GENETIC AND NUTRITIONAL VARIATION IN THE FOLATE-MEDIATED ONE-CARBON METABOLIC NETWORK AND CARDIOVASCULAR DISEASE (CVD) RISK

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Susan Marie Wernimont

January 2011

GENETIC AND NUTRITIONAL VARIATION IN THE FOLATE-MEDIATED ONE-CARBON METABOLIC NETWORK AND CARDIOVASCULAR DISEASE (CVD) RISK

Susan Marie Wernimont, Ph. D.

Cornell University 2011

The aim of this research was to investigate the role of genetic and nutritional variation within the folate-mediated one-carbon network in relation to cardiovascular disease risk. The enzymes serine hydroxymethyltransferase 1 (gene name *SHMT1*) and methylenetetrahydrofolate reductase (gene name *MTHFR*) regulate key reactions in folate-mediated one-carbon metabolism. We investigated the effect of the *SHMT1* rs1979277 SNP and the *SHMT1* rs1979277 – *MTHFR* rs1801133 interaction in two epidemiologic cohorts. In the Nurses' Health Study, the *MTHFR* rs1801133 variant genotypes were associated with an increased CVD risk, and there was an interaction between *SHMT1* and *MTHFR* such that the association of *MTHFR* rs1801133 *CT* genotype (vs. *CC;* the *TT* genotype could not be evaluated) was stronger in the presence of the *SHMT1* rs1979277 *TT* genotype. In the Health Professionals Follow-Up Study, the *MTHFR* rs1801133 genotype was not associated with CVD risk nor was there an interaction with *SHMT1* rs1979277. Next, using data from the Normative Aging Study, 330 SNPs in 52 genes were studied in relation to cardiovascular disease biomarkers. Using a nominal significance threshold of $P \leq 0.005$, 20 SNPs were associated with homocysteine, 8 with Alu methylation, and 1 with LINE-1 methylation. Using a more stringent false discovery rate threshold, SNPs in *FTCD*, *SLC19A1*, and *SLC19A3* genes were associated with plasma homocysteine, gene x vitamin B-6 interactions were identified for Alu and LINE-1 methylation, and epistatic

interactions involving the *MTHFR* rs1801133 SNP were identified for the plasma homocysteine phenotype. Finally, the SNPs were prospectively evaluated for their association with cardiovascular disease in a U.S. population studied prior to mandatory folate fortification. Using a nominal significance threshold of $P \leq 0.005$, 8 SNPs were associated with CVD risk. Using a more stringent false discovery rate threshold, a polymorphism in the *GGH* gene was associated with reduced CVD risk. A gene x folate interaction was identified (*MAT2B*) and two gene x vitamin B-12 interactions were identified (*BHMT* and *SLC25A32*). Hypotheses related to *SHMT1* were explored and significant gene x gene interactions were identified. Overall, genetic variation in folate-mediated one-carbon metabolism, other than the well-known effects of the *MTHFR 677 C→T* rs1801133, is predictive of cardiovascular disease risk.

BIOGRAPHICAL SKETCH

Susan Marie Wernimont lived the first 18 yrs of her life in Carroll, Iowa, a small Midwestern farming community where her family has lived for over 100 yrs. Sue was the 6th of 8 children, and preferred to spend her time either outdoors or in the kitchen, which still is true today. Sue attended college at Iowa State University, where she majored in Agricultural Biochemistry, intending to become a veterinarian. After taking an elective course in nutrition, however, she knew she had found a way to link her love of food and cooking with her interests in biochemistry, and she added Nutritional Science as a second major. In 2000, Sue completed a study abroad experience at the University of Otago in Dunedin, New Zealand. Upon returning from New Zealand, Sue began her M.S. in Nutritional Biochemistry at Cornell University, where she worked in the lab of Dr. Andre Bensadoun, studying lipid metabolism and lipoprotein lipase turnover. After completing her M.S., Sue chose to pursue her interest in applied nutrition and population health, and completed her dietetic internship at the University of Houston, subsequently accepting a position in clinical nutrition at the Lahey Clinic Medical Center in Burlington, MA. Sue enjoyed clinical and applied nutrition, but ultimately moved back to upstate NY to obtain her Ph.D. at Cornell University. Prior to matriculating at Cornell in the fall of 2006, Sue completed a 500 mile bike ride across Iowa. She loves to travel and has completed multiple rock-climbing and mountaineering trips throughout the United States, including Washington state, Idaho, New Mexico, Texas, Arkansas, Wisconsin, Minnesota, New York, and Maine. In her free time, Sue enjoys cooking, knitting, reading, biking, camping, hiking, cross-country skiing, and kayaking.

ACKNOWLEDGMENTS

I would like to thank my committee, especially my advisor, Dr. Patricia A. Cassano, as well as Drs. Andrew G. Clark, Martin T. Wells, Patrick J. Stover, and Martha H. Stipanuk for their guidance in the development of my research project. I would also like to thank the NIH T32 training grant program, T32DK007158 (SMW), from the National Institute of Diabetes, Digestive and Kidney Diseases, the Division of Nutritional Sciences at Cornell University, and the American Dietetic Association, all of whom provided support for my graduate studies. Genotyping services for this work were provided by the Johns Hopkins University under federal contract number (N01-HV-48195) from the National Heart, Lung, and Blood Institute. I would like to thank the Cornell Statistical Consulting Unit for their assistance, as well as the current and former members of the Cassano research group for creating an atmosphere conducive to the development of my research project. Finally, I would like to thank my family and friends for their support.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF TABLES (Continued)

LIST OF TABLES (Continued)

LIST OF ABBREVIATIONS

% 5-meC, percentage of methylated cytosines
AdoHcy, S-adenosylhomocysteine
AdoMet, S-adenosylmethionine
AHCY, Adenosylhomocysteinase
AHCYL1, Adenosylhomocysteinase-like 1
AHCYL2, Adenosylhomocysteinase-like 2, KIAA0828
ALDH1L1, Aldehyde dehydrogenase 1 family, member L1
AMT, Aminomethyltransferase
ATIC, 5-aminoimidazole-4-carboxamide ribonucleotide formyltransferase/IMP cyclohydrolase
BHMT, Betaine-homocysteine S-methyltransferase
CAD, coronary artery disease
CBS, Cystathionine-beta-synthase
CELF1, CUGBP, Elav-like family member 1
CEPH, Centre d'Etude du Polymorphisme Humain
CHD, coronary heart disease
CTH, Cystathionase (cystathionine gamma-lyase)
CV, coefficient of variation
CVD, cardiovascular disease
DHFR, Dihydrofolate reductase
DMGDH, Dimethylglycine dehydrogenase
DNMT1, DNA (cytosine-5-)-methyltransferase 1
DNMT3A, DNA (cytosine-5-)-methyltransferase 3 alpha
DNMT3B, DNA (cytosine-5-)-methyltransferase 3 beta
FDR, False Discovery Rate
FOLH1, Folate hydrolase (prostate-specific membrane antigen) 1
FOLR1, Folate receptor 1 (adult)
FOLR2, Folate receptor 2 (fetal)
FOLR3, Folate receptor 3 (gamma)
FPGS, Folylpolyglutamate synthase
FTCD, Formiminotransferase cyclodeaminase
FTH1, Ferritin, heavy polypeptide 1
GART, Phosphoribosylglycinamide formyltransferase, phosphoribosylglycinamide synthetase, phosphoribosylaminoimidazole synthetase
GCSH, Glycine cleavage system protein H (aminomethyl carrier)
GGH, Gamma-glutamyl hydrolase (conjugase, folylpolygammaglutamyl hydrolase)

GLDC, Glycine dehydrogenase (decarboxylating)
GNMT, Glycine N-methyltransferase
HPFS, Health Professionals Follow-Up Study
HSPA8, Heat shock 70kDa protein 8
HWE, Hardy-Weinberg equilibrium
LD, linkage disequilibrium
MAF, minor allele frequency
MARS, Methionyl-tRNA synthetase
MAT1A, Methionine adenosyltransferase I, alpha
MAT2A, Methionine adenosyltransferase II, alpha
MAT2B, Methionine adenosyltransferase II, beta
MI, myocardial infarction
MTHFD1, Methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 1, methenyltetrahydrofolate cyclohydrolase, formyltetrahydrofolate synthetase
MTHFD1L, Methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 1-like
MTHFD2, Methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 2, methenyltetrahydrofolate cyclohydrolase
MTHFR, Methylenetetrahydrofolate reductase (NADPH)
MTHFS, 5,10-methenyltetrahydrofolate synthetase (5-formyltetrahydrofolate cyclo-ligase)
MTR, 5-methyltetrahydrofolate-homocysteine methyltransferase
MTRR, 5-methyltetrahydrofolate-homocysteine methyltransferase reductase
NAS, Normative Aging Study
NHS, Nurses' Health Study
PLP, pyridoxal-5'-phosphate
SARDH, Sarcosine dehydrogenase
SHMT1, Serine hydroxymethyltransferase 1 (soluble)
SHMT2, Serine hydroxymethyltransferase 2 (mitochondrial)
SLC19A1, Solute carrier family 19 (folate transporter), member 1
SLC19A2, Solute carrier family 19 (thiamine transporter), member 2
SLC19A3, Solute carrier family 19, member 3
SLC25A32, Solute carrier family 25, member 32
SLC46A1, Solute carrier family 46 (folate transporter), member 1
SNP, single nucleotide polymorphism
TCN1, Transcobalamin I (vitamin B-12 binding protein, R binder family)

TCN2, Transcobalamin II
THF, tetrahydrofolate
TYMS, Thymidylate synthetase
UBE2I, Ubiquitin-conjugating enzyme E2I (UBC9 homolog, yeast)
UBE2N, Ubiquitin-conjugating enzyme E2N (UBC13 homolog, yeast)
VA, Veterans' Administration

CHAPTER 1

INTRODUCTION

Cardiovascular Disease (CVD)

CVD (primarily coronary heart disease and stroke) is America's leading killer and was responsible for 36.3% of all U.S. deaths in 2004 (16, 86). Many risk factors for CVD have been identified, ranging from modifiable factors such as high blood cholesterol and lack of physical activity to non-modifiable factors including heredity and male gender. Like other chronic diseases, the origins of CVD are complex and likely comprise both the independent and interactive effects of genetic and environmental factors, including nutrition.

Folate-mediated one-carbon metabolism and CVD risk

Folate and other B vitamins play key roles in biologic processes important to health, including DNA synthesis and the generation of cellular methylation potential for a variety of methylation reactions. Folate status is influenced by both dietary intake and variation in genes encoding folate-related enzymes, and altered folate status due to nutritional or genetic perturbations is associated with adverse outcomes, including birth defects, cardiovascular disease (CVD), and cancer (32)

The reactions of one-carbon metabolism, through which folate carries out its biological functions, comprise an interconnected set of pathways with reactions occurring in both the cytoplasm and the mitochondria of the cell (96). The purpose of one-carbon metabolism is to use tetrahydrofolate to chemically activate one-carbon

units for a number of important reduction or oxidation reactions (49, 97). In the cytoplasm, the one-carbon units donated by folate are used for both nucleotide synthesis (5,10-methylenetetrahydrofolate is used as a coenzyme in the conversion of deoxyuridine monophosphate (dUMP) to deoxythymidine monophosphate (dTMP), contributing to thymidylate biosynthesis, and can also be oxidized to 10-formyltetrahydrofolate for purine synthesis) and the generation of methylation potential (MTHFR catalyzes the reduction of 5,10-methylenetetrahydrofolate to 5-methyltetrahyrofolate, which then donates a methyl group for the remethylation of homocysteine to methionine, or for the synthesis of S-adenosylmethionine (AdoMet)). AdoMet is considered the "universal methyl donor" because of its role as a cofactor in over 80 methylation reactions, including the methylation of DNA, RNA, proteins (including histones), and neurotransmitters (49, 97). Even after AdoMet has donated its methyl group to these reactions and been converted to S-adenosylhomocysteine (AdoHcy), it still has the ability to regulate cellular methylation through its role as a potent inhibitor of most AdoMet-dependent methyltransferases (49). One-carbon metabolism also occurs in the mitochondria, where formate (the source of most one-carbon units) is synthesized (97).

Individuals homozygous for a relatively common polymorphism, *MTHFR 677 C→T TT*, which causes an alanine to valine amino acid change at residue 222, resulting in reduced enzyme activity, have 25% higher homocysteine concentrations when compared to individuals homozygous for the *CC* genotype (109). Several publications have examined the role of the *MTHFR 677 C→T* polymorphism in relation to cardiovascular disease risk. In 2002, two meta-analyses were published. The first, by Wald et al, involved case-control studies examining the prevalence of the *MTHFR 677 C→T TT* genotype (72 studies) and prospective studies of homocysteine and CVD events (20 studies) (108). Wald and colleagues computed a summary odds

ratio of 1.21 (95% CI:1.06-1.39, p=0.006) for ischemic heart disease comparing the *MTHFR 677 C→T TT* vs. *CC* genotypes (48 studies); and a summary odds ratio of 1.31 (95% CI: 0.66-4.13) for stroke, comparing *MTHFR 677 C→T TT* vs. *CC* (7 studies) (108). The authors concluded that lowering homocysteine levels 3 μmol/l would reduce the risk of ischemic heart disease by 16% and stroke by 24%. The Homocysteine Studies Collaboration, published the same year, reported that a 25% lower usual homocysteine level (about 3 μmol/l) was associated with an 11% reduction in ischemic heart disease risk (OR 0.89; 95% CI: 0.83 – 0.96) and a 19% reduction in stroke risk (OR 0.81; 95% CI: 0.69 – 0.95), after adjustment for known CVD risk factors and regression dilution bias in prospective studies (1). A third meta analysis, also published in 2002, authored by Klerk et al and the *MTHFR* Studies Collaboration, involved 40 case-control studies of the *MTHFR 677C→T* polymorphism and risk of CHD (52), and found that individuals with the *MTHFR 677 C→T TT* genotype have a 16% higher odds of CHD compared to those with the *CC* genotype (OR 1.16; 95% CI: 1.05 – 1.28) (52). In 2005, Lewis et al conducted a meta-analysis of a total of 80 case-control and prospective studies of the association between the *MTHFR 677 C→T* variant and myocardial infarction, coronary artery occlusion, or both, which found that the *MTHFR 677 C→T TT* genotype vs. the *CC* genotype was associated with an increase of 2.2 μmol/l homocysteine and a 14% higher risk of ischemic heart disease (OR 1.14, 95% CI:1.05-1.24) (58). Although the interpretation of the findings in these publications differed (52, 58, 108), Wald et al in 2006 concluded that the results of these meta-analyses were consistent with a 16% lower risk of ischemic heart disease for a 3 μmol/l decrease in homocysteine (109). In addition, two meta-analyses focusing on ischemic stroke, a form of CVD, were published. The first reviewed data from 120 case-control candidate gene association studies of ischemic stroke and found a statistically significant association between

MTHFR 677 C→T and ischemic stroke (OR 1.24, 95% CI: 1.08-1.42, p=0.002) comparing *TT* to *CC*/*CT* genotypes (15). The second meta-analysis examined 31 retrospective case-control or cross-sectional studies or prospective cohort studies of *MTHFR 677 C→T* polymorphism in relation to stroke risk (27). This analysis found that the OR for stroke/transient ischemic attack (TIA) associated with the *MTHFR TT* vs. *CC* genotype was 1.37 (95% CI: 1.15-1.64, p<0.001). A graded dose response relationship was observed for increasing numbers of T alleles (OR 1.17 for *CT*, 95% CI: 1.09-1.26, p<0.001) and when outcomes were restricted to those events confirmed by imaging, an almost identical relation was observed.

However, a meta-analysis conducted in 2009 indicated that CVD risk associations involving other variants in the one carbon metabolic pathway have been far less well-studied (114). An examination of the role of the genetic variants in the one-carbon pathway as a whole is of interest as it is not currently known how a more complete description of genetically-driven alterations in folate metabolism influences CVD risk, thus leaving an important gap in the evidence base.

Folate-related CVD Biomarkers

A sulfur-containing amino acid, homocysteine plays an important role in the metabolism of the amino acid methionine. Homocysteine levels are governed by the reactions of the folate one-carbon metabolic network, an interconnected series of reactions that serve to transfer one-carbon units for the purpose of amino acid or nucleotide metabolism (96), and are also influenced by B vitamin status, particularly folate and to some extent, vitamin B-12 (45). Methylenetetrahydrofolate reductase (MTHFR) is an enzyme which plays a key role in the methionine cycle of the one carbon metabolic pathway, and converts 5,10-methylenetetrahydrofolate to 5-

methyltetrahydrofolate, which then serves as a cofactor for the remethylation of homocysteine to methionine, a reaction catalyzed by the B-12 dependent enzyme, methionine synthase (MTR), which produces tetrahydrofolate and methionine (96). As previously described, the *MTHFR 677 C→T* rs1801133 genetic variant, which results in an enzyme with reduced activity, has been associated with increased homocysteine levels in several meta-analyses (52, 58, 108).

Elevated plasma homocysteine, a sulfur-containing amino acid byproduct of folate metabolism, is a marker of disturbed folate-mediated one-carbon metabolism, and is associated with an increased risk of CVD, based on evidence from prospective studies, meta-analyses, and genetic studies (58, 87, 100, 108). In the past ten years, homocysteine has sparked interest as a modifiable CVD risk factor because blood homocysteine concentrations are amenable to change by alterations in dietary B-vitamin intake, a simple intervention.

Thus, supplementation with key nutrients from the folate-mediated one-carbon pathway, such as folate and B-12, leads to lower circulating homocysteine concentrations (45), and is hypothesized to lead ultimately to reductions in CVD risk. The observational studies described above led to the conduct of randomized clinical trials to test the hypothesis in an experimental design. Whether homocysteine represents an independent risk factor for CVD, or is simply a marker of CVD risk, has been vigorously debated in the literature, stimulated in part by results from three recently published randomized trials of B-vitamin supplementation and the occurrence or recurrence of CVD (the Heart Outcomes Prevention Evaluation (HOPE) 2 trial (65), the Norwegian Vitamin (NORVIT) trial (14), and the Vitamin Intervention for Stroke Prevention (VISP) trial (101)). The trials published to date have failed to show that supplementation led to a reduced risk of CVD despite achieved reductions in homocysteine levels. Wald et al suggest that these randomized trials lack the

statistical power needed to detect the expected modest reductions in CVD resulting from B-vitamin supplementation and that the relatively short follow-up (from 2 – 5 years) may be insufficient, given that the Kaplan-Meier estimates published in the HOPE 2 paper begin to show modest risk reduction only in the 3rd and 4th years of follow-up (65). The consistency in the cohort studies and the genetic studies, as well as the fact that among the genetic polymorphism studies, those with the greatest difference in homocysteine had the greatest difference in CVD risk (109) continue to support a causal role for homocysteine in the development of CVD. Gillies and Krul note that there are many examples of discordant results between epidemiologic studies and clinical trials, perhaps because the simplification of complex biological systems in order to generate testable clinical hypotheses may exceed the ability of these hypotheses to inform us about the biological systems from which they arose (39); this may indeed be the case for the homocysteine trials.

The association of homocysteine with CVD is hypothesized to be mediated, in part, by changes in DNA methylation status (48). DNA methylation is an epigenetic modification that involves the addition of methyl groups to cytosine to form 5-methylcytosine. Folate-mediated one-carbon metabolism is linked to DNA methylation status through regulation of AdoMet, the universal methyl donor, as well as through the activity of enzymes involved in methylation reactions, such as the DNA methyltransferases (82, 105). Changes in methylation influence gene expression, cellular differentiation and development, preservation of chromosomal integrity, and X chromosome inactivation, and methylation status is associated with the risk of CVD and some cancers (7, 48, 82, 104).

Comprising about 27% of the human genome, LINE-1 and Alu elements are common, highly methylated transposable elements that correlate with genome-wide DNA methylation in some studies (25, 120). LINE-1 promoters contain CpG islands

that are typically highly methylated, and Alu elements contain about one-third of all human CpG methylation sites (25). Changes in LINE-1 and Alu element methylation are hypothesized to have functional consequences on the expression of nearby genes (25). Atherosclerosis is characterized by DNA hypomethylation and transposable element methylation levels are associated with heart disease, stroke, and total mortality (5, 82); these findings contribute to interest in global genomic DNA methylation as a potential biomarker of cardiovascular disease risk.

Most previous work investigating variation in genes playing a role folate-mediated one-carbon metabolism in relation to homocysteine and genomic methylation focused on a small number of candidate genes. Given the interconnectedness of the one-carbon pathway, less proximal enzymes and genes may be important, yet few studies have attempted to evaluate a more comprehensive gene set. An investigation of genetic variation within the network of genes representing folate-mediated one-carbon metabolism in relation to CVD, homocysteine, and key methylation phenotypes is needed.

Nutritional Genomics

The literature contains many examples of diseases that are associated with genetic polymorphisms that may affect nutrition requirements, including CVD, neural tube defects, Down syndrome, colon cancer, hemochromatosis, and asthma (98). The *MTHFR 677 C→T* change decreases the affinity of the MTHFR enzyme for its cofactor riboflavin (vitamin B-2), rendering the protein less enzymatically active. Carriers of the alternate allele have an increased risk for neural tube defects and CVD, but a decreased risk for colon cancer. To ameliorate these effects, carriers must increase their consumption of dietary folate (6). However, disentangling the

relationships between genotype and diet in order to formulate appropriate nutrition recommendations is not always straightforward. Derivation of specific dietary recommendations for chronic disease prevention and management must account for underlying genetic variation while maintaining disease-specificity and balancing competing interests both within and among individuals (39, 98).

Previous work has identified several pairs of SNPs in genes encoding folate-regulating enzymes that catalyze linked steps in the one-carbon metabolic pathway, which have an interactive effect on CVD risk. Lim et al demonstrated an interaction between the *MTHFR 677C→T* and *SHMT1 1420 C→T* polymorphisms that resulted in a significantly increased CVD risk in Normative Aging Study participants. *SHMT1* reversibly converts serine and THF to glycine and 5,10-methylene THF. MTHFR then converts 5,10-methylene THF to 5-methylTHF, which is required for the remethylation of homocysteine to methionine (40). Specifically, among men with *SHMT1 1420 C→T TT* genotype, the odds ratio for CVD for *MTHFR 677 C→T CT* and *TT* genotypes (compared to *MTHFR CC* genotype, no loss of function) was 3.6 and 10.6, respectively (62). Additionally, using the same Normative Aging Study data, Raiszadeh et al observed a gene-gene interaction between *MTHFD1* and *MTHFR*, which also catalyze linked steps in the one carbon pathway, such that the increased risk associated with the *MTHFR 677C→T CT* and *TT* polymorphism (compared to *MTHFR CC* genotype, no loss of function) was observed only among men with the *MTHFD1 GA/AA* genotype (odds ratios 1.2 and 1.6, respectively), but the *MTHFR* SNP had little or no association with CVD risk in persons with the *MTHFD1 GG* genotype. In fact, *MTHFR 677 CT* genotype was associated with a protective effect in those with the *MTHFD1 GG* genotype (odds ratio 0.8) (83). The *MTHFD1* enzyme uses the one carbon units produced by the mitochondria (as formate) to convert THF to 10-formylTHF, which then enters the cytoplasmic one-

carbon folate pool (97). The *MTHFD1* enzyme also converts 5,10-methenylTHF to 5,10-methyleneTHF which serves as a substrate for *SHMT1* as described above (96). These examples of gene-gene interactions observed within linked steps in the one-carbon network suggest that the framework of a metabolic pathway provides an important context for studying genetic variation.

We therefore propose to use the tools of genome wide association studies, informed by the underlying biology of one carbon metabolism, to describe the relations among a full set of genetic variants in members of the one-carbon metabolic pathway and the ways in which these variants influence CVD risk, in terms of both individual and interactive effects. In essence, the proposed study pursues a candidate network approach, including an evaluation of gene-gene and gene-nutrient interactions within the network.

Conclusion

Determining the role of genetic and nutritional factors in mediating CVD risk is of central importance in understanding not only the pathogenesis of CVD, but in developing appropriate and effective intervention strategies for CVD prevention. There is a promising body of evidence to support the significance of variation in the folate metabolic network, based on prospective cohort studies as well as genetic studies examining subsets of the folate network. There are, however, no studies to date that evaluate the role of genetic and nutritional variation in the folate network in its entirety. Such strategies are important for informing future dietary recommendations aimed at improving public health and preventing disease, as the risk of disease may be elevated in specific population subsets defined by genetic and nutritional characteristics. The lack of a complete understanding of the contribution of

genetic and nutritional variation within the folate-mediated one-carbon network to CVD risk represents an important gap in the literature and one that will be directly addressed by studies proposed herein.

CHAPTER 2

POLYMORPHISMS IN SERINE HYDROXYMETHYLTRANSFERASE 1 AND METHYLENETETRAHYDROFOLATE REDUCTASE INTERACT TO INCREASE CARDIOVASCULAR DISEASE RISK IN HUMANS[1]

2.1 Abstract

The enzymes serine hydroxymethyltransferase 1 (gene name *SHMT1*) and methylenetetrahydrofolate reductase (gene name *MTHFR*) regulate key reactions in folate-mediated one-carbon metabolism. Common genetic variants with the potential to influence disease risk exist in both genes. A prior report from the Normative Aging Study indicated no association of the *SHMT1* rs1979277 SNP with CVD, but a strong gene-gene interaction was detected with *MTHFR* rs1801133. We investigated the effect of the *SHMT1* rs1979277 SNP and the *SHMT1* rs1979277 – *MTHFR* rs1801133 interaction in two epidemiologic cohort studies. In the Nurses' Health Study, the *MTHFR* rs1801133 variant genotypes were associated with an increased CVD risk, and there was an interaction between *SHMT1* and *MTHFR* such that the association of *MTHFR* rs1801133 *CT* genotype (vs. *CC;* the *TT* genotype could not be evaluated) was stronger in the presence of the *SHMT1* rs1979277 *TT* genotype (OR 4.34, 95% CI 1.2, 16.2, P=0.049). In the Health Professionals Follow-Up Study, the *MTHFR* rs1801133 genotype was not associated with CVD risk nor was there an interaction with *SHMT1* rs1979277. The association of genetic variation in the *SHMT1* gene, alone and in interaction with *MTHFR*, in relation to CVD risk is relatively

[1] Reproduced with permission from the Journal of Nutrition, accepted for publication October, 2010.

understudied at the population level, and results in the Nurses' Health Study confirmed a past report of gene-gene interaction, which is consistent with mechanisms suggested by basic science studies.

2.2 Introduction

Folate, a B vitamin fortified in the U.S. food supply since 1998, is involved in many cellular processes that influence disease risk. As tetrahydrofolate polyglutamates, folate coenzymes carry and activate one-carbon units for use in metabolic pathways including generation of methylation potential and synthesis of purine and pyrimidine nucleotides (32). Evidence from basic science studies suggests genetic variants that alter the function of folate coenzymes have the potential to impact disease risk. The SHMT1 enzyme plays a key role in the folate metabolic pathway by catalyzing the interconversion of serine and tetrahydrofolate to glycine and 5,10-methylenetetrahydrofolate (5,10-methyleneTHF) (40). The negligible concentration of free intracellular folate implies competition for the product of this reaction, 5,10-methyleneTHF, thus, SHMT1 is poised to mediate the flow of one-carbon units through thymidylate synthesis and the methionine cycle; both pathways are highly sensitive to folate status (8, 42). The MTHFR enzyme, which catalyzes a metabolic step linked to SHMT1, functions to convert 5,10 methyleneTHF to 5-methylTHF, thereby committing one-carbon units to methylation reactions at the expense of thymidylate synthesis (21, 32). A common genetic variant in the *SHMT1* gene, rs1979277 (*1420 C→T*), results in a modified protein (L474F). While this mutation does not affect catalytic activity, it impairs SHMT1 nuclear transport and subsequent thymidylate synthesis, and results in accumulation of the altered SHMT1 protein in

the cytoplasm, where it may inhibit cellular methylation reactions by sequestering 5-methyl-THF (2, 42, 116). A common genetic variant in the *MTHFR* gene, rs1801133 (*677 C→T*) results in an altered protein (A222V) with impaired ability to partition one carbon units to the remethylation pathway (21, 35). Because the MTHFR and SHMT1 enzymes catalyze linked metabolic steps, the accumulation of the variant SHMT1 L474F protein in the cytoplasm and sequestration of 5-methylTHF, coupled with the impaired 5-methylTHF synthetic activity of the MTHFR A222V protein suggests the potential for an interaction between the *SHMT1* rs1979277 and *MTHFR* rs1801133 genotypes. Such an interaction could have important metabolic consequences on cellular remethylation reactions and/or nucleotide synthesis, and ultimately on disease risk.

The association of the *MTHFR* rs1801133 variant with CVD risk has been well-studied (52, 58, 108); the increased risk of heart disease associated with the *MTHFR* rs1801133 *TT* vs. *CC* genotype is most pronounced in unfortified populations (52). Few studies have examined genetic variation in *SHMT1* in relation to CVD risk and the majority of publications investigating genetic variation in *SHMT1* focus on cancer

(28 of 31 studies published prior to 8/2009). One of 3 studies of the *SHMT1*—CVD association (4, 41, 62) investigated the *SHMT1-MTHFR* gene-gene interaction in relation to CVD risk (62), and the risk associated with the *MTHFR* rs1801133 *CT* and *TT* genotypes (vs. *CC*) was stronger in the subgroup of men with the *SHMT1* rs1979277 *TT* genotype.

Given the key role of the SHMT1 and MTHFR enzymes in the generation of cellular methylation potential and nucleotide biosynthesis, further studies examining the *MTHFR-SHMT1* interaction in relation to CVD risk are needed. Thus, a candidate gene association study was conducted to investigate the hypothesis that the *SHMT1*

rs1979277 and *MTHFR* rs1801133 genotypes interact to increase heart disease risk in the Health Professionals Follow-Up Study (HPFS) and the Nurses' Health Study (NHS). Since the association between these genotypes and CVD risk may be mediated by homocysteine and/or modified by B vitamins, these questions were also investigated.

2.3 Methods

To compute expected *SHMT1* genotype frequencies in European-ancestry populations, a literature search was conducted in 8/2009, using standardized methods. Genetic association studies reporting the association of *SHMT1* variants and chronic disease were identified on the PubMed search engine at the National Center for Biotechnology Information (74). Studies reporting original research on humans published in English were included if they reported the frequency of the *SHMT1* rs1979277 *TT* genotype and studied any of the following outcomes: CVD, coronary artery disease, coronary heart disease (CHD), and cancers (including the pre-cancerous condition, colorectal adenoma). 18 case-control studies conducted in European-ancestry populations were identified (11, 20, 29, 41, 53, 57, 61-64, 73, 75, 78, 93, 94, 106, 118, 121), and using data from controls only (total n=12,883), genotype counts were extracted and used to compute expected *SHMT1* genotype frequencies in European-ancestry populations.

Data from two cohort studies comprised the study population. The HPFS enrolled 51,529 male health professionals aged 40 y to 75 y in 1986, and the nested case-control study of CHD has been previously described (81). Cases comprised 266 men with incident non-fatal myocardial infarction or fatal CHD; case events occurred between the date of blood draw (about 1994 for all participants providing blood

samples) and the return of the year 2000 questionnaire (81). Using risk-set sampling, 532 controls matched for age, smoking status and date of blood sampling, were randomly selected from the subgroup of participants free of CVD at the time of cases' diagnosis. Only case-control sets in which all men had successful genotyping on both *SHMT1* rs1979277 and *MTHFR* rs1801133 were included (245 cases, 474 controls comprising 229 triads (1 case: 2 controls) and 16 pairs). Blood samples were analyzed for total cholesterol (81) and red blood cell folate (90) as described elsewhere.

The NHS enrolled 121,700 female registered nurses aged 30 y to 55 y in 1976, and a nested case-control study comprising 227 incident cases of nonfatal myocardial infarction (MI) and fatal CHD occurring between 1990 and 1998 was conducted (77). All women provided a blood sample prior to the case accrual period, thus in the pre-folate fortification era; controls were matched to each case on age, smoking, month of blood draw, fasting status, and reported problems with blood drawing, and were randomly selected using risk-set sampling from participants with the matching criteria and free of CHD at the time of case diagnosis (77). Only case-control sets in which all women had successful genotyping on both *SHMT1* rs1979277 and *MTHFR* rs1801133 were included (227 cases, 425 controls comprising 198 triads (1 case: 2 controls) and 29 pairs). Blood samples were analyzed for total cholesterol (77), red blood cell folate, plasma folate, vitamin B-6, vitamin B-12 and homocysteine, as previously described (90).

Genotyping for the studied polymorphisms in both cohorts, *SHMT1* rs1979277 (*1420 C→T)* and *MTHFR* rs1801133 *(677 C→T),* was conducted as part of a larger genotyping effort (53). Laboratory personnel were unaware of case-control status and genotyping was repeated to exclude errors when genotype distributions were found to be out of Hardy-Weinberg equilibrium (HWE). The median genotyping success rate

across all polymorphisms assayed was 95%, and concordance between the 10% quality control samples and genotyped variants was 100%.

SAS software (SAS Institute Inc., Cary, NC) was used for all statistical analyses. Genotype frequencies in controls were compared with those expected in HWE and Monte Carlo estimates of the exact P-values for the disequilibrium tests were computed (10,000 permutations). Conditional logistic regression analysis was used; adjusted regression models considered known risk factors for CHD. To test for the *SHMT1* rs1979277 – *MTHFR* rs1801133 gene-gene interaction, and to test for gene-nutrient and gene-gene-nutrient interactions, product terms involving the relevant genes and/or nutrients were included in regression models. The *SHMT1* rs1979277 SNP was coded as recessive, and the *MTHFR* rs1801133 SNP was coded using dummy variables to allow for non-linear associations. Nutrients and metabolites related to folate metabolism (red blood cell folate, plasma folate, plasma vitamins B-6, B-12, and homocysteine) were considered as mediators and/or effect modifiers. An alpha level of 0.05 was used for main effects and HWE tests, an alpha level of 0.15 was used for interactions, and point estimates and 95% confidence intervals are shown. The study was approved by the Cornell University Committee on Human Subjects.

2.4 Results

In the Health Professionals Follow-Up Study, the frequency of the *MTHFR* rs1801133 *TT* genotype was higher in controls (13.3%) than cases, and the prevalence in controls was greater than the expected frequency of 10.7% based on a past meta-analysis (52). In the NHS cohort, the frequency of the *MTHFR* rs1801133 *TT* genotype was similar between cases and controls, and in the range expected. The

expected *SHMT1* rs1979277 genotype frequencies were computed using data on controls from 18 studies conducted in European-ancestry populations, as follows: 48.9% *CC* (95% CI 48.1, 49.8), 41.0% *CT* (95% CI 40.2, 41.9), and 10.0% *TT* (95% CI 9.5, 10.5). In comparison, the prevalence of the *SHMT1* rs1979277 *TT* genotype in HPFS controls (12.0%) was high, while the NHS genotype prevalence was closer to expected (**Table 2.1**). In the NHS, cross-classification of the two genotypes revealed an absence of the double variant homozygote genotype (rs1979277 *TT* and rs1801133 *TT*) in the cases. Genotype distributions for *SHMT1* rs1979277 and *MTHFR* rs1801133 did not differ significantly from HWE expectations in either cohort.

Genotype—Outcome Associations

Based on the biological hypothesis that the *SHMT1* rs1979277 *TT* genotype leads to increased cytoplasmic protein levels and therefore less remethylation of homocysteine to methionine, the first analysis goal was to establish an *MTHFR*—heart outcomes association given well-known effects of common variants in *MTHFR* on homocysteine accumulation. Paradoxically, the *MTHFR* rs1801133 *T* allele was associated with a lower risk of CHD in the HPFS, and estimates of association were similar in simple models adjusting for demographics only, and in models fully adjusting for multiple cardiovascular risk factors. In the NHS, the *MTHFR* rs1801133 *T* allele was associated with an increased risk of CHD, although there was no gradient in risk per *MTHFR* allele. Thus, women with the *MTHFR CT* genotype had a 30% increased risk of CHD (vs. *CC* genotype: OR 1.30, 95% CI 0.9, 1.8), and women with the *MTHFR TT* genotype had a 22% increased risk (vs. *CC* genotype: OR 1.22, 95% CI 0.7, 2.1).

Following the hypothesis, the next analysis tested for a genotype-genotype interaction between *MTHFR* rs1801133 and *SHMT1* rs1979277. In the HPFS, where

Table 2.1. Genotype frequencies for *SHMT1* rs1979277 *(1420 C→T)* and *MTHFR* rs1801133 (*677 C→T)* in the Health Professionals Follow-Up Study, 1994-2000 (n=719), and the Nurses' Health Study, 1990-1998 (n=652)[1]

		SHMT1 rs1979277 (*1420 C→T*)			*MTHFR* rs1801133 (*677 C→T*)		
	Genotype:	*CC*	*CT*	*TT*	*CC*	*CT*	*TT*
			%			*%*	
NHS	Cases (n=227)	46.3	46.7	7.1	43.2	46.7	10.1
	Controls (n=425)	48.5	43.1	8.5	48.9	41.2	9.9
HPFS	Cases (n=245)	49.0	44.9	6.1	42.0	48.2	9.8
	Controls (n=474)	45.6	42.4	12.0	39.0	47.7	13.3

[1]Values are percentages for cases and controls, respectively, and are presented separately for each study.

Table 2.2. The relation of *MTHFR* rs1801133 *677 C→T* genotype with CVD risk stratified by *SHMT1* rs1979277 *1420 C→T* genotype in the Health Professionals Follow-Up Study (HPFS) and the Nurses' Health Study (NHS)[1]

Cohort:	NHS[2]		HPFS[3]	
	SHMT1 genotype:			
	CC/CT	*TT*	*CC/CT*	*TT*
MTHFR genotype:	1.18	4.34	1.0	0.46
MTHFR 677 C→T CT vs. CC	0.8, 1.7	1.2, 16.2	0.7, 1.4	0.1, 1.7
MTHFR 677 C→T TT vs. CC	1.25 0.7, 2.2	N/A	0.68 0.4, 1.2	0.69 0.2, 3.1

[1]Values are odds ratios,95% confidence intervals, and P values computed from conditional logistic regression models.
[2]n=227 cases. Unadjusted. No double homozygous variants among the cases.
[3] n=245 cases. Unadjusted.

no main effects of the *MTHFR* rs1801133 genotype were evident, the risk associated with the *MTHFR* rs1801133 genotype did not differ by *SHMT1* rs1979277 genotype (P=0.47 for interaction, liklihood ratio test statistic = 1.5, 2 df; **Table 2.2** and **Table A2.1**). However, in the NHS the risk associated with the *MTHFR* rs1801133 *CT* genotype varied by *SHMT1* rs1979277 genotype (P=0.049 for interaction, liklihood ratio test statistic = 3.9, 1 df; Table 2 and Table A1.1). In women with the *SHMT1* rs1979277 *TT* genotype, the risk of CHD in the *MTHFR* rs1801133 *CT* genotype group was ~4 times the risk in the *MTHFR* rs1801133 *CC* genotype group (OR 4.34, 95% CI 1.2, 16.2). In contrast, in women with the *SHMT1* rs1979277 *CC/CT* genotype, there was little or no association of the *MTHFR* rs1801133 *CT* (vs. *CC*) genotype with CHD (OR 1.18, 95% CI 0.8, 1.7). Risk associated with the *MTHFR* rs1801133 *TT* genotype in women with the *SHMT1* rs1979277 *TT* genotype could not be estimated because there were no cases with the double homozygote genotype. In both the Health Professionals Follow-Up Study and the Nurses' Health Study, minimally and fully adjusted models yielded very similar estimates of association for genotype.

Potential mediation of the gene-disease association by homocysteine was evaluated only in the NHS where homocysteine data were available. Only the *SHMT1 1420 C→T TT* genotype was associated with homocysteine in the NHS, and this association was not diminished after adjusting for the association of the *MTHFR* variant and the *SHMT1-MTHFR* interaction (β=2.35, P<0.04). Similarly, there was little or no difference in regression coefficients for genotype associations estimated in models further adjusted for homocysteine.

Given that folate intake may modify associations with CVD risk for at least one of the polymorphisms under study, the role of folate was explored. In the HPFS, there was little or no difference in RBC folate between cases and controls (no other

folate-related biomarker data were available). In regression models, RBC folate was not predictive of CHD, and including RBC folate in regression models did not change either the coefficient for the SNP or the coefficient for the gene—gene interaction. The occurrence of cases in the HPFS spanned January, 1998 when mandatory folate fortification of the U.S. food supply took effect; about half of the outcomes occurred prior to January, 1998. Stratifying regression models by date of case event did not meaningfully change model coefficients. No statistically significant two-way interactions (gene-nutrient) were observed between RBC folate and the *SHMT1* rs1979277 or *MTHFR* rs1801133 genotypes, and there were no significant 3-way interaction terms (gene-gene-nutrient; data not shown).

In the NHS, in addition to RBC folate, plasma concentrations of vitamin B-6, vitamin B-12, folate, and homocysteine were measured. Case accrual for NHS events ended in 1998, thus, all cases occurred prior to mandatory folate fortification in the U.S. There was little or no difference in folate-related biomarkers between cases and controls, with the exception of vitamin B-6 (the mean in controls was higher, P=0.08). RBC folate, plasma folate, and plasma vitamin B-6 were not associated with CHD in models with or without adjusting for homocysteine, and adding biomarkers to regression models had little or no effect on coefficients for the SNP or the interaction. No two-way gene-nutrient interactions involving any of the polymorphisms or biomarkers were identified, and there were no significant 3-way interactions (gene-gene-nutrient; data not shown).

2.5 Discussion

The published literature investigating the role of the *SHMT1* rs1979277 *1420 C→T* genotype in chronic disease risk focuses mainly on cancer risk, and evidence for

genotype associations with cardiovascular disease (CVD) is limited to 3 studies, underscoring the need for further population-level studies of *SHMT1* and CVD. While no prior studies reported a direct association of the *SHMT1* rs1979277 SNP with CVD, a strong interaction between the *SHMT1* rs1979277 and *MTHFR* rs1801133 genotypes was identified, which is consistent with emerging findings in basic science studies of functional effects of the SNP. Thus, the association of the *MTHFR* rs1801133 *CT* and *TT* genotypes (vs. *CC*) with CVD risk are stronger in people with the *SHMT1* rs1979277 *TT* genotype (62).

