The 677C→T variant of *MTHFR* is the major genetic modifier of biomarkers of folate status in a young, healthy Irish population

Barry Shane,[1] Faith Pangilinan,[2] James L Mills,[3] Ruzong Fan,[4] Tingting Gong,[4] Cheryl D Cropp,[5] Yoonhee Kim,[5] Per M Ueland,[6] Joan E Bailey-Wilson,[5] Alexander F Wilson,[5] Lawrence C Brody,[2] and Anne M Molloy[7]

[1]Department of Nutritional Sciences and Toxicology, University of California, Berkeley, CA; [2]Medical Genomics and Metabolic Genetics Branch, National Human Genome Research Institute, Bethesda, MD; [3]Division of Intramural Population Health Research, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, Bethesda, MD; [4]Department of Biostatistics, Bioinformatics, and Biomathematics, Georgetown University Medical Center (GUMC), Washington, DC; [5]Computational and Statistical Genomics Branch, National Human Genome Research Institute, Bethesda, MD; [6]Department of Clinical Science, University of Bergen and Haukeland University Hospital, Bergen, Norway; and [7]Department of Clinical Medicine, School of Medicine, Trinity College Dublin, Dublin, Ireland

ABSTRACT

Background: Genetic polymorphisms can explain some of the population- and individual-based variations in nutritional status biomarkers.

Objective: We sought to screen the entire human genome for common genetic polymorphisms that influence folate-status biomarkers in healthy individuals.

Design: We carried out candidate gene analyses and genome-wide association scans in 2232 young, healthy Irish subjects to evaluate which common genetic polymorphisms influence red blood cell folate, serum folate, and plasma total homocysteine.

Results: The 5,10-methylenetetrahydrofolate reductase (*MTHFR*) 677C→T (rs1801133) variant was the major genetic modifier of all 3 folate-related biomarkers in this Irish population and reached genome-wide significance for red blood cell folate ($P = 1.37 \times 10^{-17}$), serum folate ($P = 2.82 \times 10^{-11}$), and plasma total homocysteine ($P = 1.26 \times 10^{-19}$) concentrations. A second polymorphism in the *MTHFR* gene (rs3753584, $P = 1.09 \times 10^{-11}$) was the only additional *MTHFR* variant to exhibit any significant independent effect on red blood cell folate. Other *MTHFR* variants, including the 1298A→C variant (rs1801131), appeared to reach genome-wide significance, but these variants shared linkage disequilibrium with *MTHFR* 677C→T and were not significant when analyzed in *MTHFR* 677CC homozygotes. No additional non-*MTHFR* modifiers of red blood cell or plasma folate were detected. Two additional genome-wide significant modifiers of plasma homocysteine were found in the region of the dipeptidase 1 *(DPEP1)* gene on chromosome 16 and the Twist neighbor B (*TWISTNB*) gene on chromosome 7.

Conclusions: The *MTHFR* 677C→T variant is the predominant genetic modifier of folate status biomarkers in this healthy Irish population. It is not necessary to determine *MTHFR* 677C→T genotype to evaluate folate status because its effect is reflected in concentrations of standard folate biomarkers. The *MTHFR* 1298A→C variant had no independent effect on folate status biomarkers. To our knowledge, this is the first genome-wide association study report on red blood cell folate and the first report of an association between homocysteine and *TWISTNB*. *Am J Clin Nutr* 2018;108:1334–1341.

Keywords: biomarkers, folate, homocysteine, vitamin B-12, methylenetetrahydrofolate reductase, 1-carbon metabolism, Trinity Student Study

INTRODUCTION

Serum (or plasma) and red blood cell folate concentrations are the most commonly used biomarkers for assessing folate status (1). Serum folate concentrations are influenced by recent folate intake as well as by fasting status. Red blood cell folate concentrations are considered a better indicator of long-term folate status because red blood cell folate stores are established during erythropoiesis and the lifetime of the mature circulating red blood cell is ∼120 d (2). Plasma total homocysteine is used as a functional biomarker of folate status, although

Supported by the Intramural Research Programs of the National Human Genome Research Institute (NHGRI) and the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development of the NIH (contract N01HD33348), the Health Research Board, Dublin, and USDA Multi-State Research Project W3002.

Supplemental Figures 1 and 2 and Supplemental Tables 1–4 are available from the "Supplementary data" link in the online posting of the article and from the same link in the online table of contents at https://academic.oup.com/ajcn/.

Address correspondence to BS (e-mail: bandie@berkeley.edu).

Abbreviations used: *CHMP1A*, chromatin modifying protein 1A isoform; *DPEP1*, dipeptidase 1; *FOLR,* folate receptor; *FPGS*, folylpolyglutamate synthetase; GWAS, genome-wide association study; LD, linkage disequilibrium; MTHFR, 5,10-methylenetetrahydrofolate reductase; *SLC*, solute carrier; SNP, single nucleotide polymorphism; *TWISTNB*, Twist neighbor B.

Received April 16, 2018. Accepted for publication July 19, 2018.

First published online October 19, 2018; doi: https://doi.org/10.1093/ajcn/nqy209.

plasma homocysteine concentrations are also influenced by other vitamins (vitamin B-12 and vitamin B-6), by methionine intake, and by renal function (3).

A number of studies have shown that folate status indicators can be influenced by common genetic polymorphisms and such polymorphisms are likely to be partly responsible for the variation in biomarker concentrations in individuals (4–9). For example, subjects homozygous for the 677T allele variant (rs1801133) of the 5,10-methylenetetrahydrofolate reductase (*MTHFR*) gene have reduced serum and red blood cell folate concentrations whereas plasma homocysteine concentrations are increased in these individuals, at least in those with poorer folate or riboflavin status (4, 6, 10).

Previous genome-wide association studies (GWASs) have reported significant associations between several genetic variants, including the *MTHFR* 677C→T variant, and plasma homocysteine concentrations (11–18). One deep-sequencing project identified the *MTHFR* 677C→T variant as a genome-wide determinant of serum folate (19) whereas other GWAS efforts were unable to detect any genome-wide signals for serum folate (11, 12, 20). A GWAS of red blood cell folate concentrations has, to our knowledge, not previously been reported. In this study, we investigated genetic influences on commonly used indicators of folate status in a healthy, young population.

