NIH Public Access
Author Manuscript
Circ Cardiovasc Genet. Author manuscript; available in PMC 2010 January 1.

Published in final edited form as:
Circ Cardiovasc Genet. 2009 January 1; 2(2): 142–150. doi:10.1161/CIRCGENETICS.108.829804.

Novel Associations of CPS1, MUT, NOX4 and DPEP1 with Plasma Homocysteine in a Healthy Population: A Genome Wide Evaluation of 13,974 Participants in the Women's Genome Health Study

Guillaumé Pare, MD, MSc, FRCP(C)[a,*,b], **Daniel I. Chasman, PhD**[a,*,b], **Alexander N. Parker, PhD**[c], **Robert R.Y. Zee, PhD, MPH**[a,b], **Anders Mälarstig, PhD**[d], **Udo Seedorf, PhD**[e], **Rory Collins, MBBS, MSc, FRCP**[f], **Hugh Watkins, MD, PhD, FRCP**[g], **Anders Hamsten, MD, PhD, FRCP**[d], **Joseph P. Miletich, MD**[c], and **Paul M. Ridker, MD, MPH**[a,b]

[a] Center for Cardiovascular Disease Prevention, Harvard Medical School, Boston, MA, 02215, USA

[b] Donald W Reynolds Center for Cardiovascular Research, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, 02215, USA

[c] Amgen, Inc., Cambridge, MA, 02139, USA

[d] Atherosclerosis Research Unit, Department of Medicine Solna, Karolinska Institutet, Stockholm, Sweden

[e] Leibniz-Institut für Arteriosklerosеforschung an der Universität Münster, Münster, Germany

[f] Clinical Trial Service Unit and Epidemiological Studies Unit, University of Oxford, UK

[g] Department of Cardiovascular Medicine, University of Oxford, UK and the PROCARDIS Consortium, www.procardis.org

Abstract

Background—Homocysteine is a sulfur amino acid whose plasma concentration has been associated with the risk of cardiovascular diseases, neural tube defects and loss of cognitive function in epidemiological studies. While genetic variants of *MTHFR* and *CBS* are known to influence homocysteine concentration, common genetic determinants of homocysteine remain largely unknown.

Methods and Results—To address this issue comprehensively, we performed a genome wide association analysis, testing 336,469 SNPs in 13,974 healthy Caucasian women. While we confirm association with *MTHFR* (1p36.22; rs1801133; $p=8.1 \times 10^{-35}$) and *CBS* (21q22.3; rs6586282; $p=3.2 \times 10^{-10}$), we found novel associations with *CPS1* (2q34; rs7422339; $p=1.9 \times 10^{-11}$), *MUT* (6p12.3; rs4267943; $p=2.0 \times 10^{-9}$), *NOX4* (11q14.3; rs11018628; $p=9.6 \times 10^{-12}$) and *DPEP1* (16q24.3; rs1126464; $p=1.2 \times 10^{-12}$). The associations at *MTHFR*, *DPEP1* and *CBS* were replicated in an independent sample from the PROCARDIS study, whereas the association at *CPS1* was only replicated among the women.

Correspondence to Guillaume Pare, MD, Center for Cardiovascular Disease Prevention, Brigham and Women's Hospital, 900 Commonwealth Avenue East, 3rd Floor, Boston, MA 02215. Tel 617 525-6857 FAX 617 734-1508 gpare@rics.bwh.harvard.edu.
*These authors contributed equally

DISCLOSURES
None.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Conclusions—These associations offer new insights into the biochemical pathways involved in homocysteine metabolism and provide opportunities to better delineate the role of homocysteine in health and disease.

Keywords

Genetics; Metabolism; Amino Acids

INTRODUCTION

Homocysteine is a non-protein-forming sulfur amino acid produced during the catabolism of methionine. Its concentration is tightly regulated and kept at low levels through catabolism by either remethylation or transsulfuration. The small amount of homocysteine found in plasma is the result of a cellular export mechanism that complements the remethylation and transsulfuration pathways in maintaining low intra-cellular concentration of this potentially cytotoxic and pro-oxidant amino acid[1–3]. Plasma homocysteine levels are influenced by genetic as well as environmental factors, such as age, sex, smoking status, intake of folate and intake of B vitamins. Homocysteine concentration has been epidemiologically correlated with the risk of cardiovascular diseases[4, 5], neural tube defects [6] and loss of cognitive functions [7].