Using data from nested case-control studies in the population-based, prospective NHS and HPFS cohorts, the gene-gene interaction hypothesis was tested. The NHS findings were consistent with associations previously reported in the Normative Aging Study (NAS): risk associated with the *MTHFR* rs1801133 *T* allele was stronger in women with the *SHMT1* rs1979277 *TT* genotype. The absence of double homozygote cases in the NHS precluded estimating the association of *MTHFR* rs1801133 *TT* genotype with CHD risk in women with the *SHMT1* rs1979277 *TT* genotype. The lack of double homozygote cases may indicate the selective loss of such individuals from the cohort, but this scenario is highly unlikely given observed association sizes. An alternative and more likely explanation is that observed genotype frequencies are a chance phenomenon. While the fraction of all NHS cohort members who provided blood samples, and the fraction who developed CHD and were included in this study, should be random with respect to genotype, it is possible that a chance event led to the absence of the *MTHFR* rs1801133 *TT* / *SHMT1* rs1979277 *TT* double homozygote genotype class. Indeed, given the prevalence of the genotypes for each SNP in European-ancestry populations (expect 10.0% *SHMT1 TT* and expect 10.7% *MTHFR TT* in European-Ancestry populations), only 3 double homozygous individuals are expected in the 249 incident NHS cases under the null hypothesis of no

association with CHD. The three cells that can be estimated from the NHS data are each consistent with past findings (62).

The findings in the HPFS are inconsistent with summary estimates from prior meta-analyses regarding the association of the *MTHFR* rs1801133 genotype on CVD risk. In the HPFS, the *MTHFR* rs1801133 *T* allele was associated with a decreased risk of CVD, thus opposite in direction to findings from prior meta-analyses (52, 58, 108). The HPFS findings are also inconsistent with the past report (62) and the NHS findings reported herein with regard to the gene-gene interaction between *MTHFR* rs1801133 and *SHMT1* rs1979277. It is unlikely that gender influences the findings given that the past report of the interaction was in the Normative Aging Study, an all-male cohort. It is also unlikely that genotyping errors led to the paradoxical findings in the HPFS: an independent genotyping effort involving a larger sample from the NHS and HPFS case-control studies confirmed the same associations with CHD risk reported herein, based on imputed genotypes for *MTHFR* rs1801133 and *SHMT1* rs1979277; in this expanded subset, the imputed *MTHFR* rs1801133 *TT* genotype frequency was also 13.3% (data not shown). The differences in findings in the HPFS are also not likely to be due to differences in average folate or homocysteine levels. The mean folate and homocysteine levels in the cohorts are virtually identical, differing by less than 10% in all comparisons (data not shown). A possible explanation for the unexpected findings is random error.

A recent Institute of Medicine report on Nutrigenomics states, “There is, today, an unprecedented opportunity to use foods and food components to aid in achieving the genetic potential of humans, improve the overall performance of humans, and reduce the risk for chronic disease” (72). Given that the evidence base needed to make nutrition recommendations tailored to individual genotype is far from complete, this study investigated the interaction of nutritional status with the genetic variants studied,

and carefully considered the gene x nutrient and gene x gene x nutrient interactions. No gene–nutrient or gene-gene-nutrient interactions involving *SHMT1* rs1979277 and/or *MTHFR* rs1801133 were evident in either the NHS or the HPFS. In the HPFS, blood markers of folate status were collected prior to the introduction of mandatory folate fortification, although approximately half of the case events occurred after fortification. In both the HPFS and the NHS, nutritional status was not assessed at a uniform time relative to disease occurrence in cases. Future work should carefully consider the timing of nutrient measurements, and the role of folate fortification in the interpretation of findings.

While prior studies of the association of *MTHFR* rs1801133 *677 C→T* estimate the *TT* genotype (vs. *CC*) is associated with a 14-21% excess risk of CVD in the 10% of the population who have the *TT* genotype (52, 58, 108), our findings suggest that the risk associated with *MTHFR* rs1801133 *677 C→T T* allele is limited to the subgroup with the *SHMT1* rs1979277 *1420 C→T TT* genotype, comprising about 4.5% of the population studied. Similarly, in a prior report (62), risk associated with the *MTHFR* rs1801133 *677 C→T CT* (vs. *CC*) genotype was limited to the subgroup with the *SHMT1* rs1979277 *1420 C→T TT* genotype, comprising about 5% of the population studied. Furthermore, CHD risk in men with the *MTHFR* rs1801133 *677 C→T TT* genotype was 10 times the risk of men with the rs1801133 *CC* genotype in the 1% of men who were also *SHMT1* rs1979277 *1420 C→T TT* genotype (62). A prior report from the NHS found only moderate increases in CHD risk associated with *MTHFR 677 C→T* (rs1801133) variant genotypes (90), but the authors did not consider the *MTHFR* rs1801133 – *SHMT1* rs1979277 interaction investigated herein. Given that specific genotype and nutrition combinations may be required to confer risk, future studies wishing to further clarify the role of folate-related genes in chronic

disease will need to consider the possibility that risk may be stronger in population subgroups.

The effect of the *SHMT1* rs1979277 and *MTHFR* rs1801133 polymorphisms on protein structure and function supports the population-level interaction observed in the NAS and the NHS cohorts. Woeller et al established that the *SHMT1* rs1979277 polymorphism, which results in an L474F amino acid change, weakens the interaction between ubiquitin conjugating enzyme 9 (UBC9) and SHMT1 and prevents the addition of small ubiquitin-like modifiers (SUMOylation) *in vitro* (116), inhibiting nuclear transport (2), and thereby resulting in the accumulation of the variant SHMT1 protein in the cytoplasm. An increase in cytoplasmic levels of the SHMT1 protein impairs homocysteine remethylation and binds the same substrate as MTHFR, 5-methylTHF (42); therefore, an accumulation of SHMT1 in the cytoplasm is hypothesized to exacerbate the reduced enzyme activity of the variant A222V MTHFR protein (116). This proposed biological mechanism for the gene-gene interaction observed in the NAS and NHS suggests a pathogenic effect of the polymorphisms mediated by direct effects of homocysteine and/or effects on cellular methylation potential. Neither study demonstrated that the gene-gene interaction was mediated by homocysteine, but homocysteine may be an incomplete marker of cellular methylation potential or a single homocysteine measurement may not be an adequate reflection of either long-term homocysteine, homocysteine at the time of the event, or rate of change in homocysteine, depending on which aspect is most informative of risk. Additional biomarker data, including longitudinal measurements of homocysteine beginning prior to the event (84), and/or measures of other remethylation pathway biomarkers, such as S-adenosylhomocysteine (107), may be more informative.

While cells contain a second serine hydroxymethyltransferase, encoded by the *SHMT2* gene, which exhibits functional redundancy with the SHMT1 enzyme under

investigation here (2), evidence suggests the functional redundancy is incomplete. The *SHMT2* gene encodes two transcripts – an SHMT2 protein that localizes only to the mitochondria, and an SHMT2α protein that, lacking a mitochondrial targeting sequence, localizes to the cytoplasm and nucleus (2). While the SHMT2α protein provides some functional redundancy to the SHMT1 protein, studies of SHMT1 -/- mice confirm persistent aberrations in folate-related metabolism (2, 66), arguing against complete compensation for loss of SHMT1 function. Thus, there is some justification for the penetrance of the *SHMT1* rs1979277 SNP despite expression of SHMT2 protein.

This study directly addresses an important gap in the literature by investigating the relation of the *SHMT1* rs1979277 *TT* genotype to heart disease and takes advantage of existing data from two nested case-control studies conducted in large epidemiologic prospective cohort studies. A weakness was the inability to test the *MTHFR TT* / *SHMT1 TT* strata of the gene-gene interaction in the NHS cohort because of the complete absence of double homozygote cases. However, the *MTHFR—SHMT1* interaction that was detected agreed in direction and magnitude with a previous report (62). The study reported herein had a well-defined hypothesis linked to findings in basic science, and focused on only two genetic variants, the associated gene-gene and gene-nutrient interactions and mediation by a folate-related biomarker.

In summary, in the Nurses' Health Study, the association of the *MTHFR* rs1801133 genotype was modified by *SHMT1* rs1979277 genotype and associations were consistent in magnitude and direction with previously published findings. The biological plausibility for an interaction between *MTHFR* rs1801133 and *SHMT1* rs1979277 is strong, yet a review of published literature identified only one prior publication investigating the association of the *MTHFR* rs1801133 / *SHMT1* rs1979277 interaction in relation to CVD. Further studies are warranted, but future

work must consider folate fortification and the range of folate nutrition in the population, as well as the timing and collection of data on dietary intake, nutrition status, and folate pathway biomarkers, and include a more complete evaluation of genetic variation across the network of folate-related genes.

2.6 Acknowledgments

P.A.C., E.B.R., and D.J.H. designed research; P.A.C., P.J.S., E.B.R., D.J.H., S.M.W., and W.T. conducted research; S.M.W., P.A.C., F.R., and W.T. analyzed data; S.M.W., P.A.C., and P.J.S. wrote the paper. P.A.C. had primary responsibility for all work and final content. All authors read and approved the final manuscript.

APPENDIX

TABLE A2.1

CONDITIONAL LOGISTIC REGRESSION MODELS OF THE RELATION OF *SHMT1* rs1979277 *(1420 C→T)* AND *MTHFR* rs1801133 *(677 C→T)* GENOTYPES TO CORONARY HEART DISEASE RISK, NURSES' HEALTH STUDY (NHS), 1990-1998, AND HEALTH PROFESSIONALS FOLLOW-UP STUDY (HPFS), 1994-2000

	NHS[2,3]			HPFS[4]		
Variables	β	SE	P	β	SE	P
SHMT1 1420 C→T (TT vs. CC/CT)	-0.98	0.56	0.08	-0.40	0.47	0.39
MTHFR 677 C→T (CT vs. CC)	0.17	0.18	0.35	0.005	0.17	0.98
MTHFR 677 C→T (TT vs. CC)	0.22	0.29	0.44	-0.39	0.30	0.19
Interaction between *SHMT1 TT* and MTHFR *CT*	1.30	0.69	0.06	-0.77	0.69	0.26
Interaction between *SHMT1 TT* and MTHFR *TT*	N/A	N/A	N/A	0.014	0.83	0.99

[1]Values from conditional logistic regression models are β, beta-coefficient; SE, standard error; and P, P value.
[2]n=652 cases and controls, unadjusted model.
[3]Due to lack of double homozygous variants in cases, the second interaction term could not be estimated.
[4]n=719 cases and controls**,** unadjusted model.

CHAPTER 3

FOLATE NETWORK GENETIC VARIATION, PLASMA HOMOCYSTEINE, AND GLOBAL GENOMIC METHYLATION CONTENT

3.1 Abstract

Sequence variants in genes functioning in folate-mediated one-carbon metabolism are hypothesized to lead to changes in levels of homocysteine and DNA methylation, which, in turn, are associated with risk of cardiovascular disease. 330 SNPs in 52 genes were studied in relation to plasma homocysteine and global genomic methylation. SNPs were selected based on functional effects and to achieve gene coverage, and assayed on the Illumina Goldengate platform. Age-, smoking-, and nutrient-adjusted genotype—phenotype associations were estimated in regression models. Using a nominal $P \leq 0.005$ threshold for statistical significance, 20 SNPs were associated with homocysteine, 8 were associated with Alu methylation, and 1 was associated with LINE-1 methylation. Using a more stringent false discovery rate threshold, SNPs in *FTCD*, *SLC19A1*, and *SLC19A3* genes were associated with plasma homocysteine, gene x vitamin B-6 interactions were identified for both Alu and LINE-1 methylation, and epistatic interactions involving the *MTHFR* rs1801133 SNP were identified for the plasma homocysteine phenotype. Pleiotropy involving the *MTHFD1L* and *SARDH* genes for both plasma homocysteine and Alu methylation phenotypes was identified. No single gene was associated with all three phenotypes, and the set of most statistically significant SNPs predictive of homocysteine or Alu or LINE-1 methylation was unique to each phenotype. Genetic variation in folate-

mediated one-carbon metabolism, other than the well-known effects of the *MTHFR 677 C→T* rs1801133, is predictive of cardiovascular disease biomarkers.

3.2 Introduction

Folate and other B vitamins play key roles in biologic processes important to health, including DNA synthesis and the generation of cellular methylation potential for a variety of methylation reactions. Folate status is influenced by both dietary intake and variation in genes encoding folate-related enzymes, and altered folate status due to nutritional or genetic perturbations is associated with adverse outcomes, including birth defects, cardiovascular disease (CVD), and cancer (32).

Elevated plasma homocysteine, a sulfur-containing amino acid by-product of folate metabolism, is a marker of disturbed folate-mediated one-carbon metabolism, and is associated with an increased risk of CVD in prospective, meta-analysis, and genetic studies (58, 87, 100, 108). Homocysteine levels are modulated by nutrition, particularly folate and vitamin B-12 (45), and by genetic variants, including a well-studied SNP in the methylenetetrahydrofolate reductase gene (*MTHFR* 677 *C→T*; rs1801133) (52).

The association of homocysteine with CVD is hypothesized to be mediated, in part, by changes in DNA methylation (48). Folate-mediated one-carbon metabolism is linked to DNA methylation status through regulation of S-adenosylmethionine, the universal methyl donor, as well as through the activity of enzymes involved in methylation reactions, such as the DNA methyltransferases (82, 105). Changes in DNA methylation influence gene expression, cellular differentiation and development,

preservation of chromosomal integrity, and X chromosome inactivation, and methylation status is associated with the risk of CVD and some cancers (7, 48, 82, 104).

LINE-1 and Alu elements are abundant, transposable elements whose methylation status has been shown to be highly correlated with genome-wide DNA methylation in some studies (25) (120). LINE-1 promoters contain CpG dinucleotides that are typically highly methylated, and Alu elements contain about one-third of all human CpG methylation sites (25). Changes in LINE-1 and Alu element methylation are hypothesized to have functional consequences on the expression of nearby genes. Atherosclerosis is characterized by global DNA hypomethylation and transposable element methylation levels are associated with heart disease, stroke, and total mortality; in fact, we recently showed that reduced LINE-1 methylation was associated with increased incidence of ischemic heart disease and stroke within the Normative Aging Study (5, 82). These findings contribute to interest in LINE-1, Alu, or global genomic DNA methylation as potential biomarkers of cardiovascular disease risk.

Most previous work investigating variation in genes contributing to folate-mediated one-carbon metabolism in relation to homocysteine and genomic methylation phenotypes focused on a small number of candidate genes. Given the interconnectedness of the one-carbon pathway, less proximal enzymes and genes may be important, yet few studies have attempted to evaluate a comprehensive set of folate-related genes.

To investigate the genetic and nutritional predictors of homocysteine and methylation phenotypes, this candidate gene study examined variation across the network of genes representing folate-mediated one-carbon metabolism in relation to homocysteine and key methylation phenotypes. 330 single nucleotide polymorphisms

(SNPs) in 52 genes with a role in folate-mediated one-carbon metabolism were studied. The selection of the set of genes, the SNP markers, and the nutrients that were examined in this study was designed to represent the full functional variation of the folate-mediated one carbon metabolic pathway.

3.3 Methods

Study Population

The Normative Aging Study (NAS) was established by the Veterans' Administration (VA) in 1961, and 2,280 men aged 21-81 years (mean age of 42 y at study entry) were enrolled in the study on the basis of health criteria; details have been described elsewhere (9, 28). The rate of continued participation of NAS men over the follow-up period from 1961 through 1998 was excellent, with <1% attrition for all causes. As of June 1998, just prior to the homocysteine phenotype measurements, 543 participants (24%) were deceased and ~1600 men (70%) were actively participating in follow-up visits (mean age of 70 y). Men in the NAS were primarily non-Hispanic White males (>98% of the total NAS population) and the small number of participants in other racial/ethnic groups precluded separate analyses. The analyses described herein focus on non-Hispanic white males using data from the subset of men (~ 700) with measurements of homocysteine and global genomic DNA methylation (Alu and LINE-1).

This study was approved by the following: Brigham and Women's Hospital Human Subjects committee, Veterans' Administration R&D committee, Harvard School of Public Health, and the Cornell University Committee on Human Subjects.

DNA Extraction

Genomic DNA was extracted from stored frozen buffy coat of 7 ml whole blood using the QIAamp DNA Blood Kit (QIAGEN, Valencia, CA). The REPLI-g whole genome amplification kit (QIAGEN) was used to amplify genomic DNA when quantity was insufficient for genotyping. Whole-genome amplified samples were quality checked using Taqman genotyping (119) and poor-performing samples were excluded. Of 1,304 participant samples submitted for genotyping, 54.4% were genomic, 45.5% were whole-genome amplified.

SNP Selection

52 genes that contribute to folate-mediated one-carbon metabolism were identified (**Table A3.1**). SNP selection encompassed 2 kb on either side of the gene to include promoter and/or regulatory region variants. Four databases were used for SNP selection: the National Center for Biotechnology Information (NCBI) dbSNP and PubMed databases, (http://www.ncbi.nlm.nih.gov/), the Applied Biosystems Incorporated (ABI) SNPBrowser website (http://www.allsnps.com), and the Illumina Assay Design Tool (www.illumina.com). SNPs identified from literature searches and functional variants (non-synonymous coding region SNPs and promoter/regulatory region SNPs) were selected preferentially. Next, gene coverage considerations assessed linkage disequilibrium (LD) across the gene and physical coverage of the gene. Adjacent SNPs were selected such that the decay of maximum linkage disequilibrium between the 2 was no more than 33% (≤0.9 LDU between adjacent SNPs, where 1 LDU represents the decay of LD between two SNPs by about 37% of its maximum value when fitted to the Malecot model (67)) whenever possible to ensure sufficient SNP density to adequately represent the LD characteristics of the gene. GoldenGate SNP validation status was considered at each step, and SNPs with a minor allele frequency (MAF) ≥5% in European-ancestry populations were selected

where possible although exceptions were made for SNPs with prior evidence of putative function or when no SNPs with MAF ≥5% were available. A total of 384 SNPs were selected, including SNPs that were intentionally redundant to provide coverage in the event that key SNPs of interest failed in the genotyping assay. To simultaneously evaluate genetic variation within the network as a whole in a single model, a subset of 52 non-redundant SNPs was selected to represent the most likely functional variant with the highest MAF that could be selected for each of the 52 genes.

SNP Genotyping

384 SNPs were submitted to the Center for Inherited Disease Research at the Johns Hopkins University for genotyping via an Illumina GoldenGate custom genotyping panel. Genotype frequencies in controls were compared with those expected in Hardy-Weinberg equilibrium (HWE) and tested with Monte Carlo permutation estimates of exact P-values for HWE using 10,000 permutations. Of the 384 SNPs originally submitted, 54 were ultimately excluded for assay failure (46), monomorphic genotype data (4), minor allele frequencies less than 1% (3), and genotype frequencies out of HWE (1), leaving 330 SNPs available for analysis (**Table A3.2**). Both blind duplicates and HapMap Centre d'Etude du Polymorphisme Humain (CEPH) control samples with known genotypes were included, and reproducibility rates were excellent (99.99% for blind duplicates, 99.83% for HapMap CEPH controls).

Covariates

Extensive previously collected data on study participants includes physical measurements, lifestyle factors, and blood assays. Since the time of enrollment participants have had clinical examinations at 3- to 5-year intervals, with a response rate > 90% for mailed questionnaires. Fasting plasma samples were drawn at the VA

field site and stored at -80 °C. Plasma samples were transferred to the Jean Mayer USDA Human Nutrition Research Center on Aging, where they were analyzed; the time between blood draw and analysis averaged 1.7 ± 1.2 y as previously described (102). Plasma nutrient biomarkers were assayed in an unselected subset of stored blood samples. Plasma folate, vitamin B-6 (as pyridoxal-5'-phosphate; PLP) and vitamin B-12 were assayed, and the methods for these measurements have been previously described; coefficients of variation (CV) were uniformly excellent, as follows: 4.3% for folate, 5.0% for vitamin B-6, and 4.7% for vitamin B-12 (102).

Phenotype Assessment

Plasma total homocysteine was assayed in the same unselected subset of stored blood samples as plasma folate, vitamin B-6, and vitamin B-12. Methods for the determination of plasma total homocysteine were published, and the CV for the assay was 4.0% (102). The analysis of transposon DNA methylation has been reported in prior publications (Baccarelli, 2009) (13). Briefly, Alu and LINE-1 transposons were assayed in bisulfite-treated blood leukocyte genomic DNA using highly quantitative polymerase chain reaction –pyrosequencing technology. The degree of methylation was expressed for both Alu and LINE-1 as the percentage of methylated cytosines over the sum of methylated and unmethylated cytosines (% 5-meC). Each marker was tested in three replicates, and the average was used in statistical analyses.

Statistical Analyses

Restricted maximum likelihood and ordinary least squares regression models evaluated the relation between SNPs and the plasma homocysteine and global DNA methylation phenotypes. In a first step, additive, dominant, recessive, and overdominant genetic models of inheritance were tested for each SNP, and the model yielding the most statistically significant result in the unadjusted single SNP analysis was chosen as the best model going forward. An exception was made for the *MTHFR*

rs1801133 SNP, where prior evidence supported model-free dummy variable coding. Previous work in this cohort demonstrated no population substructure (115), thus no adjustments were made. All regression models were adjusted for age, smoking status, and nutrient residuals (portion of variation in nutrient not directly predicted by SNP), and an extended model also adjusted for the *MTHFR 677 C→T* (rs1801133) variant. Further models tested the interaction of each genotype with the rs1801133 SNP in relation to the homocysteine outcome, and later with nutrients in relation to all outcomes. Where specific genetic models of inheritance led to sparse data for main effects analyses, additive coding was used as the default. A small number of SNPs could not be evaluated in an interaction with *MTHFR* rs1801133, due to extensive data sparsity.

Regression coefficients with a nominal $P \leq 0.005$ were flagged to obtain a liberal set of associations for comparison with other studies and consideration for further follow up. Subsequent analyses used the False Discovery Rate (FDR) multiple testing correction of Benjamini and Hochberg (10) to adjust P-values, with a q-value significance threshold of 0.05, indicating that we expect less than 5% of tests flagged by this criterion to be false positives. For interactions, a nominal significance threshold of $P \leq 0.02$ was used, with an FDR q value significance threshold of 0.2, unless otherwise specified.

To investigate the joint role of nutrition and genetic variation, analyses of the SNP—phenotype association considered the possibility that nutrient biomarkers may affect this association by: 1) contributing to variation ("noise") in the phenotype due to causal pathways that do not involve the SNP under consideration, 2) mediating the association of the SNP with the phenotype, and 3) modifying the SNP—phenotype association. Because the set of genes under consideration encode enzymes that function as an interconnected network, it was important to account for variation in the

phenotype due to causes other than the SNP under consideration; for example, homocysteine is well-known to respond to folate levels, which may be influenced by a variety of factors including diet and several genetic variants. Nutrient residuals were calculated and used in models to account for variation in the nutrient biomarker that was not directly associated with the SNP of interest. Second, nutrient biomarkers may mediate the SNP-phenotype association and comparing models adjusting for the full variation in a nutrient to unadjusted models assessed the extent of mediation. Finally, nutrients may modify the SNP—phenotype association (for example, the association of *MTHFR 677 C→T* rs1801133 with homocysteine is modified by blood folate levels (47)). To assess effect modification, product terms between the SNP and the nutrient biomarker residual were included in models. Interactions were captured in a single model term, therefore, significance of the interaction was assessed by the P value for the interaction term (except for *MTHFR* rs1801133, which was dummy-coded; thus, the interaction was assessed through the Likelihood Ratio Test P value). To facilitate description of the genotype—nutrient interaction, a standardized approach was used, as follows: significant SNP—nutrient interactions were evaluated at 3 levels of the centered, log transformed nutrients: the 10^{th} percentile ("low nutrient levels"), the 50^{th} percentile ("median nutrient levels"), and the 90^{th} percentile ("high nutrient levels").

All statistical analyses were conducted with SAS v. 9.2 (SAS, Cary, NC).

3.4 Results

Measurements of the homocysteine phenotype, the Alu element methylation phenotype, and the LINE-1 methylation phenotype were available for 760, 628 and 621 participants, respectively. All participants had genotype data, and a subset of 533 men had data on all three phenotypes. Rather than limit data analysis to the subset of

533 men, each analysis included the maximum number possible. The three phenotype groups had similar genotype frequencies for the *MTHFR 677 C→T* rs1801133 *TT* genotype, but differed by age and hence differed slightly on age-related variables (**Table 3.1**). The *MTHFR 677 C→T* rs1801133 *TT* genotype prevalence in the largest phenotype group, the plasma homocysteine group, was 12.2%, calculated on data from over 1500 chromosomes; this

prevalence is higher than a sample of 120 chromosomes from the HapMap CEPH population with a *TT* frequency of 6.7%, but closely matches the *TT* frequency in a sample of 5064 chromosomes from North American 'control' participants (from case-control studies of heart outcomes) (52). In exploratory regression models, prior to assessing the genotype—phenotype associations, age and current smoking status were associated with homocysteine ($P \leq 0.001$), age was associated with Alu ($P \leq 0.005$), and current smoking was associated with LINE-1 ($P = 0.055$). Plasma folate, vitamin B-6, and vitamin B-12 were associated with homocysteine ($P \leq 0.005$), plasma vitamin B-6 was associated with Alu methylation ($P \leq 0.05$), and the nutrition biomarkers had little or no association with LINE-1.

Models exploring the SNP—phenotype association were adjusted for age, smoking, and nutrient residuals. Adjustment for age and smoking made little difference to the coefficients for each SNP. The set of SNPs represented in the most significant associations was nearly identical with or without adjustment for nutrient residuals. When models were further adjusted for the *MTHFR 677 C→T* (rs1801133) variant, the SNP regression coefficients were about the same as in models without the rs1801133 term, thus this term was not included in final models. The top SNP hits for the three phenotypes were relatively common variants, and nearly all had MAF $\geq$ 13%, but the set of top SNP hits was unique to each phenotype (**Tables 3.2-3.4** and

Table 3.1. Characteristics of Normative Aging Study participants, 1961-2001, with measurements on three phenotypes[1]

	Plasma homocysteine[2]	Global genomic DNA methylation (Alu elements)[3]	Global genomic DNA methylation (LINE-1 elements)[3]
	N=760	N=628	N=621
Age at phenotype measurement	68.6 (7.3)	72.5 (6.8)	72.5 (6.8)
Education – college graduate or higher (%)	26.8	28.6	28.7
White (%)	100	100	100
Baseline BMI (kg/m^2)	25.9 (2.9)	25.9 (2.9)	25.9 (2.9)
Cigarette smoking[4]			
Current (%)	6.7	5.4	5.5
Former (%)	63.0	63.4	63.5
Never (%)	30.3	31.2	31.1
Alcohol intake (% consuming ≥2 drinks/day)	12.4	13.7	13.5
Baseline diabetes diagnosis (%)	0.13	0.16	0.16
Baseline systolic blood pressure (mm Hg)	122.1 (12.7)	121.7 (12.5)	121.6 (12.6)
MTHFR 677 C→T (rs1801133) *TT* genotype (%)	12.2	12.3	12.6
Plasma folate (ng/ml)[5]	10.4 (5.7)	17.2 (15.1)	17.0 (14.9)
Plasma vitamin B-6 (pmol/ml)	84.9 (85.3)	104.6 (96.1)	104.8 (96.7)
Plasma vitamin B-12 (pg/ml)	458.9 (190.6)	512.8 (371.2)	514.9 (373.8)
Plasma total homocysteine (nmol/ml)	10.6 (3.7)	11.0 (4.2)	10.9 (4.2)
Global DNA methylation in Alu elements (%)		26.3 (1.1)	
Global DNA methylation in LINE elements (%)			76.9 (1.9)

[1]Mean (standard deviation) unless otherwise indicated
[2] N for homocysteine group ranges from 730 to 760 for variables in table
[3]Men in the two global genomic methylation groups were very similar, and dates of marker collection were nearly identical. N for Alu group ranges from 618 to 628, and N for LINE-1 group ranges from 611 to 621.
[4]Smoking status was assessed using most recent data prior to phenotype measurement.
[5]Plasma measures of folate, vitamin B-6, and vitamin B-12 were collected prior to the initiation of folate fortification for men included in the homocysteine group, and were collected after the initiation of folate fortification for men included in the two global genomic methylation groups.

Table 3.2. The most statistically significant associations (P≤0.005) between single nucleotide polymorphisms and the plasma homocysteine phenotype for men in the Normative Aging Study[1,5]

Gene Name	rs#	Nominal P	Effect[3]	Chr.	Coded allele	Coded allele frequency	Genetic Model	SNP Type
FTCD	rs2277820	3.09E-04[2]	7.22%	21	T	26%	Overdominant	Intronic
SLC19A1	rs1051266	4.16E-04[2]	5.04%	21	A	44%	Additive	Coding nonsynonymous
SLC19A1	rs1131596	4.31E-04[2]	5.03%	21	C	44%	Additive	5' region
SLC19A3	rs13007334	4.61E-04[2]	6.89%	2	C	46%	Overdominant	Intronic
SLC19A1	rs4819130	5.65E-04[2]	4.94%	21	C	44%	Additive	Intronic
MTHFD1L	rs11754661[4]	1.51E-03	49.98%	6	A	7%	Recessive	Intronic
DNMT1	rs2228611	2.42E-03	-6.44%	19	G	49%	Dominant	Coding synonymous
ALDH1L1	rs3772424	2.52E-03	6.14%	3	A	20%	Dominant	Intronic
GGH	rs4617146	2.55E-03	5.31%	8	T	19%	Additive	Intronic
CELF1	rs4752843	2.74E-03	-5.72%	11	C	14%	Additive	Intronic
SLC19A1	rs12482346	3.02E-03	4.08%	21	T	44%	Additive	Intronic
SLC19A1	rs2297291	3.39E-03	6.07%	21	A	41%	Dominant	Intronic
TCN2	rs4820886	3.41E-03	-17.30%	22	G	13%	Recessive	Intronic
TCN2	rs9621049	3.41E-03	-17.30%	22	T	13%	Recessive	Coding nonsynonymous
GLDC	rs7848919	3.52E-03	5.70%	9	G	32%	Dominant	3' region
SARDH	rs2502741[6]	3.60E-03	6.60%	9	A	50%	Dominant	Intronic
SLC19A1	rs1051298	3.68E-03	3.98%	21	T	44%	Additive	3' region
CBS	rs6586282	4.06E-03	-5.80%	21	T	18%	Overdominant	Intronic
FOLH1	rs202673	4.08E-03	-19.30%	11	G	14%	Recessive	Intronic
MTHFD1	rs1950902[6]	4.19E-03	-5.18%	14	T	16%	Additive	Coding nonsynonymous

[1]Model adjusted for age, smoking, residuals of plasma folate, vitamin B-6, and vitamin B-12; forward strand allele shown.
[2]Adjusted P values reached False Discovery Rate significance threshold of 0.05.
[3]Effect is shown as percent change in plasma homocysteine levels.
[4]Sparse data (fewer than 5 individuals per category) for some genotype categories.
[5]No SNPs map to more than one gene.
[6]Lower quality SNP.

Table 3.3. The most statistically significant associations (P≤0.005) between single nucleotide polymorphisms and the global genomic DNA methylation phenotype (Alu elements) for men in the Normative Aging Study[1,4,6]

Gene	rs#	Nominal P	Effect (s.d.)[3]	Chr	Coded allele	Coded allele frequency	Genetic Model	Type
GNMT	rs1051218[5]	2.14E-04	-0.57	6	T	3%	Dominant	3' region
DNMT3B	rs2424914	2.16E-03	0.30	20	G	45%	Recessive	Intronic
SLC25A32	rs3134297[5]	2.20E-03	0.65	8	C	20%	Recessive	5' region
DNMT3B	rs2424922	2.21E-03	0.30	20	C	45%	Recessive	Coding synonymous
DNMT3B	rs6058891	2.21E-03	0.30	20	C	45%	Recessive	Coding synonymous
MTHFD1L	rs1738574	2.39E-03	0.24	6	T	45%	Overdominant	Intronic
AHCYL2	rs1665105	3.87E-03	0.16	7	T	44%	Additive	3' region
SARDH	rs129886	3.92E-03	-0.60	9	T	19%	Recessive	3' region

[1]Model adjusted for age, smoking, and residuals of plasma folate, plasma vitamin B-6, and plasma vitamin B-12; forward strand allele shown; s.d.: standard deviation.
[2]No adjusted P values reached False Discovery Rate significance threshold of 0.05.
[3]Effect represents fold change in Alu element methylation standard deviation.
[4]No sparse data (fewer than 5 individuals per category) for any genotype categories of these SNPs.
[5]SNP maps to more than one gene (rs1051218 also maps to *PEX6*, rs3134297 also maps to *WDSOF1/DCAF13*).
[6]No lower quality SNPs.

Table 3.4. The most statistically significant association (P≤0.005) between single nucleotide polymorphisms and the global genomic DNA methylation phenotype (LINE-1 elements) for men in the Normative Aging Study[1,2]

Gene	rs#	Nominal P	Effect (s.d.)[3]	Chr	Coded allele	Coded allele frequency	Genetic Model	Type
MTHFR	rs12121543	4.29E-03	0.48	1	A	24%	Recessive	Intronic

[1]Model adjusted for age, smoking, and residuals of plasma folate, plasma vitamin B-6, and plasma vitamin B-12; forward strand allele shown; s.d.: standard deviation.
[2]No adjusted P values reached False Discovery Rate significance threshold of 0.05.
[3]Effect represents fold change in LINE-1 element methylation standard deviation.

Figure 3.1). A nominal significance threshold of P $\leq$ 0.005 was chosen as the threshold that effectively separated the top hits.

Total Plasma Homocysteine Phenotype

In analyses of the homocysteine phenotype, of the 20 SNPs with a nominal P ≤0.005, five SNPs were also significant at the FDR threshold (FDR-adjusted P≤0.05) (Table 3.2). These 5 SNPs comprise 3 genes: formiminotransferase cyclodeaminase (*FTCD*; 1 SNP, intronic), solute carrier family 19 (folate transporter), member 1 (*SLC19A1*, 3 hits, representing coding nonsynonymous, 5' region, and intronic variants), and solute carrier family 19, member 3 (*SLC19A3*, 1 SNP, intronic). Genetic variation in all 5 SNPs was positively associated with plasma homocysteine levels, and effects were similar in direction and magnitude; variant genotypes were associated with a 4.9-7.2% higher plasma total homocysteine compared to the referent genotype. In each case, the association of the genotype with homocysteine was mediated in part by nutrients; when plasma folate and vitamin B-6 or B-12 biomarkers were added to the models, the regression coefficients were reduced by 29% for *FTCD* rs2277820, by 43% for *SLC19A1* rs1051266, rs1131596, and rs4819130, and by 34% for *SLC19A3* rs13007334 (data not shown). The model containing a nonredundant set of 3 of the top 5 FDR-significant SNPs (*FTCD* rs2277820, *SLC19A3* rs13007334, *SLC19A1* rs1051266) explained 3.6% of the variation in plasma homocysteine beyond the variation explained in the model containing age, smoking, and folate, B-6, and B-12 residuals (data not shown), and the set of 3 SNPs made a statistically significant contribution (LRT for 3 SNPs = 17.6, p=0.0005, 3df); the coefficients for each SNP were similar to coefficients obtained in single SNP models. Genetic variation in the folate network as represented by a subset of 52 SNPs was significantly predictive of plasma homocysteine levels (likelihood ratio test for the model with vs. without set of 52 SNPs= 114.4; P value 1.39E-06, 52 degrees of freedom); after simultaneous

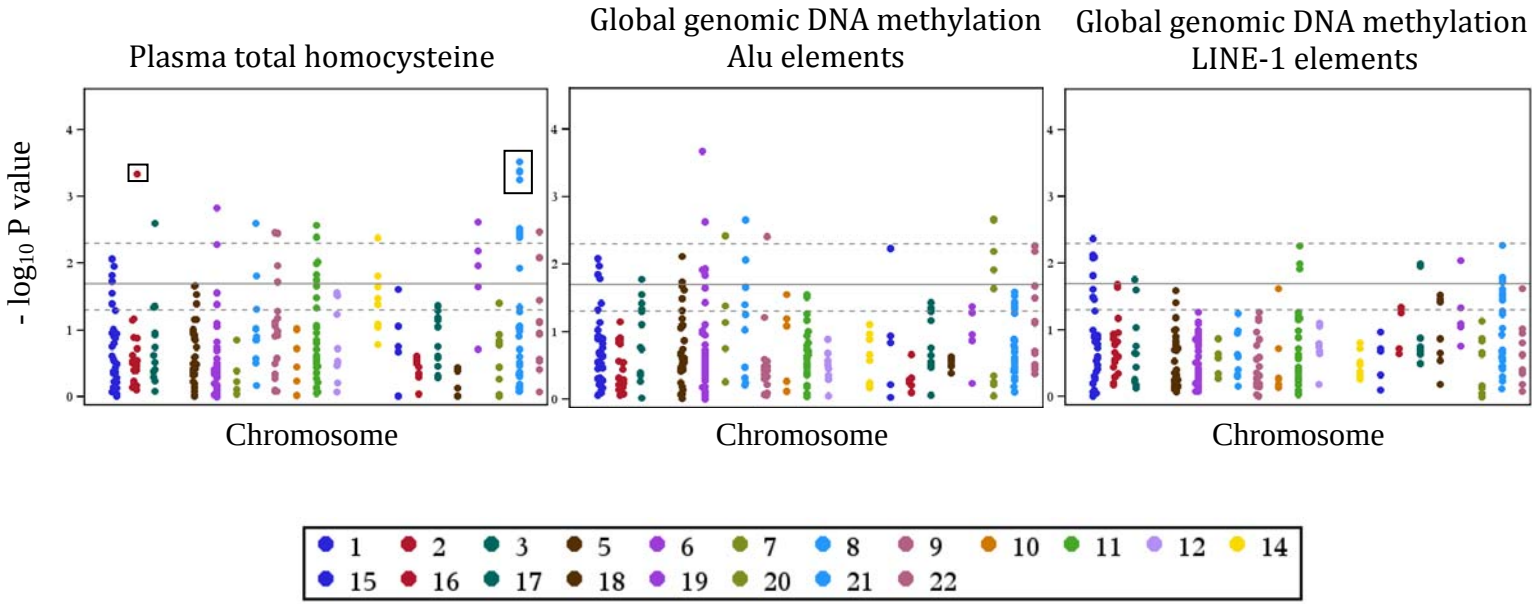

Figure 3.1. Manhattan plot: Folate-related SNPs as predictors of plasma total homocysteine and global genomic DNA methylation phenotypes. Models adjusted for age, smoking status, and folate, vitamin B-6, and vitamin B-12 residuals. Horizontal lines represent nominal P values of 0.05 (lower dashed line), 0.02 (center solid line) and 0.005 (upper dashed line). Boxes indicate SNPs that reached False Discovery Rate significance.

adjustment for the other 51 SNPs in the nonredundant set, the *SLC19A1* rs1051266 SNP was the top hit (nominal P= 0.0024). The set of 52 nonredundant SNPs together explained 14.3% of the variation in homocysteine above and beyond that explained by age and smoking. The genes represented in the most significant hits were similar between the single SNP models and the simultaneous model of 52 SNPs. The top 3 nominally significant hits in the simultaneous model were *SLC19A1* rs1051266, *TCN2* rs9621049, and *MTHFR* rs1801133. There was little or no difference between the coefficients for these SNPs in the simultaneous compared to the single SNP models, although for all three SNPs, p values were lower in the single SNP models as expected.

In models investigating interactions between each SNP and *MTHFR 677 C→T* rs1801133, 18 interaction terms were FDR-significant (FDR-adjusted P≤0.1) for the total plasma homocysteine phenotype (**Table A3.3**). Plasma folate and vitamins B-6 and B-12 are cofactors for enzymes involved in the reactions of the one-carbon metabolic network, and these nutrients are predicted to modify the SNP—phenotype association. In further analyses assessing SNP—nutrient interactions (folate, B-6, or B-12), no interaction coefficients reached FDR-significance levels (FDR-adjusted P≤0.2) (**Table A3.4**).

Global Genomic DNA Methylation Phenotype: Alu elements

In analyses of the Alu element methylation phenotype, 8 SNPs were statistically significant with a nominal P≤0.005, however, none were statistically significant at the FDR threshold (FDR-adjusted P≤0.05) (Table 3.3). There was little or no mediation of the association by nutrients or plasma total homocysteine levels (data not shown). Genetic variation in the folate network as represented by a subset of 52 SNPs together explained 8.4% of the variation in Alu element methylation above and beyond that explained by age and smoking; however, this model was not

significantly predictive of Alu methylation levels as assessed by the likelihood ratio test (test statistic for the model with vs. without 52 SNPs: 20.4; P value 0.99, 52 degrees of freedom). The genes represented in the most significant hits differed somewhat between the single SNP models and the simultaneous model of 52 SNPs; the top 3 nominally significant hits in the simultaneous model were *FOLR2* rs514933, *AHCYL2* rs1665105, and *MTHFR* rs1801133. The regression coefficient for the top hit from the simultaneous model (*FOLR2* rs514933) was increased 45% compared to the coefficient from the single SNP model, but there was little or no difference in the regression coefficients for *AHCYL2* rs166105 and *MTHFR* rs1801133.

There were no SNP—nutrient interactions with folate or B-12 that reached FDR thresholds for statistical significance (FDR-adjusted P≤0.2) (**Table A3.5**). Three SNPs had a statistically significant interaction with plasma vitamin B-6 (nominal P≤0.02) (Table A3.6); these interactions involved 3 intronic SNPs in 2 genes: aminomethyltransferase (*AMT*, rs1464567 and rs1464566) and DNA (cytosine-5-)-methyltransferase 3 beta (*DNMT3B*, rs1883729). Comparing men in the *AMT* rs1464567 *CC*/*CG* genotype group to the *GG* genotype, the mean Alu element methylation was 0.4 SD higher at low B-6, 0.1 SD higher at median B-6, and 0.4 SD lower at high B-6. Comparing men in the *AMT* rs1464566 *GG*/*GA* genotype group to the *AA* genotype, the mean Alu element methylation was 0.4 SD higher at low B-6, 0.1 SD higher at median B-6, and 0.3 SD lower at high B-6. Comparing men in the *DNMT3B* rs1883729 *AA* genotype group to the *AG*/*GG* genotype, the mean Alu element methylation was 0.1 SD lower at low B-6, 0.3 SD higher at median B-6, and 0.8 SD higher at high B-6.