METHODS

Subjects

The Trinity Student Study enrolled students attending the University of Dublin, Trinity College between February 2003 and 2004. In total, 2524 subjects met the eligibility criteria which included age between 18 and 28 y, no current serious medical condition, and Irish ethnicity based on grandparents' origins. Ethical approval was obtained from the Dublin Federated Hospitals Research Ethics Committee, which is affiliated with Trinity College Dublin, and subjects gave written informed consent. The study was also reviewed by the Office of Human Subjects Research at the US NIH. Further details relating to the Trinity Student Study have been published elsewhere (21–24). Fifteen subjects with no questionnaire data and 1 duplicate sample were excluded, leaving 2508 valid participants whose samples were assigned to analysis. Participants were not required to fast. Each participant gave 30 mL of blood for preparation of EDTA-plasma or serum and red blood cell lysates. All samples were processed within 3 h of collection and aliquots were stored at −80°C. Blood samples and questionnaire data were coded and made anonymous before analysis. A study population flow chart is shown in **Supplemental Figure 1**.

Biochemical analyses

Serum folate and red blood cell folate were measured by microbiological assays as previously described (25). The between-assay CVs were <11.0%. Plasma total homocysteine was assayed by Bevital (www.bevital.no) via an automated isotope-dilution gas chromatography–mass spectrometry method (26). Details of methodologies for the measurement of other metabolites noted in **Table 1** have been described previously (22, 27). The between-day CVs were ≤8.1%.

Genetic approach and analysis

In the GWAS analyses, genotyping and quality control of genotype data were performed as previously described (27). Briefly, DNA for each sample was extracted from peripheral blood. Eighteen subjects were excluded because of insufficient DNA on extraction, and genome-wide single nucleotide polymorphism (SNP) genotyping was carried out on 2490 subjects with the use of Illumina 1M HumanOmni1-Quad_v1-0_B chips. These arrays include 758,443 genotyped SNPs.

Because of concerns that correction factors required in the GWAS might exclude the possibility of finding genome-wide physiologically relevant markers, we decided to initially carry out a candidate gene analysis through the use of a cleaned-up subset of the GWAS data. We selected 76 SNPs to tag common variation in 9 genes which have been shown, or might be expected, to influence folate status and/or metabolism. The genes selected are involved in folate transport: the intestinal folate transporter [*PCFT*/solute carrier family 46 member 1 *(SLC46A1)*], the reduced folate carrier (*RFC/SLC19A1*), the mitochondrial folate transporter (*MFTC/SLC25A32*), and the folate receptor (*FOLR1*); tissue folate accumulation and retention: folylpolyglutamate synthetase (*FPGS*) and folate hydrolase (*FOLH1*); and folate interconversion, metabolism, and catabolism: *MTHFR*, methenyltetrahydrofolate synthetase (*MTHFS*), and dihydrofolate reductase (*DHFR*) (**Supplemental Table 1**). Because plasma total homocysteine concentrations are influenced by folate status, the same candidate gene variants were used to analyze homocysteine concentrations, although these variants would only represent a subset of genes that might be expected to influence homocysteine. We also evaluated the GWAS data to investigate whether additional gene variants that were not part of our candidate gene study influenced folate status biomarkers. In all analyses, SNPs were tested for association with biomarkers via the simple linear regression model executed in PLINK version 1.7 under the assumption of an additive genetic model.

Red blood cell folate, serum folate, and plasma total homocysteine concentrations had nonnormal distributions. In the candidate gene analyses, inverse normal rank transformation was applied to make them satisfy the normality assumption before examination of associations with metabolite variables through linear regression analysis. The inverse normal rank−transformed metabolic data were regressed on each SNP for association analysis with adjustment for gender, serum vitamin B-12 concentrations, and use of folic acid supplements. Bonferroni corrections for these tests were based on the number of SNPs tested for all candidate genes (significance set at $P = 0.05/76$).

In the total GWAS data analyses, SNPs were again tested for association with red blood cell and serum folate and plasma total homocysteine via the simple linear regression model under the assumption of an additive genetic model. In this case log-transformed data adjusted for age and gender were used. The GWAS genome-wide significance level was set at $P < 5 \times 10^{-8}$.

We also performed metabolite statistical analyses using SAS software (version 9.4; SAS Institute Inc., Cary, NC). Independent Student's *t* tests were used to determine statistically significant differences between groups. Pearson correlation coefficients were used to test associations between metabolites. Kruskal-Wallace and Mann-Whitney analyses were used for data that exhibited nonnormal distributions.

TABLE 1
Characteristics of the Trinity Student Study population (2232 subjects)

	Men			Women			
	n	Mean	SD	*n*	Mean	SD	*P*[1]
Age, y	922	22.6	1.7	1310	22.3	1.6	<0.0001
BMI, kg/m^2	903	23.4	2.9	1306	22.7	3.1	<0.0001
Alcohol intake, g/d	921	33.3	24.8	1309	18.5	15.6	<0.0001
Creatinine, μmol/L	921	75.4	11.7	1307	59.6	10.0	<0.0001
C-reactive protein, mg/L	804	1.40	3.60	1237	1.58	2.58	0.2189
Hemoglobin, g/dL	922	15.4	1.3	1310	13.4	1.6	<0.0001
Serum ferritin, μg/L	865	66.5	40.3	1282	26.8	22.1	<0.0001
Serum folate, nmol/L	921	32.4	17.2	1310	35.1	19.5	0.0006
Red blood cell folate, nmol/L	922	1099	416	1309	1058	437	0.0274
Serum vitamin B-12, pmol/L	922	350	138	1310	319	148	<0.0001
Serum holotranscobalamin II, pmol/L	913	64.2	29.2	1307	56.1	31.4	<0.0001
Plasma homocysteine, μmol/L	922	9.42	3.38	1310	8.12	2.58	<0.0001
Plasma methylmalonic acid, μmol/L	908	0.18	0.07	1302	0.19	0.09	0.0851
Folic acid users in past week	180 of 922 (19.5%)			352 of 1310 (26.9%)			0.0001[2]
Folic acid intake among users, μg/d	180	200[3]	100–300[4]	352	200[3]	100–286[4]	0.378[5]

[1]Gender differences. An independent *t* test was used for all gender comparisons except folic acid supplements per week and folic acid users in the past week, as noted.
[2]Chi-square test.
[3]Median.
[4]IQR intake among folic acid users.
[5]Mann-Whitney *U* test.