Despite heritability estimates ranging from 25% to 44% [8, 9], relatively little is known of the genetic determinants of homocysteine levels. Linkage scans have revealed linkage signals at 11q23, 12q14, 13q31 and 16q12, although these remain un-validated to date [8, 9]. Rare homozygous defects in genes encoding for enzymes of homocysteine metabolism (i.e. *CBS* and *MTHFR*) lead to dramatically increased homocysteine concentration and premature occlusive vascular disease [1]. However, few common polymorphisms have been unequivocally associated with homocysteine concentration. Among these, the strongest is the *MTHFR* SNP rs1801133 (C677T) correlated with reduced enzymatic activity and higher homocysteine levels [10]. *MTHFR* encodes the enzyme methylenetetrahydrofolate reductase, which catalyses the conversion of 5,10-methylenetetrahydrofolate to 5-methyltetrahydrofolate, a cosubstrate for homocysteine remethylation to methionine by methionine synthase. Along with *MTHFR*, genetic variations at loci such as *CBS*[11, 12] (cystathionine beta-synthase), *MTR* 13 (5-methyltetrahydrofolate-homocysteine methyltransferase) and *MTRR* 14 (5-methyltetrahydrofolate-homocysteine methyltransferase reductase) have also been reported to influence homocysteine concentration. In an effort to identify new genetic variants influencing homocysteine concentration, we performed a genome wide association analysis, testing 336,469 SNPs in 13,974 healthy women. Because homocysteine is a direct intermediary of metabolic pathways, this offers the possibility of improving our knowledge not only of homocysteine metabolism but also of the genetic architecture of metabolic traits in general. Furthermore, finding novel genetic associations with homocysteine concentration will provide opportunities to better delineate the role of homocysteine in health and disease.

METHODS

WGHS Study Participants

The study population derived from the Women's Genome Health Study (WGHS)[15]. Briefly, participants in the WGHS include American women from the Women's Health Study (WHS) with no prior history of cardiovascular disease, cancer, or other major chronic illness who also provided a baseline blood sample at the time of study enrollment from which genomic DNA was extracted[16]. Individuals with prevalent diabetes were excluded from analysis. The study was approved by the institutional review board of the Brigham and Women's Hospital.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

EDTA blood samples were obtained at the time of enrollment and stored in vapor phase liquid nitrogen (−170 °C). The concentration of homocysteine was determined using an enzymatic assay on the Hitachi 917 analyzer (Roche Diagnostics) using reagents and calibrators from Catch, Inc[17]. The sensitivity of the homocysteine assay is 0.24 μmol/L and its coefficient of variation is 1.8% at 15.6 μmol/L. Participants also completed a 131-item semiquantitative food frequency questionnaire at baseline [18] from which intake of folate, vitamin B_6 and vitamin B_{12} was evaluated. Nutrient intake assessments made by use of the food frequency questionnaire used have been previously shown to be valid and reliable[19, 20], with correlation coefficients of 0.49–0.55 for estimated folate and B-vitamin intake with this questionnaire method compared with measured plasma concentrations in apparently healthy female professionals similar to our study participants[21, 22].

Genotyping

DNA samples were genotyped with the Infinium II technology from Illumina. Either the HumanHap300 Duo-Plus chip or the combination of the HumanHap300 Duo and I-Select chips was used. In either case, the custom content was identical and consisted of candidate SNPs chosen without regard to allele frequency to increase coverage of genetic variation with impact on biological function including metabolism, inflammation or cardiovascular diseases. Specifically, all UCSC coding non-synonymous SNPs, UCSC splice-site SNPs, UCSC SNPs with phenotype annotation, and SNPs from targeted cardiovascular diseases, diabetes, venous thromboembolism, blood pressure and inflammation candidate genes were selected. Furthermore, all SNPs with functional effects according to Online Mendelian Inheritance in Man (OMIM) were included, irrespective of genic region. Both OMIM and UCSC databases were accessed in December 2006. Genotyping at 318,237 HumanHap300 Duo SNPs and 45,571 custom content SNPs was attempted, for a total of 363,808 SNPs. Genetic context for all annotations are derived from human genome build 36.1 and dbSNP build 126.

SNPs with call rates <90% were excluded from further analysis. Likewise, all samples with percentage of missing genotypes higher than 2% were removed. Among retained samples, SNPs were further evaluated for deviation from Hardy-Weinberg equilibrium using an exact method[23] and were excluded when the P-value was lower than 10^{-5}. Samples were further validated by comparison of genotypes at 44 SNPs that had been previously ascertained using alternative technologies. SNPs with minor allele frequency >1% in Caucasians were used for analysis. After quality control, 307,805 HumanHap300 Duo SNPs and 28,664 custom content SNPs were left, for a total of 336,469 SNPs with a mean call rate of 99.6%.