Global Genomic DNA methylation Phenotype: LINE-1 elements

In models considering the main effect of SNPs on the LINE-1 methylation phenotype, no associations reached the FDR-significance threshold (FDR-adjusted

P≤0.05; Table 3.4). Genetic variation in the folate network as represented by a subset of 52 SNPs was significantly predictive of LINE-1 methylation levels (likelihood ratio for the model with vs. without 52 SNPs: 76.5; P value 0.015, 52 degrees of freedom) and the *AMT* rs8897 SNP had the lowest nominal P value (P=0.0024). The set of 52 nonredundant SNPs together explained 11.0% of the variation in LINE-1 element methylation above and beyond that explained by age and smoking. The genes represented in the most statistically significant hits for LINE-1 were generally similar between the single SNP models and the simultaneous model of 52 SNPs; the top 3 nominally significant hits in the simultaneous model were *AMT* rs8897, *DNMT1* rs2228612, and *CBS* rs1788484. The regression coefficient for the top hit in the simultaneous model (*AMT* rs8897) increased by 28% compared to the coefficient from the single SNP model, but there was little or no difference in the regression coefficients for *DNMT1* rs2228612 and *CBS* rs1788484.

There were no SNP—nutrient interactions for folate or B-12 that reached FDR-significance levels (FDR-adjusted P≤0.2) (**Table A3.6**). An interaction of plasma B-6 with 1 SNP was significant at the FDR threshold of P≤0.2 (rs17080689, an intronic SNP in methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 1-like, *MTHFD1L*) (Table A3.7), suggesting that the relation of the SNP to LINE-1 methylation varied according to plasma levels of vitamin B-6. Thus, comparing participants with the *MTHFD1L* rs17080689 *CA* genotype to the *CC*/*AA* genotype group, mean LINE-1 element methylation was 0.6 SD higher at low B-6, 0.2 SD higher at median B-6, and 0.4 SD lower at high B-6.

3.5 Discussion

We investigated sequence variation in a network of candidate genes involved in one-carbon metabolism in relation to plasma total homocysteine and two measures of global genomic DNA methylation (Alu and LINE-1 elements). Genes, SNPs, and related nutrients were carefully chosen to represent the full functional variation of the folate-mediated one carbon metabolism pathway. The analysis considered all 330 SNPs in tests for main effects, epistatic interactions between each SNP and the *MTHFR* rs1801133 SNP in relation to the plasma homocysteine outcome, and interactions between each SNP and plasma measures of folate, vitamins B-6 and B-12. FDR-significant main effects were identified for plasma total homocysteine and Alu element methylation phenotypes, FDR-significant interactions with *MTHFR* rs1801133 were identified for the homocysteine outcome, and FDR significant SNP—B-6 interactions were identified for Alu and LINE-1 element methylation phenotypes.

Genes involved in absorption and transport had the most significant associations with the homocysteine phenotype; consistent with this finding, about 30-40% of the association was mediated through plasma folate and vitamin B-6 and B-12 levels. Mitochondrial metabolism, methylation/ homocysteine pathways, cytoplasmic metabolism, nuclear metabolism, and B-12 metabolism were also represented in the top hits. For the Alu-element methylation phenotype, the top hits were in genes playing a role in mitochondrial metabolism, nuclear metabolism, and methylation/homocysteine metabolism. For the global genomic methylation LINE-1 phenotype, the top SNP was in a gene in the methylation/homocysteine pathway of one-carbon metabolism. There was no evidence that plasma folate or vitamins B-6 or B-12 mediated the association of SNPs with the methylation phenotypes.

The set of genes represented in the top hits was unique to each phenotype, although pleiotropy was identified among the plasma homocysteine and Alu element methylation outcomes, in which both the *MTHFD1L* and sarcosine dehydrogenase, *SARDH,* genes were among the most significant associations. When a subset of nonredundent SNPs (52 SNPs) was included in a single model, the genes identified as top hits for each phenotype were generally similar to genes identified in the single SNP models for each phenotypes, suggesting independent associations.

Plasma total homocysteine phenotype

SLC19A1. There were FDR-significant associations between 3 SNPs in the *SLC19A1* gene and plasma total homocysteine levels, and each SNP had about the same direction and magnitude of association with homocysteine. Thus, each copy of the coding nonsynonymous rs1051266 *A* allele was associated with a 5.0% increase in plasma homocysteine; similarly, each copy of the 5'region rs1131596 *C* allele was associated with a 5.0% increase in plasma homocysteine and, each copy of the intronic rs4819130 *C* allele was associated with a 4.9% increase in plasma homocysteine. HapMap LD plots indicate the linkage disequilibrium across the *SLC19A1* gene is high, and the results for all three *SLC19A1* SNPs are similar suggesting that SNP hits in this gene may represent a single effect. The *SLC19A1* gene encodes a ubiquitously expressed plasma membrane transporter responsible for the uptake of reduced folates, folate analogues (such as methotrexate), and phosphate esters of thiamine (37). This enzyme has also been detected in some mitochondrial membranes and may play a role in intracellular folate distribution (37). Consistent with the functions of the SLC19A1 protein, about half (~43%) of the association of these three *SLC19A1* SNPs with homocysteine was mediated by plasma folate and vitamins B-6/B-12; thus, the SNP is proposed to lead to changes in the nutrient levels that in turn leads to changes in homocysteine. The *SLC19A1* rs1051266 SNP (formerly known as rs61510559 or

c.80A→G) results in a histidine to arginine change at amino acid 27, and, in prior reports, the variant was associated with blood folate levels, although the direction of effect was not consistent (33, 95), and with increased risk of intracranial aneurysm (88), but not with homocysteine (33, 41) or abdominal aortic aneurysm (41). Expression of the transporter appears to be regulated by folate status (high folate levels, low transporter expression) (37). The *C* allele of the 5' region *SLC19A1* rs1131596 SNP was associated with reduced RBC folate levels in coronary artery disease patients and also with decreased SLC19A1 protein expression (17, 18). Together, these findings suggest that genetic variation in *SLC19A1* influences levels of folate-related biomarkers mediated by changes in folate and B-6/B-12 levels. Although SLC19A1 is not known to directly transport vitamins B-6/B-12, there was additional mediation through these nutrients, above and beyond mediation by folate, which may reflect changes in folate network flux that allow B-6/B-12 to compensate for altered folate transport.

FTCD. The intronic rs2277820 SNP in the *FTCD* gene was associated with plasma total homocysteine. The *CT* genotype group was 7.2% higher on plasma total homocysteine compared to the *CC*/*TT* genotype group. *FTCD* encodes a bifunctional Golgi-membrane associated enzyme involved in histidine and purine catabolism, resulting in the production of 5,10-methenyl-tetrahydrofolate (THF), which enters the cytoplasmic one-carbon pool (32). Based on HapMap LD patterns across the *FTCD* gene, it is possible that the association with the intronic rs2277820 SNP is a proxy for variation elsewhere in the gene – mutations in portions of the *FTCD* gene important for folate-binding, dimerization, or cyclodeaminase activity have been associated with autosomal recessive disorders of folate metabolism, characterized by significantly reduced formiminotransferase activity, absent cyclodeaminase activity, mental and physical retardation and metabolic disturbances (43). Consistent with the role of the

FTCD enzyme in production of 5,10-methenyl-THF, approximately one-third (29%) of the association between rs2277820 and homocysteine was mediated through plasma folate and vitamins B-6/B-12.

SLC19A3. A single FDR-significant association was identified between the intronic rs13007334 SNP in *SLC19A3* and plasma total homocysteine. The *CT* genotype group was 6.9% higher on plasma total homocysteine compared to the *CC*/*TT* genotype group. The *SLC19A3* gene belongs to the folate transporter family and encodes a ubiquitously expressed thiamine transporter (37); indeed SLC19A1 is also capable of transporting thiamine or phosphate esters of thiamine (37). Although SLC19A3 is not known to transport folate or vitamins B-6/B-12, about one-third (34%) of the association between the rs13007334 SNP and homocysteine was mediated by plasma folate and vitamins B-6/B-12. No prior reports link *SLC19A3* to biochemical or disease phenotypes, and a biological basis for the link to thiamine metabolism could not be identified.

Proportion of variability explained. The variability in homocysteine explained by the model containing the set of the 3 most significant nonredundant SNP hits was 3.6%, a small proportion of the estimated >50% heritability in homocysteine, based on twin studies (76, 92). The proportion of variation in homocysteine explained by these 3 SNPs is similar to the proportion explained by two other determinants of homocysteine, age and smoking (84), which together explained 3.5%. Together, a set of 52 nonredundant SNPs explained 14.3% of the variation in plasma homocysteine above and beyond that explained by age and smoking and was significantly associated with the homocysteine phenotype based on the likelihood ratio test.

Interactions. Previous meta-analyses have identified important determinants of homocysteine levels, including the *MTHFR 677 C→T* rs1801133 variant, intake of folate and vitamin B-12, and to a lesser extent vitamin B-6 (52) (45). In this study,

there were 18 statistically significant interactions between studied SNPs and *MTHFR 677 C→T* rs1801133 (Table A3.3), and the most highly significant of these involved *DHFR* rs12517451. Within the strata of individuals with the rs12517451 *AG* genotype, men with 1 copy of the rs1801133 *T* allele had plasma homocysteine levels 1.9% higher than the men with no copies, and men with 2 copies of the rs1801133 *T* allele had plasma homocysteine levels 4.2% higher than men with no copies.

In contrast to the epistatic interactions, there were no significant interactions between studied SNPs and plasma folate, vitamin B-6, or vitamin B-12 for the plasma homocysteine phenotype. The null results may be due to an overly conservative FDR significance threshold, network compensation for genetic and nutritional stresses, or inadequate power to evaluate interactions involving low MAF SNPs. An additional explanation for this finding may be that the folate status for men in the NAS, although measured prior to introduction of mandatory folate fortification, was high: the mean plasma folate was 10.4 ng/ml, significantly higher than the prefortification serum folate mean of 5.8 ng/ml in non-Hispanic Whites reported by NHANES III (1988-1994), although not as high as the post-fortification mean of 14.8 ng/ml measured in 1999-2000 (80). SNP—nutrient interactions may be blunted in this range of folate status. For example, the *MTHFR* rs1801133 SNP, which is expected to interact with folate in predicting the homocysteine phenotype, had a nonsignificant interaction in these data (p= 0.16). At low plasma folate (4.36 ng folate/ml), men with the rs1801133 *TT* genotype had a 16% higher plasma homocysteine compared to the *CC* genotype; at median plasma folate (9.53 ng folate/ml) the *TT* group had a 4.4% higher homocysteine, while at high plasma folate (17.23 ng folate/ml) the *TT* group had a 3.6% lower homocysteine. Although these findings are consistent with direction of the well-described *MTHFR* 677—folate interaction, the magnitude is slightly smaller than previously reported (47, 108).

A cluster of SNP—vitamin B-6 interactions was noted for variants in the *CBS* gene, but the P values for these interaction terms were about 0.1 and did not reach thresholds set prior to the analysis. The smallest P value was 0.086 for the interaction of vitamin B-6 with a variant in the *CBS* gene (rs1788484) and 0.216 for an interaction of vitamin B-6 with a variant in the *SHMT1* gene (rs643333). These findings are consistent with previous work showing that the human SHMT1 and CBS proteins bind B-6 tightly (K_d 850 nM and 700 nM, respectively) and effects of increasing vitamin B-6 availability on SHMT1 quantity and activity as well as cellular AdoMet levels were greatest under stringent B-6 deprivation conditions; thus, interactions between vitamin B-6 and genetic variants in the *SHMT1* and *CBS* genes may only be evident with very low vitamin B-6 status (79, 99). A systematic review of literature published prior to August, 2009 revealed 1 statistically significant interaction between genetic variation in *SHMT1* (rs1979277) and B-6 (53).

Global genomic DNA methylation phenotype (Alu elements)

There were no FDR-significant main effect associations observed for the Alu element methylation phenotype. The set of 52 nonredundant SNPs explained 8.4% of the variation in Alu element methylation above and beyond that explained by age and smoking but this set was not significantly predictive of Alu methylation based on the likelihood ratio test. None of the SNP—folate or SNP—vitamin B-12 interaction terms reached FDR significance in predicting Alu element methylation. Because the Alu phenotype was measured after the introduction of mandatory folate fortification in the U.S. findings may be limited by this timing. However, three FDR-significant SNP—vitamin B-6 interactions were identified, including two intronic SNPs in the *AMT* gene (rs1464567 and rs1464566) and one intronic SNP in the *DNMT3B* gene (rs1883729). In men with the *AMT* rs1464567 *CC*/*CG* genotype (compared to the *GG* genotype), mean Alu element methylation was 0.4 SD higher at low B-6, 0.1 SD

higher at median B-6, and 0.4 SD lower at high B-6. Similarly, in men with the *AMT* rs1464566 *GG*/*GA* genotype (compared to the *AA* genotype), mean Alu element methylation was 0.4 SD higher at low B-6, 0.1 SD higher at median B-6, and 0.3 SD lower at high B-6. For men with the *DNMT3B* rs1883729 *AA* genotype (vs. *AG/GG* genotype), mean Alu element methylation was 0.1 SD lower at low B-6, 0.3 SD lower at median B-6, and 0.8 SD higher at high B-6. The *AMT* gene encodes an enzyme that functions as an aminomethyltransferase (the so-called "T protein") in the vitamin B-6-dependent mitochondrial glycine cleavage system, which functions to catabolize glycine while at the same time synthesizing serine and 5,10-methyleneTHF, and is a major route for the provision of 1 carbon units to the cellular one-carbon pool as formate(56). Although the AMT enzyme itself is not vitamin B-6-dependent, through the glycine cleavage system it interacts closely with the vitamin B-6-dependent glycine dehydrogenase (decarboxylating) enzyme, (the so-called "P protein," which is encoded by the *GLDC* gene). B-6 interactions involving SNPs in *GLDC* were among the top nominally-significant hits for the homocysteine and Alu methylation phenotypes, although these interaction terms did not reach FDR-significance. The *DNMT3B* gene encodes a DNA methyltransferase enzyme that is localized to the nucleus and developmentally regulated. In mice, Dnmt3b knockouts are embryonic lethal while in humans, *DNMT3B* mutations have been shown to cause immunodeficiency, chromosomal instabilities, and facial abnormalities (ICF) syndrome (110). Further, DNMT3B expression is associated with cancer (85, 103, 110). A recent publication links the DNMT3B enzyme to histone ubiquitin ligase complexes, suggesting it may be important for repressive chromatin formation (91). DNMT3B functions to establish de novo methylation patterns, including the methylation of CpG dinucleotides within transposons located near chromosome centromeres (103, 110). Although cell culture studies have not supported Alu

elements as DNMT3B targets (22, 110) there is evidence that B vitamins may interact with DNA methylation enzymes. In both *in vitro* and *in vivo* models, DNMT3b protein levels were down-regulated by B vitamin deficiency (deficiency of folate, B-6, and B-12 together), and further, *de novo* methylation was suppressed both *in vitro* and *in vivo* under conditions of B vitamin deficiency (36). Finally, cell culture work has demonstrated that cellular AdoMet levels were markedly decreased in response to lowered B-6 concentrations in culture medium (79), consistent with the direction of effect observed for the *DNMT3B*-B-6 interaction.

Global genomic DNA methylation phenotype (LINE-1 elements)

There were no FDR-significant associations observed for the LINE-1 methylation phenotype. The set of 52 nonredundant SNPs explained 11% of the variation in LINE-1 element methylation above that explained by age and smoking and this set of SNPs was significantly predictive of LINE-1 element methylation based on the likelihood ratio test. There were no FDR-significant interactions between SNPs and folate or vitamin B-12; however, the LINE-1 phenotype was measured after the introduction of mandatory folate fortification in the U.S. which may have limited findings. A single SNP—vitamin B-6 interaction was significant at the FDR threshold: for the intronic rs17080689 in the *MTHFD1L* gene, comparing men with the *CA* genotype to the *AA/CC* genotype, mean LINE-1 element methylation was 0.6 SD higher at low B-6, 0.2 SD higher at median B-6, and 0.4 SD lower at high B-6. While the *MTHFD1L* gene is not vitamin B-6-dependent, the gene encodes a monofunctional mitochondrial tetrahydrofolate synthase and functions as part of the cellular folate metabolic network, downstream from the vitamin B-6-dependent glycine cleavage system, contributing to production of THF and formate (23). Intronic variation in *MTHFD1L* was associated with CVD in a large-scale genome-wide association study (112).

3.6 Conclusion

Strengths of the present study include investigation of a large cohort with homocysteine data collected prior to the introduction of mandatory folate-fortification in the U.S. Thus, the study population is well-suited to investigate the network of folate-related genes in relation to such quantitative phenotypes. SNP selection for the genotyping assay reflected coverage of genes based on function, linkage, and physical coverage, which led to markers that captured functional variants in addition to tagging of variation elsewhere in the gene. In a systematic approach to the data analysis, we identified the best genetic models for each SNP, then tested single SNPs, a reduced set of non-redundant SNPs, the interaction of each SNP with *MTHFR* 677 rs1801133, and the interaction of each SNP with folate, vitamin B-6 and vitamin B-12. Findings were corrected for multiple comparisons and findings based on the FDR threshold were discussed in more detail. Weaknesses of the study include the fact that methylation (but not homocysteine) measures were collected after the introduction of mandatory folate fortification in the U.S., which may have attenuated associations, particularly gene-nutrient interactions, which might have been present in a population with more variation in B-vitamin status. Available nutrient measures included folate and vitamins B-6 and B-12; however, information on additional nutrients such as choline would have allowed a more complete evaluation of gene-nutrient interactions. Although AdoHcy and/or the AdoMet/AdoHcy ratio are likely to be more sensitive indicators of vascular disease risk than homocysteine (51, 107), these biomarkers were not measured in the NAS cohort study. Finally, due to genotyping failure, some key variants could not be analyzed, ex., rs6922269 in *MTHFD1L* (112).

The most significant hits for the homocysteine and methylation outcomes reflected genes involved in the generation of one-carbon units, including *SLC19A1*

and *FTCD*. Because a unique set of genes was identified among the most significant hits for each outcome, and because the top hits could not be predicted on the basis of projected impact on cellular methylation potential, this work suggests that not all folate effects are mediated through AdoMet/AdoHcy and that beyond the well-described *MTHFR* rs1801133 SNP, polymorphisms in other genes make important contributions to homocysteine and global genomic DNA methylation phenotypes. Furthermore, some associations are sensitive to nutritional status of B vitamins. Future work should continue to include a broad evaluation of one-carbon network genetic and nutritional variation in unfortified or pre-fortification populations and extend these findings for CVD biomarkers to investigate CVD phenotypes directly.

3.7 Conflict of Interest

The authors declare that they have no conflict of interest.

3.8 Ethical Standards

All experiments complied with the current laws of the country in which they were performed.

3.9 Acknowledgements

The authors wish to thank Ms. Ellen M. Smith for her contributions to preliminary analyses of this work. The project was supported in part by the Center for Vertebrate Genomics (PAC), the Presidents Council for Cornell Women (PAC), and the Bronfenbrenner Life Course Center (PAC), all at Cornell University; by USDA (CSREES) grant 2003-34324-13135 (subproject to PAC); by NIEHS grant

R01ES015172 (JS); and by T32DK007158 (SMW), National Institute of Diabetes, Digestive and Kidney Diseases. Dr. Baccarelli receives salary support from New Investigator funding from the HSPH-NIEHS Center for Environmental Health (ES000002). Genotyping services were provided (to PAC) by the Johns Hopkins University under federal contract number (N01-HV-48195) from the National Heart, Lung, and Blood Institute. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Diabetes And Digestive And Kidney Diseases, the National Heart, Lung, and Blood Institute, or the National Institutes of Health.

APPENDIX

TABLE A3.1

THE 52 GENES IN THE FOLATE-MEDIATED ONE-CARBON PATHWAY

Gene Symbol	Gene Name	Entrez GeneID
AHCY	Adenosylhomocysteinase	191
AHCYL1	Adenosylhomocysteinase-like 1	10768
AHCYL2	Adenosylhomocysteinase-like 2, KIAA0828	23382
ALDH1L1	Aldehyde dehydrogenase 1 family, member L1	10840
AMT	Aminomethyltransferase	275
ATIC	5-aminoimidazole-4-carboxamide ribonucleotide formyltransferase/IMP cyclohydrolase	471
BHMT	Betaine-homocysteine S-methyltransferase	635
CBS	Cystathionine-beta-synthase	875
CTH	Cystathionase (cystathionine gamma-lyase)	1491
CELF1	CUGBP, Elav-like family member 1	10658
DHFR	Dihydrofolate reductase	1719
DMGDH	Dimethylglycine dehydrogenase	29958
DNMT1	DNA (cytosine-5-)-methyltransferase 1	1786
DNMT3A	DNA (cytosine-5-)-methyltransferase 3 alpha	1788
DNMT3B	DNA (cytosine-5-)-methyltransferase 3 beta	1789
FOLH1	Folate hydrolase (prostate-specific membrane antigen) 1	2346
FOLR1	Folate receptor 1 (adult)	2348
FOLR2	Folate receptor 2 (fetal)	2350
FOLR3	Folate receptor 3 (gamma)	2352
FPGS	Folylpolyglutamate synthase	2356
FTCD	Formiminotransferase cyclodeaminase	10841
FTH1	Ferritin, heavy polypeptide 1	2495
GART	Phosphoribosylglycinamide formyltransferase, phosphoribosylglycinamide synthetase, phosphoribosylaminoimidazole synthetase	2618
GCSH	Glycine cleavage system protein H (aminomethyl carrier)	2653
GGH	Gamma-glutamyl hydrolase (conjugase, folylpolygammaglutamyl hydrolase)	8836
GLDC	Glycine dehydrogenase (decarboxylating)	2731
GNMT	Glycine N-methyltransferase	27232
HSPA8	Heat shock 70kDa protein 8	3312
MARS	Methionyl-tRNA synthetase	4141
MAT1A	Methionine adenosyltransferase I, alpha	4143
MAT2A	Methionine adenosyltransferase II, alpha	4144
MAT2B	Methionine adenosyltransferase II, beta	27430
MTHFD1	Methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 1, methenyltetrahydrofolate cyclohydrolase, formyltetrahydrofolate synthetase	4522
MTHFD1L	Methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 1-like	25902
MTHFD2	Methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 2, methenyltetrahydrofolate cyclohydrolase	10797
MTHFR	Methylenetetrahydrofolate reductase (NAD(P)H)	4524
MTHFS	5,10-methenyltetrahydrofolate synthetase (5-formyltetrahydrofolate cyclo-ligase)	10588
MTR	5-methyltetrahydrofolate-homocysteine methyltransferase	4548

TABLE A3.1 (Continued)

Gene Symbol	Gene Name	Entrez GeneID
MTRR	5-methyltetrahydrofolate-homocysteine methyltransferase reductase	4552
SARDH	Sarcosine dehydrogenase	1757
SHMT1	Serine hydroxymethyltransferase 1 (soluble)	6470
SHMT2	Serine hydroxymethyltransferase 2 (mitochondrial)	6472
SLC19A1	Solute carrier family 19 (folate transporter), member 1	6573
SLC19A2	Solute carrier family 19 (thiamine transporter), member 2	10560
SLC19A3	Solute carrier family 19, member 3	80704
SLC25A32	Solute carrier family 25, member 32	81034
SLC46A1	Solute carrier family 46 (folate transporter), member 1	113235
TCN1	Transcobalamin I (vitamin B-12 binding protein, R binder family)	6947
TCN2	Transcobalamin II	6948
TYMS	Thymidylate synthetase	7298
UBE2I	Ubiquitin-conjugating enzyme E2I (UBC9 homolog, yeast)	7329
UBE2N	Ubiquitin-conjugating enzyme E2N (UBC13 homolog, yeast)	7334

TABLE A3.2

330 FOLATE-RELATED SNPS ASSAYED IN MEN IN THE NORMATIVE AGING STUDY[1-5]

Gene	SNP Name	Gene	SNP Name	Gene	SNP Name
AHCY	rs819133	*BHMT*	rs558133	*DMGDH*	rs1805073
AHCY	rs819159	*BHMT*	rs585800	*DMGDH*	rs532964
AHCY	rs819155	*CBS*	rs706209	*DMGDH*	rs2253262
AHCY	rs13043752	*CBS*	rs2124458	*DMGDH*	rs644191
AHCY	rs819148	*CBS*	rs760124	*DNMT1*	rs8101626
AHCY	rs1205366	*CBS*	rs6586281	*DNMT1*	rs11880388
AHCY	AHCYC34T	*CBS*	rs6586282	*DNMT1*	rs2228611
AHCY	rs819146	*CBS*	rs11203172	*DNMT1*	rs2228612[1]
AHCYL1	rs333079	*CBS*	rs234704	*DNMT1*	rs2162560
AHCYL1	rs3768480	*CBS*	rs1801181	*DNMT3A*	rs11695471
AHCYL1	rs2298116	*CBS*	rs2014564	*DNMT3A*	rs6546045
AHCYL1	rs720917	*CBS*	rs1789953	*DNMT3A*	rs7578575
AHCYL1	rs186724	*CBS*	rs12329764	*DNMT3A*	rs6733868
AHCYL2	rs6467233	*CBS*	rs234705	*DNMT3A*	rs11678631
AHCYL2	rs4731569	*CBS*	rs234706	*DNMT3A*	rs1550117
AHCYL2	rs7788327	*CBS*	rs234709	*DNMT3B*	rs6058869
AHCYL2	rs6467244	*CBS*	rs234711	*DNMT3B*	rs1883729
AHCYL2	rs1665105	*CBS*	rs11701048	*DNMT3B*	rs2424914
ALDH1L1	rs4646760	*CBS*	rs1788484	*DNMT3B*	rs6058891
ALDH1L1	rs4646750	*CTH*	rs648743	*DNMT3B*	rs2424922
ALDH1L1	rs4646745	*CTH*	rs663465	*DNMT3B*	rs6058896
ALDH1L1	rs3772424	*CTH*	rs681475	*FOLH1*	rs16906158
ALDH1L1	rs3772414	*CTH*	rs535112	*FOLH1*	rs202673
ALDH1L1	rs2305230	*CTH*	rs663649	*FOLH1*	rs664584
ALDH1L1	rs11715574	*CTH*	rs1021737	*FOLH1*	rs202676
ALDH1L1	rs1868138	*CUGBP1*	rs7102372	*FOLR1*	rs2071010
ALDH1L1	rs1823213	*CUGBP1*	rs2242081	*FOLR1*	rs9282688
AMT	rs11922013	*CUGBP1*	rs7933019	*FOLR2*	rs514933
AMT	rs1464567	*CUGBP1*	rs4752843	*FOLR2*	rs2298444
AMT	rs1464566	*DHFR*	rs12517451	*FOLR3*	rs7926875
AMT	rs8897	*DHFR*	rs1677666	*FOLR3*	rs11235449
ATIC	rs7585489	*DHFR*	rs1650723	*FPGS*	rs10106
ATIC	rs2372536	*DHFR*	rs2560424	*FPGS*	rs41319447
ATIC	rs3821353	*DHFR*	rs1643659	*FPGS*	rs4451422
ATIC	rs1997059	*DHFR*	rs1643650	*FTCD*	rs16978930
ATIC	rs4672768	*DHFR*	rs836822	*FTCD*	rs2277820
BHMT	rs16876512	*DHFR*	rs1650697	*FTH1*	rs1800009
BHMT	rs7700970	*DHFR*	rs380691	*FTH1*	rs17185413

TABLE A3.2 (Continued)

Gene	SNP Name	Gene	SNP Name	Gene	SNP Name
BHMT	rs506500	*DHFR*	rs1382540	*FTH1*	rs1801621
BHMT	rs567754	*DMGDH*	rs28326	*FTH1*	rs17156609
BHMT	rs10037045	*DMGDH*	rs1805074	*FTH1*	rs2073588
GART	rs8971	*MAT1A*	rs1819684	*MTHFD1L*	rs538017
GART	rs1804385	*MAT1A*	rs17677908	*MTHFD1L*	rs9478162
GART	rs8788[2]	*MAT2A*	rs1446667	*MTHFD1L*	rs9478908
GART	rs2027592	*MAT2A*	rs2028900	*MTHFD1L*	rs7770982
GART	rs4817580	*MAT2A*	rs1078004	*MTHFD1L*	rs17080689
GCSH	rs8177940	*MAT2A*	rs2043675	*MTHFD1L*	rs1076746
GCSH	rs8177876	*MAT2B*	rs10515861	*MTHFD1L*	rs509474
GCSH	rs4889233	*MAT2B*	rs6882306	*MTHFD1L*	rs7746991
GCSH	rs11866124	*MAT2B*	rs4869087	*MTHFD1L*	rs6910267
GCSH	rs1563072	*MAT2B*	rs17061795	*MTHFD1L*	rs17354394
GGH	rs11995525	*MAT2B*	rs7733775	*MTHFD1L*	rs1047665
GGH	rs11545078	*MTHFD1*	rs1076991	*MTHFD2*	rs7340453
GGH	rs4617146	*MTHFD1*	rs8003379	*MTHFD2*	rs1667627
GGH	rs3780126	*MTHFD1*	rs1950902	*MTHFR*	rs1537516
GGH	rs12544045	*MTHFD1*	rs17751556	*MTHFR*	rs4846049
GLDC	rs7848919	*MTHFD1*	rs3783728	*MTHFR*	rs13306556
GLDC	rs3902970	*MTHFD1*	rs11627525	*MTHFR*	rs1801131
GLDC	rs1929933	*MTHFD1*	rs8003567	*MTHFR*	rs12121543
GLDC	rs4237166	*MTHFD1*	rs8012229	*MTHFR*	rs1994798
GLDC	rs10975681	*MTHFD1*	rs2281603	*MTHFR*	rs6541003
GLDC	rs4629927	*MTHFD1L*	rs7765521	*MTHFR*	rs1801133
GLDC	rs7049056	*MTHFD1L*	rs11754661	*MTHFR*	rs17421462
GLDC	rs1821892	*MTHFD1L*	rs4869953	*MTHFR*	rs17367629
GLDC	rs2118653	*MTHFD1L*	rs17349743	*MTHFR*	rs3737965
GLDC	rs1755617	*MTHFD1L*	rs803422	*MTHFS*	rs8923
GNMT	rs11752813	*MTHFD1L*	rs997429	*MTHFS*	rs7177659
GNMT	rs2296805	*MTHFD1L*	rs2295084	*MTHFS*	rs6495450
GNMT	rs2296804	*MTHFD1L*	rs803456	*MTHFS*	rs2733106
GNMT	rs736158	*MTHFD1L*	rs803455	*MTHFS*	rs2586167
GNMT	rs1051218	*MTHFD1L*	rs803454	*MTR*	rs16834388
HSPA8	rs4936770	*MTHFD1L*	rs12201472	*MTR*	rs3754255
HSPA8	rs1461496	*MTHFD1L*	rs4869954	*MTR*	rs10925260
HSPA8	rs11218941	*MTHFD1L*	rs6902664	*MTR*	rs1805087
HSPA8	rs1136141	*MTHFD1L*	rs1474787	*MTR*	rs2275566
MARS	rs899653	*MTHFD1L*	rs4869955	*MTR*	rs2229276[3]
MARS	rs496245	*MTHFD1L*	rs742832	*MTR*	rs1131449[4]

TABLE A3.2 (Continued)

Gene	SNP Name	Gene	SNP Name	Gene	SNP Name
MARS	rs1678537	*MTHFD1L*	rs803447	*MTR*	rs1050996
MAT1A	rs1985908	*MTHFD1L*	rs803446	*MTRR*	rs2966952
MAT1A	rs2993763	*MTHFD1L*	rs803466	*MTRR*	rs1801394
MAT1A	rs10788546	*MTHFD1L*	rs1738574	*MTRR*	rs7730643
MAT1A	rs1143694	*MTHFD1L*	rs524732	*MTRR*	rs161869
MTRR	rs1532268	*SLC19A1*	rs1051266	*TCN2*	rs5749132
MTRR	rs3776465	*SLC19A1*	rs1131596[5]	*TCN2*	rs9606756
MTRR	rs162036	*SLC19A1*	rs4819130	*TCN2*	rs740233
MTRR	rs2303081	*SLC19A2*	rs6656822	*TCN2*	rs757874
MTRR	rs10380	*SLC19A2*	rs1983546	*TCN2*	rs4820021
MTRR	rs1802059	*SLC19A2*	rs17518769	*TCN2*	rs9621049
MTRR	rs8659	*SLC19A2*	rs2038024	*TCN2*	rs4820886
SARDH	rs129886	*SLC19A3*	rs13025803	*TCN2*	rs4820887
SARDH	rs129932	*SLC19A3*	rs11694828	*TCN2*	rs2301957
SARDH	rs129891	*SLC19A3*	rs13007334	*TCN2*	rs2301958
SARDH	rs756682	*SLC19A3*	rs17438244	*TCN2*	rs4820889
SARDH	rs2073817	*SLC25A32*	rs1061196	*TCN2*	rs10418
SARDH	rs2502741	*SLC25A32*	rs3098260	*TYMS*	rs2853533
SARDH	rs4979632	*SLC25A32*	rs17803441	*TYMS*	rs502396
SARDH	rs2073815	*SLC25A32*	rs3098243	*TYMS*	rs2612095
SHMT1	rs12952556	*SLC25A32*	rs3134297	*TYMS*	rs2853543
SHMT1	rs1979276	*SLC46A1*	rs2239908	*TYMS*	rs16948305
SHMT1	rs1979277	*SLC46A1*	rs2239907	*TYMS*	rs699517
SHMT1	rs2273028	*SLC46A1*	rs17719944	*TYMS*	rs2790
SHMT1	rs17806489	*TCN1*	rs17154234	*UBE2I*	rs9926094
SHMT1	rs4924750	*TCN1*	rs34324219	*UBE2I*	rs909915
SHMT1	rs2461838	*TCN1*	rs519221	*UBE2I*	rs8052688
SHMT1	rs643333	*TCN1*	rs557564	*UBE2I*	rs11248868
SHMT2	rs28365862	*TCN1*	rs2000613	*UBE2N*	rs4020454
SHMT2	rs7301155	*TCN1*	rs34528912	*UBE2N*	rs7311222
SLC19A1	rs1051298	*TCN1*	rs526934	*UBE2N*	rs7309933
SLC19A1	rs12482346	*TCN2*	rs5749131	*UBE2N*	rs7300607
SLC19A1	rs2297291	*TCN2*	rs5753231	*UBE2N*	rs1483003

[1]Formerly known as rs8111085; [2]Formerly known as rs9984077; [3]Formerly known as rs16834521; [4]Formerly known as rs10737812; [5]Formerly known as rs3177999.

TABLE A3.3

INTERACTIONS WITH *MTHFR 677 C→T* (rs1801133) WITH THE MOST STATISTICALLY SIGNIFICANT ASSOCIATION (LIKELIHOOD RATIO TEST P≤0.025) WITH THE PLASMA HOMOCYSTEINE PHENOTYPE[1,2]

Gene	rs#	Genetic Model	*MTHFR* genotype	Beta coefficient	LRT P	Chr	Coded allele	Coded allele frequency (%)	Type
DHFR	rs12517451	Overdominant	CT	0.0081	0.021	5	A	24.7%	3'
			TT	0.018					
GGH	rs11995525	Overdominant	CT	-0.012	0.021	8	A	26.9%	Intronic
			TT	-0.022					
DHFR	rs1650697[3,4]	Overdominant	CT	0.012	0.021	5	T	24.6%	5'
			TT	0.016					
CBS	rs234706[3]	Overdominant	CT	0.018	0.021	21	A	35.0%	Synonymous
			TT	0.0056					
TCN1	rs526934	Overdominant	CT	0.011	0.022	11	G	27.5%	Intronic
			TT	-0.015					
DNMT1	rs2162560	Dominant	CT	-0.011	0.022	19	A	41.9%	Intronic
			TT	-0.025					
AMT	rs8897[3,4]	Dominant	CT	-0.0029	0.022	3	T	26.8%	3'
			TT	0.033					
DNMT1	rs8101626	Dominant	CT	0.00051	0.022	19	G	42.5%	Intronic
			TT	-0.025					
SLC19A2	rs1983546[3]	Dominant	CT	0.0056	0.023	1	C	37.0%	Intronic
			TT	-0.029					
MTR	rs1050996	Dominant	CT	0.0079	0.023	1	G	41.2%	3'
			TT	0.023					
TCN2	rs2301957[3,5]	Overdominant	CT	-0.027	0.023	22	T	41.0%	Intronic
			TT	-0.010					
MTHFD1L	rs803422	Dominant	CT	-0.021	0.023	6	T	28.9%	Intronic
			TT	0.016					
DNMT3A	rs7578575[3]	Overdominant	CT	-0.026	0.023	2	A	30.6%	Intronic
			TT	-0.022					
BHMT	rs506500[3]	Dominant	CT	-0.029	0.023	5	T	30.7%	Intronic
			TT	-0.015					

TABLE A3.3 (Continued)

Gene	rs#	Genetic Model	*MTHFR* genotype	Beta coefficient	LRT P	Chr	Coded allele	Coded allele frequency (%)	Type
BHMT	rs7700970	Dominant	CT	0.0092	0.024	5	T	29.8%	Intronic
			TT	-0.036					
UBE2I	rs9926094[3]	Dominant	CT	0.0069	0.025	16	C	13.2%	Intronic
			TT	0.0049					
UBE2I	rs909915[3]	Dominant	CT	0.0094	0.025	16	T	13.0%	Intronic
			TT	-0.00019					
DNMT1	rs2228611[3]	Dominant	CT	0.0094	0.025	19	G	49.3%	Synonymous
			TT	0.015					

[1] Model adjusted for age and smoking; forward strand allele shown; gene and rs# represent SNP interacting with *MTHFR* rs1801133; interactions among SNPs within the same gene were not considered.
[2]False Discovery Rate-adjusted P values reached significance threshold of 0.025 for all tests presented.
[3]Sparse data (fewer than 5 individuals per category) for some genotype combinations.
[4]SNP maps to more than one gene (rs1650697 also maps to *MSH3*, rs8897 also maps to *NICN1*).
[5]Lower quality SNP.

TABLE A3.4

NUTRIENT INTERACTIONS WITH THE MOST STATISTICALLY SIGNIFICANT ASSOCIATION ($P \leq 0.02$) WITH THE PLASMA HOMOCYSTEINE (THC) PHENOTYPE[1-

Folate interactions

Gene	rs#	Nominal P	Beta coefficient	Chr	Coded allele	Coded allele frequency (%)	Genetic Model	Type
MTRR	rs7730643	1.87E-03	-0.38	5	G	19%	Recessive	Intronic
DHFR	rs1650697[4]	2.07E-03	-0.13	5	T	25%	Overdominant	5' region
DHFR	rs12517451	2.30E-03	-0.12	5	A	25%	Overdominant	3' region
SARDH	rs129891	2.55E-03	-0.09	9	A	34%	Additive	Intronic
DMGDH	rs532964	1.47E-02	-0.10	5	T	46%	Dominant	Coding nonsynonymous
Vitamin B-6 interactions								
GGH	rs4617146	1.22E-03	0.09	8	T	19%	Additive	Intronic
GLDC	rs1929933	1.67E-03	0.16	9	G	31%	Recessive	Intronic
FTH1	rs17185413[4]	2.16E-03	0.20	11	C	23%	Recessive	3' region
GNMT	rs11752813	4.67E-03	0.09	6	G	46%	Overdominant	5' region
MAT1A	rs1819684	9.63E-03	0.11	10	T	7%	Dominant	Intronic
SARDH	rs756682	1.22E-02	0.08	9	G	36%	Dominant	Intronic
SARDH	rs2073817	1.29E-02	0.08	9	A	36%	Dominant	Coding nonsynonymous
AMT	rs8897[4]	1.34E-02	0.08	3	T	27%	Dominant	5' region
MTHFD1L	rs7765521[5]	1.40E-02	-0.08	6	A	47%	Overdominant	Intronic
SARDH	rs129891	1.70E-02	0.05	9	A	34%	Additive	Intronic
MTHFR	rs1801133 CT	1.83E-02	-0.033	1	T	35%	Dummy	Coding nonsynonymous
	rs1801133 TT		0.0053					
Vitamin B-12 interactions								
MTHFD1L	rs9478162	2.46E-03	0.33	6	A	23%	Recessive	Intronic
MTHFD1L	rs9478908	2.89E-03	0.32	6	G	25%	Recessive	Intronic
DNMT3A	rs11678631	6.42E-03	0.17	2	A	49%	Recessive	Intronic
GGH	rs4617146	7.26E-03	-0.12	8	T	19%	Additive	Intronic
BHMT	rs16876512	9.46E-03	0.67	5	T	10%	Recessive	5' region
CTH	rs663465	9.75E-03	0.15	1	G	42%	Dominant	5' region
CTH	rs648743	9.89E-03	0.15	1	C	42%	Dominant	5' region
GNMT	rs2296805	1.66E-02	0.19	6	T	41%	Recessive	Intronic
ALDH1L1	rs3772424	1.71E-02	-0.13	3	A	20%	Dominant	Intronic

TABLE A3.4 (Continued)

Gene	rs#	Nominal P	Beta coefficient	Chr	Coded allele	Coded allele frequency (%)	Genetic Model	Type
Vitamin B-12 interactions								
GNMT	rs2296804[4]	1.78E-02	0.18	6	G	41%	Recessive	Intronic

[1]Model adjusted for age, smoking, and residuals of plasma folate, plasma vitamin B-6, and plasma vitamin B-12; forward strand allele shown; rs# represents SNP involved in interaction.
[2]No False Discovery Rate-adjusted P values reached significance threshold of 0.2.
[3]No sparse data (fewer than 5 individuals per category) for any genotype categories of these SNPs.
[4]SNP maps to more than one gene (rs1650697 also maps to *MSH3;* rs17185413 also maps to *BEST1*, rs8897 also maps to *NICN1;* rs2296804 also maps to *PEX6*).
[5]Lower quality SNP.