RESULTS

Characteristics of the selected population

High-quality genotyping data were obtained for 2438 of the 2490 samples genotyped. Two hundred and six subjects were excluded for various reasons including sex discrepancies, abnormal sex chromosomes, siblings, cousins, and other detected relatedness, low call rate, no phenotype, or outlier. The characteristics of the final study sample (2232 subjects) are shown in Table 1. Folic acid supplements were consumed by 532 participants in the week before the study (23.3%). No subjects had medical conditions that would affect folate absorption at the time of the study.

Red blood cell and serum folate were significantly correlated (Pearson's $r = 0.61$; $P < 0.0001$). Both serum and red blood cell folate were significantly associated with plasma homocysteine ($r = -0.38$; $P < 0.0001$ and $r = -0.33$; $P < 0.0001$, respectively). Red blood cell folate and plasma homocysteine concentrations were significantly lower in women than in men ($P = 0.027$ and $P < 0.0001$, respectively) whereas serum folate was ~8% higher in women ($P = 0.0006$). Plasma total vitamin B-12 concentrations were significantly correlated with red blood cell folate ($r = 0.21$; $P < 0.0001$) and serum folate ($r = 0.18$; $P < 0.0001$) and negatively correlated with plasma homocysteine ($r = -0.28$; $P < 0.0001$).

Candidate gene analysis of effects of genetic polymorphism on folate status indicators

In the candidate gene analyses 76 SNPs from the GWAS chips were analyzed in 2232 subjects. The results were adjusted for gender, serum vitamin B-12 concentrations, and for whether or not subjects received folic acid supplements in the week before blood collection. In order to remove the effect of the *MTHFR* 677C→T variant (rs1801133), the analyses were repeated in 989 subjects who were homozygous for the C allele.

Red blood cell folate

For each SNP analyzed individually, red blood cell folate was significantly positively correlated with serum vitamin B-12 ($P = 1.5–4.2 \times 10^{-23}$, depending on SNP), with whether or not subjects consumed folate supplements ($P = 0.3–1.6 \times 10^{-11}$), and with gender ($P = 0.005–0.008$, lower in women). **Table 2** shows the SNPs that significantly influenced red blood cell folate concentrations after adjustment for multiple analyses. The most significant SNP was the rs1801133 677C→T variant of *MTHFR* ($P = 3.15 \times 10^{-17}$) with lower red blood cell folate in subjects with the minor T allele. Nine other SNPs, all in the *MTHFR* gene, also remained significant after Bonferroni adjustment, including the rs1801131 1298A→C variant, which has been reported to exhibit biological effects distinct from the *MTHFR* 677 variant (28, 29), but their *P* values were ≥5 orders of magnitude higher than for rs1801133. Biological effects of the *MTHFR* 1298 variant are somewhat controversial because the *MTHFR* 677 and 1298 variants share significant linkage disequilibrium (LD, D' = 0.997, $r^2 = 0.218$) and may be reporting on the same association signal. Consistent with this, the minor *MTHFR* 1298C allele was associated with higher red blood cell folate, i.e., the opposite effect to the *MTHFR* 677T allele. The other significant SNPs were located in the *MTHFR* gene region which exhibits high LD (**Supplemental Figure 2**). The significant association of these SNPs with red blood cell folate most likely represents a single association signal driven largely by *MTHFR* rs1801133.

TABLE 2
Candidate gene SNPs that significantly influenced red blood cell folate concentrations after adjusting for multiple analyses[1]

SNP	Gene	*P* SNP[2]	*P* SNP in *MTHFR* 677CC subjects[3] ($n = 989$)
rs1801133	*MTHFR*	3.15×10^{-17}	
rs3753584	*MTHFR*	5.88×10^{-12}	9.87×10^{-4}
rs4846052	*MTHFR*	2.16×10^{-9}	0.332
rs3737966	*MTHFR*	2.92×10^{-9}	0.342
rs4846054	*MTHFR*	3.95×10^{-9}	0.287
rs1537516	*MTHFR*	6.18×10^{-7}	0.020
rs1801131	*MTHFR*	1.14×10^{-6}	0.217
rs17037397	*MTHFR*	1.96×10^{-6}	0.010
rs13306567	*MTHFR*	1.21×10^{-4}	0.081
rs1476413	*MTHFR*	1.43×10^{-4}	0.464
SNP pairs[4]			
rs1801133	*MTHFR*	3.93×10^{-11}	
rs3753584	*MTHFR*	1.03×10^{-5}	
rs1801133	*MTHFR*	2.74×10^{-12}	
rs1801131	*MTHFR*	0.32	

[1]Data from 2232 subjects were adjusted for gender, serum vitamin B-12 concentration, and folic acid supplements. SNPs were tested with the use of a simple linear regression model executed in PLINK version 1.7 under the assumption of an additive genetic model. *MTHFR*, 5,10-methylenetetrahydrofolate reductase; SNP, single nucleotide polymorphism.

[2]Bonferroni threshold for SNPs: $P = 6.6 \times 10^{-4}$ (0.05/76).

[3]Bonferroni threshold for SNPs in *MTHFR* 677CC subjects: $P = 6.7 \times 10^{-4}$ (0.05/75).

[4]Bonferroni threshold adjustment for SNP pairs: $P = 1.75 \times 10^{-5}$ (0.05/(76 × 75/2)).

The analysis of the *MTHFR* SNPs was repeated as SNP pairs with rs1801133 (*MTHFR* 677) to attempt to adjust for the possible effect of LD with rs1801133. When we did this, only rs3753584 remained significant with a *P* value of 1.03×10^{-5}, as opposed to a *P* value of 3.93×10^{-11} for rs1801133 in the same analysis (Table 2). Finally, to eliminate any possibility that significant association was due to LD with rs1801133, we repeated the individual single SNP analyses in 989 subjects who were homozygous for the *MTHFR* 677C allele (Table 2). None of the 75 SNPs analyzed in these subjects was significant after Bonferroni adjustment and most were nonsignificant even without adjustment (i.e., $P > 0.05$). However, a strong signal was still seen for rs3753584 and its *P* value was only slightly above that required for significance. It may not have reached significance in *MTHFR* 677CC subjects because of the smaller number of subjects. Subjects expressing the minor G allele of rs3753584 had significantly increased red blood cell folate concentrations (**Table 3**). However, this variant shares significant LD with rs1801133 and the rs3753584 G variant is only seen on a *MTHFR* 677C background (Table 3, Supplemental Figure 2). Part of the increased red blood cell folate noted in rs3753584 G subjects was due to the absence of *MTHFR* 677T alleles. However, a significant increase in red blood cell folate was also observed with the rs3753584 G allele in *MTHFR* 677CC subjects (Table 3) and the rs3753584 variant does appear to be a second independent modifier of folate concentrations.