Population Stratification

Because population stratification can result in inflated type I error in genome wide association analysis, a principal component analysis using 1443 ancestry informative SNPs was performed using PLINK[24] in order to confirm self-reported ancestry. Briefly, these SNPs were chosen based on Fst > 0.4 in HapMap populations (YRB, CEU, CHB+JPT) and inter-SNP distance at least 500 kb in order to minimize linkage disequilibrium. Different ethnic groups were clearly distinguished with the two first components. Based on this analysis, 55 participants were excluded from further evaluation as they did not cluster with other Caucasians, leaving 13,974 for the current study population. Two additional steps were taken to rule out the possibility that residual stratification within Caucasians was responsible for the associations observed. First, association analysis was done with correction by genomic control. Second, a principal component analysis[25] was performed in previously identified Caucasians (only) using 64,221 SNPs chosen to have pair-wise linkage disequilibrium lower than r^2=0.2. The first ten components were then used as covariates in the association analysis. As adjustment by these covariates did not change the conclusions, we present analysis among Caucasian participants without further correction for sub-Caucasian ancestry.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Association Analysis

To identify common genetic variants influencing homocysteine levels, we first attempted to discover which loci significantly contributed to homocysteine concentration. Log-transformed plasma concentration of homocysteine was adjusted for age, body mass index, smoking status, hormone replacement therapy use, folate intake, vitamin B_6 intake and vitamin B_{12} intake using a linear regression model in R. These covariates were chosen as they are potential environmental confounders of homocysteine levels, based on available literature[26–29]. This was done to reduce the impact of clinical covariates on homocysteine variance. The adjusted homocysteine values were then tested for association with SNP genotypes by linear regression in PLINK[24], assuming an additive contribution of each minor allele. Beta coefficients of natural log-transformed adjusted homocysteine values are given using the major allele as the reference allele (major and minor alleles are presented in Supplementary Table 1). A conservative P-value cut-off of 5×10^{-8} was used to correct for the maximum of 1,000,000 independent statistical tests thought to correspond to all the common genetic variation of the human genome [30].

Once any locus with genome wide significance was identified, a forward selection linear multiple regression model was used to further define the extent of the genetic association. Briefly, all genotyped SNPs within 100 kb of a SNP with genome wide association (i.e. $P < 5\times10^{-8}$) and passing quality control requirements were tested for possible incorporation into a multiple regression model. In stepwise fashion, a SNP was added to the model if it had the smallest P-value among all the SNPs not yet included in the model and if it had a P-value was lower than 5×10^{-8}. SNPs selected by this algorithm were also used in haplotype analysis using WHAP [24], as implemented in PLINK [24].

PROCARDIS Validation Sample

A subset of 840 postinfarction patients participating in the PROCARDIS study for whom homocysteine data were available was used for replication of findings in the WGHS. Detailed descriptions of the PROCARDIS population can be found elsewhere[31, 32]. In brief, cases included in the present report had a diagnosis of myocardial infarction at or before the age of 65 years and reported having European ancestry. Recruitmentwas carried out in Germany, Italy, Sweden and the UK. The PROCARDIS protocol was approved by the Ethics Committees of the participating institutions and all subjects gave written informed consent.

DNA was extracted and quantified as described[31]. Genotyping of a genome-wide set of 1,054,559 SNPs was performed using the Illumina Infinium II Human 1M bead chip at two national centres (Centre National de Génotypage, Paris, France and the SNP Technology Platform, Uppsala, Sweden). Subjects with % genotype calls <95% were excluded from further analysis as were male samples with significant X-chromosomal heterozygosity and samples with a possible population bias, as indicated from the identity by state cluster analysis. Also, when relatedness between subjects was suspected (up to 2nd degree relationships), one individual from the pair was removed.

EDTA Blood samples were obtained in the fasting state using EDTA tubes supplemented with 4 g/l sodium fluoride as an inhibitor of erythrocyte metabolism. Total homocysteine (free and protein bound) concentration was determined immediately after thawing frozen plasma samples which had been stored at − 80 °C by fluorescence high-performance liquid chromatography (HPLC) using reagents and calibrators from Chromsystems (Munich, Germany). Isocratic reversed-phase chromatography was performed on a Kontron (Neufahrn, Germany) liquid chromatograph interfaced with a model RF-535 fluorescence detector (Shimadzu, Kyoto, Japan) by using excitation and emission wavelengths of 385 and 515 nm, respectively. The coefficients of variation within and between days for the assay were ≤ 4.8%.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Association analysis in the validation sample was performed on log-transformed plasma concentration of homocysteine, using a linear regression model that adjusted for age, sex, body mass index, smoking status and country of origin. Other linear models used included either of the MTHFR rs1801133 and MTHFR rs17350396 SNPs or both the rs1801133 and rs17350396 SNPs. For the analysis of sex-specific SNP-homocysteine associations, a general linear model was used comparing the main effects of SNPs on homocysteine adjusted for age, body mass index, smoking status and country of origin within gender subgroups. Statistical analyses were performed in SPSS 16.0 for Windows®.