TABLE A3.5

NUTRIENT INTERACTIONS WITH THE MOST STIATISTICALLY SIGNIFICANT ASSOCIATION ($P \leq 0.02$) WITH THE GLOBAL GENOMIC DNA METHYLATION PHENOTYPE (ALU ELEMENTS)[1]

Folate interactions

Gene	rs#	Nominal P	Beta coefficient	Chr	Coded allele	Coded allele frequency (%)	Genetic Model	Type
SLC25A32	rs3098243	1.58E-03	0.47	8	C	43%	Overdominant	Intronic
SARDH	rs2502741[5]	2.78E-03	-0.44	9	A	50%	Overdominant	Intronic
ALDH1L1	rs4646760	1.05E-02	-0.68	3	G	35%	Recessive	Intronic
SLC46A1	rs17719944[3]	1.08E-02	-2.82	17	G	7%	Recessive	Intronic
GART	rs8788	1.17E-02	0.38	21	C	26%	Overdominant	Coding nonsynonymous
GLDC	rs4629927	1.57E-02	-0.36	9	G	41%	Overdominant	Intronic
DNMT3A	rs7578575	1.82E-02	0.35	2	A	31%	Overdominant	Intronic
Vitamin B-6 interactions								
AMT	rs1464567	5.27E-04[2]	-0.47	3	C	42%	Dominant	Intronic
AMT	rs1464566	7.97E-04[2]	-0.45	3	G	42%	Dominant	Intronic
DNMT3B	rs1883729	1.70E-03[2]	0.52	20	A	40%	Recessive	Intronic
TCN2	rs5749131	3.19E-03	-0.39	22	A	42%	Dominant	5' region
DNMT3B	rs2424922	5.19E-03	0.43	20	C	45%	Recessive	Coding synonymous
DNMT3B	rs6058891	5.19E-03	0.43	20	C	45%	Recessive	Coding synonymous
DNMT3B	rs2424914	5.31E-03	0.43	20	G	45%	Recessive	Intronic
BHMT	rs10037045	5.58E-03	0.36	5	T	29%	Dominant	Intronic
CELF1	rs2242081	6.69E-03	0.27	11	C	46%	Additive	Intronic
DNMT3B	rs6058869[5]	9.38E-03	0.45	20	T	39%	Recessive	5' region
TCN1	rs519221	1.24E-02	-0.59	11	T	27%	Recessive	Intronic
TCN1	rs557564[5]	1.24E-02	-0.59	11	T	27%	Recessive	Intronic
GLDC	rs1755617	1.33E-02	0.36	9	T	24%	Overdominant	Intronic
TCN1	rs34528912	1.46E-02	0.55	11	T	5%	Additive	Coding nonsynonymous
AMT	rs11922013	1.50E-02	0.25	3	C	30%	Additive	Intronic
ALDH1L1	rs1868138	1.79E-02	0.32	3	T	20%	Overdominant	Intronic
SLC19A3	rs11694828	1.91E-02	-0.31	2	A	45%	Overdominant	Intronic

TABLE A3.5 (Continued)

Gene	rs#	Nominal P	Beta coefficient	Chr	Coded allele	Coded allele frequency (%)	Genetic Model	Type
Vitamin B-12 interactions								
FPGS	rs10106[5]	1.28E-02	0.44	9	G	40%	Overdominant	3' region
FPGS	rs4451422	1.44E-02	0.43	9	G	40%	Overdominant	3' region
TYMS	rs699517[4]	1.48E-02	0.75	18	T	32%	Recessive	3' region
AHCYL2	rs1665105	1.54E-02	0.30	7	T	44%	Additive	3' region

[1]Model adjusted for age, smoking, and residuals of plasma folate, plasma vitamin B-6, and plasma vitamin B-12; forward strand allele shown; rs# represents SNP involved in interaction.
[2]False Discovery Rate-adjusted P values reached significance threshold of 0.2.
[3]Sparse data (fewer than 5 individuals per category) for some genotype categories of this SNP.
[4]SNP maps to more than one gene (rs699517 also maps to *ENOSF1*).
[5]Lower quality SNP.

TABLE A3.6 NUTRIENT INTERACTIONS WITH THE MOST STATISTICALLY SIGNIFICANT ASSOCIATION (P≤0.02) WITH THE GLOBAL GENOMIC DNA METHYLATION PHENOTYPE (LINE-1 ELEMENTS)[1,3,5]

Folate interactions

Gene	rs#	Nominal P	Beta coefficient	Chr	Coded allele	Coded allele frequency (%)	Genetic Model	Type
FOLH1	rs664584	1.53E-02	-0.52	11	A	24%	Additive	Intronic
BHMT	rs16876512	1.55E-02	-3.55	5	T	10%	Recessive	5' region
MAT1A	rs1985908	1.56E-02	0.62	10	C	32%	Dominant	3' region
DMGDH	rs28326	1.96E-02	0.58	5	G	18%	Additive	Intronic
Vitamin B-6 interactions								
MTHFD1L	rs17080689	5.40E-05[2]	-1.10	6	C	11%	Overdominant	Intronic
FTH1	rs1800009[4]	3.59E-03	0.65	11	G	34%	Overdominant	3' region

[1]Model adjusted for age, smoking, and residuals of plasma folate, plasma vitamin B-6, and plasma vitamin B-12; forward strand allele shown; rs# represents SNP involved in interaction.
[2]False Discovery Rate-adjusted P values reached significance threshold of 0.2.
[3]No sparse data (fewer than 5 individuals per category) for any genotype categories of these SNPs.
[4]SNP maps to more than one gene (rs1800009 also maps to *BEST1*).
[5]No lower quality SNPs.

CHAPTER 4

FOLATE NETWORK GENETIC VARIATION AND CARDIOVASCULAR DISEASE RISK: A NETWORK ASSOCIATION STUDY

4.1 Abstract

Sequence variants in genes functioning in folate-mediated one-carbon metabolism, particularly the *MTHFR* rs1801133 *677 C→T* polymorphism, have been associated with cardiovascular disease risk, however, most work has focused on genes related to homocysteine. 330 SNPs in 52 genes were prospectively evaluated for their association with cardiovascular disease in a U.S. population studied prior to the initiation of mandatory folate fortification. SNPs were selected based on functional effects and gene coverage and assayed on the Illumina Goldengate platform. Age- and smoking-adjusted genotype—phenotype associations were estimated in regression models. Using a nominal P≤0.005 threshold for statistical significance, 8 SNPs were associated with CVD risk in single locus analyses. Using a more stringent false discovery rate threshold, a polymorphism in the *GGH* gene was associated with reduced CVD risk. A gene x folate interactions was identified (*MAT2B*) and two gene x vitamin B-12 interactions were identified (*BHMT* and *SLC25A32*). Biological hypotheses related to *SHMT1* were explored and significant gene x gene interactions were identified for *TYMS* x *UBE2N*, *FTH1* x *CELF1*, and *TYMS* x *MTHFR*. A significant association with CVD risk involving variation in *MTHFD1L* was also identified. Finally, an interaction involving *CELF1* x *MTR* was identified as the most statistically significant from among all pair-wise network interactions. This study suggests that variation in genes other than *MTHFR* and those directly involved in

homocysteine metabolism, are associated with CVD risk and supports a role for *SHMT1* and nuclear folate metabolism, including the thymidylate biosynthesis pathway, in relation to CVD.

4.2 Introduction

Heart disease and stroke were responsible for about one-third of U.S. deaths in 2004 (16, 86), and there is extensive evidence on the contribution of dietary and lifestyle factors to risk (60). Many risk factors for cardiovascular disease (CVD) have been identified, including modifiable factors such as high blood cholesterol and lack of physical activity and non-modifiable factors such as heredity and male gender, but the origins of CVD are complex and the interactive effects of genetic and environmental factors, including nutrition, play an important role (26).

Folate and other B vitamins contribute to biologic processes important to health, including DNA synthesis and repair, and the generation of cellular methylation potential for a variety of methylation reactions. Folate status is influenced by dietary intake and by variation in genes encoding folate-related enzymes, and altered folate status is associated with adverse outcomes, including birth defects, cardiovascular disease (CVD), and cancer (32)

While several candidate genes contributing to folate mediated one-carbon metabolism have been thoroughly investigated in relation to CVD risk, most notably the *MTHFR 677 C→T* rs1801133 SNP, other genes in the one carbon metabolic pathway, particularly those not directly involved in homocysteine metabolism, are understudied in relation to CVD. A systematic review and meta-analysis was conducted in 2008 to evaluate the evidence for the association between genetic variation in a set of 52 genes involved in the folate-mediated one-carbon metabolic

pathway and cardiovascular disease risk (114). Overall, folate network genes involved in transsulfuration, cytoplasmic metabolism, and vitamin B-12 metabolism have been the focus of most work relating to CVD, while folate network genes involved in absorption and transport, mitochondrial metabolism, and nuclear folate metabolism are less studied in relation to CVD. For example, serine hydroxymethyltransferase 1 (soluble), *SHMT1*, encodes an enzyme that produces 5,10-methylenetetrahydrofolate (5,10-methyleneTHF), a key intermediate in the folate metabolic pathway(40), and has been shown to govern the competing flow of one-carbon units through thymidylate synthesis and methionine cycles (8, 42). However, despite epidemiologic evidence for an interaction between *SHMT1* rs1979277 and the *MTHFR 677 C→T* rs1801133 variant that is strongly predictive of vascular disease risk (62)(Wernimont et al, 2010, in press), only a handful of published studies have evaluated this gene in relation to CVD risk (Wernimont et al, 2010, in press). Another example is methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 1-like (*MTHFD1L*), which was identified in 2007 by the Wellcome Trust Case Control Consortium genome-wide association study of CAD; a significant association was reported for the *MTHFD1L* rs6922269 genetic variant(112). *MTHFD1L* encodes a monofunctional tetrahydrofolate synthase that participates in mitochondrial folate metabolism, an aspect of folate metabolism that is far less well-studied than the methylation portion of the folate network.

Thus, enzymes and genes less proximal to homocysteine metabolism may play important roles in predicting CVD, yet few studies evaluate a comprehensive set of genes in the folate pathway, in spite of an awareness of the importance of multiple genetic variants contributing jointly to increased disease risk (34). An investigation of genetic variation within this network of genes is essential to a complete understanding of how genetically-driven alterations in folate metabolism influence CVD risk.

Determining the role of genetic and nutritional factors in mediating CVD risk is ultimately the foundation for developing strategies for CVD prevention (26). This study investigated 330 single nucleotide polymorphisms (SNPs) in 52 genes with a role in folate-mediated one-carbon metabolism. The set of genes, SNP markers, and nutrients examined in this study were selected to represent the full functional variation of the folate-mediated one carbon metabolic pathway and prospectively evaluated for their association with cardiovascular disease in a U.S. population studied prior to the initiation of mandatory folate fortification.

4.3 Methods

Study Population

The Normative Aging Study (NAS) was established by the Veterans' Administration (VA) in 1961, and at that time 2,280 men aged 21-81 years (mean age of 42 y at study entry) were enrolled in the study on the basis of health criteria (9, 28). The rate of continued participation of NAS men over the follow-up period from 1961 through 1998 was excellent, with <1% attrition for all causes. As of June 1998, just prior to the end of the follow-up period for the present study, 543 participants (24%) were deceased and ~1600 men (70%) were actively participating in follow-up visits (mean age of 70 y). Men in the NAS were primarily non-Hispanic White males (>98% of the total NAS population) and the small number of participants in other racial/ethnic groups precluded separate analyses. Thus, this study focused on non-Hispanic white males using data on 1,131 men successfully genotyped.

This study was approved by the following: Brigham and Women's Hospital Human Subjects committee, Veterans' Administration R&D committee, and the Cornell University Committee on Human Subjects.

DNA Extraction

Genomic DNA from participants was extracted from stored frozen buffy coat of 7 ml whole blood using the QIAamp DNA Blood Kit (QIAGEN, Valencia, CA). The REPLI-g whole genome amplification kit (QIAGEN) was used to amplify genomic DNA when quantity was insufficient for genotyping. Whole-genome amplified samples were quality checked using Taqman genotyping (120) and poor-performing samples were excluded. Of 1,304 participant samples submitted for genotyping, 54.4% were genomic, 45.5% were whole-genome amplified.

SNP Selection

52 genes that contribute to folate-mediated one-carbon metabolism were identified (**Table A4.1**). SNP selection encompassed 2 kb on either side of the gene to include promoter and/or regulatory region variants. Four databases were used for SNP selection: the National Center for Biotechnology Information (NCBI) dbSNP and PubMed databases, (http://www.ncbi.nlm.nih.gov/), the Applied Biosystems Incorporated (ABI) SNPBrowser website (http://www.allsnps.com), and the Illumina Assay Design Tool (www.illumina.com). SNPs identified from literature searches and functional variants (non-synonymous coding region SNPs and promoter/regulatory region SNPs) were selected preferentially. Whenever possible, adjacent SNPs were selected such that the decay of maximum linkage disequilibrium between the 2 was no more than 33% (≤0.9 LDU between adjacent SNPs, where 1 LDU represents the decay of LD between two SNPs by about 37% of its maximum value when fitted to the Malecot model (67)) to ensure sufficient SNP density to adequately represent the LD characteristics of the gene. GoldenGate SNP validation status was considered at each step, and SNPs with a minor allele frequency (MAF) ≥5% in European-ancestry populations were selected where possible although exceptions were made for SNPs with prior evidence of putative function or when no

SNPs with MAF ≥5% were available. A total of 384 SNPs were selected, including SNPs that were intentionally redundant to provide coverage in the event that key SNPs of interest failed in the genotyping assay. To simultaneously evaluate genetic variation within the network as a whole, a subset of 52 non-redundant SNPs was identified with 1 SNP per gene comprising the most likely functional variant with the highest MAF that could be selected for each gene.

SNP Genotyping

384 SNPs were submitted to the Center for Inherited Disease Research at the Johns Hopkins University for genotyping via an Illumina GoldenGate custom genotyping panel. Genotype frequencies in controls were compared with those expected in Hardy-Weinberg equilibrium (HWE) and tested with a Fisher test using Monte Carlo permutation estimates of exact P-values for HWE (10,000 permutations). Of the 384 SNPs originally submitted, 54 were ultimately excluded because of assay failure (46), monomorphic genotype data (4), minor allele frequency <1% (3), or genotype frequencies out of HWE (1), leaving 330 SNPs for analysis (**Table A4.2**). Both blind duplicates and HapMap Centre d'Etude du Polymorphisme Humain (CEPH) control samples with known genotypes were included to assess assay quality; reproducibility rates were excellent (99.99% for blind duplicates, 99.83% for HapMap CEPH controls).

Covariates

Extensive previously collected data on male participants of the Normative Aging Study includes physical measurements, lifestyle factors, and blood biomarkers. Since the time of enrollment participants have had clinical examinations at 3- to 5-year intervals, with a response rate > 90% for mailed questionnaires. Fasting plasma samples were drawn at the VA field site and stored at -80 °C prior to transfer to the Jean Mayer USDA Human Nutrition Research Center on Aging, where they were

analyzed; the time between blood draw and analysis averaged 1.7 ± 1.2 y as previously described (102). Plasma biomarkers were assayed in an unselected subset of stored blood samples. Measures of plasma folate, vitamin B-6 (as pyridoxal-5'-phosphate; PLP) and vitamin B-12 are available for the majority of men with genotype data. Methods for these measurements have been previously described and coefficients of variation (CV) were excellent, with a value of 4.3% for folate, 5.0% for vitamin B-6, and 4.7% for vitamin B-12 (102). Plasma total homocysteine was assayed; methods were previously described, and the assay CV was 4.0% (102).

Up to three measurements of homocysteine, folate, vitamin B-6, and vitamin B-12 were available prior to 12/31/98, the end of follow-up, which coincides approximately with the start of mandatory folate fortification. The intraclass correlation coefficients for log-transformed homocysteine, folate, vitamin B-6, and vitamin B-12 confirmed that a high proportion of the total variance was between-person variability (vs. within person), thus, the first available biomarker measure was used in models to best represent nutrient status in the pre-mandatory folate fortification era, and to maximize case number since we only included cases with nutrients measured prior to disease onset. All biomarkers were log-transformed to adjust for skewness and centered for best performance in regression models.

Phenotype Assessment

Follow-up time for study participants was estimated as the time from study baseline to the first occurrence of: a CVD event (angina, fatal and nonfatal MI or other coronary heart disease, and fatal and nonfatal stroke) or death from a competing cause. For all other men, time of follow-up was time on study through to 12/31/1998, the end of follow-up for all study analyses.

Ascertainment of CVD events within the NAS has been described (62, 111). Briefly, the possible occurrence of nonfatal MI and angina was ascertained by

participant report, and then confirmed by a cardiologist. Similarly, the occurrence of a nonfatal stroke was ascertained by participant report, and confirmed by a neurologist. Fatal myocardial infarction and fatal stroke among men with no prior history of CVD were ascertained through death certificates, which were reviewed by a cardiologist to assign cause of death using International Classification of Disease (ICD) codes. CVD events comprised the following diagnosis (ICD, version 8) codes: heart disease (including angina and myocardial infarction) 410-414.9; and stroke 436.9, 430, 431. Ascertainment of fatal CVD events was nearly 100% (89, 111). In all instances, the confirming physician was unaware of the participant's exposure status (genotype). From 1964 through 1998, 381 cases of fatal or nonfatal CVD occurred among the 1,131 men (31,871 person-years of follow-up) with successful Illumina genotyping.

Statistical Analyses

Time to event analysis evaluated the relation of single locus markers to CVD risk, estimated as time between entry into the study and end of follow-up, where end of follow-up was date of first CVD event for men with the outcome, and either death from competing cause or right truncation at 12/31/98 for men with no CVD event. Genetic models of inheritance (additive, dominant, recessive, overdominant) were tested for each SNP and the model with the most statistically significant association with CVD risk in the unadjusted single SNP analysis was chosen as the best model going forward; an exception was made for the *MTHFR 677 C→T* rs1801133 SNP, where prior evidence supports a dummy coded model that evaluates the *CT* and *TT* genotypes separately against the *CC* reference category (35, 62, 108). Previous work in this cohort demonstrated no population substructure (115), thus no adjustments were made. All regression models were adjusted for age and smoking status, and a further model also adjusted for the *MTHFR 677 C→T* rs1801133 genotype. The interaction of each genotype with folate and vitamins B-6 and B-12 was tested; if

specific genetic models of inheritance led to sparse data in interaction models then additive coding was used as the default. Gene-gene interactions were also tested, using model-free dummy variable coding in all cases.

Regression coefficients with a nominal $P \leq 0.005$ were flagged to obtain a liberal set of associations for comparison with other studies and consideration for further follow up. Subsequent analyses used the False Discovery Rate (FDR) multiple testing correction of Benjamini and Hochberg (10) to adjust P-values, with a q-value significance threshold of 0.1, indicating that we expect less than 10% of tests flagged by this criterion to be false positives. For interaction effects statistical significance thresholds were less stringent: a nominal P value of $P \leq 0.02$ was used, and the FDR q value significance threshold was 0.2.

To investigate the joint role of nutrition and genetic variation, analyses of the SNP—phenotype association considered the possibility that nutrient biomarkers may mediate or modify the SNP—phenotype association. To assess mediation by a nutrient, models adjusting for the nutrient were compared to unadjusted models. To assess effect modification, product terms between the SNP and the nutrient biomarker were included in models. The statistical significance of the interaction was assessed by the P value for the interaction term. To facilitate evaluation of the genotype—nutrient interaction, a standardized approach was used, as follows: SNP—nutrient interactions were evaluated at 3 levels of the centered, log-transformed nutrients: the 10^{th} percentile ("low nutrient"), the 50^{th} percentile ("median nutrient"), and the 90^{th} percentile ("high nutrient").

All statistical analyses were conducted with SAS v. 9.2 (SAS, Cary, NC), unless otherwise specified.

4.4 Results

Among 1,131 men with genotype data on 330 SNPs, 381 CVD cases occurred; 14% were myocardial infarction only, 17% angina only, 4% other coronary heart disease only, 8% stroke only; the remaining cases had multiple diagnoses. A subset of 735 men had measurements for plasma nutrient biomarkers (folate, vitamins B-6 and B-12, and homocysteine). Among study participants the *MTHFR 677 C→T* rs1801133 *TT* genotype frequency was 12.8%, representing over 2200 chromosomes; this frequency is higher than a sample of 120 chromosomes from the HapMap CEPH population in which the *TT* frequency was 6.7%, but similar to the *TT* frequency in a sample of 5064 chromosomes from North American coronary heart disease controls in a meta-analysis (52) (**Table 4.1**). In exploratory regression models, prior to assessing the genotype—phenotype associations, baseline age was associated with CVD risk ($P \leq 0.0001$), as was baseline smoking status and change in smoking status over follow-up ($P \leq 0.002$). When modeled individually, plasma folate ($P=0.9$) and vitamin B-12 ($P=0.7$) had little or no association with CVD, and although vitamin B-6 ($P=0.077$) and homocysteine ($P=0.054$) had some association in unadjusted models, the associations were attenuated in age- and smoking- adjusted models.

All models were adjusted for age at study baseline and smoking status (two variables, baseline smoking status and change in smoking status over follow-up). Further adjusting for diabetes or hypertension (baseline status and change over follow-up) made little difference overall and identified the same set of most significant SNPs, although some minor changes to regression coefficients and P values were noted. The set of most significant SNPs, and the coefficient estimates and standard errors, were nearly identical in age- and smoking-adjusted models restricted to the subset of men

Table 4.1. Baseline characteristics of Normative Aging Study participants, 1961-1998

	N=1131[1]
Age	40.2 (7.8)
Education – college graduate or higher (%)	28.1
White (%)	100
BMI (kg/m^2)	25.9 (2.9)
Obese, BMI ≥ 30 kg/m^2 (%)	8.5
Cigarette smoking (%)	
Current	30.1
Former	37.5
Never	32.4
Alcohol intake, ≥2 drinks/day (%)	13.4
Fasting glucose (mg/dl)	98.2 (10.0)
Diabetes diagnosis (%)	0.18
Systolic blood pressure (mm Hg)	122.5 (12.7)
Diastolic blood pressure (mm Hg)	76.6 (8.6)
Hypertension diagnosis (%)	2.5
Total cholesterol (mg/dl)	203.1 (42.8)
MTHFR 677 C→T (rs1801133) *TT* genotype (%)	12.8
Plasma folate (ng/ml)[2]	10.4 (5.7)
Plasma PLP (pmol/ml)	84.8 (85.0)
Plasma B-12 (pg/ml)	458.0 (190.2)
Plasma homocysteine (nmol/ml)	10.6 (3.7)
Average follow-up time (years)	28.2 (6.7)

[*]mean (standard deviation) unless otherwise indicated
[1]N varies by characteristic as some measurements available on only a subset of data.
[2]Plasma measures of folate, vitamin B-6, vitamin B-12, and homocysteine were collected prior to the initiation of folate fortification.

with biomarker data (N=735), and further adjusting these models for plasma nutrient biomarkers made no difference to the findings. Adjusting for the *MTHFR 677 C→T* (rs1801133) variant had little or no effect on the coefficients for other SNPs.

Single Locus Marker Associations

The top SNP hits for the CVD phenotype were common variants with MAF typically $\geq$ 13% (**Table 4.2).** A nominal significance threshold of P $\leq$ 0.005 effectively separated the most significant results **(Figure 4.1**).

Of the 8 SNPs with P≤0.005, 1 SNP, gamma-glutamyl hydrolase (*GGH*) rs12544045, a 5' region marker, was the most significant hit, with a nominal P=0.00055 and a False Discovery Rate-adjusted P=0.091 (Table 4.2). The *AG* genotype was associated with a 32% reduction in CVD risk vs. the combined *AA/GG* reference genotypes (odds ratio (OR) 0.68; 95% confidence interval (CI) 0.5, 0.8). The second most significant hit was *MTHFS* rs7177659; the *AC* genotype was associated with a 29% reduction in CVD risk vs. the combined *AA/CC* reference category (OR 0.71; 95% CI 0.6, 0.9). The well-known methylenetetrahydrofolate reductase (*MTHFR*) rs1801133 coding nonsynonymous SNP was associated with risk: *CT* (vs. *CC*) was associated with a 6% increase in CVD risk (OR 1.06; 95% CI 0.8, 1.3), and *TT* (vs. *CC*) was associated with a 71% increase in CVD risk (OR 1.71; 95% CI 1.3, 2.3).

Of the top 8 SNPs, two SNP—CVD associations (rs6586282 and rs6586281, both in the cystathionine-beta-synthase gene, *CBS*) were substantially mediated by nutritional status. Regression coefficients were reduced by about 36% when adjusted

Table 4.2. Genotypes with the most statistically significant association (P≤0.005) with the CVD phenotype[1,3,4,5]

Gene	rs#	Estimate	Std Error	Odds Ratio	Nominal P	Chr.	Coded allele	Coded allele frequency	Genetic Model	SNP Type
GGH	rs12544045	-0.38	0.11	0.68	5.51E-04[2]	8	A	29%	Overdominant	5' region
MTHFS	rs7177659	-0.34	0.11	0.71	1.18E-03	15	A	47%	Overdominant	Intronic
MTHFR	rs1801133 *CT* rs1801133 *TT*	0.06 0.54	0.12 0.15	1.06 1.71	1.48E-03[6]	1	T	35%	Dummy variable	Coding nonsynonymous
TCN1	rs17154234	0.49	0.16	1.62	2.40E-03	11	C	5%	Overdominant	3' region
UBE2I	rs11248868	0.34	0.11	1.40	3.33E-03	16	G	13%	Overdominant	Intronic
CBS	rs6586282	-0.32	0.11	0.73	3.35E-03	21	T	18%	Additive	Intronic
CBS	rs6586281	-0.35	0.12	0.71	3.55E-03	21	A	18%	Dominant	Intronic
UBE2I	rs909915	0.33	0.11	1.39	4.02E-03	16	T	13%	Overdominant	Intronic

[1]Model adjusted for age and smoking; forward strand allele shown.
[2]Adjusted P values reached False Discovery Rate significance threshold of 0.1.
[3]No sparse data (fewer than 5 individuals per category) for genotype categories of these SNPs.
[4]No SNPs map to more than one gene.
[5]No lower quality SNPs.
[6]P value for *MTHFR* rs1801133 SNP only came from Likelihood Ratio Test for model with vs. without *CT* and *TT* genotype terms (Chi Sq. 13.033, 2df)

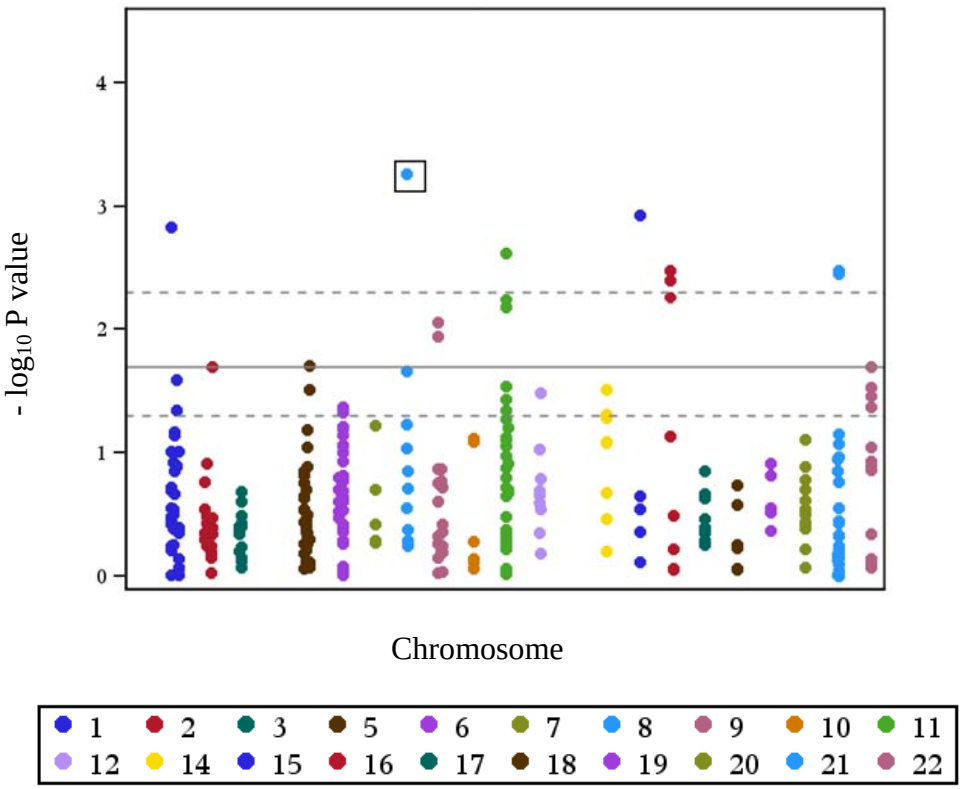

Figure 4.1. Manhattan plot: Folate-related SNPs as predictors of cardiovascular disease risk. Models adjusted for age and smoking status. Horizontal lines represent nominal P values of 0.05 (lower dashed line), 0.02 (center solid line) and 0.005 (upper dashed line). Box indicates SNP that reached False Discovery Rate significance.

for plasma folate, but further adjustment for plasma vitamins B-6, B-12, and homocysteine had no effect on the regression coefficients (**Table A4.3**). None of the associations between the other 6 SNPs and CVD risk was mediated by nutrient biomarkers or homocysteine. For some SNPs, associations were strengthened after adding biomarker terms to the model. The *MTHFS* rs7177659—CVD association increased by 60% after adjusting for plasma nutrients and homocysteine, the *TCN1* rs17154234—CVD association increased 17-19%, and the *MTHFR* rs1801133 *CT* genotype association increased 163% after adding biomarker terms, while the *TT* genotype association increased 21%.

Genetic variation in the folate-mediated one-carbon network as represented by a subset of 52 SNPs was predictive of CVD risk (Likelihood Ratio Test for the model with vs. without 52 SNPs =760.5; P value 6.19E-126, 53 degrees of freedom). In the simultaneous model, the *MTHFR* rs1801133 *TT* genotype was the top hit, and the association was increased compared to the single SNP model (simultaneous model OR 1.96, 95% CI 1.4, 2.7; nominal P= <0.0001). The top gene hits were similar between the single SNP models and the simultaneous model of 52 SNPs, and the rs11545078 SNP in *GGH*, the gene containing the most significant hit in the single SNP analysis, was the second most significant hit in the simultaneous model. Overall, regression estimates were similar between the two approaches, with one exception: the association of markers on the *CBS* gene with CVD risk was completely attenuated after adjusting for other SNPs.

Sensitivity analyses considered only those men with complete follow-up data (men who did not miss scheduled NAS follow-up visits by more than 1 yr), and regression coefficients and standard errors were similar to models including all participants. When the outcome was limited to early-onset CVD (age of onset < 55 years) the associations for some top hits were strengthened, while other associations

were attenuated (**Table A4.4**). The regression coefficient for the *GGH* rs12544045 SNP, the top hit in the original analysis, increased by 215%, and it remained the top hit in the early-onset analysis (nominal p value 0.00029, FDR-adjusted p value 0.095). The OR associated with *GGH* rs12544045 (OR 0.3; 95% CI 0.2, 0.6) was consistent with a 70% reduction in early-onset CVD risk in men with the *AG* genotype (vs. combined *AA/GG* genotypes). Similarly, the association of *MTHFS* rs7177659 with CVD risk increased by 86.3% (OR 0.53; 95% CI 0.6, 0.9; nominal P=0.017). *MTHFR* rs1801133 associations were consistent in direction with the model for all endpoints, but less statistically significant in the early-onset CVD analysis (P=0.34 for early onset CVD vs. P=0.0015 for all CVD). Finally, markers in the *UBE2I* gene, the ubiquitin-conjugating enzyme E2I (rs11248868, rs909915, and rs9926094), the ferritin, heavy polypeptide 1, *FTH1*, gene (rs2073588 and rs1801621), and the solute carrier family 19 (thiamine transporter), member 2, *SLC19A2*, gene (rs17518769 and rs2038024) were also represented in the set of top 10 hits for early-onset CVD (all with nominal P≤0.012). For two *UBE2I* SNPs, which were among the top 8 hits in the full CVD analysis, the strength of the association with CVD approximately doubled in the early-onset CVD analysis.

Gene-Nutrient Interactions

Plasma folate and vitamins B-6 and B-12 are cofactors for enzymes involved in the reactions of the one-carbon metabolic network, thus, the nutrients may modify the SNP—phenotype association. In further analyses assessing SNP—nutrient interactions (folate, B-6, or B-12), one folate interaction term reached FDR-significance levels (adjusted P≤0.2) (**Table A4.5**) for the rs6882306 variant in the methionine adenosyltransferase II, beta (*MAT2B*) gene. No vitamin B-6 interaction terms reached FDR-significance (adjusted P≤0.2), but 2 vitamin B-12 interaction terms were significant at FDR thresholds for the rs585800 SNP in betaine-

homocysteine S-methyltransferase (*BHMT*) and the rs1061196 SNP in solute carrier family 25, member 32 (*SLC25A32*).

Gene-Gene Interactions

In models investigating interactions between each SNP and *MTHFR 677 C→T* rs1801133, there were 8 interactions with a nominal P≤0.02; however, none reached the *a priori* FDR significance threshold of adjusted P≤0.2 (**Table A4.6**). The most statistically significant interaction involved the rs2118653 SNP in *GLDC*, with OR for *MTHFR TT* (vs. *CC*) of 2.29 (95% CI 1.5, 3.4) within the *GLDC* rs2118653 *GA* strata, and an OR for *MTHFR TT* (vs. *CC*) genotype of 1.86 (95% CI 1, 3.5) within the *GLDC* rs2118653 *AA* strata.

Given the interconnectedness of folate-mediated one-carbon metabolism, gene-gene interactions within the network are likely to be relevant to CVD risk. Thus, an analysis of all pairwise gene-gene interactions within the network was conducted. For this step, dummy variable coding was utilized for all genotypes to avoid making any assumptions about the correct genetic model. Of 52,832 unique pairwise interactions tested, 58 were associated with CVD risk at P ≤0.001 (**Table A4.7**). Using an FDR significance threshold of P≤0.2, none of the 58 associations reached the threshold. The most statistically significant interaction was between *CELF1* (formerly known as *CUGBP1*) rs7933019 and *MTR* rs16834388 (nominal P=0.00012, FDR-adjusted P=0.77). In men with the *CELF1* rs7933019 *GG* genotype, *MTR* rs16834388 *CT* and *TT* genotypes were associated with an increased risk of CVD (by 30% and 73%, respectively, vs. *MTR* rs16834388 *CC* genotype). In the other genotype classes of *CELF1* rs7933019 (*CG* and *CC* genotypes), the MTR—CVD association was reversed, thus genotype was inversely associated with CVD (**Table 4.3**).

Table 4.3. Odds ratios and 95% confidence intervals for network-wide gene-gene interaction most significantly associated with cardiovascular disease risk[1-4]

Gene	Genotype	OR	95% CI	OR	95% CI	OR	95% CI	LRT P	FDR P
CELF1 rs7933019:		*GG*		*CG*		*CC*		0.00012	0.77
MTR rs16834388	*GT vs.GG*	1.30	0.9, 1.8	1.03	0.7, 1.4	0.39	0.2, 0.8		
	TT vs.GG	1.73	1.1, 2.7	1.00	0.6, 1.7	0.13	0.0, 0.4		

[*]Odds ratios, 95% confidence intervals
[1]Model adjusted for age and smoking; forward strand allele shown.
[2]None of interaction estimates above were based on sparse data (fewer than 5 individuals per cell for any cells of the SNP by SNP frequency table).
[3]No SNPs map to more than one gene.
[4]No lower quality SNPs.

Beyond the discovery-based analyses presented above, these data presented the opportunity to directly explore hypotheses arising from basic science research and prior population-level studies. A reported interaction between the *MTHFR* and *SHMT* genes (62)(Wernimont et al, 2010, in press) was replicated here, and mediation by folate, vitamins B-6 and B-12, and homocysteine was explored. Adjusting for homocysteine strengthened the interaction (Likelihood Ratio Test P=0.010); thus, in men with the *SHMT* rs1979277 *TT* genotype, men with the MTHFR rs1801133 *CT* genotype had a 2.4-fold increased risk of CVD (vs. rs1801133 *CC*), and the risk increased to 10.4-fold when models were adjusted for homocysteine; little or no association of *MTHFR* genotype was evident in men with the *SHMT* rs1979277 *CT/CC* genotype, and this was not changed in homocysteine-adjusted models. Similarly, in men with the *SHMT* rs1979277 *TT* genotype, men with the MTHFR rs1801133 *TT* genotype had a 3-fold increased risk of CVD (vs. rs1801133 *CC*), and the risk increased to 15.5-fold when models were adjusted for homocysteine.

Previous *in vitro* studies have demonstrated protein-protein interactions between SHMT1 and UBC9 (yeast homolog of human UBE2I), resulting in the SUMOylation and nuclear localization of SHMT1, along with the other 2 components of thymidylate synthesis: TYMS and DHFR (2, 3, 116); UBC13 (yeast homolog of human UBE2N) has also been shown to be an SHMT1 binding partner, although the significance of this interaction is not yet fully understood (116). Other studies point to a role for CELF1 (formerly known as CUGBP1) and FTH1 in interacting with the SHMT1 mRNA to influence SHMT1 expression levels (117). The SHMT1 enzyme governs the competing flow of one carbon units to thymidylate biosynthesis or remethylation reactions (42), and thus modifications in protein expression levels or subcellular location may affect CVD risk by influencing the distribution of one-carbon units to competing reactions. Consistent with this hypothesis, 2 SNPs in *UBE2I*

(rs11248868 and rs909915) were among the most significant SNP hits for CVD overall (Table 4.2), and these SNPs, together with SNPs in *FTH1* (rs2073588 and rs1801621), were also among the most significant SNP hits for early-onset CVD (Table A4.4). The hypothesis was further explored by investigating all pairwise interactions among SNPs in the *UBE2I, UBE2N, DHFR, TYMS,* and *SHMT1* genes in relation to CVD risk. Of 451 unique pairwise interactions tested, 8 interactions had a nominal P≤0.02 (**Table A4.8**). The top hit was for the interaction of *UBE2N* rs7300607 and *TYMS* rs502396 (**Table 4.4**); in men with the *TYMS* rs502396 *CC* genotype, the *UBE2N* rs7300607 *CC* genotype (vs. *TT*) was associated with a 3-fold increased risk of CVD. Second, all pairwise interactions among SNPs in the *FTH1, CELF1,* and *SHMT1* genes were investigated in relation to CVD risk. Of 92 unique pairwise interactions tested, 16 interactions had a nominal P≤0.02 (**Table A4.9**). The top hit was for the interaction of *FTH1* rs17156609 and *CELF1* rs7933019 (Table 4.3); in men with the *FTH1* rs17156609 *GA* genotype, *CELF1* rs7933019 *CC* genotype (vs. *GG*) was associated with a 3-fold increased risk of CVD.

A landmark GWAS, the Wellcome Trust Case Control Consortium, previously identified an intronic SNP (rs6922269) in the *MTHFD1L* gene, a mitochondrial folate metabolism gene, which was associated with CAD risk (112). Based on HapMap LD plots, this SNP is located in a tightly linked 14 kb block in the *MTHFD1L* gene. Although the intronic rs6922269 SNP, along with the nearby nonsynonymous rs9767752 SNP, both failed in genotyping , 6 other SNPs successfully tagged this region. Previous work in the NAS identified joint genotypes involving 2 SNPs within that 14 kb region that best captured the association with CVD risk (113) and those SNPs

Table 4.4. Odds ratios, 95% confidence intervals for gene-gene interactions most significantly associated with cardiovascular disease risk in tests of three hypotheses[1-4]

Hypothesis:	Gene	Genotype	OR[5]	95% CI[5]	OR	95% CI	OR	95% CI	LRT P
Nuclear one-carbon metabolism	*TYMS* rs502396:		*TT*		*TC*		*CC*		0.0034
	UBE2N rs7300607	*TC vs.TT*	0.77	0.5, 1.2	1.20	0.8, 1.7	3.60	1.7, 7.6	
		CC vs.TT	1.29	0.8, 2.1	1.21	0.8, 1.9	3.13	1.4, 7.0	
SHMT1 expression	*FTH1* rs17156609:[6]		*GG*		*GA*				0.0074
	CELF1 rs7933019	*CG vs.GG*	1.08	0.9, 1.3	0.35	0.1, 1.5			
		CC vs.GG	0.92	0.6, 1.4	3.27	1.2, 8.7			
Mitochondrial one-carbon metabolism	*TYMS* rs502396:		*TT*		*CT*		*CC*		0.010
	MTHFR rs12121543	*AC vs.CC*	0.93	0.6, 1.4	1.22	0.9, 1.7	0.88	0.5, 1.4	
		AA vs.CC	0.21	0.1, 0.7	0.67	0.3, 1.4	1.83	0.9, 3.9	

[1]Model adjusted for age and smoking; forward strand allele shown.
[2]Presented P values were not False Discovery Rate-adjusted in this hypothesis-driven analysis.
[3]None of interaction estimates above were based on sparse data (fewer than 5 individuals per cell for any cells of the SNP by SNP frequency table).
[4]No lower quality SNPs.
[5]Odds ratios and 95% confidence intervals
[6]SNP maps to more than one gene: rs17156609 also maps to bestrophin 1 (*BEST1*).

(rs1474787 and rs803447) were further explored. Both SNPs were associated with CVD in the single SNP analysis, but the nominal P values were not below the thresholds set *a priori*. Exploring joint genotypes involving these SNPs revealed a highly statistically significant association with CVD (**Table 4.5)**; within the rs1474787 *GA* genotype strata, men with the rs803447 *TT* genotype had a CVD risk 3.63 times that of men with the rs803447 *CC* genotype, (interaction P=0.077). Adjusting for homocysteine strengthened this interaction; thus, within the rs1474787 *GA* genotype strata, men with the rs803447 *TT* genotype had a CVD risk 6 times that of men with the rs803447 *CC* genotype (interaction P=0.077).

Because mitochondrial folate metabolism produces formate, a major source of cytoplasmic one-carbon units (32), we hypothesized interactions may exist among *MTHFD1L, SHMT1* (as a determinant of one-carbon unit distribution among the remethylation and thymidylate synthesis pathways (42), and two downstream enzymes in each pathway, *MTHFR* and *TYMS*. Of 377 unique pairwise interactions tested, 4 interactions had a nominal P≤0.02 (**Table A4.10**). The top hit was for the interaction of *TYMS* rs502396 and *MTHFR* rs12121543 (Table 4.4); in men with the *TYMS* rs502396 *CC* genotype, the *MTHFR* rs12121543 *AA* genotype (vs. *CC*) was associated with a 2-fold increased risk of CVD.