Serum folate

For each SNP analyzed individually, serum folate was significantly positively correlated with serum vitamin B-12 ($P = 1.0$–2.5×10^{-20}, depending on the SNP), with whether or not subjects consumed folate supplements ($P = 0.9$–1.7×10^{-15}), and with gender ($P = 0.003$–0.004, higher in women).

Table 4 shows the SNPs that significantly influenced serum folate concentrations after adjustment for multiple analyses. The most significant SNP was the rs1801133 677C→T variant of *MTHFR* ($P = 3.01 \times 10^{-11}$) with lower serum folate in subjects with the minor T allele. Four other SNPs, all in *MTHFR* and including rs3753584, remained significant after Bonferroni adjustment, although their *P* values were 6 orders of magnitude higher than for rs1801133. When analyzed in subjects homozygous for the C allele of rs1801133, none of these SNPs was statistically significant (Table 4). Two other *MTHFR* SNPs did achieve significance when analyzed in *MTHFR* 677CC subjects (rs1476413 and rs2066471, see Table 4) but neither was significant when initially analyzed in the total population.

Although serum folate was highly correlated with red blood cell folate, fewer significant SNPs were detected in the serum folate analysis than in the red blood cell folate analysis, and *P* values were higher. This probably reflected the greater variation in serum folate and also the fact that the serum samples collected in this study were nonfasting.

Plasma total homocysteine

Table 5 shows the SNPs that significantly influenced plasma homocysteine concentrations after adjustment for multiple analyses. Again, the most significant SNP was the rs1801133 677C→T variant of *MTHFR* ($P = 1.64 \times 10^{-16}$) with higher plasma homocysteine in subjects with the minor T allele. Two other SNPs in the *MTHFR* gene also remained significant after Bonferroni adjustment, but with much higher *P* values than for rs1801133. Neither of these SNPs was significant when analyzed in *MTHFR* 677CC subjects. The *MTHFR* 1298 variant (rs1801131) had no significant effect on homocysteine concentrations when analyzed in the total population or in *MTHFR* 677CC subjects. Of the 75 SNPs analyzed in *MTHFR* 677CC subjects, an SNP in *FPGS* had the lowest *P* value, although this did not reach statistical significance after Bonferroni adjustment (Table 5), and this SNP was not significant when analyzed in the total population.

GWAS analyses of folate status indicators

The results from the GWAS data analyses are shown in **Supplemental Tables 2–4**. The data shown were analyzed with the use of log-transformed data adjusted for age and gender, but essentially identical results were found with the use of inverse rank-ordered data, as were used in the candidate gene analyses.

Red blood cell and serum folate

Supplemental Table 2 shows the 32 SNPs affecting red blood cell folate that reached genome-wide significance ($P < 5 \times 10^{-08}$). All 32 SNPs were on chromosome 1 in the region of the *MTHFR* gene. rs1801133 was the most significant ($P = 1.37 \times 10^{-17}$) and accounted for ~7% of the variation

TABLE 3
Effect of *MTHFR* rs3753584 variants on red blood cell folate concentrations in *MTHFR* 677 C→T variants (rs1801133)[1]

		Red blood cell folate (nmol/L), mean ± SD *MTHFR* rs3753584			
		AA	AG	GG	*P* value
Total	—	1040 ± 422 ($n = 1558$)	1145 ± 430 ($n = 599$)	1255 ± 467 ($n = 73$)	<0.0001[2]
MTHFR 677	CC	1096 ± 425 ($n = 554$)	1151 ± 436 ($n = 362$)	1255 ± 467 ($n = 73$)	0.0043[3]
MTHFR 677	CT	1048 ± 407 ($n = 745$)	1136 ± 422 ($n = 237$)	NA ($n = 0$)	0.0072[4]
MTHFR 677	TT	897 ± 425 ($n = 142$)	NA ($n = 0$)	NA ($n = 0$)	

[1] *MTHFR*, 5,10-methylenetetrahydrofolate reductase.
[2] Kruskal-Wallis test.
[3] Kruskal-Wallis test. In pairwise post-test comparisons that used Dunn's multiple comparison tests, only AA and GG differed significantly.
[4] Mann-Whitney U test.

in red blood cell folate concentrations (H^2a, Supplemental Table 2). The next-most significant SNP had a *P* value >4 orders of magnitude higher. None of these SNPs reached genome-wide significance when analyzed in *MTHFR* 677CC subjects. Supplemental Table 3 shows similar data for serum folate. The highest-ranked SNP was rs1801133 ($P = 2.82 \times 10^{-11}$) and the 4 other lower-ranked SNPs that reached genome-wide significance were not significant when analyzed in *MTHFR* 677CC subjects. As was found in the candidate gene analysis, *P* values for SNPs were higher for serum folate compared with red blood cell folate.

Plasma total homocysteine

Supplemental Table 4 shows the 29 SNPs affecting plasma homocysteine that reached genome-wide significance ($P < 5 \times 10^{-8}$). Again, the most significant SNP was rs1801133 ($P = 1.26 \times 10^{-19}$), which accounted for ~4% of the variation in plasma homocysteine concentrations. An additional 22 chromosome-1 SNPs in the vicinity of the *MTHFR* gene had *P* values ≥8 orders of magnitude higher than rs1801131 but still reached genome-wide significance. None of these SNPs had any significant effect on homocysteine concentrations in *MTHFR* 677CC subjects.

Additional genome-wide significant SNPs were detected on chromosome 7 in the vicinity of the Twist neighbor B (*TWISTNB*) gene and on chromosome 16 in the region of the chromatin modifying protein 1A isoform (*CHMP1A*) gene (Supplemental Table 4). A strong signal was retained in the chromosome 16 region in *MTHFR* 677CC subjects, although the *P* value increased to above that representing genome-wide significance, possibly reflecting the reduced size of the *MTHFR* 677CC population. The chromosome 7 signal around the *TWISTNB* gene was considerably muted in *MTHFR* 677CC subjects.

