
MTHFS	rs622506	IVS2-1411T>G	*G*	0.0121	1.15 (1.03–1.28)
MTHFD1	rs2281603	IVS27-380T>C	*T*	0.0138	0.85 (0.75–0.97)
GART	rs9984077	Ex11+195G>A	*A*	0.0138	1.17 (1.03–1.34)
UNG2	rs231622	IVS1+5C>T	*T*	0.0141	1.16 (1.03–1.30)
MTHFD2	rs7587117	IVS29-27A>G	*G*	0.0146	1.15 (1.03–1.29)

carbon metabolism genes. The pathway-level *P* value for the one-carbon metabolism pathway, based on the 82 gene-level *P* values using ARTP methodology, was not statistically significant ($P = 0.89$). All of these analyses included all cases and controls, irrespective of folate intake.

We then conducted a stratified analysis of individual SNP associations within quartiles of dietary folate (Fig. 1). Consistent with our hypothesis that one-carbon metabolism gene signals might be substantially stronger in persons with low folate intake, the smallest *P* value was observed with rs244072 (MAF = 10%) in the adenosine deaminase gene (*ADA*; nominal $P = 2.37 \times 10^{-5}$, OR = 2.33 per allele, 95% CI = 1.60–3.50; nonrisk/risk allele: *T*/*C*). For this SNP, the OR per allele decreased to 1.1 to 1.2 in the intermediate quartiles and to 1.0 in the highest quartile of folate intake (Table 3). The tests for interaction between rs244072 and quartiles of dietary folate had *P* of 0.029 (6 df test) and 0.01 (1 df test). The next top 4 SNPs from the stratified analyses were in the folate receptor 1 gene (*FOLR1*), *TYMS*, cysteine dioxygenase gene (*CDO1*), and methionine adenosyltransferase II, β gene (*MAT2B*; Table 3). Interestingly, as with rs244072 in *ADA*, the SNPs in *FOLR1* and *TYMS* also showed notable associations with risk only in the lowest quartile of dietary folate, and the SNP in *CDO1* showed notable associations with risk only in the lowest 2 quartiles.

To correct for multiple comparisons involved in the above stratified association tests for individual SNPs, we used a permutation method that takes into account both the multiple SNPs and multiple strata used in the analyses. The simulated 5% significance threshold for our analysis was at $P = 6.7 \times 10^{-5}$. Only rs244072 in *ADA* was found to remain statistically significant in the lowest quartile of folate intake, with an adjusted *P* of 0.019 (Fig. 1B). After correction, none of the other 881 SNPs were statistically significantly associated with risk in any of the quartiles of folate intake.

The ARTP method was then used to assess gene- and pathway-level significance in the lowest dietary folate quartile by combining association signals from SNPs within each of the 82 one-carbon metabolism genes. The pathway-level *P* value was statistically significant among subjects in the lowest quartile of folate intake (pathway-level $P = 0.038$). The genes contributing to this significance included *ADA* (gene-level $P = 0.001$), *CDO1* (gene-level $P = 0.0073$), and *FOLR1* (gene-level $P = 0.0076$) with gene-level *P* values smaller than 0.01. Other notable genes with gene-level *P* values smaller than 0.05 were: serine hydroxymethyltransferase (*SHMT1*; gene-level $P = 0.031$), cysteine sulfinic acid decarboxylase (*CSAD*; gene-level $P = 0.0327$), and *TYMS* (gene-level $P = 0.042$). In applying the ARTP method in each of the other quartiles of dietary folate, results did not reveal any statistically significant gene- or pathway-level *P* values (Figs. 1C–E).