4.5 Discussion

We investigated sequence variation in a network of candidate genes involved in one-carbon metabolism in relation to CVD risk in a time-to-event analysis. Genes, SNPs, and related nutrients were carefully chosen to represent the full functional variation of the folate-mediated one carbon metabolism pathway. The analysis considered all 330 SNPs in tests for main effects, interactions with the *MTHFR*

Table 4.5. Odds ratios and 95% confidence intervals for *MTHFD1L* 14 kb region genotypes most significantly associated with cardiovascular disease risk[1-4]

Gene	Genotype	OR*	95% CI*	OR	95% CI	OR	95% CI	LRT P
Unadjusted								
MTHFD1L rs1474787:			*GG*		*GA*		*AA*	0.077
MTHFD1L rs803447[4]	*CT vs.CC*	1.15	0.8, 1.7	1.24	0.8, 1.8	1.01[2]	0.1, 7.5	
	TT vs.CC	1.15	0.8, 1.7	3.63	1.8, 7.5			
Homocysteine Adjusted								
MTHFD1L rs1474787:			*GG*		*GA*		*AA*	0.077
MTHFD1L rs803447[4]	*CT vs.CC*	1.32	0.8, 2.1	1.62	1.0, 2.7	1.21[2]	0.2, 9.2	
	TT vs.CC	1.37	0.8, 2.3	6.14	2.5, 15.3			

*Odds ratios, 95% confidence intervals
[1]Model adjusted for age and smoking; forward strand allele shown.
[2]Sparse data (fewer than 5 individuals) in this cell of the SNP by SNP frequency table.
[3]No SNPs map to more than one gene.
[4]Lower quality SNP.

rs1801133 SNP, and a simultaneous evaluation of network variation. The interaction of each SNP with plasma measures of folate, vitamins B-6 and B-12 was also considered. FDR-significant main effects were identified, as well as FDR significant SNP—folate and SNP—vitamin B-12 interactions and nominally significant gene-gene interactions were identified in a more thorough investigation of 3 hypotheses about the network.

The study identified several novel results. Genetic variation in genes related to folate absorption and transport, such as the gamma-glutamyl hydrolase gene (*GGH*), were more significantly predictive of CVD risk than any other genetic polymorphism in the folate-mediated one-carbon network including *MTHFR* rs1801133, and effect sizes were doubled in analyses of early-onset CVD. Genes in this category of folate metabolism are not well-studied in relation to CVD risk. The simultaneous evaluation of genetic variation in 52 genes of folate-mediated one-carbon metabolism showed stronger associations with CVD risk than the *MTHFR* rs1801133 polymorphism considered alone, confirming the importance of multiple genes within the network. An in-depth evaluation of nuclear folate metabolism revealed significant epistasis predictive of CVD risk, and genes that regulate *SHMT1* expression interact to influence risk of CVD. The findings underscore the importance of mitochondrial metabolism to the folate pathway association with CVD risk and epistasis predictive of CVD risk exists between the thymidylate synthesis and remethylation pathways. Finally, gene-nutrient interactions, beyond the well-known *MTHFR* rs1801133-folate interaction, are strongly associated with CVD risk, including *MAT2B* and folate, *BHMT* and vitamin B-12, and *SLC25A32* and vitamin B-12.

In studying the association of single locus markers with CVD risk, genes in absorption and transport, purine metabolism, the methylation/homocysteine pathway, nuclear metabolism, and B-12 metabolism were represented in the top hits. Other than

SNPs in the *CBS* gene, there was little evidence that plasma folate or vitamins B-6 or B-12 or homocysteine mediated the association of SNPs with CVD risk; in fact, regression coefficients for several SNPs, including the *MTHFR* rs1801133 variant, as well as SNPs in the *MTHFS* and *TCN1* genes, strengthened after models were adjusted for plasma folate, B-6, B-12, and homocysteine (Table A4.3). When restricted to early onset cases only (onset age <55 yrs), main effects for SNPs in the *GGH*, *MTHFS*, and *UBE2I* genes strengthened, consistent with a true effect of these genes on risk.

Single Locus Marker Associations

GGH. An FDR-significant association was identified between the intronic rs12544045 SNP in *GGH* and CVD risk. The *AG* genotype group had a 32% lower risk of CVD compared to the *AA*/*GG* genotype group. The *GGH* gene encodes an enzyme that functions to hydrolyze polyglutamate residues from intracellular folate to facilitate folate export (46). The SNP is located in the 5' regulatory region of the gene; cell culture studies identified several other polymorphisms in the 5' region of this gene associated with increased promoter activity (19), and one of these SNPs, *GGH -124T→G*, was associated with a dose-dependent increase in DNA uracil content in humans (30). Since the SNP may affect cellular folate export and therefore measures of folate status, a reasonable assumption is that the SNP—CVD association might be mediated by plasma folate; however, in analyses adjusted for plasma folate, the coefficient for *GGH* rs12544045 decreased only slightly. Given our current understanding of the role of *GGH* in contributing to cellular folate export, there are several possible explanations for this finding. First, the marker of folate status available, plasma folate, may not represent the full extent of mediation of the association of *GGH* rs12544045 with CVD, either because it does not reflect folate status in the compartment relevant to CVD risk, or it does not reflect folate status

during the window of time relevant for CVD development, or our understanding of the role of *GGH* in contributing to CVD risk is incomplete.

MTHFR. There was a nominally significant association between the *MTHFR 677 C→T* rs1801133 SNP and CVD risk; the rs1801133 *CT* (vs. *CC*) genotype was associated with a 6% increase in CVD risk (OR 1.06; 95% CI 0.8, 1.3), and the rs1801133 *TT* (vs. *CC*) genotype was associated with a 71% increase (OR 1.71; 95% CI 1.3, 2.3). The *MTHFR* gene encodes an enzyme that catalyzes the conversion of 5,10-methyleneTHF to 5-methylTHF, which supplies a one-carbon unit for methylation reactions including the remethylation of homocysteine to methionine, as well as the generation of S-adenosylmethionine (AdoMet), the universal methyl donor (32) and is well-studied in relation to CVD risk (15, 27, 52, 58, 108). The findings in the present study are consistent with these prior reports.

Gene-Nutrient Interactions

The *MTHFR* rs1801133 polymorphism results in the production of a thermally labile enzyme with reduced activity (35); the effect of the SNP varies by folate status (47, 52). A gene-nutrient interaction between *MTHFR 677 C→T* rs1801133 and folate was not statistically significant in these data (P= 0.90) perhaps due to the relatively high folate status among the men in this cohort study (mean plasma folate: 10.4 ng/ml, significantly higher than the prefortification serum folate mean of 5.8 ng/ml in non-Hispanic Whites reported by NHANES III (1988-1994), albeit lower than the post-fortification mean of 14.8 ng/ml (80). SNP—nutrient interactions may be blunted in this range of folate status. When the dataset was restricted to men with folate levels below the 25th percentile (6.20 ng/ml, 187 men/59 CVD events), the interaction was more evident, and marginally statistically significant. The observed effect sizes were consistent with the magnitude and direction of the well-known *MTHFR 677*-folate interaction (15, 27, 52, 58, 108).

The association of variants in folate-related genes with CVD risk, including *MTHFR 677 C→T* rs1801133, is hypothesized to be mediated by homocysteine, which is sensitive to B vitamin status, including folate, vitamin B-12 and to a lesser extent, vitamin B-6(52) (45); thus, gene-nutrient interactions may exist such that the association of a given polymorphism with CVD risk may be modified by nutrient status. In this study, there were no FDR-significant interactions between studied SNPs and vitamin B-6. However, FDR-significant interactions were identified for plasma folate (methionine adenosyltransferase II, beta, *MAT2B*, rs6882306, intronic) and vitamin B-12 (betaine-homocysteine S-methyltransferase, *BHMT*, rs585800, 3' region; and, solute carrier family 25, member 32, *SLC25A32*, rs1061196, 3' region). The lack of additional highly significant gene-nutrient interactions may reflect limited power for gene-nutrient interactions, or the lack of men with low B vitamin status, even within this pre-fortification cohort of U.S. men.

MAT2B encodes the regulatory subunit of the methionine adenosyltransferase (MAT) enzyme that catalyzes the synthesis of S-adenosylmethionine from methionine and ATP (Entrez Gene). S-adenosylmethionine (or AdoMet) is the universal methyl donor for numerous intracellular methylation reactions. At high folate levels (17.27 ng folate/ml) men with the *MAT2B* rs6882306 *CC* genotype had 91% lower CVD risk (vs. combined *CT/TT*); at median folate levels (9.62 ng folate/ml) men with the *MAT2B* rs6882306 *CC* genotype had 42% lower CVD risk (vs. combined *CT/TT*); at low folate levels (4.38 ng folate/ml) men with the *MAT2B* rs6882306 *CC* genotype had a CVD risk that was 6.93 times that of the *CT/TT* reference group. The rs6882306 SNP is an intronic SNP that may tag functional variation elsewhere in the gene; HapMap LD plots are incomplete, but suggest high linkage. While no previous papers report an interaction between *MAT2B* polymorphisms and folate, a recent study identified variants in a closely related gene, methionine adenosyltransferase I, alpha

(*MAT1A*), that decreased risk of hypertension and DNA damage and increased risk of stroke, independently of the *MTHFR 677* rs1801133 polymorphism (55); *MAT1A* interactions with folate were predictive of homocysteine (low folate and the variant predicted high homocysteine), and consistent with the direction of effect seen for the *MAT2B*-folate interaction in relation to CVD risk.

The *BHMT* gene encodes a cytosolic enzyme that catalyzes the transfer of a methyl group from betaine to homocysteine, yielding dimethylglycine and methonine (Entrez Gene). This pathway provides an alternate route for the remethylation of homocysteine to methionine compared to the vitamin B-12-dependent methionine synthase pathway and may account for up to half of the homocysteine remethylation capacity of the cell in some tissues (59). The rs585800 SNP is a 3' region polymorphism and thus may have a regulatory function; at high B-12 levels (678.89 pg B-12/ml), CVD risk for the *BHMT* rs585800 *TA* genotype class was 36% lower compared to the *AA/TT* reference group. At median B-12 levels (418.62 pg B-12/ml), CVD risk for the *BHMT* rs585800 *TA* genotype class was 19% higher compared to *AA/TT*, and at low B-12 levels (259.29 pg B-12/ml), CVD risk for the *BHMT* rs585800 *TA* genotype class was 2.19 times higher compared to the *AA/TT* reference group. While the BHMT enzyme is not known to bind vitamin B-12, previous work has identified regions of homology with bacterial vitamin B-12-dependent methionine synthases (38). In rats, hepatic BHMT activity was not sensitive to vitamin B-12 deficiency(31), however, epidemiologic work in humans suggests that enhancing methionine-synthase dependent remethylation of homocysteine down regulates BHMT reactions (44).

The *SLC25A32* gene encodes a folate transporter that shuttles folate from the cytoplasm into the mitochondria (Entrez Gene). The rs1061196 polymorphism is a 3' region variant and thus may have a regulatory function. At low B-12 levels (259.29

pg/ml), CVD risk in men with the *SLC25A32* rs1061196 *AG* genotype was 55% lower compared to the *AA/AG* reference group; at median B-12 levels (418.62 pg/ml) CVD risk in men with the *SLC25A32* rs1061196 *AG* genotype was 17% lower compared to *AA/AG*; at high B-12 levels (678.89 pg/ml) CVD risk in men with the *SLC25A32* rs1061196 *AG* genotype was 1.56 times higher than the *AA/AG* reference group. No prior reports link *SLC25A32* to biochemical or disease phenotypes, and a biological basis for the link to vitamin B-12 metabolism could not be identified.

Gene-Gene Interactions

No FDR-significant interactions were identified between studied SNPs and *MTHFR 677 C→T* rs1801133. Furthermore, of all pairwise interactions within the network, no interactions surpassed the FDR significance threshold of $P \leq 0.2$.

In exploring hypotheses derived from prior findings, we focused on the *SHMT1* intersection in the folate network. Previous work (62) (Wernimont et al, 2010, in press) identified a gene-gene interaction between the *MTHFR* rs1801133 and the rs1979277 variant in the serine hydroxymethyltransferase 1 (soluble) gene, *SHMT1*, in which the association of the *MTHFR* rs1801133 *CT* and *TT* genotypes was much stronger among individuals who also carried the *SHMT1* rs1979277 *TT* genotype (62)(Wernimont et al, 2010, in press). The interaction was replicated here in an age- and smoking-adjusted model, and suggests that appropriate modeling of the *MTHFR* rs1801133 SNP is critical to detecting the gene-gene interaction, which exists only within a specific population subgroup.

Finally, pleiotropic effects were examined across outcomes. Previous work in the Normative Aging Study identified sets of SNPs most significantly associated with biomarkers of CVD risk, including plasma homocysteine and global genomic DNA methylation (Wernimont, in preparation). Of the 8 nominally significant hits for the CVD analysis, only a single pleiotropic SNP hit was identified, *CBS* rs6586282, which

was also among the most significant hits for the plasma homocysteine phenotype (nominal P value for CVD was 3.3 E-03, nominal P value for homocysteine was 4.06 x E-03). However, this SNP was not among the most significant hits for LINE-1 or ALU methylation; in fact, there were no SNP hits shared among the most significant hits for CVD and either the LINE-1 or ALU methylation phenotypes. The overdominant genetic model was the most statistically significant model in the homocysteine analysis, and the *CBS* rs6586282 *CT* (vs. combined *CC/TT*) genotype was associated with a 5.8% reduction in plasma homocysteine. Consistent with this, the *T* allele was associated with a reduction in CVD risk of 29%, in an additive model (selected as the best code for CVD analysis). The *CBS* rs6586282 polymorphism was the only SNP in the CVD analysis mediated by nutrition, which is consistent with the association of this polymorphism with both homocysteine and CVD. While the *CBS* rs6586282 SNP was the only SNP hit replicated across any of the 4 phenotypes studied, there were genes that had associations with multiple phenotypes; *GGH* was nominally significant for plasma homocysteine and FDR-significant for CVD, and *MTHFR* was nominally significant for CVD and LINE-1 methylation.

4.6 Conclusion

Strengths of the present study include a complete investigation of genetic variation in the folate metabolic network. Furthermore, the analysis was conducted within a large cohort with CVD outcome data collected prior to the introduction of mandatory folate-fortification in the U.S.; thus, the study population is well-suited to investigate the network of folate-related genes in relation to CVD risk. SNP selection for the genotyping assay reflected coverage of genes based on function, linkage, and physical coverage, which led to markers that captured functional variants in addition

to tagging variation elsewhere in the gene. In a systematic approach to the data analysis, we identified the best genetic models for each SNP, tested single SNPs, a reduced set of non-redundant SNPs, the interaction of each SNP with *MTHFR* 677 rs1801133, and the interaction of each SNP with folate, vitamin B-6 and vitamin B-12. Findings were corrected for multiple comparisons and findings based on the FDR threshold were discussed in more detail. Hypotheses based on basic science findings were explored. Weaknesses of the study include the fact that, despite restricting data to the pre-folate fortification time period, the level of folate nutrition within the population was relatively high compared to other measures of pre-fortification folate status in the U.S., which may have attenuated associations, particularly gene-nutrient interactions, that would have been present in a population with more variation in B-vitamin status.

This study suggests that variation in genes other than *MTHFR* and those directly involved in homocysteine metabolism, are associated with CVD risk. For example, a gene involved in folate absorption and transport, *GGH*, was more significantly associated with CVD than any of the other 51 genes evaluated, including *MTHFR* and *CBS*. This study also supports a role for mitochondrial metabolism in predicting CVD risk, and although a previously identified polymorphism in *MTHFD1L* (112) could not be studied directly, nearby variants showed strong associations with CVD risk. The epidemiologic findings presented here replicated molecular biology studies of interactions among the SHMT1, UBE2I, UBE2N, DHFR, and TYMS proteins resulting in the cell cycle-dependent nuclear localization of the thymidylate biosynthesis pathway, and further established that epistatic interaction among genes encoding these proteins is predictive of CVD risk. Polymorphisms in genes such as *FTH1* and *CELF1*, whose products have been shown to influence expression of *SHMT1*, a key regulator of the distribution of one-carbon units, were

evaluated and epistatic interactions among these genes were found to predict CVD risk. These findings support a role for *SHMT1* and nuclear folate metabolism, including the thymidylate biosynthesis pathway, in relation to CVD. This investigation of genetic variation within the folate network, including absorption and transport, mitochondrial, and nuclear metabolism, alters our current understanding of the relation between the folate-mediated one-carbon network and CVD. To date, the focus has been on the role of homocysteine, but the findings in this study suggest important contributions to risk through other aspects of folate metabolism. These findings may help explain the results of randomized controlled trials focused on CVD risk reduction through homocysteine lowering, which have experienced limited success (70, 71). A thorough understanding of the role of folate network genetic and nutritional variation in relation to CVD is important, particularly in the context of the numerous unfortified populations around the world.

APPENDIX

TABLE A4.1

THE 52 GENES IN THE FOLATE-MEDIATED ONE-CARBON PATHWAY

Gene Symbol	Gene Name	Entrez GeneID
AHCY	Adenosylhomocysteinase	191
AHCYL1	Adenosylhomocysteinase-like 1	10768
AHCYL2	Adenosylhomocysteinase-like 2, KIAA0828	23382
ALDH1L1	Aldehyde dehydrogenase 1 family, member L1	10840
AMT	Aminomethyltransferase	275
ATIC	5-aminoimidazole-4-carboxamide ribonucleotide formyltransferase/IMP cyclohydrolase	471
BHMT	Betaine-homocysteine S-methyltransferase	635
CBS	Cystathionine-beta-synthase	875
CTH	Cystathionase (cystathionine gamma-lyase)	1491
CELF1	CUGBP, Elav-like family member 1	10658
DHFR	Dihydrofolate reductase	1719
DMGDH	Dimethylglycine dehydrogenase	29958
DNMT1	DNA (cytosine-5-)-methyltransferase 1	1786
DNMT3A	DNA (cytosine-5-)-methyltransferase 3 alpha	1788
DNMT3B	DNA (cytosine-5-)-methyltransferase 3 beta	1789
FOLH1	Folate hydrolase (prostate-specific membrane antigen) 1	2346
FOLR1	Folate receptor 1 (adult)	2348
FOLR2	Folate receptor 2 (fetal)	2350
FOLR3	Folate receptor 3 (gamma)	2352
FPGS	Folylpolyglutamate synthase	2356
FTCD	Formiminotransferase cyclodeaminase	10841
FTH1	Ferritin, heavy polypeptide 1	2495
GART	Phosphoribosylglycinamide formyltransferase, phosphoribosylglycinamide synthetase, phosphoribosylaminoimidazole synthetase	2618
GCSH	Glycine cleavage system protein H (aminomethyl carrier)	2653
GGH	Gamma-glutamyl hydrolase (conjugase, folylpolygammaglutamyl hydrolase)	8836
GLDC	Glycine dehydrogenase (decarboxylating)	2731
GNMT	Glycine N-methyltransferase	27232
HSPA8	Heat shock 70kDa protein 8	3312
MARS	Methionyl-tRNA synthetase	4141
MAT1A	Methionine adenosyltransferase I, alpha	4143
MAT2A	Methionine adenosyltransferase II, alpha	4144
MAT2B	Methionine adenosyltransferase II, beta	27430
MTHFD1	Methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 1, methenyltetrahydrofolate cyclohydrolase, formyltetrahydrofolate synthetase	4522
MTHFD1L	Methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 1-like	25902
MTHFD2	Methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 2, methenyltetrahydrofolate cyclohydrolase	10797
MTHFR	Methylenetetrahydrofolate reductase (NAD(P)H)	4524
MTHFS	5,10-methenyltetrahydrofolate synthetase (5-formyltetrahydrofolate cyclo-ligase)	10588
MTR	5-methyltetrahydrofolate-homocysteine methyltransferase	4548

TABLE A4.1 (Continued)

Gene Symbol	Gene Name	Entrez GeneID
MTRR	5-methyltetrahydrofolate-homocysteine methyltransferase reductase	4552
SARDH	Sarcosine dehydrogenase	1757
SHMT1	Serine hydroxymethyltransferase 1 (soluble)	6470
SHMT2	Serine hydroxymethyltransferase 2 (mitochondrial)	6472
SLC19A1	Solute carrier family 19 (folate transporter), member 1	6573
SLC19A2	Solute carrier family 19 (thiamine transporter), member 2	10560
SLC19A3	Solute carrier family 19, member 3	80704
SLC25A32	Solute carrier family 25, member 32	81034
SLC46A1	Solute carrier family 46 (folate transporter), member 1	113235
TCN1	Transcobalamin I (vitamin B-12 binding protein, R binder family)	6947
TCN2	Transcobalamin II	6948
TYMS	Thymidylate synthetase	7298
UBE2I	Ubiquitin-conjugating enzyme E2I (UBC9 homolog, yeast)	7329
UBE2N	Ubiquitin-conjugating enzyme E2N (UBC13 homolog, yeast)	7334

TABLE A4.2

330 FOLATE-RELATED SNPS ASSAYED IN MEN IN THE NORMATIVE AGING STUDY[1-5]

Gene	SNP Name	Gene	SNP Name	Gene	SNP Name
AHCY	rs819133	*BHMT*	rs558133	*DMGDH*	rs1805073
AHCY	rs819159	*BHMT*	rs585800	*DMGDH*	rs532964
AHCY	rs819155	*CBS*	rs706209	*DMGDH*	rs2253262
AHCY	rs13043752	*CBS*	rs2124458	*DMGDH*	rs644191
AHCY	rs819148	*CBS*	rs760124	*DNMT1*	rs8101626
AHCY	rs1205366	*CBS*	rs6586281	*DNMT1*	rs11880388
AHCY	AHCYC34T	*CBS*	rs6586282	*DNMT1*	rs2228611
AHCY	rs819146	*CBS*	rs11203172	*DNMT1*	rs2228612[1]
AHCYL1	rs333079	*CBS*	rs234704	*DNMT1*	rs2162560
AHCYL1	rs3768480	*CBS*	rs1801181	*DNMT3A*	rs11695471
AHCYL1	rs2298116	*CBS*	rs2014564	*DNMT3A*	rs6546045
AHCYL1	rs720917	*CBS*	rs1789953	*DNMT3A*	rs7578575
AHCYL1	rs186724	*CBS*	rs12329764	*DNMT3A*	rs6733868
AHCYL2	rs6467233	*CBS*	rs234705	*DNMT3A*	rs11678631
AHCYL2	rs4731569	*CBS*	rs234706	*DNMT3A*	rs1550117
AHCYL2	rs7788327	*CBS*	rs234709	*DNMT3B*	rs6058869
AHCYL2	rs6467244	*CBS*	rs234711	*DNMT3B*	rs1883729
AHCYL2	rs1665105	*CBS*	rs11701048	*DNMT3B*	rs2424914
ALDH1L1	rs4646760	*CBS*	rs1788484	*DNMT3B*	rs6058891
ALDH1L1	rs4646750	*CTH*	rs648743	*DNMT3B*	rs2424922
ALDH1L1	rs4646745	*CTH*	rs663465	*DNMT3B*	rs6058896
ALDH1L1	rs3772424	*CTH*	rs681475	*FOLH1*	rs16906158
ALDH1L1	rs3772414	*CTH*	rs535112	*FOLH1*	rs202673
ALDH1L1	rs2305230	*CTH*	rs663649	*FOLH1*	rs664584
ALDH1L1	rs11715574	*CTH*	rs1021737	*FOLH1*	rs202676
ALDH1L1	rs1868138	*CUGBP1*	rs7102372	*FOLR1*	rs2071010
ALDH1L1	rs1823213	*CUGBP1*	rs2242081	*FOLR1*	rs9282688
AMT	rs11922013	*CUGBP1*	rs7933019	*FOLR2*	rs514933
AMT	rs1464567	*CUGBP1*	rs4752843	*FOLR2*	rs2298444
AMT	rs1464566	*DHFR*	rs12517451	*FOLR3*	rs7926875
AMT	rs8897	*DHFR*	rs1677666	*FOLR3*	rs11235449
ATIC	rs7585489	*DHFR*	rs1650723	*FPGS*	rs10106
ATIC	rs2372536	*DHFR*	rs2560424	*FPGS*	rs41319447
ATIC	rs3821353	*DHFR*	rs1643659	*FPGS*	rs4451422
ATIC	rs1997059	*DHFR*	rs1643650	*FTCD*	rs16978930
ATIC	rs4672768	*DHFR*	rs836822	*FTCD*	rs2277820
BHMT	rs16876512	*DHFR*	rs1650697	*FTH1*	rs1800009
BHMT	rs7700970	*DHFR*	rs380691	*FTH1*	rs17185413

TABLE A4.2 (Continued)

Gene	SNP Name	Gene	SNP Name	Gene	SNP Name
BHMT	rs506500	*DHFR*	rs1382540	*FTH1*	rs1801621
BHMT	rs567754	*DMGDH*	rs28326	*FTH1*	rs17156609
BHMT	rs10037045	*DMGDH*	rs1805074	*FTH1*	rs2073588
GART	rs8971	*MAT1A*	rs1819684	*MTHFD1L*	rs538017
GART	rs1804385	*MAT1A*	rs17677908	*MTHFD1L*	rs9478162
GART	rs8788[2]	*MAT2A*	rs1446667	*MTHFD1L*	rs9478908
GART	rs2027592	*MAT2A*	rs2028900	*MTHFD1L*	rs7770982
GART	rs4817580	*MAT2A*	rs1078004	*MTHFD1L*	rs17080689
GCSH	rs8177940	*MAT2A*	rs2043675	*MTHFD1L*	rs1076746
GCSH	rs8177876	*MAT2B*	rs10515861	*MTHFD1L*	rs509474
GCSH	rs4889233	*MAT2B*	rs6882306	*MTHFD1L*	rs7746991
GCSH	rs11866124	*MAT2B*	rs4869087	*MTHFD1L*	rs6910267
GCSH	rs1563072	*MAT2B*	rs17061795	*MTHFD1L*	rs17354394
GGH	rs11995525	*MAT2B*	rs7733775	*MTHFD1L*	rs1047665
GGH	rs11545078	*MTHFD1*	rs1076991	*MTHFD2*	rs7340453
GGH	rs4617146	*MTHFD1*	rs8003379	*MTHFD2*	rs1667627
GGH	rs3780126	*MTHFD1*	rs1950902	*MTHFR*	rs1537516
GGH	rs12544045	*MTHFD1*	rs17751556	*MTHFR*	rs4846049
GLDC	rs7848919	*MTHFD1*	rs3783728	*MTHFR*	rs13306556
GLDC	rs3902970	*MTHFD1*	rs11627525	*MTHFR*	rs1801131
GLDC	rs1929933	*MTHFD1*	rs8003567	*MTHFR*	rs12121543
GLDC	rs4237166	*MTHFD1*	rs8012229	*MTHFR*	rs1994798
GLDC	rs10975681	*MTHFD1*	rs2281603	*MTHFR*	rs6541003
GLDC	rs4629927	*MTHFD1L*	rs7765521	*MTHFR*	rs1801133
GLDC	rs7049056	*MTHFD1L*	rs11754661	*MTHFR*	rs17421462
GLDC	rs1821892	*MTHFD1L*	rs4869953	*MTHFR*	rs17367629
GLDC	rs2118653	*MTHFD1L*	rs17349743	*MTHFR*	rs3737965
GLDC	rs1755617	*MTHFD1L*	rs803422	*MTHFS*	rs8923
GNMT	rs11752813	*MTHFD1L*	rs997429	*MTHFS*	rs7177659
GNMT	rs2296805	*MTHFD1L*	rs2295084	*MTHFS*	rs6495450
GNMT	rs2296804	*MTHFD1L*	rs803456	*MTHFS*	rs2733106
GNMT	rs736158	*MTHFD1L*	rs803455	*MTHFS*	rs2586167
GNMT	rs1051218	*MTHFD1L*	rs803454	*MTR*	rs16834388
HSPA8	rs4936770	*MTHFD1L*	rs12201472	*MTR*	rs3754255
HSPA8	rs1461496	*MTHFD1L*	rs4869954	*MTR*	rs10925260
HSPA8	rs11218941	*MTHFD1L*	rs6902664	*MTR*	rs1805087
HSPA8	rs1136141	*MTHFD1L*	rs1474787	*MTR*	rs2275566
MARS	rs899653	*MTHFD1L*	rs4869955	*MTR*	rs2229276[3]
MARS	rs496245	*MTHFD1L*	rs742832	*MTR*	rs1131449[4]

TABLE A4.2 (Continued)

Gene	SNP Name	Gene	SNP Name	Gene	SNP Name
MARS	rs1678537	*MTHFD1L*	rs803447	*MTR*	rs1050996
MAT1A	rs1985908	*MTHFD1L*	rs803446	*MTRR*	rs2966952
MAT1A	rs2993763	*MTHFD1L*	rs803466	*MTRR*	rs1801394
MAT1A	rs10788546	*MTHFD1L*	rs1738574	*MTRR*	rs7730643
MAT1A	rs1143694	*MTHFD1L*	rs524732	*MTRR*	rs161869
MTRR	rs1532268	*SLC19A1*	rs1051266	*TCN2*	rs5749132
MTRR	rs3776465	*SLC19A1*	rs1131596[5]	*TCN2*	rs9606756
MTRR	rs162036	*SLC19A1*	rs4819130	*TCN2*	rs740233
MTRR	rs2303081	*SLC19A2*	rs6656822	*TCN2*	rs757874
MTRR	rs10380	*SLC19A2*	rs1983546	*TCN2*	rs4820021
MTRR	rs1802059	*SLC19A2*	rs17518769	*TCN2*	rs9621049
MTRR	rs8659	*SLC19A2*	rs2038024	*TCN2*	rs4820886
SARDH	rs129886	*SLC19A3*	rs13025803	*TCN2*	rs4820887
SARDH	rs129932	*SLC19A3*	rs11694828	*TCN2*	rs2301957
SARDH	rs129891	*SLC19A3*	rs13007334	*TCN2*	rs2301958
SARDH	rs756682	*SLC19A3*	rs17438244	*TCN2*	rs4820889
SARDH	rs2073817	*SLC25A32*	rs1061196	*TCN2*	rs10418
SARDH	rs2502741	*SLC25A32*	rs3098260	*TYMS*	rs2853533
SARDH	rs4979632	*SLC25A32*	rs17803441	*TYMS*	rs502396
SARDH	rs2073815	*SLC25A32*	rs3098243	*TYMS*	rs2612095
SHMT1	rs12952556	*SLC25A32*	rs3134297	*TYMS*	rs2853543
SHMT1	rs1979276	*SLC46A1*	rs2239908	*TYMS*	rs16948305
SHMT1	rs1979277	*SLC46A1*	rs2239907	*TYMS*	rs699517
SHMT1	rs2273028	*SLC46A1*	rs17719944	*TYMS*	rs2790
SHMT1	rs17806489	*TCN1*	rs17154234	*UBE2I*	rs9926094
SHMT1	rs4924750	*TCN1*	rs34324219	*UBE2I*	rs909915
SHMT1	rs2461838	*TCN1*	rs519221	*UBE2I*	rs8052688
SHMT1	rs643333	*TCN1*	rs557564	*UBE2I*	rs11248868
SHMT2	rs28365862	*TCN1*	rs2000613	*UBE2N*	rs4020454
SHMT2	rs7301155	*TCN1*	rs34528912	*UBE2N*	rs7311222
SLC19A1	rs1051298	*TCN1*	rs526934	*UBE2N*	rs7309933
SLC19A1	rs12482346	*TCN2*	rs5749131	*UBE2N*	rs7300607
SLC19A1	rs2297291	*TCN2*	rs5753231	*UBE2N*	rs1483003

[1]Formerly known as rs8111085; [2]Formerly known as rs9984077; [3]Formerly known as rs16834521; [4]Formerly known as rs10737812; [5]Formerly known as rs3177999.

TABLE A4.3 MEDIATION OF MAIN EFFECT ASSOCIATIONS WITH CVD RISK BY PLASMA BIOMARKERS[1-3]

		Model 1		Model 2			Model 3			Model 4		
Gene	**rs#**	**Estimate**	**P**	**Estimate**	**P**	**% change**	**Estimate**	**P**	**% change**	**Estimate**	**P**	**% change**
GGH	rs12544045	-0.38	0.0006	-0.36	0.013	-6%	-0.37	0.0103	-3%	-0.38	0.0093	-2%
MTHFS	rs7177659	-0.34	0.0012	-0.55	<.0001	59%	-0.54	0.0001	57%	-0.55	<.0001	62%
MTHFR	rs1801133 *CT*	0.06	0.6	0.14	0.35	134%	0.15	0.33	144%	0.16	0.29	163%
	rs1801133 *TT*	0.54	0.0003	0.62	0.0013	15%	0.66	0.0007	22%	0.65	0.0008	21%
TCN1	rs17154234	0.49	0.0024	0.57	0.005	17%	0.57	0.0048	18%	0.58	0.0045	19%
UBE2I	rs11248868	0.34	0.0033	0.32	0.029	-4%	0.34	0.023	1%	0.34	0.024	0%
CBS	rs6586282	-0.32	0.0033	-0.2	0.14	-36%	-0.2	0.14	-36%	-0.2	0.14	-37%
CBS	rs6586281	-0.35	0.0035	-0.22	0.14	-36%	-0.22	0.15	-36%	-0.22	0.15	-37%
UBE2I	rs909915	0.33	0.004	0.3	0.047	-10%	0.31	0.037	-5%	0.31	0.039	-6%

[1]Model 1 adjusted for age and smoking only; Model 2 adjusted for age, smoking, and folate; Model 3 adjusted for age, smoking, folate, vitamin B-6, and vitamin B-12; Model 4 adjusted for age, smoking, folate, vitamin B-6, vitamin B-12, and homocysteine.
[2]Nominal P values shown.
[3]% change reflects change in regression estimate from Model 1.

TABLE A4.4 GENOTYPES WITH THE MOST STATISTICALLY SIGNIFICANT ASSOCIATION (P≤0.005) WITH THE EARLY-ONSET CVD PHENOTYPE[1,3,5]

Gene	rs#	Estimate	Std Error	OR	Nominal P	Chr.	Coded allele	Coded allele frequency	Genetic Model	SNP Type
GGH	rs12544045	-1.21	0.33	0.30	2.89E-04[2]	8	A	29%	Overdominant	5' region
UBE2I	rs11248868	0.79	0.27	2.20	3.49E-03	16	G	13%	Overdominant	Intronic
FTH1	rs2073588	2.12	0.73	8.31	3.63E-03	11	G	5%	Recessive	5' region
UBE2I	rs909915	0.76	0.27	2.14	4.49E-03	16	T	13%	Overdominant	Intronic
FTH1	rs1801621[3,4]	2.88	1.02	17.87	4.71E-03	11	C	3%	Recessive	3' region

[1]Model adjusted for age and smoking; forward strand allele shown.
[2]Adjusted P values reached False Discovery Rate significance threshold of 0.1.
[3]Sparse data (fewer than 5 individuals per category) for genotype categories of this SNP.
[4]SNP maps to more than one gene: *FTH1* rs1801621 also maps to bestrophin 1 (*BEST1*).
[5]No lower quality SNPs.

TABLE A4.5 NUTRIENT INTERACTIONS WITH THE MOST STATISTICALLY SIGNIFICANT ASSOCIATION (P≤0.02) WITH CARDIOVASCULAR DISEASE RISK[1,2,4,5]

Folate interactions

Gene	rs#	Estimate	Std Error	Nominal P	Chr	Coded allele	Coded allele frequency (%)	Genetic Model	Type
MAT2B	rs6882306	-3.16	0.91	5.04E-04[2]	5	C	16%	Recessive	Intronic
SLC19A2	rs17518769[3]	-7.23	2.70	7.40E-03	1	A	9%	Recessive	Intronic
Vitamin B-6 interactions									
BHMT	rs585800	-0.56	0.20	4.69E-03	5	T	27%	Overdominant	3' region
DNMT3A	rs11695471	-0.56	0.20	4.94E-03	2	A	32%	Dominant	Intronic
BHMT	rs506500	-0.50	0.20	1.09E-02	5	T	31%	Overdominant	Intronic
SLC19A2	rs17518769[3]	-5.13	2.07	1.33E-02	1	A	9%	Recessive	Intronic
BHMT	rs558133	-0.48	0.20	1.48E-02	5	G	32%	Overdominant	Intronic
SLC19A2	rs6656822	0.50	0.21	1.62E-02	1	T	26%	Overdominant	Intronic
Vitamin B-12 interactions									
BHMT	rs585800	-1.28	0.37	5.39E-04[2]	5	T	27%	Overdominant	3' region
SLC25A32	rs1061196	1.30	0.39	9.01E-04[2]	8	A	23%	Overdominant	3' region
BHMT	rs558133	-1.09	0.36	2.42E-03	5	G	32%	Overdominant	Intronic
BHMT	rs506500	-1.07	0.36	2.76E-03	5	T	31%	Overdominant	Intronic
AMT	rs1464566[4]	1.11	0.40	5.08E-03	3	G	42%	Dominant	Intronic
MTHFD1L	rs997429	1.03	0.39	8.23E-03	6	A	18%	Dominant	Intronic
MTR	rs16834388	-0.91	0.36	1.09E-02	1	T	37%	Dominant	5' region
FOLR1	rs2071010	1.12	0.45	1.35E-02	11	A	5%	Overdominant	Intronic
GART	rs8788	0.89	0.38	1.75E-02	21	C	26%	Overdominant	Coding nonsynonymous
AMT	rs1464567[4]	0.62	0.26	1.85E-02	3	C	42%	Additive	Intronic

[1]Model adjusted for age and smoking; forward strand allele shown.
[2]Adjusted P values reached False Discovery Rate significance threshold of 0.2.
[3]Sparse data (fewer than 5 individuals per category) for some genotype categories of this SNP.
[4]SNP maps to more than one gene: rs1464566 also maps to nicolin 1(*NICN1*).
[5]No lower quality SNPs.

TABLE A4.6

ALL *MTHFR 677 C→T (*rs1801133) PAIRWISE INTERACTIONS WITH LIKELIHOOD RATIO TEST P≤0.02[1,2]

Gene 1	SNP 1	G'type 1[3]	SNP Type 1	Gene 2	SNP 2	G'type 2[3]	SNP Type 2	Estimate	LRT	LRT df	LRT P	FDR P
MTHFR	rs1801133	T/T	Nonsynonymous					0.65	12.30	2	0.0021	0.35
MTHFR	rs1801133	C/T	Nonsynonymous					0.13				
				FPGS	rs41319447	C/T	3'	0.61				
MTHFR	rs1801133	T/T	Nonsynonymous	*FPGS*	rs41319447	C/T	3'	-13.08				
MTHFR	rs1801133	C/T	Nonsynonymous	*FPGS*	rs41319447	C/T	3'	-1.01				
MTHFR	rs1801133	T/T	Nonsynonymous					0.42	16.71	4	0.0022	0.35
MTHFR	rs1801133	C/T	Nonsynonymous					0.04				
				TYMS	rs2853533	C/C	Intronic	0.40				
				TYMS	rs2853533	G/C	Intronic	-0.10				
MTHFR	rs1801133	T/T	Nonsynonymous	*TYMS*	rs2853533	C/C	Intronic	-12.81				
MTHFR	rs1801133	C/T	Nonsynonymous	*TYMS*	rs2853533	C/C	Intronic	0.37				
MTHFR	rs1801133	T/T	Nonsynonymous	*TYMS*	rs2853533	G/C	Intronic	0.78				
MTHFR	rs1801133	C/T	Nonsynonymous	*TYMS*	rs2853533	G/C	Intronic	0.10				
MTHFR	rs1801133	T/T	Nonsynonymous					0.05	15.06	4	0.0046	0.49
MTHFR	rs1801133	C/T	Nonsynonymous					-0.20				
				GLDC	rs2118653	A/A	Intronic	-0.44				
				GLDC	rs2118653	G/A	Intronic	-0.21				
MTHFR	rs1801133	T/T	Nonsynonymous	*GLDC*	rs2118653	A/A	Intronic	0.57				
MTHFR	rs1801133	C/T	Nonsynonymous	*GLDC*	rs2118653	A/A	Intronic	0.90				
MTHFR	rs1801133	T/T	Nonsynonymous	*GLDC*	rs2118653	G/A	Intronic	0.78				
MTHFR	rs1801133	C/T	Nonsynonymous	*GLDC*	rs2118653	G/A	Intronic	0.13				
MTHFR	rs1801133	T/T	Nonsynonymous					0.98	13.92	4	0.0076	0.60
MTHFR	rs1801133	C/T	Nonsynonymous					0.10				

TABLE A4.6 (Continued)

Gene 1	SNP 1	G'type 1[3]	SNP Type 1	Gene 2	SNP 2	G'type 2[3]	SNP Type 2	Estimate	LRT	LRT df	LRT P	FDR P
				SLC46A1	rs2239908	G/G	3'	0.33				
				SLC46A1	rs2239908	A/G	3'	0.11				
MTHFR	rs1801133	T/T	Nonsynonymous	*SLC46A1*	rs2239908	G/G	3'	-1.74				
MTHFR	rs1801133	C/T	Nonsynonymous	*SLC46A1*	rs2239908	G/G	3'	-0.38				
MTHFR	rs1801133	T/T	Nonsynonymous	*SLC46A1*	rs2239908	A/G	3'	-0.38				
MTHFR	rs1801133	C/T	Nonsynonymous	*SLC46A1*	rs2239908	A/G	3'	0.10				
MTHFR	rs1801133	T/T	Nonsynonymous					0.91	12.87	4	0.012	0.66
MTHFR	rs1801133	C/T	Nonsynonymous					0.06				
				SLC46A1	rs2239907	T/T	3'	0.26				
				SLC46A1	rs2239907	C/T	3'	0.06				
MTHFR	rs1801133	T/T	Nonsynonymous	*SLC46A1*	rs2239907	T/T	3'	-1.65				
MTHFR	rs1801133	C/T	Nonsynonymous	*SLC46A1*	rs2239907	T/T	3'	-0.29				
MTHFR	rs1801133	T/T	Nonsynonymous	*SLC46A1*	rs2239907	C/T	3'	-0.26				
MTHFR	rs1801133	C/T	Nonsynonymous	*SLC46A1*	rs2239907	C/T	3'	0.13				
MTHFR	rs1801133	T/T	Nonsynonymous					0.52	12.67	4	0.013	0.66
MTHFR	rs1801133	C/T	Nonsynonymous					0.19				
				SARDH	rs756682	G/G	Intronic	0.32				
				SARDH	rs756682	A/G	Intronic	-0.02				
MTHFR	rs1801133	T/T	Nonsynonymous	*SARDH*	rs756682	G/G	Intronic	-0.18				
MTHFR	rs1801133	C/T	Nonsynonymous	*SARDH*	rs756682	G/G	Intronic	-1.35				
MTHFR	rs1801133	T/T	Nonsynonymous	*SARDH*	rs756682	A/G	Intronic	0.13				
MTHFR	rs1801133	C/T	Nonsynonymous	*SARDH*	rs756682	A/G	Intronic	0.03				
MTHFR	rs1801133	T/T	Nonsynonymous					0.36	12.42	4	0.015	0.66
MTHFR	rs1801133	C/T	Nonsynonymous					0.22				

TABLE A4.6 (Continued)

Gene 1	SNP 1	G'type 1[3]	SNP Type 1	Gene 2	SNP 2	G'type 2[3]	SNP Type 2	Estimate	LRT	LRT df	LRT P	FDR P
				FOLR3	rs11235449	A/A	3'	-0.11				
				FOLR3	rs11235449	G/A	3'	-0.24				
MTHFR	rs1801133	T/T	Nonsynonymous	*FOLR3*	rs11235449	A/A	3'	0.73				
MTHFR	rs1801133	C/T	Nonsynonymous	*FOLR3*	rs11235449	A/A	3'	-0.68				
MTHFR	rs1801133	T/T	Nonsynonymous	*FOLR3*	rs11235449	G/A	3'	0.13				
MTHFR	rs1801133	C/T	Nonsynonymous	*FOLR3*	rs11235449	G/A	3'	-0.07				
MTHFR	rs1801133	T/T	Nonsynonymous					0.97	11.73	4	0.019	0.68
MTHFR	rs1801133	C/T	Nonsynonymous					0.54				
				MTHFD1L	rs7765521	A/A	Intronic	0.10				
				MTHFD1L	rs7765521	G/A	Intronic	0.41				
MTHFR	rs1801133	T/T	Nonsynonymous	*MTHFD1L*	rs7765521	A/A	Intronic	-0.10				
MTHFR	rs1801133	C/T	Nonsynonymous	*MTHFD1L*	rs7765521	A/A	Intronic	-0.63				
MTHFR	rs1801133	T/T	Nonsynonymous	*MTHFD1L*	rs7765521	G/A	Intronic	-0.87				
MTHFR	rs1801133	C/T	Nonsynonymous	*MTHFD1L*	rs7765521	G/A	Intronic	-0.68				

[1]Interactions among SNPs within the same gene were not considered.
[2]No FDR-adjusted Likelihood Ratio Test P values reached significance threshold of P≤0.02.
[3]Genotype of SNPs 1 and 2, respectively.