The reduced *TWISTNB* signal in 677CC subjects was somewhat surprising. Because one possible explanation could be an interaction between the *MTHFR* gene on chromosome 1 and the *TWISTNB* gene on chromosome 7, we further analyzed the effects of the *TWISTNB* (chromosome 7) and *CHMP1A*

TABLE 4
Candidate gene SNPs that significantly influenced serum folate concentrations after adjusting for multiple analyses[1]

SNP	Gene	*P* SNP[2]	*P* SNP in *MTHFR* 677CC subjects[3]
rs1801133	*MTHFR*	3.01×10^{-11}	—
rs9651118	*MTHFR*	3.45×10^{-5}	0.0455
rs13306567	*MTHFR*	1.55×10^{-4}	0.0547
rs17037397	*MTHFR*	3.25×10^{-4}	0.146
rs3753584	*MTHFR*	6.61×10^{-4}	0.521
rs1476413	*MTHFR*	0.471	1.39×10^{-4}
rs2066471	*MTHFR*	0.128	5.58×10^{-4}

[1] Data were adjusted for gender, serum vitamin B-12 concentration, and folic acid supplements. SNPs were tested via a simple linear regression model executed in PLINK version 1.7 under the assumption of an additive genetic model. Bonferroni adjustment for SNP pairs ($P = 1.75 \times 10^{-5}$ [0.05/(76 × 75/2)]) was not satisfied for any SNP after adjusting for the effect of the *MTHFR* rs1801133 variant. *MTHFR*, 5,10-methylenetetrahydrofolate reductase; SNP, single nucleotide polymorphism.
[2] Bonferroni threshold for SNPs: $P = 6.6 \times 10^{-4}$ (0.05/76).
[3] Bonferroni threshold for SNPs in *MTHFR* 677CC subjects: $P = 6.7 \times 10^{-4}$ (0.05/75). 989 subjects.

TABLE 5
Candidate gene SNPs that significantly influenced plasma total homocysteine concentrations after adjusting for multiple analyses[1]

SNP	Gene	*P* SNP[2]	*P* SNP in *MTHFR* 677CC subjects[3]
rs1801133	*MTHFR*	1.64×10^{-16}	
rs9651118	*MTHFR*	7.58×10^{-7}	0.100
rs17037397	*MTHFR*	3.87×10^{-6}	0.061
rs1801131	*MTHFR*	0.074	0.036
rs10760503	*FPGS*	0.052	0.005

[1] Data were adjusted for gender, serum vitamin B-12 concentration, and folic acid supplements. SNPs were tested via a simple linear regression model executed in PLINK version 1.7 under the assumption of an additive genetic model. *FPGS*, folylpolyglutamate synthetase; *MTHFR*, 5,10-methylenetetrahydrofolate reductase; SNP, single nucleotide polymorphism.
[2] Bonferroni threshold for SNPs: $P = 6.6 \times 10^{-4}$ (0.05/76).
[3] Bonferroni threshold for SNPs in *MTHFR* 677CC subjects: $P = 6.7 \times 10^{-4}$ (0.05/75). 989 subjects.

TABLE 6
Effect of *TWISTNB* (rs17354370) and *CHMP1A* (rs164746) variants on homocysteine concentrations in each *MTHFR* 677 C→T genotype group (rs1801133)[1]

		TWISTNB				*CHMP1A*			
		Homocysteine (μmol/L), mean ± SD				Homocysteine (μmol/L), mean ± SD			
		AA	AG	GG	*P* value[2]	CC	CT	TT	*P* value[2]
MTHFR	CC	8.4 ± 2.0 (n = 469)	8.1 ± 2.1 (n = 422)	7.8 ± 1.8 (n = 98)	0.0047	7.9 ± 1.8 (n = 319)	8.2 ± 2.1 (n = 474)	8.7 ± 2.2 (n = 196)	5.76×10^{-5}
MTHFR	CT	8.9 ± 2.9 (n = 486)	8.3 ± 2.3 (n = 415)	8.0 ± 1.8 (n = 82)	2.35×10^{-5}	8.4 ± 3.2 (n = 313)	8.5 ± 2.3 (n = 455)	8.9 ± 2.2 (n = 213)	5.67×10^{-5}
MTHFR	TT	11.0 ± 5.7 (n = 142)	10.3 ± 5.8 (n = 99)	9.2 ± 3.2 (n = 18)	0.35	9.8 ± 3.6 (n = 72)	11.1 ± 6.3 (n = 138)	10.4 ± 5.7 (n = 49)	>0.05

[1] *CHMP1A*, chromatin modifying protein 1A isoform; *MTHFR*, 5,10-methylenetetrahydrofolate reductase; *TWISTNB*, Twist neighbor B.
[2] The *P* values are from the *F* test of ANOVA.

(chromosome 16) variants on homocysteine concentrations in subjects with each of the different *MTHFR* 677 variants (**Table 6**). The minor allele (G) of the *TWISTNB* (rs17354370) variant decreased homocysteine concentrations in all *MTHFR* 677 genotypes. This effect, which appeared to be quantitatively smaller than the change in homocysteine caused by the minor 677T allele of *MTHFR*, was most significant in *MTHFR* 677CT individuals. The lack of a significant effect in *MTHFR* 677TT individuals appeared to be due to the reduced number of subjects and the greater variation of homocysteine concentrations in these subjects. Consequently, these data do not support an interaction between the *TWISTNB* and *MTHFR* SNPs, but do not exclude the possibility. The minor T allele of the *CHMP1A* (rs164746) variant was associated with increased plasma homocysteine concentrations in all 3 *MTHFR* 677 genotypes. This effect was not statistically significant in *MTHFR* 67TT subjects, possibly due to the smaller number of subjects and increased variation in this group (Table 6).

DISCUSSION

In this study we explored genetic modifiers of 3 folate status biomarkers in a young, healthy Irish population. Previous small-scale studies have shown that a common polymorphism in the *MTHFR* gene (rs1801133, 677C→T) lowers serum (plasma) and red blood cell folate concentrations. However, multiple GWAS analyses, with the exception of a large-scale study on the Icelandic population (19), have failed to find genome-wide significant markers that influence plasma folate concentrations and, in some cases, have failed to find any suggestion of an effect of *MTHFR* gene variants. To our knowledge, a red blood cell folate GWAS analysis has not previously been reported.