Statement of Responsibility

The authors had full access to the data and take responsibility for its integrity. All authors have read and agree to the manuscript as written.

RESULTS

Clinical characteristics of the 13,974 healthy women in the WGHS sample are provided in Table 1. Results of the genome wide association analysis of log-transformed adjusted homocysteine concentration are presented in Table 2. Twenty-nine SNPs at six different loci – *MTHFR*, *CPS1*, *MUT*, *NOX4*, *DPEP1* and *CBS* - had association P-values lower than our previously defined genome wide significance threshold of 5×10^{-8}: 15 SNPs at the *MTHFR* locus; 1 SNP at the *CPS1* locus; 1 SNP at the *MUT* locus; 6 SNPs at the *NOX4* locus; 3 SNPs at the *DPEP1* locus; 1 SNP at the *CBS* locus. Genetic context of all these loci is presented in Figure 1 along with the $-\log_{10}$ transformed P-values. In addition, 8 SNPs had association P-values just above this threshold level (between 10^{-6} and 5×10^{-8}), including the *AKAP13* SNP rs2061821 on chromosome 15q25.3 (the remaining 7 SNPs were in the vicinity of one of the six previously identified loci). Among these SNPs, only rs4267943 (*MUT*; 6p12.3) deviated from Hardy-Weinberg equilibrium (p=0.00006), but visual inspection of the raw genotyping signal for this SNP did not reveal any obvious artifact. Furthermore, other SNPs at this locus (rs2501976, rs6458690 and rs9473558; see Table 2) did not deviate from Hardy-Weinberg equilibrium while being significantly associated with homocysteine level (P-values ranging from 2.1×10^{-7} to 5.2×10^{-8}) and in linkage disequilibrium with rs4267943 ($r^2 > 0.95$ for all three pairwise comparisons).

To further define the extent of genetic associations at the 6 loci with association P-value lower than 5×10^{-8}, we applied a forward model selection algorithm to each of them in order to identify SNPs non-redundantly associated with homocysteine. Briefly, 97 SNPs at *MTHFR*, 26 at *CPS1*, 20 at *MUT*, 32 at *NOX4*, 35 at *DPEP1* and 34 at *CBS* were initially assessed for possible inclusion in a multiple linear regression model. Using a P-value cut-off of 5×10^{-8}, 7 SNPs were selected, representing the lead SNP at each of the 6 loci considered plus one SNP (rs17350396) at the *MTHFR* locus (see Table 3). Interestingly, this later SNP (rs17350396; *MTHFR*) was marginally significant in univariable analysis (p=0.008), illustrating that its inclusion in the model and significant association are conditional on the genotypes at rs1801133 (*MTHFR*). Multiple regression beta coefficients and P-values of the 7 SNPs selected are shown in Table 3. Illustrated in Figure 2 are the quantile-quantile plots of association P-values before and after adjusting homocysteine concentration for the combined effect of these 7 SNPs. Among these SNPs, only rs1801133 (*MTHFR*) showed evidence (p=0.002) for non-additive effects of the minor allele as judged by a likelihood ratio test comparing the additive regression model to an alternative genotype model with an additional degree of freedom. Specifically, the association tended toward a recessive genetic model, with mean log-transformed adjusted homocysteine values of − 0.023 (N=6246) for major allele homozygotes, 0.011 (N=6109) for heterozygotes and 0.087 (N=1619) for minor allele homozygotes. However, according to the Bayes Information Criteria, the additive model still provided the best fit to the data as compared

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

to purely recessive or dominant models. Importantly, use of a 2 degrees of freedom model did not change the result of the model selection algorithm and the genetic effect is therefore assumed to be additive for the remainder of this report. No gene-gene interaction was observed between any of the model selected SNPs.

The 2 SNPs at the *MTHFR* locus selected by our algorithm were also used in haplotype analysis (Table 4). The estimate of the proportion of variance attributable to haplotypes, as well as their regression coefficients, is consistent with the linear model of these same SNPs, reinforcing the adequacy of a multiple regression model to explain the association (as compared to the haplotype analysis). Linkage disequilibrium between these two *MTHFR* SNPs was 0.097 for r^2 and 0.978 for D'. The 2 SNPs at *MTHFR* collectively explained 1.3% of the total variance in homocysteine concentration, whereas the *CPS1* SNP (rs7422339) explained 0.3%, the *MUT* SNP (rs4267943) 0.3%, the *NOX4* SNP (rs11018622) 0.3%, the *DPEP1* SNP (rs1126464) 0.3% and the *CBS* SNP (rs6586282) 0.2%. In comparison, clinical covariates accounted for 8.7% of the variance (Table 5), and together the candidate loci and the clinical variables accounted for 11.3% of total variance.