To investigate the sensitivity of the association of rs244072 in *ADA*—which was the only statistically significant SNP after multiple testing correction, we repeated the analysis using 6 additional measures of folate intake. The results were consistent with USDA prefortification folate (Table 4); for each measure of folate intake, statistically significant associations for rs244072 (*ADA*) were seen only in the lowest quartile of intake; the nominal *P* values ranged from 6.45×10^{-5} to 0.015 and per-allele ORs ranged from 1.55 to 2.12. When using plasma folate levels, measured with a microbiologic assay for a subset of 500 cases and 500 controls, the OR for rs244072 was elevated among persons in the lowest 3 quartiles (per-allele OR = 1.41–1.87; Table 4). However, none of the associations attained statistical significance, possibly because of the reduced sample size. In addition, we also considered different cut points to define low folate intake for the USDA prefortification folate variable. Stratifying at the 20th, 15th, 10th, and 5th percentiles of the prefortification estimate of folate intake, we found ORs were consistently increased in the low folate subpopulation. Interestingly, the per-allele OR modestly increased as the definition of low folate became more stringent; the OR per allele increased from 2.4 for individuals in the lowest 20% of folate intake to 3.0 for individuals in the lowest 5%.

To assess the biological consistency of the association of rs244072 with advanced colorectal adenoma risk, we