In our candidate gene analysis, 10 SNPs in the *MTHFR* gene significantly affected red blood cell folate after adjustment for multiple analyses, with the *MTHFR* 677C→T (rs1801133) variant giving the strongest signal by far (Table 2). After adjusting for LD of the additional 9 positive *MTHFR* SNPs with rs1801133 by doing SNP pair analyses, rs3753584 was the only additional variant that remained significant. Of note, the signal for the *MTHFR* 1298A→C variant (rs1801131) completely disappeared.

Standard analyses that use a simple linear regression model and assume an additive genetic effect to separate the effects of 2 SNPs through the use of pair analysis may not be very accurate when comparing some variants such as *MTHFR* 677C→T. The T allele changes an alanine (A) in the *MTHFR* protein to a valine (V), which decreases the affinity of its FAD cofactor. Loss of the flavin cofactor causes conversion of the active dimeric enzyme to a less stable monomeric form, and increased turnover of the protein. Binding of the folate substrate stabilizes the bound FAD on the enzyme. Consequently, the increased loss of enzyme activity observed in subjects homozygous for the 677T variant, and also metabolic effects such as lowered serum and red blood cell folate and increased plasma homocysteine, are seen only in individuals with poorer folate and riboflavin status. Changes are seen in subjects heterozygous for the *MTHFR* 677C→T variant, but the extent of these changes is less than half that observed in 677T homozygotes (10). Therefore, the additive genetic model, where 1 T allele would have half the effect of 2 T alleles, does not model the actual situation that well. Because the active *MTHFR* enzyme is a dimer, one-quarter of enzyme in heterozygous subjects would be A/A dimers (as in 677C homozygotes), half would be A/V dimers, and one-quarter V/V dimers (as in 677T homozygotes). If the A/V dimeric form is as stable as the A/A dimer, the effects of homozygosity for 677T would be 4 times that of the heterozygote. Because of this, we repeated the analysis using only subjects homozygous for the *MTHFR* 677C variant, because this eliminates any effect of the T allele. The intronic rs3753584 variant was the only additional candidate *MTHFR* SNP to exhibit a significant effect on red blood cell folate when studied in *MTHFR* 677C homozygotes. The increase in red blood cell folate associated with the minor rs3753584 G allele could be explained by increased MTHFR expression, but we have no direct evidence for this. Over 30 SNPs with genome-wide significance were found in the GWAS analysis of red blood cell folate. All were in the region of the *MTHFR* gene and the *MTHFR* 677C→T (rs1801133) variant exhibited the strongest signal (Supplemental Table 2). None of the SNPs was close to exhibiting genome-wide significance in *MTHFR* 677C homozygotes. Thus, rs1801133 is the major genetic modifier of red blood cell folate concentrations in this population.

Similar results were seen in the candidate gene and GWAS analyses for serum folate, although a more limited number of significant SNPs, all in the *MTHFR* gene region, were found. Again, the *MTHFR* 677C→T (rs1801133) variant gave the strongest signal and none of the other SNPs was close to being

significant in *MTHFR* 677C homozygotes. It is likely that the smaller number of initial hits in the serum folate analysis than in the red blood cell folate analysis reflects the fact that the serum samples were nonfasting, but it may also reflect the fact that redblood cell folate is a more stable biomarker of folate status than serum or plasma folate.

Although we are unaware of any previous GWAS analysis of red blood cell folate, several analyses of plasma folate genetic modifiers have been published (11, 12, 19, 20). The only study to report genome-wide significant SNPs found 2 significant modifiers: rs1801133 (*MTHFR* 677C→T) and rs652197, a SNP in the *FOLR3* gene, encoding 1 of the folate receptors on chromosome 11 (19). We were unable to detect a significant signal for the folate receptor *FOLR1*, which is in close proximity to *FOLR3*, in our candidate gene analysis or any signal on chromosome 11 in our GWAS analysis. Other GWASs did not find any significant signal for plasma folate but did find a modest association in the fidgetin (*FIGN)* gene region on chromosome 2 (11) or in the myelin transcription factor 1 like (*MYT1L)* gene region on chromosome 2 (20), neither of which we were able to confirm. Modest signals for various *MTHFR* SNPs, including rs3737965 (11) and rs1999594 (12), have been reported. We did find a genome-wide significant association for rs1999594 in our serum folate GWAS (Supplemental Table 3) but it was not significant in subjects homozygous for the *MTHFR* 677C variant. It is likely that the reason why only 2 GWASs found genome-wide significance for rs1801133 as a modifier of serum or plasma folate is that the populations studied [Icelandic (19) and Irish (current study)] were genetically more homogeneous than in the other studies.

The *MTHFR* 677C→T (rs1801133) variant was the most significant signal in the candidate gene analysis of plasma homocysteine concentrations. Two other SNPs in *MTHFR* also reached significance but their signals were completely lost when analyzed in *MTHFR* 677C homozygotes. Twenty-nine SNPs reached genome-wide significance in our homocysteine GWAS analysis. Twenty-three were in the *MTHFR* region including rs1801133, which gave the strongest signal. None of the chromosome 1 *MTHFR* SNPs were significant in *MTHFR* 677C homozygotes, indicating they were reporting on rs1801133. No effect of the *MTHFR* 1298A→C variant (rs1801131) on homocysteine concentrations was detected.

rs1801133 has previously been reported to be a genome-wide significant modifier of plasma homocysteine in a number of GWASs (11, 12, 14, 18). Hazra et al. (11) also reported that 2 other SNPs close to *MTHFR* (rs12085006 and rs1999594) had stronger associations with homocysteine than rs1801133, and these associations remained nominally statistically significant after mutual adjustment for rs1801133 effects in a multivariate regression model. We also found genome-wide associations for rs12085006 and rs1999594 in our homocysteine GWAS (Supplemental Table 4), but these associations were completely absent in subjects homozygous for the *MTHFR* 677C variant.