Replication of the model selected SNPs was attempted in the PROCARDIS sample (see Table 1 for clinical characteristics of PROCARDIS participants). First, the two *MTHFR* SNPs, rs1801133 and rs17350396, were included in a multiple regression linear model that used homocysteine concentration as the dependant variable. Both SNPs were significantly associated with 2-sided P-values of 0.0001 (Beta=0.15) and 0.02 (Beta=0.09), respectively (and consistent direction of effect). Of note, when each SNP was included separately in the linear model, only rs1801133 was associated (P=0.0001) whereas rs17350396 was non-significant (P=0.433). Because the *DPEP1* SNP rs1126464 did not pass genotyping quality control in the PROCARDIS sample, we tested for association with rs460879, the second most significant SNP at the *DPEP1* locus in WGHS and the SNP with the strongest linkage disequilibrium based on D' ($P=3.8 \times 10^{-12}$; D'=0.99 and r^2=0.25 between rs1126464 and rs460879). The association P-value was 0.001 (Beta=0.12), with consistent direction of effect. The *CBS* SNP rs6586282 was also replicated with a P-value of 0.02 (Beta=− 0.07), again with consistent direction of effect. Finally, while the *CPS1* SNP rs7422339 was non-significant when men (N=658) and women (N=161) were considered together, the association was significant in the women considered on their own (P=0.009; Beta=0.2; direction of effect consistent) with a sex-SNP interaction P-value of 0.0003. No other sex interaction was noted. Associations with the *MUT* SNP rs4267943 and *NOX4* SNP rs11018628 were non-significant (P>0.05).

DISCUSSION

Four loci – *MTHFR*, *CBS*, *DPEP1* and *CPS1* – have been identified and confirmed in this report for association with homocysteine levels, with the *CPS1* association being sex-specific. While genetic variants of *MTHFR* and *CBS* are known to influence homocysteine metabolism, the other associations are novel. *DPEP1* (dipeptidase 1) is a kidney membrane enzyme that is highly expressed in the proximal convoluted tubules [33]. It hydrolyzes a variety of dipeptides and is implicated in renal metabolism of glutathione and its conjugates, such as leukotrienes [34]. Because DPEP1 deficiency leads to increased urinary excretion of cysteine [35, 36], a precursor of homocysteine, we hypothesize that the association between the *DPEP1* coding non-synonymous (E351Q) SNP rs1126464 and homocysteine concentration could be the result of changes in the renal handling of amino acids. While *DPEP1* offers a clear hypothesis for the observed genetic association, the high level of linkage disequilibrium in this region precludes the exclusion of other genes as mechanistically linked to homocysteine metabolism (see Figure 1-E).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Carbamoylphosphate synthetase I (*CPS1*) is a nuclear encoded mitochondrial matrix enzyme that catalyses the first and rate-limiting step of the hepatic urea cycle. The hepatic urea cycle is responsible for the elimination of ammonia in the form of urea as well as the synthesis of arginine, a precursor of the potent vasodilatator nitric oxide. CPS1 synthesizes carbamoylphosphate from bicarbonate, ATP and ammonia using a cofactor N-acetylglutamate. Its genetic deficiency results in a rare autosomal recessive disease characterized by episodes of hyperammonemia in the neonatal period, with elevated plasma glutamine and low or absent citrulline [37]. The *CPS1* SNP rs7422339 associated with homocysteine in a sex-specific manner in our study encodes the substitution of asparagine for threonine (T1405N) in the region critical for N-acetyl-glutamate binding and results in 20–30% higher enzymatic activity [38]. This variation has been shown to influence nitric oxide metabolite concentrations and vasodilation following agonist stimulation [39]. Furthermore, the same variant has been associated with the risk of pulmonary hypertension in the newborn [40] as well as the risk of veno-occlusive disease after bone marrow transplantation [38–41]. Interestingly, nitrous oxide irreversibly inactivates the cytosolic enzyme methionine synthase by oxidizing enzyme-bound vitamin B_{12} [42, 43]. Methionine synthase is a vitamin B_{12} dependant enzyme that catalyzes the synthesis of methionine and methyltetrahydrofolate from homocysteine and tetrahydrofolate, the entry point of homocysteine into the remethylation pathway. Consistent with this pathway, humans and laboratory animals subjected to nitrous oxide have elevated levels of plasma homocysteine [44–48]. The *MUT* and *NOX4* loci were associated with homocysteine levels in the WGHS sample, but these associations were not replicated in the PROCARDIS sample. While rs4267943 is a non-synonymous coding SNP (G63R) in the gene *CENPQ* (centromere protein Q), the most likely candidate gene for its association with homocysteine is *MUT*. *CENPQ* itself encodes for a protein responsible for proper kinetochore function and mitotic progression [49]. Although we cannot exclude an effect on *CENPQ*, we note that rs4267943 is approximately 10 kb upstream of the *MUT* transcription start site and is in high linkage disequilibrium ($r^2 > 0.95$) with other *MUT* genetic variants, including a non-synonymous coding SNP rs9473558 (R532H; $P=2.1 \times 10^{-7}$). *MUT* encodes for the mitochondrial enzyme methylmalonyl-Coa mutase and as such, catalyses the isomeration of methylmalonyl-Coa into succinyl-Coa. While its catalytic activity is hardly related to homocysteine, MUT has frequently been associated with homocysteine metabolism because it is one of three vitamin B_{12} dependant enzymes, along with methionine synthase and leucine aminomutase [50]. Moreover, based on the observation that patients with vitamin B_{12} deficiency but high blood folate have higher plasma methylmalonic acid concentration (than low vitamin B_{12}, low folate patients), it has been proposed that the enzymatic activity of methionine synthase interferes with the activity of methylmalonyl-Coa mutase through diversion of vitamin B_{12} from the mitochondrion to the cytosol [51]. The genetic association observed in WGHS (but not in PROCARDIS) reinforces this hypothesis by suggesting that methylmalonyl-Coa mutase activity itself could impact on methionine synthase activity (and therefore homocysteine concentration) through a similar mechanism.