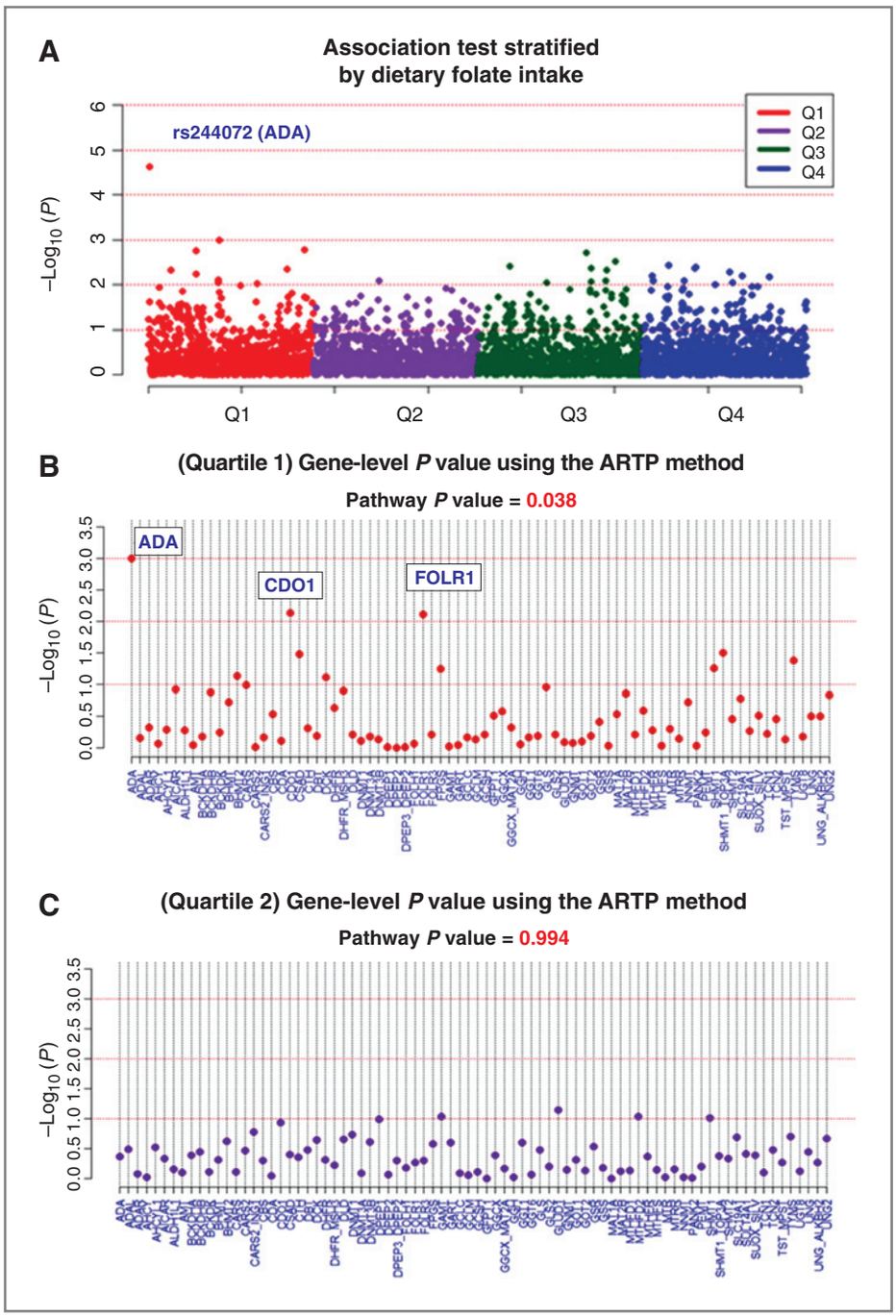

Figure 1. A, association tests between SNPs and advanced colorectal adenoma risk stratified by dietary folate. Each quartile is a different color. For example, the dots in red represent nominal *P* values for each individual one-carbon metabolism SNP for subjects in the lowest quartile of folate intake (Q1), and the dots in blue represent nominal *P* values for each individual SNP for subjects in the highest quartile of folate intake (Q4). B, gene-level *P* values using the ARTP method in lowest folate quartile. C, gene-level *P* values using the ARTP method in quartile 2.

considered its interaction with alcohol intake. High alcohol consumption is thought to modify folate homeostasis by interfering with the intestinal absorption and hepatic uptake of folate and thereby influencing folate bioavailability (31). We conducted an exploratory analysis of rs244072 (*ADA*) within quartiles of alcohol intake (g/d). Interestingly, similar to dietary folate, a statistically significant effect was observed only in the highest risk group, specifically those subjects in the highest quartile of alcohol consumption (nominal $P = 0.0009$, OR = 1.87 per allele, 95% CI = 1.28–2.73). Two-way stratified analyses by tertiles of folate intake and tertiles of alcohol intake showed that those in the highest risk group (i.e., with the lowest folate intake and highest alcohol intake) were at significantly increased risk of adenoma if they carried the *C* allele of rs244072 (nominal $P = 0.0014$, OR = 2.63 per allele, 95% CI = 1.41–4.92), while those in the lowest risk group (i.e.,