Significant signals for homocysteine were also found in our study on chromosome 16 in the region of the *CHMP1A* gene (2 SNPs) and on chromosome 7 in the region of the *TWISTNB* gene (4 SNPs) (Supplemental Table 4, Table 6). *CHMP1A* rs164746 is essentially a perfect proxy with dipeptidase 1 *(DPEP1)* rs154657 ($D' = 1$; $r^2 = 0.998$ in the Trinity Student population). *DPEP1* has been reported to be a GWAS hit for homocysteine (14, 18). *DPEP1* (EC 3.4.13.11) is a kidney membrane enzyme that hydrolyzes a variety of dipeptides and is implicated in renal metabolism of glutathione and its conjugates (30). *DPEP1* is also a genome-wide signal for plasma total cysteine in the Trinity Student Study with the minor allele associated with higher cysteine, as was found for total homocysteine (unpublished data).

The 4 SNPs detected in the *TWISTNB* region of chromosome 7 (7p21.1) have not previously been reported as modifiers of plasma homocysteine concentrations. *TWISTNB* encodes a DNA-dependent RNA polymerase that is a component of RNA polymerase I which synthesizes ribosomal RNA precursors. van Meurs et al. (18) did report a significant chromosome 7 signal but this was in the GTP-binding protein 10 (*GTPBP10*, 7q21.13) region, which is distant from *TWISTNB*, and we did not observe any signal in this region.

Our studies clearly indicate that the *MTHFR* 677C→T (rs1801133) variant is the major genetic modifier of folate status biomarkers and may be the only significant modifier of serum (plasma) and red blood cell folate concentrations, at least in young, healthy adults in the Irish population. The question arises of whether this confounds the use of these biomarkers in assessing folate status, and whether determining an individual's rs1801133 genotype would be of benefit in interpreting folate status when using these biomarkers. It has previously been established (4, 6, 10), and has been shown here, that the T allele of *MTHFR* 677 causes decreases in both serum and red blood cell folate, the latter being a proxy for tissue folate, which would indicate an impairment in folate status. Plasma homocysteine is also increased, indicative of poorer folate function in tissues. Because the poorer folate status caused by this variant is reflected in the changes in the concentrations of these biomarkers for folate status, there does not appear to be any added benefit in genotyping subjects for this variant when interpreting folate status through the use of these biomarkers. Carriers of the T allele are also responsive to folate intake (31), which indicates that a T allele–dependent low folate status is responsive to treatment, also making genetic testing unnecessary.

We acknowledge the contributions made by the participants in the Trinity Student Study.

The authors' responsibilities were as follows: BS, AMM, LCB, and JLM: designed the research; AMM, JLM, FP, and PMU: collected and analyzed the data; RF, TG, CDC, YK, JEB-W, AFW, and AMM: performed statistical analysis; BS and AMM: wrote the initial draft of the manuscript; FP, JLM, BS, and LCB: critically revised the manuscript; and all authors: had responsibility for accuracy of the final content, and read and approved the final manuscript. None of the authors reported a conflict of interest related to the study.

REFERENCES

1. Yetley EA, Pfeiffer CM, Phinney KW, Fazili Z, Lacher DA, Bailey RL, Blackmore S, Bock JL, Brody LC, Carmel R,, et al. Biomarkers of folate status in NHANES: a roundtable summary. Am J Clin Nutr 2011;94:303S–12S.
2. Chanarin I. The megaloblastic anaemias. 3rd ed. Oxford: Blackwell Scientific; 1990.
3. Refsum H, Smith AD, Ueland PM, Nexo E, Clarke R, McPartlin J, Johnston C, Engbaek F, Schneede J, McPartlin C,, et al. Facts and recommendations about total homocysteine determinations: an expert opinion. Clin Chem 2004;50:3–32.
4. Molloy AM, Daly S, Mills JL, Kirke PN, Whitehead AS, Ramsbottom D, Conley MR, Weir DG, Scott JM. Thermolabile variant of