NOX4 encodes for a recently described NADPH oxidase that is highly expressed in the kidney [52]. As such, it catalyses the formation of the free-radical superoxide using O_2 as an electron acceptor and NADPH as the donor. The exact role of NOX4 in normal physiology is yet to be determined, but NOX4 expression has recently been correlated with increased albuminuria in adiponectin deficient rats [53]. For these reasons, the most compelling explanation for the observed genetic association between the intronic *NOX4* SNP rs11018628 and plasma homocysteine concentration is regulation of renal handling of homocysteine by NOX4. This association, however, requires further validation given its failure to replicate in the PROCARDIS sample.

Despite the confirmatory nature of the *MTHFR* association, we also identified a novel *MTHFR* variant (rs17350396) that non-redundantly influenced homocysteine concentration

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

after taking into account the effect of the known *MTHFR* SNP rs1801133 (C677T; A222V), paving the way for more complete assessment of the impact of this locus on human disease. While genetic variants of *CBS* have previously been reported to influence homocysteine concentration [11, 12], the intronic *CBS* SNP rs6586282 described in this report is either in unremarkable or unknown linkage disequilibrium with these other variations. CBS catalyses the first step in the transsulfuration pathway of homocysteine catabolism[1].

Our study has potential limitations. First, even though all associations in WGHS had P-value lower than the conservative 5×10^{-8} threshold and care was taken to correct for the potential effect of population stratification, we were not able to confirm the *MUT* and *NOX4* findings in the separate PROCARDIS sample. This might be the result of the much smaller size of the PROCARDIS sample. For instance, the power to replicate the *MUT* association at a nominal level of significance of 0.05 in PROCARDIS was 27% while it was 29% for *NOX4*. Second, because of the nature of the WGHS and PROCARDIS samples, extension of these associations in non-Caucasian populations will require further testing, even if we have no *a priori* reason to believe there should be heterogeneity. Third, the overall variance explained by these associations (2.6%) is small as compared to heritability estimates of homocysteine levels[8, 9] (ranging from 25% to 44%), suggesting that many other genetic variants remain to be discovered. Fourth, the observation of a significant sex-interaction for the *CPS1* association hints at different genetic architectures of homocysteine levels in men and women. Our study might therefore have missed male-specific associations. Finally, although many of these novel associations offer tantalizing hypotheses regarding homocysteine metabolism, further biological experiments are needed to define the processes underlying them.

In this report, we found associations of homocysteine with genetic variation at the *MTHFR*, *CPS1*, *MUT*, *NOX4*, *DPEP1* and *CBS* loci, with independent replication of the *MTHFR*, *CPS1*, *DPEP1* and *CBS* findings. While associations at the *MTHFR* and *CBS* loci extend previous work done on the genetic basis of homocysteine plasma concentration, our other findings are entirely novel. It will be of considerable interest to study these associations in non-Caucasian populations, as well as test these polymorphisms for association with cardiovascular diseases, neural tube defects and cognitive decline. These hypotheses generating genetic observations pave the way for further biological studies of homocysteine in the pathophysiology of these diseases. Furthermore, these genetic observations provide new insights into the biochemical pathways involved in homocysteine metabolism.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Lynda Rose for her review of the manuscript and helpful comments.