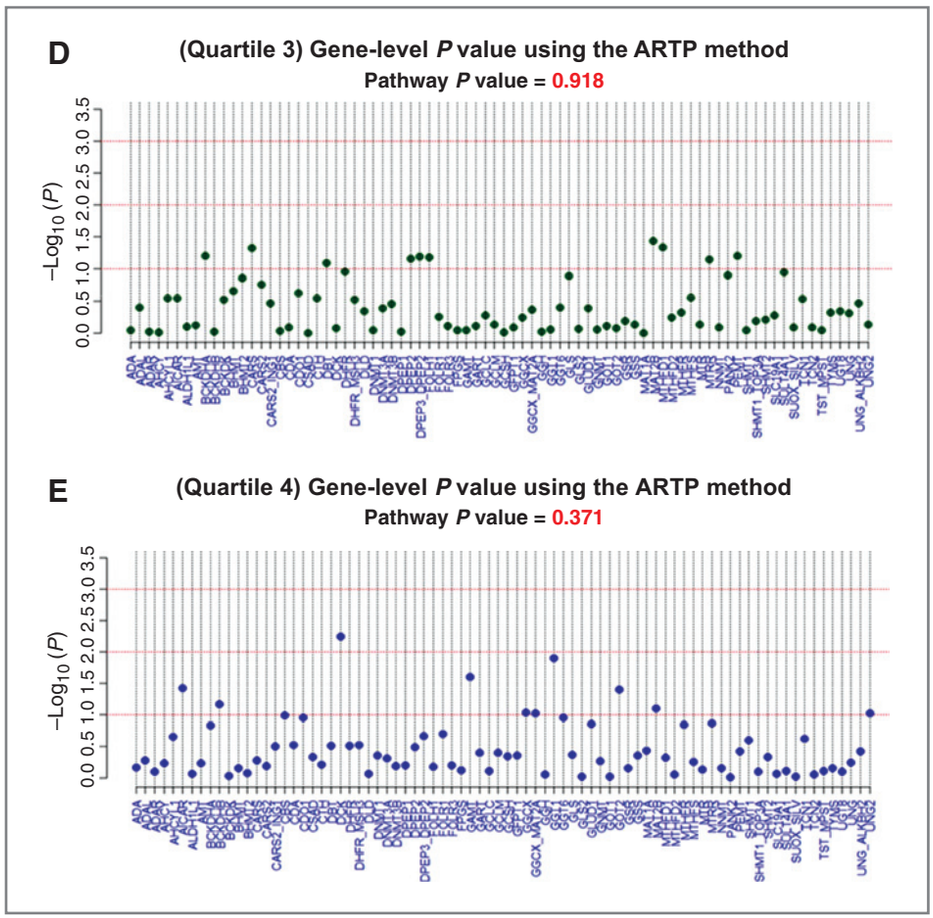

Figure 1. (*Continued*) D, gene-level *P* values using the ARTP method in quartile 3. E, gene-level *P* values using the ARTP method in quartile 4. Each dot represents a *P* value (*y*-axis) for each gene, as listed along the *x*-axis.

with the highest folate intake and lowest alcohol intake) were not at increased risk (nominal $P = 0.82$, OR = 0.93 per allele, 95% CI = 0.52–1.67).

Discussion

Our study found a statistically significant pathway-level association between common variants in one-carbon metabolism genes and advanced colorectal adenoma risk in persons with low folate intake (pathway-level $P = 0.04$). Of the 82 one-carbon metabolism genes explored, the ones contributing to this pathway-level association included *ADA*, *CDO1*, and *FOLR1* (gene-level *P* values of 0.001, 0.007, and 0.008, respectively). We found no statistically significant associations at the SNP-, gene-, or pathway-level in quartiles 2 through 4 of folate intake.

Several studies previously reported interactions between polymorphisms in one-carbon metabolism genes and folate intake. For example, evidence for an interaction was reported between the 677C>T polymorphism in the *MTHFR* gene and folate by Ulrich and colleagues (15), in which they found persons with the *TT* genotype in the lowest tertile of folate, vitamin B_{12}, or vitamin B_6 intake were at increased risk for colorectal adenoma. However, in another study by Chen and colleagues (32), no statistically significant interaction between this polymorphism and folate intake was detected with respect to colorectal adenoma risk. Interactions between 3 *TYMS* polymorphisms and folate intake have also been reported in the literature, with reduced colorectal adenoma risk observed among carriers of variant alleles who had high folate levels (33).

Our data included both of these genes. Analysis of the 12 SNPs in *MTHFR*, including the 677C>T polymorphism (rs1801133), in our study did not reveal any statistically significant SNP- or gene-level associations in the entire population or in any folate quartile under an additive genetic model. Previous reports suggested that *MTHFR* 677C>T has a recessive effect in subjects with low folate intake, that is, the *TT* genotype is associated with elevated risk compared with the *CT* and *CC* genotypes (15). Hence, we analyzed this SNP under a recessive model. The association test was not statistically significant in the entire population (nominal $P > 0.2$); however, stratified analysis revealed increased risk for the *TT* genotype in each of the lowest 2 quartiles of folate (OR = 1.45, $P = 0.13$ and OR = 1.72, $P = 0.02$, for the lowest and the second lowest quartiles of folate, respectively). In contrast, the

Table 3. Top findings from association tests in each quartile of prefortification dietary folate

Gene	SNP	Location	Risk allele	Dietary folate quartile	Nominal P[a]	OR per allele	95% CI
ADA	rs244072	IVS2-1496T>C	*C*	1	**2.37E-05**[b]	2.33	1.60–3.50
				2	0.7194	1.07	0.75–1.50
				3	0.4952	1.15	0.77–1.70
				4	0.9398	0.99	0.68–1.40
FOLR1	rs10501409	7502T>G	*T*	1	**0.0010**	1.77	1.25–2.50
				2	0.8014	1.05	0.74–1.48
				3	0.8105	0.96	0.72–1.52
				4	0.3134	1.22	0.83–1.81
TYMS	rs11664283	−6774A>G	*A*	1	**0.0016**	1.45	1.15–1.85
				2	0.0598	0.79	0.63–1.01
				3	0.3690	1.12	0.91–1.45
				4	0.9589	1.01	0.77–1.30
CDO1	rs34869	−599C>G	*C*	1	**0.0018**	1.41	1.14–1.75
				2	0.0177	1.31	1.05–1.64
				3	0.1311	1.20	0.66–1.06
				4	0.4069	0.90	0.71–1.15
MAT2B	rs4869087	IVS2-1783C>A	*C*	1	0.7900	1.03	0.81–1.30
				2	0.1579	1.18	0.94–1.50
				3	**0.0020**	0.65	0.49–0.86
				4	0.2827	1.16	0.89–1.50

NOTE: P value that is the most significant in each SNP is marked bold.