5,10-methylenetetrahydrofolate reductase associated with low red-cell folates: implications for folate intake recommendations. Lancet 1997;349(9065):1591–3.
5. Christensen B, Arbour L, Tran P, Leclerc D, Sabbaghian N, Platt R, Gilfix BM, Rosenblatt DS, Gravel RA, Forbes P,, et al. Genetic polymorphisms in methylenetetrahydrofolate reductase and methionine synthase, folate levels in red blood cells, and risk of neural tube defects. Am J Med Genet 1999;84:151–7.
6. Jacques PF, Kalmbach R, Bagley PJ, Russo GT, Rogers G, Wilson PW, Rosenberg IH, Selhub J. The relationship between riboflavin and plasma total homocysteine in the Framingham Offspring cohort is influenced by folate status and the C677T transition in the methylenetetrahydrofolate reductase gene. J Nutr 2002;132:283–8.
7. Bueno O, Molloy AM, Fernandez-Ballart JD, Garcia-Minguillan CJ, Ceruelo S, Rios L, Ueland PM, Meyer K, Murphy MM. Common polymorphisms that affect folate transport or metabolism modify the effect of the MTHFR 677C >T polymorphism on folate status. J Nutr 2016;146:1–8.
8. Stanislawska-Sachadyn A, Mitchell LE, Woodside JV, Buckley PT, Kealey C, Young IS, Scott JM, Murray L, Boreham CA, McNulty H,, et al. The reduced folate carrier (SLC19A1) c.80G>A polymorphism is associated with red cell folate concentrations among women. Ann Hum Genet 2009;73:484–91.
9. Parle-McDermott A, Mills JL, Molloy AM, Carroll N, Kirke PN, Cox C, Conley MR, Pangilinan FJ, Brody LC, Scott JM. The MTHFR 1298CC and 677TT genotypes have opposite associations with red cell folate levels. Molec Genet Metab 2006;88:290–4.
10. McNulty H, McKinley MC, Wilson B, McPartlin J, Strain JJ, Weir DG, Scott JM. Impaired functioning of thermolabile methylenetetrahydrofolate reductase is dependent on riboflavin status: implications for riboflavin requirements. Am J Clin Nutr 2002;76:436–41.
11. Hazra A, Kraft P, Lazarus R, Chen C, Chanock SJ, Jacques P, Selhub J, Hunter DJ. Genome-wide significant predictors of metabolites in the one-carbon metabolism pathway. Hum Molec Genet 2009;18:4677–87.
12. Tanaka T, Scheet P, Giusti B, Bandinelli S, Piras MG, Usala G, Lai S, Mulas A, Corsi AM, Vestrini A,, et al. Genome-wide association study of vitamin B6, vitamin B12, folate, and homocysteine blood concentrations. Am J Hum Genet 2009;84:477–82.
13. Williams SR, Yang Q, Chen F, Liu X, Keene KL, Jacques P, Chen WM, Weinstein G, Hsu FC, Beiser A,, et al. Genome-wide meta-analysis of homocysteine and methionine metabolism identifies five one carbon metabolism loci and a novel association of ALDH1L1 with ischemic stroke. PLoS Genet 2014;10(3):e1004214.
14. Pare G, Chasman DI, Parker AN, Zee RR, Malarstig A, Seedorf U, Collins R, Watkins H, Hamsten A, Miletich JP,, et al. Novel associations of CPS1, MUT, NOX4, and DPEP1 with plasma homocysteine in a healthy population: a genome-wide evaluation of 13 974 participants in the Women's Genome Health Study. Circ Cardio Genet 2009;2:142–50.
15. Lange LA, Croteau-Chonka DC, Marvelle AF, Qin L, Gaulton KJ, Kuzawa CW, McDade TW, Wang Y, Li Y, Levy S,, et al. Genome-wide association study of homocysteine levels in Filipinos provides evidence for CPS1 in women and a stronger MTHFR effect in young adults. Hum Molec Genet 2010;19:2050–8.
16. Malarstig A, Buil A, Souto JC, Clarke R, Blanco-Vaca F, Fontcuberta J, Peden J, Andersen M, Silveira A, Barlera S,, et al. Identification of ZNF366 and PTPRD as novel determinants of plasma homocysteine in a family-based genome-wide association study. Blood 2009;114:1417–22.
17. Raffield LM, Ellis J, Olson NC, Duan Q, Li J, Durda P, Pankratz N, Keating BJ, Wassel CL, Cushman M,, et al. Genome-wide association study of homocysteine in African Americans from the Jackson Heart Study, the Multi-Ethnic Study of Atherosclerosis, and the Coronary Artery Risk in Young Adults study. J Hum Genet 2018;63:327–37.
18. van Meurs JB, Pare G, Schwartz SM, Hazra A, Tanaka T, Vermeulen SH, Cotlarciuc I, Yuan X, Malarstig A, Bandinelli S,, et al. Common genetic loci influencing plasma homocysteine concentrations and their effect on risk of coronary artery disease. Am J Clin Nutr 2013;98:668–76.
19. Grarup N, Sulem P, Sandholt CH, Thorleifsson G, Ahluwalia TS, Steinthorsdottir V, Bjarnason H, Gudbjartsson DF, Magnusson OT, Sparso T,, et al. Genetic architecture of vitamin B12 and folate levels uncovered applying deeply sequenced large datasets. PLoS Genet 2013;9(6):e1003530.
20. Keene KL, Chen WM, Chen F, Williams SR, Elkhatib SD, Hsu FC, Mychaleckyj JC, Doheny KF, Pugh EW, Ling H,, et al. Genetic associations with plasma B12, B6, and folate levels in an ischemic stroke population from the Vitamin Intervention for Stroke Prevention (VISP) trial. Front Pub Health 2014;2:112.
21. Carter TC, Pangilinan F, Molloy AM, Fan R, Wang Y, Shane B, Gibney ER, Midttun O, Ueland PM, Cropp CD,, et al. Common variants at putative regulatory sites of the tissue nonspecific alkaline phosphatase gene influence circulating pyridoxal 5′-phosphate concentration in healthy adults. J Nutr 2015;145:1386–93.
22. Deac OM, Mills JL, Shane B, Midttun O, Ueland PM, Brosnan JT, Brosnan ME, Laird E, Gibney ER, Fan R,, et al. Tryptophan catabolism and vitamin B-6 status are affected by gender and lifestyle factors in healthy young adults. J Nutr 2015;145:701–7.
23. Mills JL, Carter TC, Scott JM, Troendle JF, Gibney ER, Shane B, Kirke PN, Ueland PM, Brody LC, Molloy AM. Do high blood folate concentrations exacerbate metabolic abnormalities in people with low vitamin B-12 status? Am J Clin Nutr 2011;94:495–500.
24. Stone N, Pangilinan F, Molloy AM, Shane B, Scott JM, Ueland PM, Mills JL, Kirke PN, Sethupathy P, Brody LC. Bioinformatic and genetic association analysis of microRNA target sites in one-carbon metabolism genes. PloS One 2011;6(7):e21851.
25. Molloy AM, Scott JM. Microbiological assay for serum, plasma, and red cell folate using cryopreserved, microtiter plate method. Meth Enzymol 1997;281:43–53.
26. Windelberg A, Arseth O, Kvalheim G, Ueland PM. Automated assay for the determination of methylmalonic acid, total homocysteine, and related amino acids in human serum or plasma by means of methylchloroformate derivatization and gas chromatography-mass spectrometry. Clin Chem 2005;51:2103–9.
27. Molloy AM, Pangilinan F, Mills JL, Shane B, O'Neill MB, McGaughey DM, Velkova A, Abaan HO, Ueland PM, McNulty H,, et al. A common polymorphism in HIBCH influences methylmalonic acid concentrations in blood independently of cobalamin. Am J Hum Genet 2016;98:869–82.
28. Skibola CF, Smith MT, Hubbard A, Shane B, Roberts AC, Law GR, Rollinson S, Roman E, Cartwright RA, Morgan GJ. Polymorphisms in the thymidylate synthase and serine hydroxymethyltransferase genes and risk of adult acute lymphocytic leukemia. Blood 2002;99:3786–91.
29. van der Put NM, Gabreëls F, Stevens EM, Smeitink JA, Trijbels FJ, Eskes TK, van den Heuvel LP, Blom HJ. A second common mutation in the methylenetetrahydrofolate reductase gene: an additional risk factor for neural-tube defects? Am J Hum Genet 1998;62:1044–51.
30. Kozak EM, Tate SS. Glutathione-degrading enzymes of microvillus membranes. J Biol Chem 1982;257:6322–7.
31. Crider KS, Zhu J-H, Hao L, Yang Q-H, Yang TP, Gindler J, Maneval DR, Quinlivan EP, Li Z, Bailey LB,, et al. MTHFR 677C->T genotype is associated with folate and homocysteine concentrations in a large, population-based, double-blind trial of folic acid supplementation. Am J Clin Nutr 2011;93:1365–72.