TABLE A4.7

ALL PAIRWISE SNP INTERACTIONS WITH LIKELIHOOD RATIO $P \leq 0.001$[1,2]

Gene 1	SNP 1	G'type 1[3]	SNP Type 1	Gene 2	SNP 2	G'type 2[3]	SNP Type 2	Est.[4]	LRT	LRT df	LRT P	FDR P
CELF1	rs4752843	C/C	Intronic					-0.34	25.42	4	0.000041	0.68
CELF1	rs4752843	A/C	Intronic					-0.33				
				SLC19A3	rs11694828	A/A	Intronic	-0.37				
				SLC19A3	rs11694828	G/A	Intronic	-0.08				
CELF1	rs4752843	C/C	Intronic	*SLC19A3*	rs11694828	A/A	Intronic	-11.38				
CELF1	rs4752843	A/C	Intronic	*SLC19A3*	rs11694828	A/A	Intronic	0.95				
CELF1	rs4752843	C/C	Intronic	*SLC19A3*	rs11694828	G/A	Intronic	1.97				
CELF1	rs4752843	A/C	Intronic	*SLC19A3*	rs11694828	G/A	Intronic	-0.02				
CELF1	rs4752843	C/C	Intronic					-0.31	25.01	4	0.000050	0.68
CELF1	rs4752843	A/C	Intronic					-0.31				
				SLC19A3	rs13025803	T/T	Intronic	-0.37				
				SLC19A3	rs13025803	C/T	Intronic	-0.03				
CELF1	rs4752843	C/C	Intronic	*SLC19A3*	rs13025803	T/T	Intronic	-11.39				
CELF1	rs4752843	A/C	Intronic	*SLC19A3*	rs13025803	T/T	Intronic	0.94				
CELF1	rs4752843	C/C	Intronic	*SLC19A3*	rs13025803	C/T	Intronic	1.92				
CELF1	rs4752843	A/C	Intronic	*SLC19A3*	rs13025803	C/T	Intronic	-0.06				
AHCYL1	rs186724	A/A	Intronic					-13.53	24.96	4	0.000051	0.68
AHCYL1	rs186724	G/A	Intronic					-0.27				
				FPGS	rs10106	G/G	3'	-0.08				
				FPGS	rs10106	A/G	3'	-0.38				
AHCYL1	rs186724	A/A	Intronic	*FPGS*	rs10106	G/G	3'	12.52				
AHCYL1	rs186724	G/A	Intronic	*FPGS*	rs10106	G/G	3'	-0.23				
AHCYL1	rs186724	A/A	Intronic	*FPGS*	rs10106	A/G	3'	13.96				

TABLE A4.7 (Continued)

Gene 1	SNP 1	G'type 1[3]	SNP Type 1	Gene 2	SNP 2	G'type 2[3]	SNP Type 2	Est.[4]	LRT	LRT df	LRT P	FDR P
AHCYL1	rs186724	G/A	Intronic	*FPGS*	rs10106	A/G	3'	0.33				
FOLR1	rs9282688	T/T	5'					-10.79	19.75	2	0.000051	0.68
FOLR1	rs9282688	C/T	5'					0.02				
				ALDH1L1	rs3772424	A/A	Intronic	-0.27				
				ALDH1L1	rs3772424	G/A	Intronic	-0.11				
FOLR1	rs9282688	T/T	5'	*ALDH1L1*	rs3772424	A/A	Intronic	0.00				
FOLR1	rs9282688	C/T	5'	*ALDH1L1*	rs3772424	A/A	Intronic	4.37				
FOLR1	rs9282688	T/T	5'	*ALDH1L1*	rs3772424	G/A	Intronic	0.00				
FOLR1	rs9282688	C/T	5'	*ALDH1L1*	rs3772424	G/A	Intronic	-0.51				
CELF1	rs7933019	C/C	Intronic					1.02	23.07	4	0.00012	0.77
CELF1	rs7933019	G/C	Intronic					0.25				
				MTR	rs16834388	T/T	5'	0.55				
				MTR	rs16834388	G/T	5'	0.27				
CELF1	rs7933019	C/C	Intronic	*MTR*	rs16834388	T/T	5'	-2.60				
CELF1	rs7933019	G/C	Intronic	*MTR*	rs16834388	T/T	5'	-0.55				
CELF1	rs7933019	C/C	Intronic	*MTR*	rs16834388	G/T	5'	-1.20				
CELF1	rs7933019	G/C	Intronic	*MTR*	rs16834388	G/T	5'	-0.23				
MTHFD1L	rs803456	C/C	Intronic					0.37	22.24	4	0.00018	0.77
MTHFD1L	rs803456	T/C	Intronic					-0.34				
				DHFR	rs380691	C/C	5'	-0.85				
				DHFR	rs380691	T/C	5'	-0.28				
MTHFD1L	rs803456	C/C	Intronic	*DHFR*	rs380691	C/C	5'	-0.43				
MTHFD1L	rs803456	T/C	Intronic	*DHFR*	rs380691	C/C	5'	1.21				
MTHFD1L	rs803456	C/C	Intronic	*DHFR*	rs380691	T/C	5'	-0.30				

TABLE A4.7 (Continued)

Gene 1	SNP 1	G'type 1[3]	SNP Type 1	Gene 2	SNP 2	G'type 2[3]	SNP Type 2	Est.[4]	LRT	LRT df	LRT P	FDR P
MTHFD1L	rs803456	T/C	Intronic	*DHFR*	rs380691	T/C	5'	0.68				
CELF1	rs7933019	C/C	Intronic					0.99	22.07	4	0.00019	0.77
CELF1	rs7933019	G/C	Intronic					0.25				
				MTR	rs2229276	G/G	Synonymous	0.50				
				MTR	rs2229276	A/G	Synonymous	0.24				
CELF1	rs7933019	C/C	Intronic	*MTR*	rs2229276	G/G	Synonymous	-2.54				
CELF1	rs7933019	G/C	Intronic	*MTR*	rs2229276	G/G	Synonymous	-0.58				
CELF1	rs7933019	C/C	Intronic	*MTR*	rs2229276	A/G	Synonymous	-1.18				
CELF1	rs7933019	G/C	Intronic	*MTR*	rs2229276	A/G	Synonymous	-0.24				
CELF1	rs7102372	T/T	Intronic					-0.45	19.05	3	0.00027	0.77
CELF1	rs7102372	C/T	Intronic					-0.24				
				FTCD	rs16978930	G/G	Intronic	-0.94				
				FTCD	rs16978930	A/G	Intronic	-0.13				
CELF1	rs7102372	T/T	Intronic	*FTCD*	rs16978930	G/G	Intronic	0.00				
CELF1	rs7102372	C/T	Intronic	*FTCD*	rs16978930	G/G	Intronic	3.62				
CELF1	rs7102372	T/T	Intronic	*FTCD*	rs16978930	A/G	Intronic	0.36				
CELF1	rs7102372	C/T	Intronic	*FTCD*	rs16978930	A/G	Intronic	0.94				
AHCY	rs819148	G/G	Intronic					0.34	20.89	4	0.00033	0.77
AHCY	rs819148	A/G	Intronic					-0.51				
				FTH1	rs1800009	G/G	3'	-0.57				
				FTH1	rs1800009	A/G	3'	-0.11				
AHCY	rs819148	G/G	Intronic	*FTH1*	rs1800009	G/G	3'	-11.02				
AHCY	rs819148	A/G	Intronic	*FTH1*	rs1800009	G/G	3'	1.63				
AHCY	rs819148	G/G	Intronic	*FTH1*	rs1800009	A/G	3'	-0.05				

TABLE A4.7 (Continued)

Gene 1	SNP 1	G'type 1[3]	SNP Type 1	Gene 2	SNP 2	G'type 2[3]	SNP Type 2	Est.[4]	LRT	LRT df	LRT P	FDR P
AHCY	rs819148	A/G	Intronic	*FTH1*	rs1800009	A/G	3'	0.68				
MTHFD1L	rs803447	T/T	Intronic					0.38	20.81	4	0.00034	0.77
MTHFD1L	rs803447	C/T	Intronic					0.33				
				TYMS	rs16948305	T/T	Intronic	0.47				
				TYMS	rs16948305	C/T	Intronic	0.44				
MTHFD1L	rs803447	T/T	Intronic	*TYMS*	rs16948305	T/T	Intronic	1.11				
MTHFD1L	rs803447	C/T	Intronic	*TYMS*	rs16948305	T/T	Intronic	-13.66				
MTHFD1L	rs803447	T/T	Intronic	*TYMS*	rs16948305	C/T	Intronic	-0.56				
MTHFD1L	rs803447	C/T	Intronic	*TYMS*	rs16948305	C/T	Intronic	-0.62				
DNMT3A	rs1550117	A/A	5'					-12.47	20.74	4	0.00036	0.77
DNMT3A	rs1550117	G/A	5'					-0.47				
				MTHFR	rs12121543	A/A	Intronic	-0.56				
				MTHFR	rs12121543	C/A	Intronic	-0.12				
DNMT3A	rs1550117	A/A	5'	*MTHFR*	rs12121543	A/A	Intronic	15.26				
DNMT3A	rs1550117	G/A	5'	*MTHFR*	rs12121543	A/A	Intronic	0.96				
DNMT3A	rs1550117	A/A	5'	*MTHFR*	rs12121543	C/A	Intronic	12.64				
DNMT3A	rs1550117	G/A	5'	*MTHFR*	rs12121543	C/A	Intronic	1.15				
MTHFD1	rs17751556	C/C	Intronic					-11.22	17.93	3	0.00045	0.77
MTHFD1	rs17751556	T/C	Intronic					0.03				
				GLDC	rs10975681	C/C	Intronic	0.19				
				GLDC	rs10975681	T/C	Intronic	-0.05				
MTHFD1	rs17751556	C/C	Intronic	*GLDC*	rs10975681	C/C	Intronic	0.00				
MTHFD1	rs17751556	T/C	Intronic	*GLDC*	rs10975681	C/C	Intronic	-1.82				
MTHFD1	rs17751556	C/C	Intronic	*GLDC*	rs10975681	T/C	Intronic	12.60				

TABLE A4.7 (Continued)

Gene 1	SNP 1	G'type 1[3]	SNP Type 1	Gene 2	SNP 2	G'type 2[3]	SNP Type 2	Est.[4]	LRT	LRT df	LRT P	FDR P
MTHFD1	rs17751556	T/C	Intronic	*GLDC*	rs10975681	T/C	Intronic	0.49				
AHCY	rs819146	G/G	5'					0.37	20.13	4	0.00047	0.77
AHCY	rs819146	T/G	5'					-0.47				
				FTH1	rs1800009	G/G	3'	-0.58				
				FTH1	rs1800009	A/G	3'	-0.11				
AHCY	rs819146	G/G	5'	*FTH1*	rs1800009	G/G	3'	-11.04				
AHCY	rs819146	T/G	5'	*FTH1*	rs1800009	G/G	3'	1.58				
AHCY	rs819146	G/G	5'	*FTH1*	rs1800009	A/G	3'	-0.09				
AHCY	rs819146	T/G	5'	*FTH1*	rs1800009	A/G	3'	0.64				
CELF1	rs7933019	C/C	Intronic					1.00	20.07	4	0.00048	0.77
CELF1	rs7933019	G/C	Intronic					0.27				
				MTR	rs3754255	T/T	Intronic	0.39				
				MTR	rs3754255	C/T	Intronic	0.16				
CELF1	rs7933019	C/C	Intronic	*MTR*	rs3754255	T/T	Intronic	-2.29				
CELF1	rs7933019	G/C	Intronic	*MTR*	rs3754255	T/T	Intronic	-0.64				
CELF1	rs7933019	C/C	Intronic	*MTR*	rs3754255	C/T	Intronic	-1.14				
CELF1	rs7933019	G/C	Intronic	*MTR*	rs3754255	C/T	Intronic	-0.23				
AHCY	rs819133	T/T	Intronic					0.47	20.06	4	0.00049	0.77
AHCY	rs819133	G/T	Intronic					-0.44				
				FTH1	rs1800009	G/G	3'	-0.57				
				FTH1	rs1800009	A/G	3'	-0.11				
AHCY	rs819133	T/T	Intronic	*FTH1*	rs1800009	G/G	3'	-11.15				
AHCY	rs819133	G/T	Intronic	*FTH1*	rs1800009	G/G	3'	1.56				
AHCY	rs819133	T/T	Intronic	*FTH1*	rs1800009	A/G	3'	-0.19				

TABLE A4.7 (Continued)

Gene 1	SNP 1	G'type 1[3]	SNP Type 1	Gene 2	SNP 2	G'type 2[3]	SNP Type 2	Est.[4]	LRT	LRT df	LRT P	FDR P
AHCY	rs819133	G/T	Intronic	*FTH1*	rs1800009	A/G	3'	0.65				
CELF1	rs2242081	C/C	Intronic					0.68	20.05	4	0.00049	0.77
CELF1	rs2242081	T/C	Intronic					-0.06				
				MTR	rs16834388	T/T	5'	0.35				
				MTR	rs16834388	G/T	5'	0.39				
CELF1	rs2242081	C/C	Intronic	*MTR*	rs16834388	T/T	5'	-1.63				
CELF1	rs2242081	T/C	Intronic	*MTR*	rs16834388	T/T	5'	-0.01				
CELF1	rs2242081	C/C	Intronic	*MTR*	rs16834388	G/T	5'	-1.01				
CELF1	rs2242081	T/C	Intronic	*MTR*	rs16834388	G/T	5'	-0.23				
AHCY	rs819159	A/A	Intronic					0.47	20.03	4	0.00049	0.77
AHCY	rs819159	T/A	Intronic					-0.45				
				FTH1	rs1800009	G/G	3'	-0.57				
				FTH1	rs1800009	A/G	3'	-0.11				
AHCY	rs819159	A/A	Intronic	*FTH1*	rs1800009	G/G	3'	-11.14				
AHCY	rs819159	T/A	Intronic	*FTH1*	rs1800009	G/G	3'	1.57				
AHCY	rs819159	A/A	Intronic	*FTH1*	rs1800009	A/G	3'	-0.18				
AHCY	rs819159	T/A	Intronic	*FTH1*	rs1800009	A/G	3'	0.63				
GLDC	rs4237166	C/C	Intronic					0.32	20.02	4	0.00049	0.77
GLDC	rs4237166	G/C	Intronic					0.22				
				SLC25A32	rs1061196	A/A	3'	0.72				
				SLC25A32	rs1061196	G/A	3'	0.57				
GLDC	rs4237166	C/C	Intronic	*SLC25A32*	rs1061196	A/A	3'	-2.43				
GLDC	rs4237166	G/C	Intronic	*SLC25A32*	rs1061196	A/A	3'	-0.49				
GLDC	rs4237166	C/C	Intronic	*SLC25A32*	rs1061196	G/A	3'	-1.08				

TABLE A4.7 (Continued)

Gene 1	SNP 1	G'type 1[3]	SNP Type 1	Gene 2	SNP 2	G'type 2[3]	SNP Type 2	Est.[4]	LRT	LRT df	LRT P	FDR P
GLDC	rs4237166	G/C	Intronic	*SLC25A32*	rs1061196	G/A	3'	-0.73				
MTHFD1L	rs524732	T/T	Intronic					0.11	19.90	4	0.00052	0.77
MTHFD1L	rs524732	C/T	Intronic					-0.31				
				AMT	rs11922013	C/C	Intronic	-0.90				
				AMT	rs11922013	G/C	Intronic	-0.01				
MTHFD1L	rs524732	T/T	Intronic	*AMT*	rs11922013	C/C	Intronic	-9.78				
MTHFD1L	rs524732	C/T	Intronic	*AMT*	rs11922013	C/C	Intronic	1.82				
MTHFD1L	rs524732	T/T	Intronic	*AMT*	rs11922013	G/C	Intronic	-0.26				
MTHFD1L	rs524732	C/T	Intronic	*AMT*	rs11922013	G/C	Intronic	0.22				
FTCD	rs2277820	T/T	Intronic					-0.40	19.90	4	0.00052	0.77
FTCD	rs2277820	C/T	Intronic					-0.14				
				GGH	rs11995525	A/A	Intronic	-0.49				
				GGH	rs11995525	G/A	Intronic	-0.36				
FTCD	rs2277820	T/T	Intronic	*GGH*	rs11995525	A/A	Intronic	-11.42				
FTCD	rs2277820	C/T	Intronic	*GGH*	rs11995525	A/A	Intronic	1.46				
FTCD	rs2277820	T/T	Intronic	*GGH*	rs11995525	G/A	Intronic	0.66				
FTCD	rs2277820	C/T	Intronic	*GGH*	rs11995525	G/A	Intronic	0.38				
GNMT	rs11752813	G/G	5'					0.43	19.89	4	0.00053	0.77
GNMT	rs11752813	C/G	5'					0.26				
				GLDC	rs10975681	C/C	Intronic	0.79				
				GLDC	rs10975681	T/C	Intronic	0.37				
GNMT	rs11752813	G/G	5'	*GLDC*	rs10975681	C/C	Intronic	-0.79				
GNMT	rs11752813	C/G	5'	*GLDC*	rs10975681	C/C	Intronic	-1.17				
GNMT	rs11752813	G/G	5'	*GLDC*	rs10975681	T/C	Intronic	-1.19				

TABLE A4.7 (Continued)

Gene 1	SNP 1	G'type 1[3]	SNP Type 1	Gene 2	SNP 2	G'type 2[3]	SNP Type 2	Est.[4]	LRT	LRT df	LRT P	FDR P
GNMT	rs11752813	C/G	5'	*GLDC*	rs10975681	T/C	Intronic	-0.19				
TCN2	rs4820886	G/G	Intronic					-12.17	19.84	4	0.00054	0.77
TCN2	rs4820886	T/G	Intronic					0.58				
				DNMT3A	rs11678631	A/A	Intronic	-0.03				
				DNMT3A	rs11678631	T/A	Intronic	0.13				
TCN2	rs4820886	G/G	Intronic	*DNMT3A*	rs11678631	A/A	Intronic	14.04				
TCN2	rs4820886	T/G	Intronic	*DNMT3A*	rs11678631	A/A	Intronic	-0.61				
TCN2	rs4820886	G/G	Intronic	*DNMT3A*	rs11678631	T/A	Intronic	12.27				
TCN2	rs4820886	T/G	Intronic	*DNMT3A*	rs11678631	T/A	Intronic	-0.64				
CBS	rs6586281	A/A	Intronic					-1.88	19.79	4	0.00055	0.77
CBS	rs6586281	G/A	Intronic					-0.58				
				SLC19A3	rs17438244	C/C	5'	-1.74				
				SLC19A3	rs17438244	A/C	5'	-0.03				
CBS	rs6586281	A/A	Intronic	*SLC19A3*	rs17438244	C/C	5'	-7.32				
CBS	rs6586281	G/A	Intronic	*SLC19A3*	rs17438244	C/C	5'	2.76				
CBS	rs6586281	A/A	Intronic	*SLC19A3*	rs17438244	A/C	5'	2.05				
CBS	rs6586281	G/A	Intronic	*SLC19A3*	rs17438244	A/C	5'	0.48				
FOLH1	rs16906158	C/C	Intronic					0.64	19.64	4	0.00059	0.77
FOLH1	rs16906158	T/C	Intronic					0.29				
				GGH	rs11995525	A/A	Intronic	0.41				
				GGH	rs11995525	G/A	Intronic	-0.04				
FOLH1	rs16906158	C/C	Intronic	*GGH*	rs11995525	A/A	Intronic	-12.20				
FOLH1	rs16906158	T/C	Intronic	*GGH*	rs11995525	A/A	Intronic	-1.72				
FOLH1	rs16906158	C/C	Intronic	*GGH*	rs11995525	G/A	Intronic	0.80				

TABLE A4.7 (Continued)

Gene 1	SNP 1	G'type 1[3]	SNP Type 1	Gene 2	SNP 2	G'type 2[3]	SNP Type 2	Est.[4]	LRT	LRT df	LRT P	FDR P
FOLH1	rs16906158	T/C	Intronic	*GGH*	rs11995525	G/A	Intronic	-0.90				
MTHFS	rs8923	G/G	Nonsynonymous					3.24	17.30	3	0.00061	0.77
MTHFS	rs8923	A/G	Nonsynonymous					-0.68				
				GNMT	rs2296805	T/T	Intronic	-0.21				
				GNMT	rs2296805	G/T	Intronic	0.09				
MTHFS	rs8923	G/G	Nonsynonymous	*GNMT*	rs2296805	T/T	Intronic	0.00				
MTHFS	rs8923	A/G	Nonsynonymous	*GNMT*	rs2296805	T/T	Intronic	1.46				
MTHFS	rs8923	G/G	Nonsynonymous	*GNMT*	rs2296805	G/T	Intronic	-4.66				
MTHFS	rs8923	A/G	Nonsynonymous	*GNMT*	rs2296805	G/T	Intronic	0.54				
SLC25A32	rs3098243	C/C	Intronic					0.15	19.40	4	0.00066	0.77
SLC25A32	rs3098243	T/C	Intronic					-0.10				
				MTHFD1L	rs17349743	C/C	Intronic	0.84				
				MTHFD1L	rs17349743	T/C	Intronic	-0.27				
SLC25A32	rs3098243	C/C	Intronic	*MTHFD1L*	rs17349743	C/C	Intronic	-1.08				
SLC25A32	rs3098243	T/C	Intronic	*MTHFD1L*	rs17349743	C/C	Intronic	-0.64				
SLC25A32	rs3098243	C/C	Intronic	*MTHFD1L*	rs17349743	T/C	Intronic	-0.03				
SLC25A32	rs3098243	T/C	Intronic	*MTHFD1L*	rs17349743	T/C	Intronic	0.68				
MTHFD1L	rs803466	G/G	Intronic					0.06	19.37	4	0.00066	0.77
MTHFD1L	rs803466	A/G	Intronic					-0.11				
				CELF1	rs7933019	C/C	Intronic	0.34				
				CELF1	rs7933019	G/C	Intronic	-0.24				
MTHFD1L	rs803466	G/G	Intronic	*CELF1*	rs7933019	C/C	Intronic	-0.25				
MTHFD1L	rs803466	A/G	Intronic	*CELF1*	rs7933019	C/C	Intronic	-0.92				
MTHFD1L	rs803466	G/G	Intronic	*CELF1*	rs7933019	G/C	Intronic	-0.04				

TABLE A4.7 (Continued)

Gene 1	SNP 1	G'type 1[3]	SNP Type 1	Gene 2	SNP 2	G'type 2[3]	SNP Type 2	Est.[4]	LRT	LRT df	LRT P	FDR P
MTHFD1L	rs803466	A/G	Intronic	*CELF1*	rs7933019	G/C	Intronic	0.67				
MTHFR	rs1801131	C/C	Nonsynonymous					-0.42	19.32	4	0.00068	0.77
MTHFR	rs1801131	A/C	Nonsynonymous					-0.23				
				DNMT3A	rs1550117	A/A	5'	-11.61				
				DNMT3A	rs1550117	G/A	5'	-0.61				
MTHFR	rs1801131	C/C	Nonsynonymous	*DNMT3A*	rs1550117	A/A	5'	13.32				
MTHFR	rs1801131	A/C	Nonsynonymous	*DNMT3A*	rs1550117	A/A	5'	11.60				
MTHFR	rs1801131	C/C	Nonsynonymous	*DNMT3A*	rs1550117	G/A	5'	0.93				
MTHFR	rs1801131	A/C	Nonsynonymous	*DNMT3A*	rs1550117	G/A	5'	1.20				
MTHFS	rs8923	G/G	Nonsynonymous					3.22	17.06	3	0.00069	0.77
MTHFS	rs8923	A/G	Nonsynonymous					-0.68				
				GNMT	rs2296804	G/G	Intronic	-0.19				
				GNMT	rs2296804	C/G	Intronic	0.08				
MTHFS	rs8923	G/G	Nonsynonymous	*GNMT*	rs2296804	G/G	Intronic	0.00				
MTHFS	rs8923	A/G	Nonsynonymous	*GNMT*	rs2296804	G/G	Intronic	1.44				
MTHFS	rs8923	G/G	Nonsynonymous	*GNMT*	rs2296804	C/G	Intronic	-4.63				
MTHFS	rs8923	A/G	Nonsynonymous	*GNMT*	rs2296804	C/G	Intronic	0.54				
FTH1	rs1800009	G/G	3'					-0.59	19.28	4	0.00069	0.77
FTH1	rs1800009	A/G	3'					-0.11				
				AHCY	rs1205366	T/T	Intronic	0.38				
				AHCY	rs1205366	C/T	Intronic	-0.46				
FTH1	rs1800009	G/G	3'	*AHCY*	rs1205366	T/T	Intronic	-9.72				
FTH1	rs1800009	A/G	3'	*AHCY*	rs1205366	T/T	Intronic	-0.38				
FTH1	rs1800009	G/G	3'	*AHCY*	rs1205366	C/T	Intronic	1.60				

TABLE A4.7 (Continued)

Gene 1	SNP 1	G'type 1[3]	SNP Type 1	Gene 2	SNP 2	G'type 2[3]	SNP Type 2	Est.[4]	LRT	LRT df	LRT P	FDR P
FTH1	rs1800009	A/G	3'	*AHCY*	rs1205366	C/T	Intronic	0.65				
BHMT	rs16876512	T/T	5'					-0.42	19.24	4	0.00070	0.77
BHMT	rs16876512	C/T	5'					-0.50				
				MAT1A	rs1143694	T/T	Synonymous	0.23				
				MAT1A	rs1143694	C/T	Synonymous	-0.19				
BHMT	rs16876512	T/T	5'	*MAT1A*	rs1143694	T/T	Synonymous	-10.78				
BHMT	rs16876512	C/T	5'	*MAT1A*	rs1143694	T/T	Synonymous	-0.91				
BHMT	rs16876512	T/T	5'	*MAT1A*	rs1143694	C/T	Synonymous	1.14				
BHMT	rs16876512	C/T	5'	*MAT1A*	rs1143694	C/T	Synonymous	0.88				
CBS	rs2014564	A/A	Intronic					0.02	19.24	4	0.00071	0.77
CBS	rs2014564	G/A	Intronic					-0.01				
				HSPA8	rs1136141	T/T	5'	1.55				
				HSPA8	rs1136141	C/T	5'	-0.26				
CBS	rs2014564	A/A	Intronic	*HSPA8*	rs1136141	T/T	5'	-0.76				
CBS	rs2014564	G/A	Intronic	*HSPA8*	rs1136141	T/T	5'	-2.31				
CBS	rs2014564	A/A	Intronic	*HSPA8*	rs1136141	C/T	5'	0.94				
CBS	rs2014564	G/A	Intronic	*HSPA8*	rs1136141	C/T	5'	0.42				
GNMT	rs11752813	G/G	5'					0.37	19.20	4	0.00072	0.77
GNMT	rs11752813	C/G	5'					0.48				
				BHMT	rs10037045	T/T	Intronic	0.99				
				BHMT	rs10037045	A/T	Intronic	0.44				
GNMT	rs11752813	G/G	5'	*BHMT*	rs10037045	T/T	Intronic	-1.01				
GNMT	rs11752813	C/G	5'	*BHMT*	rs10037045	T/T	Intronic	-1.90				
GNMT	rs11752813	G/G	5'	*BHMT*	rs10037045	A/T	Intronic	-0.90				

TABLE A4.7 (Continued)

Gene 1	SNP 1	G'type 1[3]	SNP Type 1	Gene 2	SNP 2	G'type 2[3]	SNP Type 2	Est.[4]	LRT	LRT df	LRT P	FDR P
GNMT	rs11752813	C/G	5'	*BHMT*	rs10037045	A/T	Intronic	-0.60				
SARDH	rs129891	A/A	Intronic					0.19	19.10	4	0.00075	0.77
SARDH	rs129891	G/A	Intronic					0.31				
				GNMT	rs11752813	G/G	5'	-0.13				
				GNMT	rs11752813	C/G	5'	0.47				
SARDH	rs129891	A/A	Intronic	*GNMT*	rs11752813	G/G	5'	-0.31				
SARDH	rs129891	G/A	Intronic	*GNMT*	rs11752813	G/G	5'	0.16				
SARDH	rs129891	A/A	Intronic	*GNMT*	rs11752813	C/G	5'	-0.22				
SARDH	rs129891	G/A	Intronic	*GNMT*	rs11752813	C/G	5'	-0.87				
TCN1	rs34528912	C/T	Nonsynonymous					-0.85	14.33	2	0.00077	0.77
				FOLR2	rs2298444	G/G	Intronic	0.32				
				FOLR2	rs2298444	A/G	Intronic	0.00				
TCN1	rs34528912	C/T	Nonsynonymous	*FOLR2*	rs2298444	G/G	Intronic	-10.62				
TCN1	rs34528912	C/T	Nonsynonymous	*FOLR2*	rs2298444	A/G	Intronic	1.43				
CBS	rs6586281	A/A	Intronic					-13.25	19.02	4	0.00078	0.77
CBS	rs6586281	G/A	Intronic					-0.24				
				MTRR	rs1801394	A/A	Nonsynonymous	-0.10				
				MTRR	rs1801394	G/A	Nonsynonymous	0.12				
CBS	rs6586281	A/A	Intronic	*MTRR*	rs1801394	A/A	Nonsynonymous	15.79				
CBS	rs6586281	G/A	Intronic	*MTRR*	rs1801394	A/A	Nonsynonymous	-0.13				
CBS	rs6586281	A/A	Intronic	*MTRR*	rs1801394	G/A	Nonsynonymous	12.59				
CBS	rs6586281	G/A	Intronic	*MTRR*	rs1801394	G/A	Nonsynonymous	-0.10				
MTHFR	rs3737965	C/T	5'					0.19	11.29	1	0.00078	0.77
				DNMT1	rs2228612	C/C	Nonsynonymous	-13.20				

TABLE A4.7 (Continued)

Gene 1	SNP 1	G'type 1[3]	SNP Type 1	Gene 2	SNP 2	G'type 2[3]	SNP Type 2	Est.[4]	LRT	LRT df	LRT P	FDR P
				DNMT1	rs2228612	T/C	Nonsynonymous	0.24				
MTHFR	rs3737965	C/T	5'	*DNMT1*	rs2228612	C/C	Nonsynonymous	0.00				
MTHFR	rs3737965	C/T	5'	*DNMT1*	rs2228612	T/C	Nonsynonymous	-13.36				
GLDC	rs2118653	A/A	Intronic					0.40	19.01	4	0.00078	0.77
GLDC	rs2118653	G/A	Intronic					0.02				
				ALDH1L1	rs1868138	T/T	Intronic	0.16				
				ALDH1L1	rs1868138	A/T	Intronic	0.42				
GLDC	rs2118653	A/A	Intronic	*ALDH1L1*	rs1868138	T/T	Intronic	-1.75				
GLDC	rs2118653	G/A	Intronic	*ALDH1L1*	rs1868138	T/T	Intronic	0.95				
GLDC	rs2118653	A/A	Intronic	*ALDH1L1*	rs1868138	A/T	Intronic	-0.88				
GLDC	rs2118653	G/A	Intronic	*ALDH1L1*	rs1868138	A/T	Intronic	-0.35				
MTHFD1	rs17751556	C/C	Intronic					0.93	16.75	3	0.00080	0.77
MTHFD1	rs17751556	T/C	Intronic					-0.13				
				MAT1A	rs17677908	G/G	5'	0.19				
				MAT1A	rs17677908	A/G	5'	-0.18				
MTHFD1	rs17751556	C/C	Intronic	*MAT1A*	rs17677908	G/G	5'	0.00				
MTHFD1	rs17751556	T/C	Intronic	*MAT1A*	rs17677908	G/G	5'	-11.25				
MTHFD1	rs17751556	C/C	Intronic	*MAT1A*	rs17677908	A/G	5'	-12.29				
MTHFD1	rs17751556	T/C	Intronic	*MAT1A*	rs17677908	A/G	5'	1.07				
GNMT	rs2296804	G/G	Intronic					0.10	18.92	4	0.00081	0.77
GNMT	rs2296804	C/G	Intronic					0.04				
				FTH1	rs2073588	G/G	3'	4.39				
				FTH1	rs2073588	T/G	3'	-0.52				
GNMT	rs2296804	G/G	Intronic	*FTH1*	rs2073588	G/G	3'	-14.28				

TABLE A4.7 (Continued)

Gene 1	SNP 1	G'type 1[3]	SNP Type 1	Gene 2	SNP 2	G'type 2[3]	SNP Type 2	Est.[4]	LRT	LRT df	LRT P	FDR P
GNMT	rs2296804	C/G	Intronic	*FTH1*	rs2073588	G/G	3'	-3.63				
GNMT	rs2296804	G/G	Intronic	*FTH1*	rs2073588	T/G	3'	-0.79				
GNMT	rs2296804	C/G	Intronic	*FTH1*	rs2073588	T/G	3'	1.04				
DNMT3A	rs6733868	C/C	Intronic					-0.18	18.91	4	0.00082	0.77
DNMT3A	rs6733868	G/C	Intronic					0.07				
				MTHFD1	rs2281603	G/G	Intronic	1.51				
				MTHFD1	rs2281603	A/G	Intronic	-0.13				
DNMT3A	rs6733868	C/C	Intronic	*MTHFD1*	rs2281603	G/G	Intronic	-0.49				
DNMT3A	rs6733868	G/C	Intronic	*MTHFD1*	rs2281603	G/G	Intronic	-1.87				
DNMT3A	rs6733868	C/C	Intronic	*MTHFD1*	rs2281603	A/G	Intronic	0.23				
DNMT3A	rs6733868	G/C	Intronic	*MTHFD1*	rs2281603	A/G	Intronic	0.16				
CBS	rs6586282	T/T	Intronic					-1.86	18.84	4	0.00084	0.77
CBS	rs6586282	C/T	Intronic					-0.56				
				SLC19A3	rs17438244	C/C	5'	-1.74				
				SLC19A3	rs17438244	A/C	5'	-0.03				
CBS	rs6586282	T/T	Intronic	*SLC19A3*	rs17438244	C/C	5'	-7.34				
CBS	rs6586282	C/T	Intronic	*SLC19A3*	rs17438244	C/C	5'	2.75				
CBS	rs6586282	T/T	Intronic	*SLC19A3*	rs17438244	A/C	5'	1.84				
CBS	rs6586282	C/T	Intronic	*SLC19A3*	rs17438244	A/C	5'	0.51				
MTHFR	rs6541003	G/G	Intronic					-0.77	18.83	4	0.00085	0.77
MTHFR	rs6541003	A/G	Intronic					-0.31				
				SLC46A1	rs2239908	G/G	3'	-1.03				
				SLC46A1	rs2239908	A/G	3'	-0.07				
MTHFR	rs6541003	G/G	Intronic	*SLC46A1*	rs2239908	G/G	3'	1.92				

TABLE A4.7 (Continued)

Gene 1	SNP 1	G'type 1[3]	SNP Type 1	Gene 2	SNP 2	G'type 2[3]	SNP Type 2	Est.[4]	LRT	LRT df	LRT P	FDR P
MTHFR	rs6541003	A/G	Intronic	*SLC46A1*	rs2239908	G/G	3'	1.13				
MTHFR	rs6541003	G/G	Intronic	*SLC46A1*	rs2239908	A/G	3'	0.36				
MTHFR	rs6541003	A/G	Intronic	*SLC46A1*	rs2239908	A/G	3'	0.20				
CELF1	rs7933019	C/C	Intronic					-0.16	16.61	3	0.00085	0.77
CELF1	rs7933019	G/C	Intronic					0.05				
				ALDH1L1	rs4646750	G/G	Nonsynonymous	-12.26				
				ALDH1L1	rs4646750	A/G	Nonsynonymous	-0.22				
CELF1	rs7933019	C/C	Intronic	*ALDH1L1*	rs4646750	G/G	Nonsynonymous	0.00				
CELF1	rs7933019	G/C	Intronic	*ALDH1L1*	rs4646750	G/G	Nonsynonymous	13.87				
CELF1	rs7933019	C/C	Intronic	*ALDH1L1*	rs4646750	A/G	Nonsynonymous	1.50				
CELF1	rs7933019	G/C	Intronic	*ALDH1L1*	rs4646750	A/G	Nonsynonymous	0.02				
FOLR2	rs514933	G/G	Intronic					-0.45	18.79	4	0.00087	0.77
FOLR2	rs514933	A/G	Intronic					0.10				
				DNMT3A	rs1550117	A/A	5'	-11.46				
				DNMT3A	rs1550117	G/A	5'	0.48				
FOLR2	rs514933	G/G	Intronic	*DNMT3A*	rs1550117	A/A	5'	0.24				
FOLR2	rs514933	A/G	Intronic	*DNMT3A*	rs1550117	A/A	5'	11.68				
FOLR2	rs514933	G/G	Intronic	*DNMT3A*	rs1550117	G/A	5'	-0.08				
FOLR2	rs514933	A/G	Intronic	*DNMT3A*	rs1550117	G/A	5'	-1.33				
SLC25A32	rs3098243	C/C	Intronic					-0.32	18.67	4	0.00091	0.77
SLC25A32	rs3098243	T/C	Intronic					-0.45				
				CBS	rs1788484	T/T	5'	-0.29				
				CBS	rs1788484	C/T	5'	-0.66				
SLC25A32	rs3098243	C/C	Intronic	*CBS*	rs1788484	T/T	5'	0.61				

TABLE A4.7 (Continued)

Gene 1	SNP 1	G'type 1[3]	SNP Type 1	Gene 2	SNP 2	G'type 2[3]	SNP Type 2	Est.[4]	LRT	LRT df	LRT P	FDR P
SLC25A32	rs3098243	T/C	Intronic	*CBS*	rs1788484	T/T	5'	0.68				
SLC25A32	rs3098243	C/C	Intronic	*CBS*	rs1788484	C/T	5'	0.50				
SLC25A32	rs3098243	T/C	Intronic	*CBS*	rs1788484	C/T	5'	1.07				
FOLH1	rs202673	G/G	Intronic					0.39	18.65	4	0.00092	0.77
FOLH1	rs202673	A/G	Intronic					-0.36				
				MAT1A	rs1143694	T/T	Synonymous	0.19				
				MAT1A	rs1143694	C/T	Synonymous	-0.19				
FOLH1	rs202673	G/G	Intronic	*MAT1A*	rs1143694	T/T	Synonymous	0.93				
FOLH1	rs202673	A/G	Intronic	*MAT1A*	rs1143694	T/T	Synonymous	-0.93				
FOLH1	rs202673	G/G	Intronic	*MAT1A*	rs1143694	C/T	Synonymous	-1.20				
FOLH1	rs202673	A/G	Intronic	*MAT1A*	rs1143694	C/T	Synonymous	0.74				
CTH	rs663649	T/T	Intronic					0.35	18.61	4	0.00094	0.77
CTH	rs663649	G/T	Intronic					-0.12				
				SLC25A32	rs1061196	A/A	3'	0.07				
				SLC25A32	rs1061196	G/A	3'	0.03				
CTH	rs663649	T/T	Intronic	*SLC25A32*	rs1061196	A/A	3'	0.88				
CTH	rs663649	G/T	Intronic	*SLC25A32*	rs1061196	A/A	3'	-0.41				
CTH	rs663649	T/T	Intronic	*SLC25A32*	rs1061196	G/A	3'	-1.35				
CTH	rs663649	G/T	Intronic	*SLC25A32*	rs1061196	G/A	3'	0.28				
MTHFD1L	rs803456	C/C	Intronic					0.55	18.60	4	0.00094	0.77
MTHFD1L	rs803456	T/C	Intronic					0.37				
				MTHFD1	rs1950902	T/T	Nonsynonymous	1.77				
				MTHFD1	rs1950902	C/T	Nonsynonymous	0.32				
MTHFD1L	rs803456	C/C	Intronic	*MTHFD1*	rs1950902	T/T	Nonsynonymous	-3.33				

TABLE A4.7 (Continued)