[a]Nominal P before adjustment for multiple comparisons.

[b]P value that is statistically significant ($P < 0.05$) after adjustment for multiple testing.

risk was not elevated for the *TT* genotype in the second highest and highest quartiles of folate intake. This finding is consistent with the previous report by Ulrich and colleagues (15).

In contrast, our results do not support the previous report indicating that *TYMS* polymorphisms modify risk of colorectal adenoma at high folate intake. In our study, *TYMS* showed a gene-level P value of 0.042, but only in the lowest quartile of dietary folate intake; no associations were found in other quartiles (gene-level $P > 0.4$).

Therefore, although our results implicate the one-carbon metabolism pathway, further studies are needed for replication and elucidation of the magnitude and direction of genetic effects for specific SNPs and their interactions with folate intake. Our statistical approach focused on an additive genetic model for testing associations between SNPs and adenoma risk, which is known to be robust under several genetic models. However, it is possible that any association signal from genetic variants under a recessive model could have been missing from our top findings. Supplementary Table S1 shows P values of all the SNPs analyzed in our data, both in the entire population and by levels of folate.

We reported that the only statistically significant SNP-level association was for rs244072 in *ADA* in the lowest folate quartile (nominal $P = 2.37 \times 10^{-5}$, adjusted $P = 0.019$), while no associations for this SNP were observed in the other quartiles. The functional relevance of SNP rs244072 is unclear, and the association could be due to other SNPs in linkage disequilibrium that have functional significance. *ADA* is an enzyme involved in DNA metabolism, specifically purine metabolism, and in the salvage pathway activity of cancerous tissues and cells (34, 35). The role of *ADA* in the one-carbon metabolism pathway is the deamination of adenosine produced through the hydrolysis of *S*-adenosylhomocysteine into adenosine and homocysteine (36). A fully functional *ADA*, facilitating the irreversible degradation and removal of adenosine, may be necessary in the setting of increased purine and pyrimidine metabolism in carcinogenesis. Previous studies have observed higher levels of *ADA* activity in cancerous colorectal tissue, as compared with normal adjacent tissue from the same subjects, and in cancerous bladder tissue, as compared with normal adjacent tissue from the same subjects and from cancer-free subjects (34, 37). Deficiencies in *ADA* have also been associated with a type of severe combined immunodeficiency disease (SCID; ref. 38).

The other top gene-level associations were observed with the folate receptor gene (*FOLR1*; gene-level $P = 0.007$) and the *CDO1* (gene-level $P = 0.008$) in the lowest quartile of folate intake. *FOLR1* may be involved in cellular

Table 4. Associations between rs244072 (*ADA*) and advanced colorectal adenoma risk in subjects in the lowest folate quartile, using different folate measures

		Per-allele OR (Nominal *P*)			
Food composition database	**Folate measure**	**Quartile 1[a] 356/350[b]**	**Quartile 2 301/349**	**Quartile 3 281/349**	**Quartile 4 274/350**
USDA	Prefortification dietary folate	2.33 (2.37E-05)	1.07 (0.72)	1.15 (0.50)	0.99 (0.94)
	Postfortification dietary folate	2.12 (9.28E-05)	1.07 (0.71)	1.01 (0.92)	1.06 (0.74)
	Total folate from diet (post-fortification) and supplements	1.61 (8.90E-03)	1.17 (0.41)	0.93 (0.73)	1.28 (0.21)
University of Minnesota Nutrition Data Systems for Research	Postfortification dietary folate[c]	1.88 (6.80E-04)	1.1 (0.61)	1.02 (0.90)	1.11 (0.58)
	Postfortification dietary folate[d]	1.89 (5.70E-04)	1.09 (0.65)	0.88 (0.57)	1.18 (0.36)
	Natural folate[e]	2.16 (6.45E-05)	0.85 (0.40)	1.26 (0.20)	1.04 (0.81)
	Synthetic folate[f]	1.55 (1.53E-02)	1.2 (0.33)	0.93 (0.74)	1.23 (0.25)
N/A	Plasma folate	1.68 (0.14)	1.87 (0.13)	1.41 (0.29)	0.93 (0.84)