FUNDING SOURCES: This study was supported by grants from the National Heart, Lung, and Blood Institute and the National Cancer Institute (Bethesda, MD, USA; grants HL080467, HL043851 and CA047988), the Donald W Reynolds Foundation (Las Vegas, NV, USA), the Doris Duke Charitable Foundation (New York, NY, USA), and the Fondation Leducq (Paris, France), with collaborative scientific and genotyping support provided by Amgen Inc (Thousand Oaks, CA, USA). The PROCARDIS genome wide association study was supported by the European Commission (grant LSHM-2007-027273), the British Heart Foundation, the Knut and Alice Wallenberg Foundation, the Swedish Medical Research Council (project 8691), the Swedish Heart-Lung Foundation, the Karolinska Institutet and the Stockholm County Council (project 560183).

References

1. Selhub J. Homocysteine metabolism. Annu Rev Nutr 1999;19:217–246. [PubMed: 10448523]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

2. Upchurch GR Jr, Welch GN, Fabian AJ, Freedman JE, Johnson JL, Keaney JF Jr, Loscalzo J. Homocyst (e)ine decreases bioavailable nitric oxide by a mechanism involving glutathione peroxidase. J Biol Chem 1997;272:17012–17017. [PubMed: 9202015]
3. Weiss N, Zhang YY, Heydrick S, Bierl C, Loscalzo J. Overexpression of cellular glutathione peroxidase rescues homocyst(e)ine-induced endothelial dysfunction. Proc Natl Acad Sci U S A 2001;98:12503–12508. [PubMed: 11606774]
4. Homocysteine Studies Collaboration. Homocysteine and risk of ischemic heart disease and stroke: a meta-analysis. Jama 2002;288:2015–2022. [PubMed: 12387654]
5. Welch GN, Loscalzo J. Homocysteine and atherothrombosis. N Engl J Med 1998;338:1042–1050. [PubMed: 9535670]
6. Mills JL, McPartlin JM, Kirke PN, Lee YJ, Conley MR, Weir DG, Scott JM. Homocysteine metabolism in pregnancies complicated by neural-tube defects. Lancet 1995;345:149–151. [PubMed: 7741859]
7. Clarke R, Smith AD, Jobst KA, Refsum H, Sutton L, Ueland PM. Folate, vitamin B12, and serum total homocysteine levels in confirmed Alzheimer disease. Arch Neurol 1998;55:1449–1455. [PubMed: 9823829]
8. Souto JC, Blanco-Vaca F, Soria JM, Buil A, Almasy L, Ordonez-Llanos J, Martin-Campos JM, Lathrop M, Stone W, Blangero J, Fontcuberta J. A genomewide exploration suggests a new candidate gene at chromosome 11q23 as the major determinant of plasma homocysteine levels: results from the GAIT project. Am J Hum Genet 2005;76:925–933. [PubMed: 15849667]
9. Vermeulen SH, van der Vleuten GM, de Graaf J, Hermus AR, Blom HJ, Stalenhoef AF, den Heijer M. A genome-wide linkage scan for homocysteine levels suggests three regions of interest. J Thromb Haemost 2006;4:1303–1307. [PubMed: 16706975]
10. Frosst P, Blom HJ, Milos R, Goyette P, Sheppard CA, Matthews RG, Boers GJ, den Heijer M, Kluijtmans LA, van den Heuvel LP, Rozen R. A candidate genetic risk factor for vascular disease: a common mutation in methylenetetrahydrofolate reductase. Nat Genet 1995;10:111–113. [PubMed: 7647779]
11. De Stefano V, Dekou V, Nicaud V, Chasse JF, London J, Stansbie D, Humphries SE, Gudnason V. Linkage disequilibrium at the cystathionine beta synthase (CBS) locus and the association between genetic variation at the CBS locus and plasma levels of homocysteine. The Ears II Group. European Atherosclerosis Research Study. Ann Hum Genet 1998;62:481–490. [PubMed: 10363126]
12. Lievers KJ, Kluijtmans LA, Heil SG, Boers GH, Verhoef P, van Oppenraay-Emmerzaal D, den Heijer M, Trijbels FJ, Blom HJ. A 31 bp VNTR in the cystathionine beta-synthase (CBS) gene is associated with reduced CBS activity and elevated post-load homocysteine levels. Eur J Hum Genet 2001;9:583–589. [PubMed: 11528503]
13. Tsai MY, Bignell M, Yang F, Welge BG, Graham KJ, Hanson NQ. Polygenic influence on plasma homocysteine: association of two prevalent mutations, the 844ins68 of cystathionine beta-synthase and A(2756)G of methionine synthase, with lowered plasma homocysteine levels. Atherosclerosis 2000;149:131–137. [PubMed: 10704624]
14. Gaughan DJ, Kluijtmans LA, Barbaux S, McMaster D, Young IS, Yarnell JW, Evans A, Whitehead AS. The methionine synthase reductase (MTRR) A66G polymorphism is a novel genetic determinant of plasma homocysteine concentrations. Atherosclerosis 2001;157:451–456. [PubMed: 11472746]
15. Ridker PM, Chasman DI, Zee RY, Parker A, Rose L, Cook NR, Buring JE. Rationale, design, and methodology of the Women's Genome Health Study: a genome-wide association study of more than 25,000 initially healthy american women. Clin Chem 2008;54:249–255. [PubMed: 18070814]
16. Ridker PM, Cook NR, Lee IM, Gordon D, Gaziano JM, Manson JE, Hennekens CH, Buring JE. A randomized trial of low-dose aspirin in the primary prevention of cardiovascular disease in women. N Engl J Med 2005;352:1293–1304. [PubMed: 15753114]
17. Zee RY, Mora S, Cheng S, Erlich HA, Lindpaintner K, Rifai N, Buring JE, Ridker PM. Homocysteine, 5,10-methylenetetrahydrofolate reductase 677C>T polymorphism, nutrient intake, and incident cardiovascular disease in 24,968 initially healthy women. Clin Chem 2007;53:845–851. [PubMed: 17332146]
18. Zhang SM, Moore SC, Lin J, Cook NR, Manson JE, Lee IM, Buring JE. Folate, vitamin B6, multivitamin supplements, and colorectal cancer risk in women. Am J Epidemiol 2006;163:108–115. [PubMed: 16339055]