Gene 1	SNP 1	G'type 1[3]	SNP Type 1	Gene 2	SNP 2	G'type 2[3]	SNP Type 2	Est.[4]	LRT	LRT df	LRT P	FDR P
MTHFD1L	rs803456	T/C	Intronic	*MTHFD1*	rs1950902	T/T	Nonsynonymous	-1.97				
MTHFD1L	rs803456	C/C	Intronic	*MTHFD1*	rs1950902	C/T	Nonsynonymous	-0.71				
MTHFD1L	rs803456	T/C	Intronic	*MTHFD1*	rs1950902	C/T	Nonsynonymous	-0.75				
MAT2B	rs7733775	A/A	Intronic					-0.18	18.60	4	0.00094	0.77
MAT2B	rs7733775	G/A	Intronic					-0.30				
				SHMT1	rs12952556	C/C	3'	-1.17				
				SHMT1	rs12952556	T/C	3'	-0.13				
MAT2B	rs7733775	A/A	Intronic	*SHMT1*	rs12952556	C/C	3'	0.47				
MAT2B	rs7733775	G/A	Intronic	*SHMT1*	rs12952556	C/C	3'	1.79				
MAT2B	rs7733775	A/A	Intronic	*SHMT1*	rs12952556	T/C	3'	0.37				
MAT2B	rs7733775	G/A	Intronic	*SHMT1*	rs12952556	T/C	3'	0.39				
AHCY	rs819155	G/G	Intronic					0.47	18.57	4	0.00095	0.77
AHCY	rs819155	A/G	Intronic					-0.42				
				FTH1	rs1800009	G/G	3'	-0.58				
				FTH1	rs1800009	A/G	3'	-0.11				
AHCY	rs819155	G/G	Intronic	*FTH1*	rs1800009	G/G	3'	-9.81				
AHCY	rs819155	A/G	Intronic	*FTH1*	rs1800009	G/G	3'	1.56				
AHCY	rs819155	G/G	Intronic	*FTH1*	rs1800009	A/G	3'	-0.46				
AHCY	rs819155	A/G	Intronic	*FTH1*	rs1800009	A/G	3'	0.63				
SLC19A3	rs13007334	C/C	Intronic					0.28	18.55	4	0.00096	0.77
SLC19A3	rs13007334	T/C	Intronic					-0.19				
				BHMT	rs10037045	T/T	Intronic	-1.10				
				BHMT	rs10037045	A/T	Intronic	0.06				
SLC19A3	rs13007334	C/C	Intronic	*BHMT*	rs10037045	T/T	Intronic	0.12				

TABLE A4.7 (Continued)

Gene 1	SNP 1	G'type 1[3]	SNP Type 1	Gene 2	SNP 2	G'type 2[3]	SNP Type 2	Est.[4]	LRT	LRT df	LRT P	FDR P
SLC19A3	rs13007334	T/C	Intronic	*BHMT*	rs10037045	T/T	Intronic	1.71				
SLC19A3	rs13007334	C/C	Intronic	*BHMT*	rs10037045	A/T	Intronic	-0.63				
SLC19A3	rs13007334	T/C	Intronic	*BHMT*	rs10037045	A/T	Intronic	0.05				
MTHFD1L	rs9478908	G/G	Intronic					-0.16	16.32	3	0.00098	0.77
MTHFD1L	rs9478908	A/G	Intronic					0.08				
				MTHFD1	rs17751556	C/C	Intronic	1.88				
				MTHFD1	rs17751556	T/C	Intronic	0.35				
MTHFD1L	rs9478908	G/G	Intronic	*MTHFD1*	rs17751556	C/C	Intronic	0.00				
MTHFD1L	rs9478908	A/G	Intronic	*MTHFD1*	rs17751556	C/C	Intronic	-13.73				
MTHFD1L	rs9478908	G/G	Intronic	*MTHFD1*	rs17751556	T/C	Intronic	-1.04				
MTHFD1L	rs9478908	A/G	Intronic	*MTHFD1*	rs17751556	T/C	Intronic	-0.43				
MTHFD1L	rs9478162	A/A	Intronic					-0.08	16.27	3	0.0010	0.77
MTHFD1L	rs9478162	G/A	Intronic					0.06				
				MTHFD1	rs17751556	C/C	Intronic	1.86				
				MTHFD1	rs17751556	T/C	Intronic	0.33				
MTHFD1L	rs9478162	A/A	Intronic	*MTHFD1*	rs17751556	C/C	Intronic	0.00				
MTHFD1L	rs9478162	G/A	Intronic	*MTHFD1*	rs17751556	C/C	Intronic	-13.70				
MTHFD1L	rs9478162	A/A	Intronic	*MTHFD1*	rs17751556	T/C	Intronic	-1.11				
MTHFD1L	rs9478162	G/A	Intronic	*MTHFD1*	rs17751556	T/C	Intronic	-0.46				
SLC19A2	rs17518769	A/A	Intronic					2.10	16.26	3	0.0010	0.77
SLC19A2	rs17518769	G/A	Intronic					0.25				
				DHFR	rs1650697	T/T	5'	0.30				
				DHFR	rs1650697	C/T	5'	0.28				
SLC19A2	rs17518769	A/A	Intronic	*DHFR*	rs1650697	T/T	5'	-13.62				

TABLE A4.7 (Continued)

Gene 1	SNP 1	G'type 1[3]	SNP Type 1	Gene 2	SNP 2	G'type 2[3]	SNP Type 2	Est.[4]	LRT	LRT df	LRT P	FDR P
SLC19A2	rs17518769	G/A	Intronic	*DHFR*	rs1650697	T/T	5'	0.23				
SLC19A2	rs17518769	A/A	Intronic	*DHFR*	rs1650697	C/T	5'	0.00				
SLC19A2	rs17518769	G/A	Intronic	*DHFR*	rs1650697	C/T	5'	-0.80				
MTRR	rs161869	T/T	Intronic					-0.51	18.45	4	0.0010	0.77
MTRR	rs161869	C/T	Intronic					-0.43				
				SARDH	rs129932	A/A	Synonymous	-1.05				
				SARDH	rs129932	G/A	Synonymous	-0.26				
MTRR	rs161869	T/T	Intronic	*SARDH*	rs129932	A/A	Synonymous	1.52				
MTRR	rs161869	C/T	Intronic	*SARDH*	rs129932	A/A	Synonymous	1.39				
MTRR	rs161869	T/T	Intronic	*SARDH*	rs129932	G/A	Synonymous	0.15				
MTRR	rs161869	C/T	Intronic	*SARDH*	rs129932	G/A	Synonymous	0.43				
TCN2	rs9606756	G/G	Nonsynonymous					0.30	18.41	4	0.0010	0.77
TCN2	rs9606756	A/G	Nonsynonymous					-0.07				
				MTHFD1L	rs17349743	C/C	Intronic	0.21				
				MTHFD1L	rs17349743	T/C	Intronic	0.03				
TCN2	rs9606756	G/G	Nonsynonymous	*MTHFD1L*	rs17349743	C/C	Intronic	-11.64				
TCN2	rs9606756	A/G	Nonsynonymous	*MTHFD1L*	rs17349743	C/C	Intronic	1.12				
TCN2	rs9606756	G/G	Nonsynonymous	*MTHFD1L*	rs17349743	T/C	Intronic	1.24				
TCN2	rs9606756	A/G	Nonsynonymous	*MTHFD1L*	rs17349743	T/C	Intronic	-0.01				
MAT2B	rs7733775	A/A	Intronic					-0.19	18.40	4	0.0010	0.77
MAT2B	rs7733775	G/A	Intronic					-0.29				
				SHMT1	rs1979277	A/A	Nonsynonymous	-1.17				
				SHMT1	rs1979277	G/A	Nonsynonymous	-0.13				
MAT2B	rs7733775	A/A	Intronic	*SHMT1*	rs1979277	A/A	Nonsynonymous	0.48				

TABLE A4.7 (Continued)

Gene 1	SNP 1	G'type 1[3]	SNP Type 1	Gene 2	SNP 2	G'type 2[3]	SNP Type 2	Est.[4]	LRT	LRT df	LRT P	FDR P
MAT2B	rs7733775	G/A	Intronic	*SHMT1*	rs1979277	A/A	Nonsynonymous	1.78				
MAT2B	rs7733775	A/A	Intronic	*SHMT1*	rs1979277	G/A	Nonsynonymous	0.39				
MAT2B	rs7733775	G/A	Intronic	*SHMT1*	rs1979277	G/A	Nonsynonymous	0.36				
MTHFS	rs2586167	T/T	5'					0.15	18.38	4	0.0010	0.77
MTHFS	rs2586167	C/T	5'					0.05				
				SLC19A3	rs17438244	C/C	5'	0.27				
				SLC19A3	rs17438244	A/C	5'	0.15				
MTHFS	rs2586167	T/T	5'	*SLC19A3*	rs17438244	C/C	5'	-13.56				
MTHFS	rs2586167	C/T	5'	*SLC19A3*	rs17438244	C/C	5'	-13.31				
MTHFS	rs2586167	T/T	5'	*SLC19A3*	rs17438244	A/C	5'	-0.54				
MTHFS	rs2586167	C/T	5'	*SLC19A3*	rs17438244	A/C	5'	0.03				

[1]Interactions among SNPs within the same gene were not considered.
[2]No FDR-adjusted Likelihood Ratio Test P values reached significance threshold of P≤0.2.
[3]Genotype of SNPs 1 and 2, respectively.
[4]Regression estimate.

TABLE A4.8

ALL *SHMT1, TS, DHFR, UBE2I,* AND *UBE2N* PAIRWISE INTERACTIONS WITH LIKELIHOOD RATIO TEST P≤0.02[1]

Gene 1	SNP 1	G'type 1[2]	SNP Type 1	Gene 2	SNP 2	G'type 2[2]	SNP Type 2	Estimate	LRT	LRT df	LRT P
UBE2N	rs7300607	C/C	Intronic					0.25	15.74	4	0.0034
UBE2N	rs7300607	T/C	Intronic					-0.26			
				TYMS	rs502396	C/C	Intronic	-1.28			
				TYMS	rs502396	T/C	Intronic	-0.38			
UBE2N	rs7300607	C/C	Intronic	*TYMS*	rs502396	C/C	Intronic	0.89			
UBE2N	rs7300607	T/C	Intronic	*TYMS*	rs502396	C/C	Intronic	1.54			
UBE2N	rs7300607	C/C	Intronic	*TYMS*	rs502396	T/C	Intronic	-0.07			
UBE2N	rs7300607	T/C	Intronic	*TYMS*	rs502396	T/C	Intronic	0.44			
TYMS	rs502396	C/C	Intronic					-1.28	15.46	4	0.0038
TYMS	rs502396	T/C	Intronic					-0.39			
				UBE2N	rs1483003	G/G	Intronic	0.22			
				UBE2N	rs1483003	A/G	Intronic	-0.24			
TYMS	rs502396	C/C	Intronic	*UBE2N*	rs1483003	G/G	Intronic	0.94			
TYMS	rs502396	T/C	Intronic	*UBE2N*	rs1483003	G/G	Intronic	-0.05			
TYMS	rs502396	C/C	Intronic	*UBE2N*	rs1483003	A/G	Intronic	1.54			
TYMS	rs502396	T/C	Intronic	*UBE2N*	rs1483003	A/G	Intronic	0.42			
UBE2N	rs7300607	C/C	Intronic					0.24	14.35	4	0.0063
UBE2N	rs7300607	T/C	Intronic					-0.19			
				TYMS	rs2612095	C/C	Intronic	-1.19			
				TYMS	rs2612095	T/C	Intronic	-0.18			
UBE2N	rs7300607	C/C	Intronic	*TYMS*	rs2612095	C/C	Intronic	1.13			
UBE2N	rs7300607	T/C	Intronic	*TYMS*	rs2612095	C/C	Intronic	1.49			
UBE2N	rs7300607	C/C	Intronic	*TYMS*	rs2612095	T/C	Intronic	-0.12			
UBE2N	rs7300607	T/C	Intronic	*TYMS*	rs2612095	T/C	Intronic	0.36			

TABLE A4.8 (Continued)

Gene 1	SNP 1	G'type 1[2]	SNP Type 1	Gene 2	SNP 2	G'type 2[2]	SNP Type 2	Estimate	LRT	LRT df	LRT P
UBE2N	rs7300607	C/C	Intronic					0.29	14.17	4	0.007
UBE2N	rs7300607	T/C	Intronic					-0.17			
				TYMS	rs2853543	G/G	Intronic	-1.16			
				TYMS	rs2853543	C/G	Intronic	-0.17			
UBE2N	rs7300607	C/C	Intronic	*TYMS*	rs2853543	G/G	Intronic	1.10			
UBE2N	rs7300607	T/C	Intronic	*TYMS*	rs2853543	G/G	Intronic	1.42			
UBE2N	rs7300607	C/C	Intronic	*TYMS*	rs2853543	C/G	Intronic	-0.20			
UBE2N	rs7300607	T/C	Intronic	*TYMS*	rs2853543	C/G	Intronic	0.36			
TYMS	rs2853543	G/G	Intronic					-1.16	13.89	4	0.008
TYMS	rs2853543	C/G	Intronic					-0.17			
				UBE2N	rs1483003	G/G	Intronic	0.26			
				UBE2N	rs1483003	A/G	Intronic	-0.15			
TYMS	rs2853543	G/G	Intronic	*UBE2N*	rs1483003	G/G	Intronic	1.15			
TYMS	rs2853543	C/G	Intronic	*UBE2N*	rs1483003	G/G	Intronic	-0.17			
TYMS	rs2853543	G/G	Intronic	*UBE2N*	rs1483003	A/G	Intronic	1.41			
TYMS	rs2853543	C/G	Intronic	*UBE2N*	rs1483003	A/G	Intronic	0.34			
TYMS	rs2612095	C/C	Intronic					-1.19	13.73	4	0.008
TYMS	rs2612095	T/C	Intronic					-0.18			
				UBE2N	rs1483003	G/G	Intronic	0.19			
				UBE2N	rs1483003	A/G	Intronic	-0.17			
TYMS	rs2612095	C/C	Intronic	*UBE2N*	rs1483003	G/G	Intronic	1.20			
TYMS	rs2612095	T/C	Intronic	*UBE2N*	rs1483003	G/G	Intronic	-0.07			
TYMS	rs2612095	C/C	Intronic	*UBE2N*	rs1483003	A/G	Intronic	1.47			
TYMS	rs2612095	T/C	Intronic	*UBE2N*	rs1483003	A/G	Intronic	0.33			

TABLE A4.8 (Continued)

Gene 1	SNP 1	G'type 1[2]	SNP Type 1	Gene 2	SNP 2	G'type 2[2]	SNP Type 2	Estimate	LRT	LRT df	LRT P
DHFR	rs380691	C/C	5'					-0.78	12.75	4	0.013
DHFR	rs380691	T/C	5'					-0.37			
				SHMT1	rs2273028	T/T	Intronic	-0.04			
				SHMT1	rs2273028	C/T	Intronic	-0.27			
DHFR	rs380691	C/C	5'	*SHMT1*	rs2273028	T/T	Intronic	0.11			
DHFR	rs380691	T/C	5'	*SHMT1*	rs2273028	T/T	Intronic	0.08			
DHFR	rs380691	C/C	5'	*SHMT1*	rs2273028	C/T	Intronic	1.05			
DHFR	rs380691	T/C	5'	*SHMT1*	rs2273028	C/T	Intronic	0.69			
DHFR	rs380691	C/C	5'					-0.79	11.76	4	0.019
DHFR	rs380691	T/C	5'					-0.34			
				SHMT1	rs1979276	A/A	3'	-0.11			
				SHMT1	rs1979276	G/A	3'	-0.29			
DHFR	rs380691	C/C	5'	*SHMT1*	rs1979276	A/A	3'	0.11			
DHFR	rs380691	T/C	5'	*SHMT1*	rs1979276	A/A	3'	0.12			
DHFR	rs380691	C/C	5'	*SHMT1*	rs1979276	G/A	3'	1.07			
DHFR	rs380691	T/C	5'	*SHMT1*	rs1979276	G/A	3'	0.63			

[1]Interactions among SNPs within the same gene were not considered.
[2]Genotype of SNPs 1 and 2, respectively.

TABLE A4.9

ALL *SHMT1*, *FTH1*, AND *CELF1* PAIRWISE INTERACTIONS WITH LIKELIHOOD RATIO TEST $P \leq 0.02$[1]

Gene 1	SNP 1	G'type 1[2]	SNP Type 1	Gene 2	SNP 2	G'type 2[2]	SNP Type 2	Estimate	LRT	LRT df	LRT P
SHMT1	rs1979276	A/A	3'					0.47	15.33	4	0.0041
SHMT1	rs1979276	G/A	3'					0.29			
				FTH1	rs1800009	G/G	3'	0.15			
				FTH1	rs1800009	A/G	3'	0.28			
SHMT1	rs1979276	A/A	3'	*FTH1*	rs1800009	G/G	3'	-13.49			
SHMT1	rs1979276	G/A	3'	*FTH1*	rs1800009	G/G	3'	-0.22			
SHMT1	rs1979276	A/A	3'	*FTH1*	rs1800009	A/G	3'	-0.92			
SHMT1	rs1979276	G/A	3'	*FTH1*	rs1800009	A/G	3'	-0.37			
SHMT1	rs2273028	T/T	Intronic					0.51	15.08	4	0.0045
SHMT1	rs2273028	C/T	Intronic					0.33			
				FTH1	rs1800009	G/G	3'	0.18			
				FTH1	rs1800009	A/G	3'	0.28			
SHMT1	rs2273028	T/T	Intronic	*FTH1*	rs1800009	G/G	3'	-13.54			
SHMT1	rs2273028	C/T	Intronic	*FTH1*	rs1800009	G/G	3'	-0.27			
SHMT1	rs2273028	T/T	Intronic	*FTH1*	rs1800009	A/G	3'	-0.84			
SHMT1	rs2273028	C/T	Intronic	*FTH1*	rs1800009	A/G	3'	-0.38			
SHMT1	rs2273028	T/T	Intronic					0.32	14.77	4	0.0052
SHMT1	rs2273028	C/T	Intronic					0.32			
				FTH1	rs17185413	C/C	3'	0.51			
				FTH1	rs17185413	T/C	3'	0.25			
SHMT1	rs2273028	T/T	Intronic	*FTH1*	rs17185413	C/C	3'	-12.82			
SHMT1	rs2273028	C/T	Intronic	*FTH1*	rs17185413	C/C	3'	-1.88			
SHMT1	rs2273028	T/T	Intronic	*FTH1*	rs17185413	T/C	3'	-0.68			
SHMT1	rs2273028	C/T	Intronic	*FTH1*	rs17185413	T/C	3'	-0.29			

TABLE A4.9 (Continued)

Gene 1	SNP 1	G'type 1[2]	SNP Type 1	Gene 2	SNP 2	G'type 2[2]	SNP Type 2	Estimate	LRT	LRT df	LRT P
SHMT1	rs1979276	A/A	3'					0.27	14.75	4	0.0053
SHMT1	rs1979276	G/A	3'					0.26			
				FTH1	rs17185413	C/C	3'	0.48			
				FTH1	rs17185413	T/C	3'	0.23			
SHMT1	rs1979276	A/A	3'	*FTH1*	rs17185413	C/C	3'	-12.78			
SHMT1	rs1979276	G/A	3'	*FTH1*	rs17185413	C/C	3'	-1.82			
SHMT1	rs1979276	A/A	3'	*FTH1*	rs17185413	T/C	3'	-0.79			
SHMT1	rs1979276	G/A	3'	*FTH1*	rs17185413	T/C	3'	-0.23			
CELF1	rs7933019	C/C	Intronic					-0.08	9.83	2	0.0074
CELF1	rs7933019	G/C	Intronic					0.08			
				FTH1	rs17156609	G/A	3'	-0.12			
CELF1	rs7933019	C/C	Intronic	*FTH1*	rs17156609	G/A	3'	1.26			
CELF1	rs7933019	G/C	Intronic	*FTH1*	rs17156609	G/A	3'	-1.14			
SHMT1	rs2461838	T/T	Intronic					0.13	13.44	4	0.0093
SHMT1	rs2461838	C/T	Intronic					0.09			
				FTH1	rs17185413	C/C	3'	0.30			
				FTH1	rs17185413	T/C	3'	0.10			
SHMT1	rs2461838	T/T	Intronic	*FTH1*	rs17185413	C/C	3'	-11.47			
SHMT1	rs2461838	C/T	Intronic	*FTH1*	rs17185413	C/C	3'	-2.20			
SHMT1	rs2461838	T/T	Intronic	*FTH1*	rs17185413	T/C	3'	-0.72			
SHMT1	rs2461838	C/T	Intronic	*FTH1*	rs17185413	T/C	3'	0.02			
CELF1	rs7102372	T/T	Intronic					0.13	11.19	3	0.011
CELF1	rs7102372	C/T	Intronic					-0.10			
				SHMT1	rs17806489	A/A	Intronic	-1.10			

TABLE A4.9 (Continued)

Gene 1	SNP 1	G'type 1[2]	SNP Type 1	Gene 2	SNP 2	G'type 2[2]	SNP Type 2	Estimate	LRT	LRT df	LRT P
				SHMT1	rs17806489	G/A	Intronic	0.22			
CELF1	rs7102372	T/T	Intronic	*SHMT1*	rs17806489	A/A	Intronic	0.00			
CELF1	rs7102372	C/T	Intronic	*SHMT1*	rs17806489	A/A	Intronic	2.57			
CELF1	rs7102372	T/T	Intronic	*SHMT1*	rs17806489	G/A	Intronic	-12.58			
CELF1	rs7102372	C/T	Intronic	*SHMT1*	rs17806489	G/A	Intronic	0.18			
FTH1	rs1800009	G/G	3'					0.15	12.59	4	0.013
FTH1	rs1800009	A/G	3'					0.26			
				SHMT1	rs12952556	C/C	3'	0.40			
				SHMT1	rs12952556	T/C	3'	0.32			
FTH1	rs1800009	G/G	3'	*SHMT1*	rs12952556	C/C	3'	-13.38			
FTH1	rs1800009	A/G	3'	*SHMT1*	rs12952556	C/C	3'	-0.75			
FTH1	rs1800009	G/G	3'	*SHMT1*	rs12952556	T/C	3'	-0.25			
FTH1	rs1800009	A/G	3'	*SHMT1*	rs12952556	T/C	3'	-0.38			
SHMT1	rs1979277	A/A	Nonsynonymous					0.19	12.43	4	0.014
SHMT1	rs1979277	G/A	Nonsynonymous					0.24			
				FTH1	rs17185413	C/C	3'	0.45			
				FTH1	rs17185413	T/C	3'	0.16			
SHMT1	rs1979277	A/A	Nonsynonymous	*FTH1*	rs17185413	C/C	3'	-11.61			
SHMT1	rs1979277	G/A	Nonsynonymous	*FTH1*	rs17185413	C/C	3'	-1.88			
SHMT1	rs1979277	A/A	Nonsynonymous	*FTH1*	rs17185413	T/C	3'	-0.54			
SHMT1	rs1979277	G/A	Nonsynonymous	*FTH1*	rs17185413	T/C	3'	-0.14			
SHMT1	rs2273028	T/T	Intronic					0.06	8.22	2	0.016
SHMT1	rs2273028	C/T	Intronic					0.08			
				FTH1	rs17156609	G/A	3'	-0.38			

TABLE A4.9 (Continued)

Gene 1	SNP 1	G'type 1[2]	SNP Type 1	Gene 2	SNP 2	G'type 2[2]	SNP Type 2	Estimate	LRT	LRT df	LRT P
SHMT1	rs2273028	T/T	Intronic	*FTH1*	rs17156609	G/A	3'	-11.87			
SHMT1	rs2273028	C/T	Intronic	*FTH1*	rs17156609	G/A	3'	0.74			
SHMT1	rs1979277	A/A	Nonsynonymous					0.39	12.09	4	0.017
SHMT1	rs1979277	G/A	Nonsynonymous					0.29			
				FTH1	rs1800009	G/G	3'	0.14			
				FTH1	rs1800009	A/G	3'	0.24			
SHMT1	rs1979277	A/A	Nonsynonymous	*FTH1*	rs1800009	G/G	3'	-13.37			
SHMT1	rs1979277	G/A	Nonsynonymous	*FTH1*	rs1800009	G/G	3'	-0.26			
SHMT1	rs1979277	A/A	Nonsynonymous	*FTH1*	rs1800009	A/G	3'	-0.73			
SHMT1	rs1979277	G/A	Nonsynonymous	*FTH1*	rs1800009	A/G	3'	-0.35			
SHMT1	rs2273028	T/T	Intronic					0.05	10.16	3	0.017
SHMT1	rs2273028	C/T	Intronic					0.13			
				FTH1	rs1801621	C/C	3'	-11.91			
				FTH1	rs1801621	T/C	3'	0.22			
SHMT1	rs2273028	T/T	Intronic	*FTH1*	rs1801621	C/C	3'	0.00			
SHMT1	rs2273028	C/T	Intronic	*FTH1*	rs1801621	C/C	3'	16.19			
SHMT1	rs2273028	T/T	Intronic	*FTH1*	rs1801621	T/C	3'	-12.07			
SHMT1	rs2273028	C/T	Intronic	*FTH1*	rs1801621	T/C	3'	-0.25			
SHMT1	rs1979277	A/A	Nonsynonymous					0.00	10.00	3	0.019
SHMT1	rs1979277	G/A	Nonsynonymous					0.11			
				FTH1	rs1801621	C/C	3'	-11.93			
				FTH1	rs1801621	T/C	3'	0.20			
SHMT1	rs1979277	A/A	Nonsynonymous	*FTH1*	rs1801621	C/C	3'	0.00			
SHMT1	rs1979277	G/A	Nonsynonymous	*FTH1*	rs1801621	C/C	3'	16.21			

TABLE A4.9 (Continued)

Gene 1	SNP 1	G'type 1[2]	SNP Type 1	Gene 2	SNP 2	G'type 2[2]	SNP Type 2	Estimate	LRT	LRT df	LRT P
SHMT1	rs1979277	A/A	Nonsynonymous	*FTH1*	rs1801621	T/C	3'	-12.01			
SHMT1	rs1979277	G/A	Nonsynonymous	*FTH1*	rs1801621	T/C	3'	-0.22			
FTH1	rs1801621	C/C	3'					-11.93	10.00	3	0.019
FTH1	rs1801621	T/C	3'					0.21			
				SHMT1	rs12952556	C/C	3'	0.00			
				SHMT1	rs12952556	T/C	3'	0.11			
FTH1	rs1801621	C/C	3'	*SHMT1*	rs12952556	C/C	3'	0.00			
FTH1	rs1801621	T/C	3'	*SHMT1*	rs12952556	C/C	3'	-12.01			
FTH1	rs1801621	C/C	3'	*SHMT1*	rs12952556	T/C	3'	16.20			
FTH1	rs1801621	T/C	3'	*SHMT1*	rs12952556	T/C	3'	-0.23			
SHMT1	rs1979276	A/A	3'					-0.02	9.97	3	0.019
SHMT1	rs1979276	G/A	3'					0.09			
				FTH1	rs1801621	C/C	3'	-11.94			
				FTH1	rs1801621	T/C	3'	0.20			
SHMT1	rs1979276	A/A	3'	*FTH1*	rs1801621	C/C	3'	0.00			
SHMT1	rs1979276	G/A	3'	*FTH1*	rs1801621	C/C	3'	16.22			
SHMT1	rs1979276	A/A	3'	*FTH1*	rs1801621	T/C	3'	-12.00			
SHMT1	rs1979276	G/A	3'	*FTH1*	rs1801621	T/C	3'	-0.21			
FTH1	rs17185413	C/C	3'					0.46	11.79	4	0.019
FTH1	rs17185413	T/C	3'					0.17			
				SHMT1	rs12952556	C/C	3'	0.20			
				SHMT1	rs12952556	T/C	3'	0.25			
FTH1	rs17185413	C/C	3'	*SHMT1*	rs12952556	C/C	3'	-11.61			
FTH1	rs17185413	T/C	3'	*SHMT1*	rs12952556	C/C	3'	-0.55			

TABLE A4.9 (Continued)

Gene 1	SNP 1	G'type 1[2]	SNP Type 1	Gene 2	SNP 2	G'type 2[2]	SNP Type 2	Estimate	LRT	LRT df	LRT P
FTH1	rs17185413	C/C	3'	*SHMT1*	rs12952556	T/C	3'	-1.81			
FTH1	rs17185413	T/C	3'	*SHMT1*	rs12952556	T/C	3'	-0.16			

[1]Interactions among SNPs within the same gene were not considered.
[2]Genotype of SNPs 1 and 2, respectively.

TABLE A4.10

ALL *MTHFD1L*, *MTHFR*, *TS*, AND *SHMT1* PAIRWISE INTERACTIONS WITH LIKELIHOOD RATIO TEST $P \leq 0.02$[1]

Gene 1	SNP 1	G'type 1[2]	SNP Type 1	Gene 2	SNP 2	G'type 2[2]	SNP Type 2	Estimate	LRT	LRT df	LRT P
MTHFD1L	rs803447	T/T	Intronic					0.38	20.81	4	0.00035
MTHFD1L	rs803447	C/T	Intronic					0.33			
				TYMS	rs16948305	T/T	Intronic	0.47			
				TYMS	rs16948305	C/T	Intronic	0.44			
MTHFD1L	rs803447	T/T	Intronic	*TYMS*	rs16948305	T/T	Intronic	1.11			
MTHFD1L	rs803447	C/T	Intronic	*TYMS*	rs16948305	T/T	Intronic	-13.66			
MTHFD1L	rs803447	T/T	Intronic	*TYMS*	rs16948305	C/T	Intronic	-0.56			
MTHFD1L	rs803447	C/T	Intronic	*TYMS*	rs16948305	C/T	Intronic	-0.62			
TYMS	rs2853533	C/C	Intronic					0.40	16.71	4	0.0022
TYMS	rs2853533	G/C	Intronic					-0.10			
				MTHFR	rs1801133	T/T	Nonsynonymous	0.42			
				MTHFR	rs1801133	C/T	Nonsynonymous	0.04			
TYMS	rs2853533	C/C	Intronic	*MTHFR*	rs1801133	T/T	Nonsynonymous	-12.81			
TYMS	rs2853533	G/C	Intronic	*MTHFR*	rs1801133	T/T	Nonsynonymous	0.78			
TYMS	rs2853533	C/C	Intronic	*MTHFR*	rs1801133	C/T	Nonsynonymous	0.37			
TYMS	rs2853533	G/C	Intronic	*MTHFR*	rs1801133	C/T	Nonsynonymous	0.10			
TYMS	rs502396	C/C	Intronic					-0.30	13.20	4	0.010
TYMS	rs502396	T/C	Intronic					-0.31			
				MTHFR	rs12121543	A/A	Intronic	-1.55			
				MTHFR	rs12121543	C/A	Intronic	-0.07			
TYMS	rs502396	C/C	Intronic	*MTHFR*	rs12121543	A/A	Intronic	2.16			
TYMS	rs502396	T/C	Intronic	*MTHFR*	rs12121543	A/A	Intronic	1.16			
TYMS	rs502396	C/C	Intronic	*MTHFR*	rs12121543	C/A	Intronic	-0.06			

TABLE A4.10 (Continued)

Gene 1	SNP 1	G'type 1[2]	SNP Type 1	Gene 2	SNP 2	G'type 2[2]	SNP Type 2	Estimate	LRT	LRT df	LRT P
TYMS	rs502396	T/C	Intronic	*MTHFR*	rs12121543	C/A	Intronic	0.27			
MTHFR	rs6541003	G/G	Intronic					-0.21	11.70	4	0.020
MTHFR	rs6541003	A/G	Intronic					-0.13			
				SHMT1	rs17806489	A/A	Intronic	-12.04			
				SHMT1	rs17806489	G/A	Intronic	0.09			
MTHFR	rs6541003	G/G	Intronic	*SHMT1*	rs17806489	A/A	Intronic	16.71			
MTHFR	rs6541003	A/G	Intronic	*SHMT1*	rs17806489	A/A	Intronic	11.49			
MTHFR	rs6541003	G/G	Intronic	*SHMT1*	rs17806489	G/A	Intronic	-0.18			
MTHFR	rs6541003	A/G	Intronic	*SHMT1*	rs17806489	G/A	Intronic	0.26			

[1]Interactions among SNPs within the same gene were not considered.
[2]Genotype of SNPs 1 and 2, respectively.

CHAPTER 5

CONCLUSION

The unifying theme of these three research projects was to understand the associations between genetic variation in the folate-mediated one-carbon network and cardiovascular disease risk. At the time these projects were undertaken, much of gene-disease association research was focused on single genetic variants or small groups of variants and few studies included an evaluation of gene or nutrient interactions; at the same time, the first discovery-oriented genome-wide association studies involving statistical tests of hundreds of thousands of genetic variants in relation to disease status were being published. Important questions in the field of folate metabolism and cardiovascular risk surrounded the role of B-vitamin supplementation for CVD prevention through homocysteine-lowering (14, 65, 101). Several recent meta-analyses had built consensus for the relation of the *MTHFR* rs1801133 genetic variant and CVD risk (15, 27, 52, 58, 108), while the bulk of the research on another important folate-related gene, *SHMT1,* focused on cancer outcomes with little evidence for associations with CVD. Few studies had investigated gene-gene interactions in the network; however, a recent publication studying the *SHMT1 1420 C→T* rs1979277 variant in relation to CVD had identified a novel interaction with *MTHFR* 677 C→T rs1801133 such that CVD risk was increased in subgroups of individuals carrying both the *SHMT1* 1979277 *TT* and *MTHFR* rs1801133 *CT* or *TT* genotypes (62).

Against this backdrop, and with the knowledge that folate related genes function as part of a highly interconnected biochemical network, we set out to investigate the role of genetic variation within the folate-mediated one-carbon network

and CVD risk. The research goals were multifold: 1) to summarize the current evidence base for the association between genetic variants in folate-mediated one-carbon metabolism and cardiovascular disease risk; 2) to investigate the hypothesis that the *SHMT1* rs1979277 and *MTHFR* rs1801133 genotypes interact to increase heart disease risk in the Health Professionals Follow-Up Study (HPFS) and the Nurses' Health Study (NHS); 3) to evaluate the association between genetic variation in the folate-mediated one-carbon network and methylation-related phenotypes in the Normative Aging Study, including main effects, gene-gene interactions, and gene-nutrient interactions; and 4) to evaluate the association between genetic variation in the folate-mediated one-carbon network and CVD risk in the Normative Aging Study, including main effects, gene-gene interactions, and gene-nutrient interactions.

A systematic review and meta-analysis was conducted in 2008 to evaluate the evidence base for the association between genetic variation in a set of 52 genes involved in the folate-mediated one-carbon metabolic pathway and cardiovascular disease risk (114). This work revealed that folate network genes involved in transsulfuration, cytoplasmic metabolism, and vitamin B12 metabolism have been the focus of most work relating to CVD, while folate network genes involved in absorption and transport, mitochondrial metabolism, and nuclear folate metabolism, have few publications relating to CVD, with many genes lacking any evidence whatsoever. Further, only variants in genes related to homocysteine metabolism (*CBS*, *MTR*, *MTRR*, and *TCN2*) had been sufficiently well-studied to allow meta-analysis, and the association with CVD differed by gene and outcome.

Polymorphisms in Serine Hydroxymethyltransferase 1 and Methylenetetrahydrofolate Reductase Interact to Increase Cardiovascular Risk in Humans

The first analysis of this dissertation was a candidate gene association study conducted to investigate the hypothesis that the *SHMT1* rs1979277 and *MTHFR* rs1801133 genotypes interact to increase heart disease risk in the paired cohorts of the Health Professionals Follow-Up Study (HPFS) and the Nurses' Health Study (NHS). Since the association between these genotypes and CVD risk may be mediated by homocysteine and/or modified by B vitamins, these questions were also investigated. The published literature investigating the role of the *SHMT1* rs1979277 *1420 C→T* genotype in chronic disease risk focuses mainly on cancer risk, and evidence for genotype associations with cardiovascular disease (CVD) is limited to 3 studies, underscoring the need for further population-level studies of *SHMT1* and CVD.

The NHS findings were consistent with associations previously reported in the Normative Aging Study (NAS): risk associated with the *MTHFR* rs1801133 *T* allele was stronger in women with the *SHMT1* rs1979277 *TT* genotype. The absence of double homozygote cases in the NHS precluded estimating the association of *MTHFR* rs1801133 *TT* genotype with CHD risk in women with the *SHMT1* rs1979277 *TT* genotype. The lack of double homozygote cases may indicate the selective loss of such individuals from the cohort, but this scenario is highly unlikely given observed association sizes. An alternative and more likely explanation is that observed genotype frequencies are a chance phenomenon. The findings in the HPFS, however, were inconsistent with summary estimates from prior meta-analyses regarding the association of the *MTHFR* rs1801133 genotype on CVD risk. In the HPFS, the *MTHFR* rs1801133 *T* allele was associated with a decreased risk of CVD, thus

opposite in direction to findings from prior meta-analyses (52, 58, 108). The HPFS findings are also inconsistent with the past report (62) and the NHS findings reported herein with regard to the gene-gene interaction between *MTHFR* rs1801133 and *SHMT1* rs1979277.

This study addressed an important gap in the literature by investigating the relation of the *SHMT1* rs1979277 *TT* genotype to heart disease. The study was based on a biologically plausible hypothesis, linked to basic science findings, and, coupled with the results of the systematic review and meta-analysis, highlights the need for a more complete evaluation of genetic variation within the folate metabolic network and cardiovascular disease risk, particularly in pre-fortification or unfortified populations, including gene-gene epistasis and gene-nutrient interactions.

Folate Network Genetic Variation, Plasma Homocysteine, and Global Genomic Methylation Content

Most previous work investigating variation in genes contributing to folate-mediated one-carbon metabolism in relation to homocysteine and genomic methylation phenotypes focused on a small number of candidate genes. Given the interconnectedness of the one-carbon pathway, less proximal enzymes and genes may be important, yet few studies have attempted to evaluate a comprehensive set of folate-related genes. Having completed an investigation of the role of polymorphisms in two key folate network genes in relation to CVD risk, we wished to explore a more complete set of genes and polymorphisms within the folate network, to better understand how genetic variation within this interconnected pathway was associated with CVD risk.

Therefore, we next investigated sequence variation in a network of candidate genes involved in one-carbon metabolism in relation to plasma total homocysteine and two measures of global genomic DNA methylation (Alu and LINE-1 elements). Genes, SNPs, and related nutrients were carefully chosen to represent the full functional variation of the folate-mediated one carbon metabolism pathway. The analysis considered all 330 SNPs in tests for main effects, interactions between each SNP and the *MTHFR* rs1801133 SNP, and interactions between each SNP and plasma measures of folate, vitamins B-6 and B-12. FDR-significant main effects were identified for the plasma total homocysteine phenotype, and FDR significant SNP—B-6 interactions were identified for Alu and LINE-1 element methylation phenotypes.

Polymorphisms in the *SLC19A1, FTCD,* and *SLC19A3* genes were most predictive of plasma homocysteine levels, suggesting the importance of pathways that function to generate one-carbon units. Together, the variability in homocysteine explained by the model containing the set of the 3 most significant nonredundant SNP hits was 3.6%, a small proportion of the estimated >50% heritability in homocysteine, based on twin studies (76, 92), but similar to the variability explained by two other determinants of homocysteine, age and smoking (84), which together explained 3.5%. In this study, there were FDR-significant interactions between studied SNPs and *MTHFR 677 C→T* rs1801133 in relation to the homocysteine outcome, but no FDR-significant interactions between studied SNPs and plasma folate, vitamin B-6, or vitamin B-12 for the plasma homocysteine phenotype, a finding which may reflect the relative high folate status of this pre-fortification population or limited power for gene-nutrient interactions.

Variation in the *GNMT* was associated with Alu element methylation levels at nominal significance levels. The biological rationale for this association is strong, as the *GNMT* gene encodes a methyltransferase that regulates cellular methylation

potential by catalyzing the conversion of S-adenosylmethionine (AdoMet) and glycine to S-adenosylhomocysteine (AdoHcy) and sarcosine. When AdoMet is low, the MTHFR enzyme produces 5-methyl-THF, which in turn binds to GNMT and prevents AdoMet catabolism. When AdoMet is high, AdoMet inhibits the production of 5-methyl-THF by MTHFR, thus activating GNMT, which catabolizes the excess AdoMet to yield AdoHcy (32). In the current study, there was little or no mediation of the *GNMT* rs1051218—Alu phenotype association by plasma folate, vitamin B-6 or vitamin B-12, but methylation status was assessed after the initiation of mandatory folate fortification in the U.S. and that limits consideration of mediation. The variability in Alu element global genomic DNA methylation explained by the rs1051218 SNP was 2.2%. Heritability estimates for DNA methylation are highly variable, ranging from 0 to 94% (12, 50), and the variability explained by the rs1051218 SNP was greater than the variation explained by age and smoking, which accounted for about 2% of variation in Alu element methylation in our models. There were no SNP—*MTHFR 677* rs1801133 interactions associated with Alu element methylation at FDR significance levels. Similarly, none of the SNP—folate or SNP—vitamin B-12 interaction terms reached FDR significance in predicting Alu element methylation. However, three FDR-significant SNP—vitamin B-6 interactions were identified, including two intronic SNPs in the *AMT* gene (rs1464567 and rs1464566) and one intronic SNP in the *DNMT3B* gene (rs1883729).

There were no FDR-significant associations observed for the LINE-1 methylation phenotype. Similarly, no SNPs had statistically significant interactions with *MTHFR 677* rs1801133, and there were no FDR-significant interactions between SNPs and folate or vitamin B-12. Because the LINE-1 phenotype was measured after the introduction of mandatory folate fortification in the U.S. findings may be limited by this timing. An interaction of plasma B-6 with 1 SNP in the *MTHFD1L* gene was

significant at the FDR threshold. While the *MTHFD1L* gene is not vitamin B-6-dependent, it functions as part of mitochondrial folate metabolism, downstream from the vitamin B-6-dependent glycine cleavage system (23). Intronic variation in *MTHFD1L* was associated with CVD in a large-scale genome-wide association study (112).

This analysis demonstrated that genes involved in absorption and transport of nutrients related to folate-mediated one-carbon metabolism had the most significant associations with the homocysteine phenotype; consistent with this finding, about 30-40% of the association was mediated through plasma folate and vitamin B-6 and B-12 levels. Mitochondrial metabolism, methylation/ homocysteine pathways, cytoplasmic metabolism, nuclear metabolism, and B-12 metabolism were also represented in the top hits. For the Alu-element methylation phenotype, the top hits were in genes playing a role in mitochondrial metabolism, nuclear metabolism, and methylation/homocysteine metabolism. For the global genomic methylation LINE-1 phenotype, the top SNP was in a gene in the methylation/homocysteine pathway of one-carbon metabolism. There was no evidence that plasma folate or vitamins B-6 or B-12 mediated the association of SNPs with the methylation phenotypes.