[a]Quartiles were calculated based on the distribution of the folate variable among the controls.
[b]The number of cases and controls.
[c]Natural + synthetic, based on microbiologic folate assays.
[d]Natural + synthetic, based on estimated bioavailability.
[e]Food folate from diet.
[f]Folic acid from diet (postfortification).

proliferation (39), and its gene product is a useful prognostic marker for certain types of cancer (40, 41). *CDO1* catalyzes the production of cysteine sulfinate from cysteine, which is metabolized from homocysteine and serine. Overexpression of *CDO1* is associated with Sézary syndrome, a type of T-cell lymphoma (39).

Our data included 7 measures of folate intake derived from a single food frequency questionnaire administered to each participant at the time of study entry based on different food composition databases. Our findings for the *ADA* gene were consistent using each measure, including an integrated measure of dietary and supplemental folate. Furthermore, in our control group, the dietary folate intake estimates were highly correlated (Pearson r ranging from 0.78 to 0.95). Therefore, our results do not seem to be sensitive to the estimates of folate intake.

Our study has several strengths. First, our study included a substantially larger number of cases and controls than previous studies of one-carbon metabolism and colorectal adenoma. Second, our analyses considered advanced colorectal adenomas because advanced adenomas are more likely to progress to colorectal cancer than other adenomas, and hence these lesions are more relevant to colorectal cancer etiology than adenomas in general. Third, although previous studies of one-carbon metabolism genes and colorectal adenomas or cancers focused mostly on individual genes in the pathway, to our knowledge, our study is the first report to systematically investigate the pathway-level association between one-carbon metabolism genes and advanced colorectal adenoma risk. Fourth, we used permutation methods to control the false positive rate accounting for multiple testing of SNPs and the ARTP method to access significance at the gene- and pathway-level by combining small association signals into a single test.

Despite the strengths described above, this study has several limitations. First, only a subset of participants had plasma folate measured at baseline, which is a more accurate measure of folate bioavailability than dietary intake. However, the biologic interaction of one-carbon metabolism genes could be primarily with dietary folate and help determine the levels of circulating folate. Second, our results are for distal advanced adenomas only. In PLCO, only sigmoidoscopy was used for screening at baseline, and we did not consider the presence or absence of proximal adenomas in the cases or controls. Distal and proximal adenomas have been shown to vary in their tumor initiation and growth rates, which may represent somewhat different pathogenic processes (40, 41). However, we know of no evidence that the role of common genetic variation in the etiology of colorectal cancer is restricted to either distal or proximal tumors. Third, this study may not be powered to detect some important gene–environment interactions because the required sample size for detecting interactions is much larger than for detecting main effects. We therefore focused on stratified analyses by folate intake and systematically applied interaction tests to our top findings.

To our knowledge, no associations between one-carbon metabolism genes and colorectal cancer have been reported in GWAS. Given our results, this is not surprising as the primary hypothesis of GWAS is that susceptible SNPs can be detected on the basis of main effects of SNPs. Hence, SNPs that interact with environmental exposures may be overlooked. In contrast, we detected associations using stratified analysis by folate intake. Our results suggest that the development and application of more powerful tests to detect interactions in GWAS studies may be very useful. These tests could be based on versions of the methods developed here.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

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 December 27, 2011; revised December 28, 2011; accepted January 11, 2012; published OnlineFirst January 17, 2012.

References

1. Jemal A, Siegel R, Ward E, Hao Y, Xu J, Thun MJ. Cancer statistics, 2009. CA Cancer J Clin 2009;59:225–49.
2. Bond JH. Colon polyps and cancer. Endoscopy 2005;37:208–12.
3. Levin B, Lieberman DA, McFarland B, Andrews KS, Brooks D, Bond J, et al. Screening and surveillance for the early detection of colorectal cancer and adenomatous polyps, 2008: a joint guideline from the American Cancer Society, the US Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology. Gastroenterology 2008;134:1570–95.
4. U.S. Preventive Services Task Force. Screening for colorectal cancer: U.S preventive services task force recommendation statement. Ann Intern Med 2008;149:627–37.
5. Bond JH. Polyp guideline: diagnosis, treatment, and surveillance for patients with colorectal polyps. Am J Gastroenterol 2000;95:3053–63.
6. Choi SW, Mason JB. Folate and carcinogenesis: an integrated scheme. J Nutr 2000;130:129–32.
7. Choi SW, Mason JB. Folate status: effects on pathways of colorectal carcinogenesis. J Nutr 2002;132:2413S–8S.
8. Prinz-Langenohl R, Fohr I, Pietrzik K. Beneficial role for folate in the prevention of colorectal and breast cancer. Eur J Nutr 2001;40:98–105.
9. Kim YI. Role of folate in colon cancer development and progression. J Nutr 2003;133:3731S–9S.
10. Hazra A, Wu K, Kraft P, Fuchs CS, Giovannucci EL, Hunter DJ. Twenty-four non-synonymous polymorphisms in the one-carbon metabolic pathway and risk of colorectal adenoma in the Nurses' Health Study. Carcinogenesis 2007;28:1510–9.
11. Giovannucci E. Epidemiologic studies of folate and colorectal neoplasia: a review. J Nutr 2002;132:2350S–5S.
12. Levine AJ, Siegmund KD, Ervin CM, Diep A, Lee ER, Frankl HD, et al. The methylenetetrahydrofolate reductase 677C T polymorphism and distal colorectal adenoma risk. Cancer Epidemiol Biomarkers Prev 2000;9:657.
13. Hubner RA, Liu JF, Sellick GS, Logan RFA, Houlston RS, Muir KR. Thymidylate synthase polymorphisms, folate and B-vitamin intake, and risk of colorectal adenoma. Br J Cancer 2007;97:1449–56.
14. Ulrich CM, Bigler J, Bostick R, Fosdick L, Potter JD. Thymidylate synthase promoter polymorphism, interaction with folate intake, and risk of colorectal adenomas. Cancer Res 2002;62:3361.
15. Ulrich CM, Kampman E, Bigler J, Schwartz SM, Chen C, Bostick R, et al. Colorectal adenomas and the C677T MTHFR polymorphism: evidence for gene-environment interaction? Cancer Epidemiol Biomarkers Prev 1999;8:659.
16. Friso S, Choi SW. Gene-nutrient interactions in one-carbon metabolism. Curr Drug Metab 2005;6:37–46.
17. Thomas D. Gene-Environment-wide association studies: emerging approaches. Nat Rev Genet 2010;11:259–72.
18. Hunter DJ. Gene-Environment interactions in human diseases. Nat Rev Genet 2005;6:287–98.
19. Chen J, Giovannucci E, Kelsey K, Rimm EB, Stampfer MJ, Colditz GA, et al. A methylenetetrahydrofolate reductase polymorphism and the risk of colorectal cancer. Cancer Res 1996;56:4862.
20. Ma J, Stampfer MJ, Giovannucci E, Artigas C, Hunter DJ, Fuchs C, et al. Methylenetetrahydrofolate reductase polymorphism, dietary interactions, and risk of colorectal cancer. Cancer Res 1997;57:1098.
21. Broderick P, Carvajal-Carmona L, Pittman AM, Webb E, Howarth K, Rowan A, et al. A genome-wide association study shows that common alleles of SMAD7 influence colorectal cancer risk. Nat Genet 2007;39:1315–7.
22. Zanke BW, Greenwood CMT, Rangrej J, Kustra R, Tenesa A, Farrington SM, et al. Genome-wide association scan identifies a colorectal cancer susceptibility locus on chromosome 8q24. Nat Genet 2007;39:989–94.