This study suggests that beyond the well-described *MTHFR* rs1801133 SNP, other SNPs make important contributions to homocysteine and global genomic DNA methylation phenotypes, and some associations are sensitive to nutritional status of B vitamins. Future work should continue to include a broad evaluation of one-carbon network genetic and nutritional variation in unfortified or pre-fortification populations and extend these findings for CVD biomarkers to investigate CVD phenotypes directly.

Folate Network Genetic Variation and Cardiovascular Disease (CVD) Risk: A Network Association Study

The previous 2 studies established that associations of folate network genetic variants with CVD risk extends beyond the *MTHFR 677 C→T* rs1801133 polymorphism, that polymorphisms in genes related to the generation of one-carbon units were predictive of methylation-related CVD biomarkers, and that gene-gene interactions predictive of CVD risk exist within the network. Therefore, goal of the final analysis was to study genetic variation in the folate metabolic network as a whole in relation to CVD risk, including gene-gene and gene-nutrient interactions.

We investigated sequence variation in a network of candidate genes involved in one-carbon metabolism in relation to CVD risk in a time-to-event analysis. Genes, SNPs, and related nutrients were carefully chosen to represent the full functional variation of the folate-mediated one carbon metabolism pathway. The analysis considered all 330 SNPs in tests for main effects, interactions with the *MTHFR* rs1801133 SNP, and a simultaneous evaluation of network variation. The interaction of each SNP with plasma measures of folate, vitamins B-6 and B-12 was also considered. FDR-significant main effects were identified, as well as FDR significant SNP—folate and SNP—vitamin B-12 interactions and nominally significant gene-gene interactions were identified in a more thorough investigation of 3 hypotheses about the network.

Variation in the *GGH* gene was associated with CVD risk at FDR-significance levels. The *GGH* gene encodes an enzyme that functions to hydrolyze polyglutamate residues from intracellular folate to facilitate folate export (46); however, the association was not substantially mediated by plasma folate. The previously described association between the *MTHFR* rs1801133 polymorphism and CVD risk was

observed in this cohort and findings, including the rs1801133-folate interaction, were consistent with previous reports; however, the relatively high folate status among the men in this cohort study (mean plasma folate: 10.4 ng/ml, significantly higher than the prefortification serum folate mean of 5.8 ng/ml in non-Hispanic Whites reported by NHANES III (1988-1994), albeit lower than the post-fortification mean of 14.8 ng/ml (80) may have blunted the strength of this interaction. No FDR-significant interactions between studied SNPs and vitamin B-6, were observed; however, FDR-significant interactions were identified for plasma folate and a polymorphism in the *MAT2B* gene, and vitamin B-12 and polymorphisms in the *BHMT* and *SLC25A32* genes. The lack of additional highly significant gene-nutrient interactions may reflect limited power for gene-nutrient interactions, or the lack of men with low B vitamin status, even within this pre-fortification cohort of U.S. men.

No FDR-significant interactions were identified between studied SNPs and *MTHFR 677 C→T* rs1801133. Furthermore, of all pairwise interactions within the network, no interactions surpassed the FDR significance threshold of $P \leq 0.2$. A previously described gene-gene interaction involving *SHMT1* and *MTHFR* was replicated in these data; the association was strengthened considerably in models adjusted for homocysteine, and attenuated in models adjusted for folate and vitamin B-12. Finally, pleiotropic effects were examined across outcomes. Of the 8 nominally significant hits for the CVD analysis, only a single pleiotropic SNP hit was identified, *CBS* rs6586282, which was also among the most significant hits for the plasma homocysteine phenotype. Genes that had associations with multiple phenotypes included *GGH* and *MTHFR.*

In a systematic approach to the data analysis, we identified the best genetic models for each SNP, tested single SNPs, a reduced set of non-redundant SNPs, the interaction of each SNP with *MTHFR* 677 rs1801133, and the interaction of each SNP

with folate, vitamin B-6 and vitamin B-12. Findings were corrected for multiple comparisons and findings based on the FDR threshold were discussed in more detail. Hypotheses based on basic science findings were explored.

This study suggests that variation in genes other than *MTHFR* and those directly involved in homocysteine metabolism, are associated with CVD risk. For example, a gene involved in folate absorption and transport, *GGH*, was more significantly associated with CVD than any of the other 51 genes evaluated, including *MTHFR* and *CBS*. This study also supports a role for mitochondrial metabolism in predicting CVD risk, and although a previously identified polymorphism in *MTHFD1L* (112) could not be studied directly, nearby variants showed strong associations with CVD risk. The epidemiologic findings presented here replicated molecular biology studies of interactions among the *SHMT1, UBE2I, UBE2N, DHFR,* and *TYMS* proteins resulting in the cell cycle-dependent nuclear localization of the thymidylate biosynthesis pathway, and further established that epistatic interaction among genes encoding these proteins is predictive of CVD risk.

Strengths of this work include a complete investigation of genetic variation in the folate metabolic network. Furthermore, the analysis was conducted within a large cohort with CVD outcome data collected prior to the introduction of mandatory folate-fortification in the U.S.; thus, the study population is well-suited to investigate the network of folate-related genes in relation to CVD risk. SNP selection for the genotyping assay reflected coverage of genes based on function, linkage, and physical coverage, which led to markers that captured functional variants in addition to tagging variation elsewhere in the gene. In a systematic approach to the data analysis, we identified the best genetic models for each SNP, tested single SNPs, a reduced set of non-redundant SNPs, the interaction of each SNP with *MTHFR* 677 rs1801133, and the interaction of each SNP with folate, vitamin B-6 and vitamin B-12. Findings were

corrected for multiple comparisons and findings based on the FDR threshold were discussed in more detail. Hypotheses based on basic science findings were explored. Weaknesses of the study include the fact that, despite restricting data to the pre-folate fortification time period, the level of folate nutrition within the population was relatively high compared to other measures of pre-fortification folate status in the U.S., which may have attenuated associations, particularly gene-nutrient interactions, that would have been present in a population with more variation in B-vitamin status.

These findings support a role for *SHMT1* and nuclear folate metabolism, including the thymidylate biosynthesis pathway, in relation to CVD. This investigation of genetic variation within the folate network, including absorption and transport, mitochondrial, and nuclear metabolism, alters our current understanding of the relation between the folate-mediated one-carbon network and CVD. To date, the focus has been on the role of homocysteine, but the findings in this study suggest important contributions to risk through other aspects of folate metabolism. These findings may help explain the results of randomized controlled trials focused on CVD risk reduction through homocysteine lowering, which have experienced limited success (70, 71). A thorough understanding of the role of folate network genetic and nutritional variation in relation to CVD is important, particularly in the context of the numerous unfortified populations around the world.

Significance of this work

This dissertation was designed to address the role of genetic variation in the folate-mediated one-carbon metabolic network as a whole in relation to CVD risk, and has shown that evaluations of networks of candidate genes and related nutrients can contribute to our understanding of chronic disease risk in ways that are not possible

through more limited evaluations of genes or nutrients involved in biochemical pathways.

The work presented in this dissertation represents several novel findings. The most significant hits for the homocysteine and methylation outcomes reflected genes involved in the generation of one-carbon units, including *SLC19A1, FTCD,* and a gene known to regulate levels of AdoMet, *GNMT*. A genetic variant in gamma-glutamyl hydrolase (*GGH*), was more significantly predictive of CVD risk than any other one-carbon network polymorphism studied, and the effect size for the *GGH* variant doubled in analyses of early-onset CVD. Genes in this category of folate metabolism are not well-studied in relation to CVD risk. An in-depth evaluation of nuclear folate metabolism revealed significant epistasis predictive of CVD risk, and genes that regulate *SHMT1* expression interact to influence risk of CVD. The findings underscore the importance of mitochondrial metabolism in the folate pathway–CVD association, and epistasis predictive of CVD risk exists between the thymidylate synthesis and remethylation pathways. Overall, genetic variants beyond the well-known *MTHFR* rs1801133 polymorphism, are strongly associated with CVD risk as well as the homocysteine and global genomic DNA methylation phenotypes, epistatic interactions predictive of CVD risk exist within the network, and some associations are sensitive to nutritional status of B vitamins.

Future directions

The work presented here represents an important contribution to our understanding of genetic variation within a network of genes and CVD risk, and suggests several possible directions for future work. First, future studies should consider utilizing the approach outlined here to evaluate genetic variation in the folate-

mediated one-carbon network in relation to phenotypes other than CVD. Previous work has identified important associations between genetic and nutritional variation in the folate-mediated one-carbon network and phenotypes including cancer, birth defects, and cognitive outcomes; it would be informative to investigate whether the same set of most significant SNPs, gene-gene, and gene-nutrient interactions identified for CVD also predict other disease outcomes when a more complete set of genetic variation is considered. Furthermore, where known gene-gene interactions exist, future evaluations of gene-disease association should explore these interactions in lieu of main effects for the relevant polymorphisms.

To help capture missing heritability and increase the proportion of variation explained by models, future work may wish to extend considerations of genetic variation beyond SNPs to include copy number variants, rare variants, haplotypes, and higher order gene-gene and gene-nutrient interactions. There are a number of copy number variants within the genes of the folate-mediated one-carbon network, such as those in the *CBS* and *TYMS* genes, that have been previously associated with disease phenotypes, and although every effort was made to capture this variation through LD, these should be included in future evaluations of network genetic variation as they may not be completely captured by SNPs (24). Future genetic association studies using large sample sizes could consider sequencing folate network genes to capture rare variants that may have strong associations with disease but are not well-represented in these data. This approach would allow complete typing of both rare and common variants, and explicitly capture haplotypes. Future work should also attempt to characterize the functional consequences of genetic variation in the *MTHFD1L* gene, particularly the 14 kb block identified in the Wellcome Trust study, with regard to cardiovascular disease risk. Where sufficient power exists, future studies may wish to investigate higher order gene-gene and gene-nutrient interactions

than could be studied here. Finally, these findings should be replicated in other cohorts, particularly unfortified populations with large ranges in B vitamin status; European populations may be a good choice.

Understanding the functional consequences of the genetic variants highlighted in this work is an important next step, and work to characterize these consequences in terms of their impact on protein structure and enzyme kinetics will be important to provide a mechanism for the epidemiologic associations identified here. For example, genes involved in absorption and transport, nuclear, and mitochondrial metabolism would be good targets for further functional characterization, as would genes involved in epistatic interactions. In addition, functional characterization of genetic variation in genes such as *SLC19A3*, which have not been well-studied thus far, would be informative; this gene belongs to the family of folate transporters but is not known to directly transport folate, vitamin B-6, or vitamin B-12 yet was among the most significant SNP hits for the homocysteine phenotype and the association was mediated by folate, B-6 and B-12. For studies investigating genetic variants in relation to cellular methylation potential, indicators of methylation potential more sensitive than homocysteine and global genomic DNA methylation, such as AdoHcy (51, 107), and gene-specific DNA methylation and expression levels, may yield even stronger associations with network variants and facilitate our understanding of the mechanism for the associations observed here.

Ultimately, more complete information on the associations between networks of genetic variants and disease risk will enable more appropriate nutrition interventions for disease prevention. However, additional work is needed to translate findings from gene-disease association studies into practice. Currently, most gene-disease association hits have not been sufficiently well-characterized to be clinically useful(68, 69). Therefore, the clinical applications of findings resulting from such

studies should be explored, including an evaluation of CVD risk prediction based on a knowledge of genetic variation in the folate network. Ultimately, the benefit of genetic testing for CVD risk prediction will need to be weighed against the costs associated with implementing testing and resulting interventions(54); however, better understanding the ways in which genetic variation and nutrient status interact to influence risk of CVD may enable the development of more effective strategies for preventing this common and costly disease.

LITERATURE CITED

1. Anonymous (2002) Homocysteine and risk of ischemic heart disease and stroke: a meta-analysis. JAMA 288: 2015-2022.

2. Anderson, D. D. & Stover, P. J. (2009) SHMT1 and SHMT2 are functionally redundant in nuclear de novo thymidylate biosynthesis. PLoS One 4: e5839.

3. Anderson, D. D., Woeller, C. F. & Stover, P. J. (2007) Small ubiquitin-like modifier-1 (SUMO-1) modification of thymidylate synthase and dihydrofolate reductase. Clin Chem Lab Med 45: 1760-1763.

4. Baccarelli, A., Cassano, P. A., Litonjua, A., Park, S. K., Suh, H., Sparrow, D., Vokonas, P. & Schwartz, J. (2008) Cardiac autonomic dysfunction: effects from particulate air pollution and protection by dietary methyl nutrients and metabolic polymorphisms. Circulation 117: 1802-1809.

5. Baccarelli, A., Wright, R., Bollati, V., Litonjua, A., Zanobetti, A., Tarantini, L., Sparrow, D., Vokonas, P. & Schwartz, J. (2010) Ischemic Heart Disease and Stroke in Relation to Blood DNA Methylation. Epidemiology (Cambridge, Mass.

6. Bailey, L. B. (2003) Folate, methyl-related nutrients, alcohol, and the MTHFR 677C-->T polymorphism affect cancer risk: intake recommendations. The Journal of nutrition 133: 3748S-3753S.

7. Baylin, S. B. (2005) DNA methylation and gene silencing in cancer. Nat. Clin Pract. Oncol. 2 Suppl 1: S4-11.

8. Beaudin, A. E. & Stover, P. J. (2009) Insights into metabolic mechanisms underlying folate-responsive neural tube defects: a minireview. Birth Defects Res. Part A Clin. Mol. Teratol. 85: 274-284.

9. Bell, B., Rose, C. L. & Damon, A. (1966) The Veterans Administration longitudinal study of healthy aging. The Gerontologist 6: 179-184.

10. Benjamini, Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society, Series B (Methodological) 57: 289-300.

11. Bentley, A. R., Raiszadeh, F., Stover, P. J., Hunter, D. J., Hankinson, S. E. & Cassano, P. A. (2010) No association between cSHMT genotypes and the risk of breast cancer in the Nurses' Health Study. European journal of clinical nutrition 64: 108-110.

12. Boks, M. P., Derks, E. M., Weisenberger, D. J., Strengman, E., Janson, E., Sommer, I. E., Kahn, R. S. & Ophoff, R. A. (2009) The relationship of DNA methylation with age, gender and genotype in twins and healthy controls. PLoS One 4: e6767.

13. Bollati, V., Schwartz, J., Wright, R., Litonjua, A., Tarantini, L., Suh, H., Sparrow, D., Vokonas, P. & Baccarelli, A. (2009) Decline in genomic DNA methylation through aging in a cohort of elderly subjects. Mech. Ageing Dev. 130: 234-239.

14. Bonaa, K. H., Njolstad, I., Ueland, P. M., Schirmer, H., Tverdal, A., Steigen, T., Wang, H., Nordrehaug, J. E., Arnesen, E. & Rasmussen, K. (2006) Homocysteine lowering and cardiovascular events after acute myocardial infarction. N. Engl. J. Med. 354: 1578-1588.

15. Casas, J. P., Hingorani, A. D., Bautista, L. E. & Sharma, P. (2004) Meta-analysis of genetic studies in ischemic stroke: thirty-two genes involving approximately 18,000 cases and 58,000 controls. Arch. Neurol. 61: 1652-1661.

16. Center for Disease Control and Prevention, D. f. H. D. a. S. P. (2006) Chronic Disease: Cardiovascular Health At-A-Glance.

17. Chatzikyriakidou, A., Georgiou, I., Voulgari, P. V., Papadopoulos, C. G., Tzavaras, T. & Drosos, A. A. (2007) Transcription regulatory polymorphism -43T>C in the 5'-flanking region of SLC19A1 gene could affect rheumatoid arthritis patient response to methotrexate therapy. Rheumatol. Int 27: 1057-1061.

18. Chatzikyriakidou, A., Vakalis, K. V., Kolaitis, N., Kolios, G., Naka, K. K., Michalis, L. K. & Georgiou, I. (2008) Distinct association of SLC19A1

polymorphism -43T>C with red cell folate levels and of MTHFR polymorphism 677C>T with plasma folate levels. Clin Biochem. 41: 174-176.

19. Chave, K. J., Ryan, T. J., Chmura, S. E. & Galivan, J. (2003) Identification of single nucleotide polymorphisms in the human gamma-glutamyl hydrolase gene and characterization of promoter polymorphisms. Gene 319: 167-175.

20. Chen, J., Kyte, C., Valcin, M., Chan, W., Wetmur, J. G., Selhub, J., Hunter, D. J. & Ma, J. (2004) Polymorphisms in the one-carbon metabolic pathway, plasma folate levels and colorectal cancer in a prospective study. Int. J. Cancer 110: 617-620.

21. Chiang, E. P., Wang, Y. C. & Tang, F. Y. (2007) Folate restriction and methylenetetrahydrofolate reductase 677T polymorphism decreases adoMet synthesis via folate-dependent remethylation in human-transformed lymphoblasts. Leukemia 21: 651-658.

22. Choi, S. H., Heo, K., Byun, H. M., An, W., Lu, W. & Yang, A. S. (2010) Identification of preferential target sites for human DNA methyltransferases. Nucleic acids research.

23. Christensen, K. E. & MacKenzie, R. E. (2008) Mitochondrial methylenetetrahydrofolate dehydrogenase, methenyltetrahydrofolate cyclohydrolase, and formyltetrahydrofolate synthetases. Vitam. Horm. 79: 393-410.

24. Cooper, G. M., Zerr, T., Kidd, J. M., Eichler, E. E. & Nickerson, D. A. (2008) Systematic assessment of copy number variant detection via genome-wide SNP genotyping. Nat. Genet 40: 1199-1203.

25. Cordaux, R. & Batzer, M. A. (2009) The impact of retrotransposons on human genome evolution. Nat. Rev Genet 10: 691-703.

26. Corella, D. & Ordovas, J. M. (2009) Nutrigenomics in cardiovascular medicine. Circ. Cardiovasc. Genet 2: 637-651.

27. Cronin, S., Furie, K. L. & Kelly, P. J. (2005) Dose-related association of MTHFR 677T allele with risk of ischemic stroke: evidence from a cumulative meta-analysis. Stroke; a journal of cerebral circulation 36: 1581-1587.

28. Damon, A., Seltzer, C. C., Stoudt, H. W. & Bell, B. (1972) Age and physique in health white veterans at Boston. Journal of gerontology 27: 202-208.

29. de Jonge R., Tissing, W. J., Hooijberg, J. H., Jansen, G., Kaspers, G. J., Lindemans, J., Peters, G. J. & Pieters, R. (2009) Polymorphisms in folate-related genes and risk of pediatric acute lymphoblastic leukemia. Blood 113: 2284-2289.

30. DeVos, L., Chanson, A., Liu, Z., Ciappio, E. D., Parnell, L. D., Mason, J. B., Tucker, K. L. & Crott, J. W. (2008) Associations between single nucleotide polymorphisms in folate uptake and metabolizing genes with blood folate, homocysteine, and DNA uracil concentrations. Am. J. Clin. Nutr. 88: 1149-1158.

31. Doi, T., Kawata, T., Tadano, N., Iijima, T. & Maekawa, A. (1989) Effect of vitamin B12-deficiency on the activity of hepatic cystathionine beta-synthase in rats. J Nutr Sci. Vitaminol. (Tokyo) 35: 101-110.

32. Fox, J. T. & Stover, P. J. (2008) Folate-mediated one-carbon metabolism. Vitam. Horm. 79: 1-44.

33. Fredriksen, A., Meyer, K., Ueland, P. M., Vollset, S. E., Grotmol, T. & Schneede, J. (2007) Large-scale population-based metabolic phenotyping of thirteen genetic polymorphisms related to one-carbon metabolism. Hum Mutat. 28: 856-865.

34. Freimer, N. B. & Sabatti, C. (2007) Human genetics: variants in common diseases. Nature 445: 828-830.

35. Frosst, P., Blom, H. J., Milos, R., Goyette, P., Sheppard, C. A., Matthews, R. G., Boers, G. J., den, H. M., Kluijtmans, L. A. et al. (1995) A candidate genetic risk factor for vascular disease: a common mutation in methylenetetrahydrofolate reductase. Nat. Genet 10: 111-113.

36. Fuso, A., Nicolia, V., Cavallaro, R. A. & Scarpa, S. (2010) DNA methylase and demethylase activities are modulated by one-carbon metabolism in Alzheimer's disease models. J Nutr Biochem.

37. Ganapathy, V., Smith, S. B. & Prasad, P. D. (2004) SLC19: the folate/thiamine transporter family. Pflugers Arch 447: 641-646.

38. Garrow, T. A. (1996) Purification, kinetic properties, and cDNA cloning of mammalian betaine-homocysteine methyltransferase. J Biol Chem 271: 22831-22838.

39. Gillies, P. J. & Krul, E. S. (2007) Using genetic variation to optimize nutritional preemption. The Journal of nutrition 137: 270S-274S.

40. Girgis, S., Nasrallah, I. M., Suh, J. R., Oppenheim, E., Zanetti, K. A., Mastri, M. G. & Stover, P. J. (1998) Molecular cloning, characterization and alternative splicing of the human cytoplasmic serine hydroxymethyltransferase gene. Gene 210: 315-324.

41. Giusti, B., Saracini, C., Bolli, P., Magi, A., Sestini, I., Sticchi, E., Pratesi, G., Pulli, R., Pratesi, C. & Abbate, R. (2008) Genetic analysis of 56 polymorphisms in 17 genes involved in methionine metabolism in patients with abdominal aortic aneurysm. J. Med. Genet. 45: 721-730.

42. Herbig, K., Chiang, E. P., Lee, L. R., Hills, J., Shane, B. & Stover, P. J. (2002) Cytoplasmic serine hydroxymethyltransferase mediates competition between folate-dependent deoxyribonucleotide and S-adenosylmethionine biosyntheses. J. Biol. Chem. 277: 38381-38389.

43. Hilton, J. F., Christensen, K. E., Watkins, D., Raby, B. A., Renaud, Y., de la Luna, S., Estivill, X., MacKenzie, R. E., Hudson, T. J. & Rosenblatt, D. S. (2003) The molecular basis of glutamate formiminotransferase deficiency. Human mutation 22: 67-73.

44. Holm, P. I., Bleie, O., Ueland, P. M., Lien, E. A., Refsum, H., Nordrehaug, J. E. & Nygard, O. (2004) Betaine as a determinant of postmethionine load total plasma homocysteine before and after B-vitamin supplementation. Arterioscler. Thromb. Vasc. Biol 24: 301-307.

45. Homocysteine Lowering Trialists' Collaboration (2005) Dose-dependent effects of folic acid on blood concentrations of homocysteine: a meta-analysis of the randomized trials. Am. J. Clin. Nutr. 82: 806-812.

46. Ifergan, I. & Assaraf, Y. G. (2008) Molecular mechanisms of adaptation to folate deficiency. Vitam. Horm. 79: 99-143.

47. Jacques, P. F., Bostom, A. G., Williams, R. R., Ellison, R. C., Eckfeldt, J. H., Rosenberg, I. H., Selhub, J. & Rozen, R. (1996) Relation between folate status, a common mutation in methylenetetrahydrofolate reductase, and plasma homocysteine concentrations. Circulation 93: 7-9.

48. Jamaluddin, M. S., Yang, X. & Wang, H. (2007) Hyperhomocysteinemia, DNA methylation and vascular disease. Clin Chem Lab Med 45: 1660-1666.

49. Jang, H., Mason, J. B. & Choi, S. W. (2005) Genetic and epigenetic interactions between folate and aging in carcinogenesis. The Journal of nutrition 135: 2967S-2971S.

50. Kaminsky, Z. A., Tang, T., Wang, S. C., Ptak, C., Oh, G. H., Wong, A. H., Feldcamp, L. A., Virtanen, C., Halfvarson, J. et al. (2009) DNA methylation profiles in monozygotic and dizygotic twins. Nat. Genet 41: 240-245.

51. Kerins, D. M., Koury, M. J., Capdevila, A., Rana, S. & Wagner, C. (2001) Plasma S-adenosylhomocysteine is a more sensitive indicator of cardiovascular disease than plasma homocysteine. Am. J. Clin. Nutr. 74: 723-729.

52. Klerk, M., Verhoef, P., Clarke, R., Blom, H. J., Kok, F. J. & Schouten, E. G. (2002) MTHFR 677C-->T polymorphism and risk of coronary heart disease: a meta-analysis. JAMA 288: 2023-2031.

53. Koushik, A., Kraft, P., Fuchs, C. S., Hankinson, S. E., Willett, W. C., Giovannucci, E. L. & Hunter, D. J. (2006) Nonsynonymous polymorphisms in genes in the one-carbon metabolism pathway and associations with colorectal cancer. Cancer Epidemiol. Biomarkers Prev. 15: 2408-2417.

54. Kraft, P. & Hunter, D. J. (2009) Genetic risk prediction--are we there yet? N Engl J Med 360: 1701-1703.

55. Lai, C. Q., Parnell, L. D., Troen, A. M., Shen, J., Caouette, H., Warodomwichit, D., Lee, Y. C., Crott, J. W., Qiu, W. Q. et al. (2010) MAT1A variants are associated with hypertension, stroke, and markers of DNA damage

and are modulated by plasma vitamin B-6 and folate. Am J Clin Nutr 91: 1377-1386.

56. Lamers, Y., Williamson, J., Gilbert, L. R., Stacpoole, P. W. & Gregory, J. F., III (2007) Glycine turnover and decarboxylation rate quantified in healthy men and women using primed, constant infusions of [1,2-(13)C2]glycine and [(2)H3]leucine. The Journal of nutrition 137: 2647-2652.

57. Lee, K. M., Lan, Q., Kricker, A., Purdue, M. P., Grulich, A. E., Vajdic, C. M., Turner, J., Whitby, D., Kang, D. et al. (2007) One-carbon metabolism gene polymorphisms and risk of non-Hodgkin lymphoma in Australia. Hum. Genet. 122: 525-533.

58. Lewis, S. J., Ebrahim, S. & Davey Smith, G. (2005) Meta-analysis of MTHFR 677C->T polymorphism and coronary heart disease: does totality of evidence support causal role for homocysteine and preventive potential of folate? Br. Med. J. 331: 1053.

59. Li, F., Feng, Q., Lee, C., Wang, S., Pelleymounter, L. L., Moon, I., Eckloff, B. W., Wieben, E. D., Schaid, D. J. et al. (2008) Human betaine-homocysteine methyltransferase (BHMT) and BHMT2: common gene sequence variation and functional characterization. Mol. Genet Metab 94: 326-335.

60. Lichtenstein, A. H., Appel, L. J., Brands, M., Carnethon, M., Daniels, S., Franch, H. A., Franklin, B., Kris-Etherton, P., Harris, W. S. et al. (2006) Summary of American Heart Association Diet and Lifestyle Recommendations revision 2006. Arterioscler. Thromb. Vasc. Biol 26: 2186-2191.

61. Lightfoot, T. J., Skibola, C. F., Willett, E. V., Skibola, D. R., Allan, J. M., Coppede, F., Adamson, P. J., Morgan, G. J., Roman, E. & Smith, M. T. (2005) Risk of non-Hodgkin lymphoma associated with polymorphisms in folate-metabolizing genes. Cancer Epidemiol Biomarkers Prev 14: 2999-3003.

62. Lim, U., Peng, K., Shane, B., Stover, P. J., Litonjua, A. A., Weiss, S. T., Gaziano, J. M., Strawderman, R. L., Raiszadeh, F. et al. (2005) Polymorphisms in cytoplasmic serine hydroxymethyltransferase and methylenetetrahydrofolate reductase affect the risk of cardiovascular disease in men. J. Nutr. 135: 1989-1994.

63. Lim, U., Wang, S. S., Hartge, P., Cozen, W., Kelemen, L. E., Chanock, S., Davis, S., Blair, A., Schenk, M. et al. (2007) Gene-nutrient interactions among determinants of folate and one-carbon metabolism on the risk of non-Hodgkin lymphoma: NCI-SEER case-control study. Blood 109: 3050-3059.

64. Lissowska, J., Gaudet, M. M., Brinton, L. A., Chanock, S. J., Peplonska, B., Welch, R., Zatonski, W., Szeszenia-Dabrowska, N., Park, S. et al. (2007) Genetic polymorphisms in the one-carbon metabolism pathway and breast cancer risk: a population-based case-control study and meta-analyses. Int. J. Cancer 120: 2696-2703.

65. Lonn, E., Yusuf, S., Arnold, M. J., Sheridan, P., Pogue, J., Micks, M., McQueen, M. J., Probstfield, J., Fodor, G. et al. (2006) Homocysteine lowering with folic acid and B vitamins in vascular disease. N. Engl. J. Med. 354: 1567-1577.

66. MacFarlane, A. J., Liu, X., Perry, C. A., Flodby, P., Allen, R. H., Stabler, S. P. & Stover, P. J. (2008) Cytoplasmic serine hydroxymethyltransferase regulates the metabolic partitioning of methylenetetrahydrofolate but is not essential in mice. J. Biol. Chem. 283: 25846-25853.

67. Maniatis, N., Collins, A., Xu, C. F., McCarthy, L. C., Hewett, D. R., Tapper, W., Ennis, S., Ke, X. & Morton, N. E. (2002) The first linkage disequilibrium (LD) maps: delineation of hot and cold blocks by diplotype analysis. Proc. Natl. Acad. Sci. U. S. A 99: 2228-2233.

68. Manolio, T. A. (2010) Genomewide association studies and assessment of the risk of disease. N Engl J Med 363: 166-176.

69. Manolio, T. A., Collins, F. S., Cox, N. J., Goldstein, D. B., Hindorff, L. A., Hunter, D. J., McCarthy, M. I., Ramos, E. M., Cardon, L. R. et al. (2009) Finding the missing heritability of complex diseases. Nature 461: 747-753.

70. Marti-Carvajal, A. J., Sola, I., Lathyris, D. & Salanti, G. (2009) Homocysteine lowering interventions for preventing cardiovascular events. Cochrane. Database. Syst. Rev CD006612.

71. Miller, E. R., III, Juraschek, S., Pastor-Barriuso, R., Bazzano, L. A., Appel, L. J. & Guallar, E. (2010) Meta-analysis of folic acid supplementation trials on

risk of cardiovascular disease and risk interaction with baseline homocysteine levels. Am J Cardiol. 106: 517-527.

72. Milner, J. Introductory Comments to the Scientific Sessions. Yaktine, Ann and Pool, Robert. 2007. The National Academies Press. Nutrigenomics and Beyond: Informing the Future Workshop Summary.
Ref Type: Report

73. Moore, L. E., Malats, N., Rothman, N., Real, F. X., Kogevinas, M., Karami, S., Garcia-Closas, R., Silverman, D., Chanock, S. et al. (2007) Polymorphisms in one-carbon metabolism and trans-sulfuration pathway genes and susceptibility to bladder cancer. Int. J. Cancer 120: 2452-2458.

74. National Center for Biotechnology Information. PubMed search engine. 2010.
Ref Type: Data File

75. Niclot, S., Pruvot, Q., Besson, C., Savoy, D., Macintyre, E., Salles, G., Brousse, N., Varet, B., Landais, P. et al. (2006) Implication of the folate-methionine metabolism pathways in susceptibility to follicular lymphomas. Blood 108: 278-285.

76. Nilsson, S. E., Read, S., Berg, S. & Johansson, B. (2009) Heritabilities for fifteen routine biochemical values: findings in 215 Swedish twin pairs 82 years of age or older. Scand. J Clin Lab Invest 69: 562-569.

77. Pai, J. K., Kraft, P., Cannuscio, C. C., Manson, J. E., Rexrode, K. M., Albert, C. M., Hunter, D. & Rimm, E. B. (2006) Polymorphisms in the CC-chemokine receptor-2 (CCR2) and -5 (CCR5) genes and risk of coronary heart disease among US women. Atherosclerosis 186: 132-139.

78. Patino-Garcia, A., Zalacain, M., Marrodan, L., San-Julian, M. & Sierrasesumaga, L. (2009) Methotrexate in pediatric osteosarcoma: response and toxicity in relation to genetic polymorphisms and dihydrofolate reductase and reduced folate carrier 1 expression. J. Pediatr. 154: 688-693.

79. Perry, C., Yu, S., Chen, J., Matharu, K. S. & Stover, P. J. (2007) Effect of vitamin B6 availability on serine hydroxymethyltransferase in MCF-7 cells. Arch. Biochem. Biophys. 462: 21-27.

80. Pfeiffer, C. M., Johnson, C. L., Jain, R. B., Yetley, E. A., Picciano, M. F., Rader, J. I., Fisher, K. D., Mulinare, J. & Osterloh, J. D. (2007) Trends in blood folate and vitamin B-12 concentrations in the United States, 1988 2004. Am J Clin Nutr 86: 718-727.

81. Pischon, T., Girman, C. J., Hotamisligil, G. S., Rifai, N., Hu, F. B. & Rimm, E. B. (2004) Plasma adiponectin levels and risk of myocardial infarction in men. JAMA 291: 1730-1737.

82. Pogribny, I. P. & Beland, F. A. (2009) DNA hypomethylation in the origin and pathogenesis of human diseases. Cell Mol. Life Sci. 66: 2249-2261.

83. Raiszadeh, F. Association of MTHFD1 1958G?A and MTHFR 677C?T polymorphisms with serum homocysteine level and cardiovascular disease: The Normative Aging Study.
Ref Type: Unpublished Work

84. Refsum, H., Smith, A. D., Ueland, P. M., Nexo, E., Clarke, R., McPartlin, J., Johnston, C., Engbaek, F., Schneede, J. et al. (2004) Facts and recommendations about total homocysteine determinations: an expert opinion. Clin. Chem. 50: 3-32.

85. Robertson, K. D., Uzvolgyi, E., Liang, G., Talmadge, C., Sumegi, J., Gonzales, F. A. & Jones, P. A. (1999) The human DNA methyltransferases (DNMTs) 1, 3a and 3b: coordinate mRNA expression in normal tissues and overexpression in tumors. Nucleic acids research 27: 2291-2298.

86. Rosamond, W., Flegal, K., Friday, G., Furie, K., Go, A., Greenlund, K., Haase, N., Ho, M., Howard, V. et al. (2007) Heart disease and stroke statistics--2007 update: a report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Circulation 115: e69-171.

87. Selhub, J. (2006) The many facets of hyperhomocysteinemia: studies from the Framingham cohorts. The Journal of nutrition 136: 1726S-1730S.

88. Semmler, A., Linnebank, M., Krex, D., Gotz, A., Moskau, S., Ziegler, A. & Simon, M. (2008) Polymorphisms of homocysteine metabolism are associated with intracranial aneurysms. Cerebrovasc. Dis. 26: 425-429.

89. Sesso, H. D., Kawachi, I., Vokonas, P. S. & Sparrow, D. (1998) Depression and the risk of coronary heart disease in the Normative Aging Study. Am J Cardiol. 82: 851-856.

90. Shai, I., Stampfer, M. J., Ma, J., Manson, J. E., Hankinson, S. E., Cannuscio, C., Selhub, J., Curhan, G. & Rimm, E. B. (2004) Homocysteine as a risk factor for coronary heart diseases and its association with inflammatory biomarkers, lipids and dietary factors. Atherosclerosis 177: 375-381.

91. Shamay, M., Greenway, M., Liao, G., Ambinder, R. F. & Hayward, S. D. (2010) De-novo DNA methyltransferase DNMT3b interacts with NEDD8 modified proteins. J. Biol. Chem.

92. Siva, A., De Lange, M., Clayton, D., Monteith, S., Spector, T. & Brown, M. J. (2007) The heritability of plasma homocysteine, and the influence of genetic variation in the homocysteine methylation pathway. QJM. 100: 495-499.

93. Skibola, C. F., Forrest, M. S., Coppede, F., Agana, L., Hubbard, A., Smith, M. T., Bracci, P. M. & Holly, E. A. (2004) Polymorphisms and haplotypes in folate-metabolizing genes and risk of non-Hodgkin lymphoma. Blood 104: 2155-2162.

94. Skibola, C. F., Smith, M. T., Hubbard, A., Shane, B., Roberts, A. C., Law, G. R., Rollinson, S., Roman, E., Cartwright, R. A. & Morgan, G. J. (2002) Polymorphisms in the thymidylate synthase and serine hydroxymethyltransferase genes and risk of adult acute lymphocytic leukemia. Blood 99: 3786-3791.

95. Stanislawska-Sachadyn, A., Mitchell, L. E., Woodside, J. V., Buckley, P. T., Kealey, C., Young, I. S., Scott, J. M., Murray, L., Boreham, C. A. et al. (2009) The reduced folate carrier (SLC19A1) c.80G>A polymorphism is associated with red cell folate concentrations among women. Ann. Hum Genet 73: 484-491.

96. Stipanuk, M. H. (2000) Biochemical and physiological aspects of human nutrition. W.B. Saunders Company, Philadelphia.

97. Stover, P. J. (2004) Physiology of folate and vitamin B12 in health and disease. Nutr. Rev. 62: S3-12.

98. Stover, P. J. (2006) Influence of human genetic variation on nutritional requirements. Am. J. Clin. Nutr. 83: 436S-442S.

99. Taoka, S., West, M. & Banerjee, R. (1999) Characterization of the heme and pyridoxal phosphate cofactors of human cystathionine beta-synthase reveals nonequivalent active sites. Biochemistry 38: 7406.

100. The Homocysteine Studies Collaboration (2002) Homocysteine and risk of ischemic heart disease and stroke: a meta-analysis. JAMA 288: 2015-2022.

101. Toole, J. F., Malinow, M. R., Chambless, L. E., Spence, J. D., Pettigrew, L. C., Howard, V. J., Sides, E. G., Wang, C. H. & Stampfer, M. (2004) Lowering homocysteine in patients with ischemic stroke to prevent recurrent stroke, myocardial infarction, and death: the Vitamin Intervention for Stroke Prevention (VISP) randomized controlled trial. JAMA 291: 565-575.

102. Tucker, K. L., Qiao, N., Scott, T., Rosenberg, I. & Spiro, A., III (2005) High homocysteine and low B vitamins predict cognitive decline in aging men: the Veterans Affairs Normative Aging Study. Am. J. Clin. Nutr. 82: 627-635.

103. Turek-Plewa, J. & Jagodzinski, P. P. (2005) The role of mammalian DNA methyltransferases in the regulation of gene expression. Cell Mol. Biol Lett 10: 631-647.

104. Turunen, M. P., Aavik, E. & Yla-Herttuala, S. (2009) Epigenetics and atherosclerosis. Biochim. Biophys. Acta 1790: 886-891.

105. Ulrey, C. L., Liu, L., Andrews, L. G. & Tollefsbol, T. O. (2005) The impact of metabolism on DNA methylation. Hum Mol. Genet 14 Spec No 1: R139-R147.

106. van den Donk, M., Visker, M. H., Harryvan, J. L., Kok, F. J. & Kampman, E. (2007) Dietary intake of B-vitamins, polymorphisms in thymidylate synthase and serine hydroxymethyltransferase 1, and colorectal adenoma risk: a Dutch case-control study. Cancer Lett. 250: 146-153.

107. Wagner, C. & Koury, M. J. (2007) S-Adenosylhomocysteine: a better indicator of vascular disease than homocysteine? Am. J. Clin. Nutr. 86: 1581-1585.

108. Wald, D. S., Law, M. & Morris, J. K. (2002) Homocysteine and cardiovascular disease: evidence on causality from a meta-analysis. Br. Med. J. 325: 1202.

109. Wald, D. S., Wald, N. J., Morris, J. K. & Law, M. (2006) Folic acid, homocysteine, and cardiovascular disease: judging causality in the face of inconclusive trial evidence. Br. Med. J. 333: 1114-1117.

110. Weisenberger, D. J., Velicescu, M., Cheng, J. C., Gonzales, F. A., Liang, G. & Jones, P. A. (2004) Role of the DNA methyltransferase variant DNMT3b3 in DNA methylation. Mol. Cancer Res 2: 62-72.

111. Weisskopf, M. G., Jain, N., Nie, H., Sparrow, D., Vokonas, P., Schwartz, J. & Hu, H. (2009) A prospective study of bone lead concentration and death from all causes, cardiovascular diseases, and cancer in the Department of Veterans Affairs Normative Aging Study. Circulation 120: 1056-1064.

112. Wellcome Trust Case Control Consortium (2007) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447: 661-678.

113. Wernimont, S. M., Clark, A. G., Stover, P. J., Wells, M. T., Litonjua, A. A., Weiss, S. T., Gaziano, J. M. & Cassano, P. A. (2010) Variation in the methylenetetrahydrofolate dehydrogenase 1-like (MTHFD1L) gene and cardiovascular disease (CVD) risk. FASEB Journal 23 (abs.).

114. Wernimont, S. M., Stover, P. J., Wells, M. T. & Cassano, P. A. (2009) A systematic overview and meta-analysis of genetic variation within the folate metabolic network and cardiovascular disease (CVD) risk. FASEB Journal 23: 724.15 (abs.).

115. Wilker, E. H., Alexeeff, S. E., Poon, A., Litonjua, A. A., Sparrow, D., Vokonas, P. S., Mittleman, M. A. & Schwartz, J. (2009) Candidate genes for respiratory disease associated with markers of inflammation and endothelial dysfunction in elderly men. Atherosclerosis 206: 480-485.

116. Woeller, C. F., Anderson, D. D., Szebenyi, D. M. & Stover, P. J. (2007) Evidence for small ubiquitin-like modifier-dependent nuclear import of the thymidylate biosynthesis pathway. J. Biol. Chem. 282: 17623-17631.

117. Woeller, C. F., Fox, J. T., Perry, C. & Stover, P. J. (2007) A ferritin-responsive internal ribosome entry site regulates folate metabolism. J Biol Chem 282: 29927-29935.

118. Xu, X., Gammon, M. D., Zhang, H., Wetmur, J. G., Rao, M., Teitelbaum, S. L., Britton, J. A., Neugut, A. I., Santella, R. M. & Chen, J. (2007) Polymorphisms of one-carbon-metabolizing genes and risk of breast cancer in a population-based study. Carcinogenesis 28: 1504-1509.

119. Yan, J., Feng, J., Hosono, S. & Sommer, S. S. (2004) Assessment of multiple displacement amplification in molecular epidemiology. Biotechniques 37: 136-3.

120. Yang, A. S., Estecio, M. R., Doshi, K., Kondo, Y., Tajara, E. H. & Issa, J. P. (2004) A simple method for estimating global DNA methylation using bisulfite PCR of repetitive DNA elements. Nucleic acids research 32: e38.

121. Zhang, Z., Shi, Q., Sturgis, E. M., Spitz, M. R. & Wei, Q. (2005) Polymorphisms and haplotypes of serine hydroxymethyltransferase and risk of squamous cell carcinoma of the head and neck: a case-control analysis. Pharmacogenet. Genomics 15: 557-564.