23. Tomlinson I, Webb E, Carvajal-Carmona L, Broderick P, Kemp Z, Spain S, et al. A genome-wide association scan of tag SNPs identifies a susceptibility variant for colorectal cancer at 8q24. 21. Nat Genet 2007;39:984–8.
24. Subar AF, Thompson FE, Smith AF, Jobe JB, Ziegler RG, Potischman N, et al. Improving food frequency questionnaires: a qualitative approach using cognitive interviewing. J Am Diet Assoc 1995;95:781–8.
25. Food. and Drug Administration, Food standards: amendment of the standards of identity for enriched grain products to require addition of folic acid. Fed Regist 1996;61:8781–807.
26. Horne DW, Patterson D. *Lactobacillus casei* microbiological assay of folic acid derivatives in 96-well microtiter plates. Clin Chem 1988;34:2357.
27. Wang SS, Gonzalez P, Yu K, Porras C, Li Q, Safaeian M, et al. Common genetic variants and Risk for HPV persistence and progression to cervical cancer. PLoS One 2010;5:e8667.
28. Yu K, Li Q, Bergen AW, Pfeiffer RM, Rosenberg PS, Caporaso N, et al. Pathway analysis by adaptive combination of P-values. Genet Epidemiol 2009;33:700–9.
29. Yu K, Li Q, Bergen AW, Pfeiffer RM, Rosenberg PS, Caporaso N, et al. Pathway analysis by adaptive combination of P-values. Genet Epidemiol 2009;33:700–9.
30. Yang HP, Gonzalez Bosquet J, Li Q, Platz EA, Brinton LA, Sherman ME, et al. Common genetic variation in the sex hormone metabolic pathway and endometrial cancer risk: pathway-based evaluation of candidate genes. Carcinogenesis 2010;31:827.
31. Halsted CH, Villanueva JA, Devlin AM, Chandler CJ. Metabolic interactions of alcohol and folate. J Nutr 2002;132:2367S–72S.
32. Chen J, Giovannucci EL, Hunter DJ. MTHFR polymorphism, methyl-replete diets and the risk of colorectal carcinoma and adenoma among US men and women: an example of gene-environment interactions in colorectal tumorigenesis. J Nutr 1999;129:560.
33. Hubner RA, Liu JF, Sellick GS, Logan RF, Houlston RS, Muir KR. Thymidylate synthase polymorphisms, folate and B-vitamin intake, and risk of colorectal adenoma. Br J Cancer 2007;97:1449–56.
34. Eroglu A, Canbolat O, Demirci S, Kocaoglu H, Eryavuz Y, Akgul H. Activities of adenosine deaminase and 5′-nucleotidase in cancerous and noncancerous human colorectal tissues. Med Oncol 2000;17:319–24.
35. Sauer J, Mason JB, Choi SW. Too much folate: a risk factor for cancer and cardiovascular disease? Curr Opin Clin Nutr Metab Care 2009;12:30–6.
36. Yi P, Melnyk S, Pogribna M, Pogribny IP, Hine RJ, James SJ. Increase in plasma homocysteine associated with parallel increases in plasma S-adenosylhomocysteine and lymphocyte DNA hypomethylation. J Biol Chem 2000;275:29318.

37. Durak I, Perk H, Kavutcu M, Canbolat O, Akyol O, Beduk Y. Adenosine deaminase, 5′nucleotidase, xanthine oxidase, superoxide dismutase, and catalase activities in cancerous and noncancerous human bladder tissues. Free Radic Biol Med 1994;16:825–31.
38. Aiuti A, Cattaneo F, Galimberti S, Benninghoff U, Cassani B, Callegaro L, et al. Gene therapy for immunodeficiency due to adenosine deaminase deficiency. N Engl J Med 2009;360:447.
39. Booken N, Gratchev A, Utikal J, Weiss C, Yu X, Qadoumi M, et al. Sezary syndrome is a unique cutaneous T-cell lymphoma as identified by an expanded gene signature including diagnostic marker molecules CDO1 and DNM3. Leukemia 2007;22:393–9.
40. Bufill JA. Colorectal cancer: evidence for distinct genetic categories based on proximal or distal tumor location. Ann Inter Med 1990;113:779.
41. Meza R, Jeon J, Renehan AG, Luebeck EG. Colorectal cancer incidence trends in the United States and United Kingdom: evidence of right-to left-sided biological gradients with implications for screening. Cancer Res 2010;70:5419.