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

19. Willett WC, Sampson L, Browne ML, Stampfer MJ, Rosner B, Hennekens CH, Speizer FE. The use of a self-administered questionnaire to assess diet four years in the past. Am J Epidemiol 1988;127:188–199. [PubMed: 3337073]
20. Salvini S, Hunter DJ, Sampson L, Stampfer MJ, Colditz GA, Rosner B, Willett WC. Food-based validation of a dietary questionnaire: the effects of week-to-week variation in food consumption. Int J Epidemiol 1989;18:858–867. [PubMed: 2621022]
21. Giovannucci E, Stampfer MJ, Colditz GA, Rimm EB, Trichopoulos D, Rosner BA, Speizer FE, Willett WC. Folate, methionine, and alcohol intake and risk of colorectal adenoma. J Natl Cancer Inst 1993;85:875–884. [PubMed: 8492316]
22. Zhang SM, Willett WC, Selhub J, Hunter DJ, Giovannucci EL, Holmes MD, Colditz GA, Hankinson SE. Plasma folate, vitamin B6, vitamin B12, homocysteine, and risk of breast cancer. J Natl Cancer Inst 2003;95:373–380. [PubMed: 12618502]
23. Wigginton JE, Cutler DJ, Abecasis GR. A note on exact tests of Hardy-Weinberg equilibrium. Am J Hum Genet 2005;76:887–893. [PubMed: 15789306]
24. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81:559–575. [PubMed: 17701901]
25. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 2006;38:904–909. [PubMed: 16862161]
26. Gambacciani M, Mannella P. Homocysteine, menopause and cardiovascular disease. Menopause Int 2007;13:23–26. [PubMed: 17448264]
27. Siri PW, Verhoef P, Kok FJ. Vitamins B6, B12, and folate: association with plasma total homocysteine and risk of coronary atherosclerosis. J Am Coll Nutr 1998;17:435–441. [PubMed: 9791839]
28. O'Callaghan P, Meleady R, Fitzgerald T, Graham I. Smoking and plasma homocysteine. Eur Heart J 2002;23:1580–1586. [PubMed: 12323157]
29. Mora S, Lee IM, Buring JE, Ridker PM. Association of physical activity and body mass index with novel and traditional cardiovascular biomarkers in women. Jama 2006;295:1412–1419. [PubMed: 16551713]
30. Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, Gibbs RA, Belmont JW, Boudreau A, Hardenbol P, Leal SM, Pasternak S, Wheeler DA, Willis TD, Yu F, Yang H, Zeng C, Gao Y, Hu H, Hu W, Li C, Lin W, Liu S, Pan H, Tang X, Wang J, Wang W, Yu J, Zhang B, Zhang Q, Zhao H, Zhao H, Zhou J, Gabriel SB, Barry R, Blumenstiel B, Camargo A, Defelice M, Faggart M, Goyette M, Gupta S, Moore J, Nguyen H, Onofrio RC, Parkin M, Roy J, Stahl E, Winchester E, Ziaugra L, Altshuler D, Shen Y, Yao Z, Huang W, Chu X, He Y, Jin L, Liu Y, Shen Y, Sun W, Wang H, Wang Y, Wang Y, Xiong X, Xu L, Waye MM, Tsui SK, Xue H, Wong JT, Galver LM, Fan JB, Gunderson K, Murray SS, Oliphant AR, Chee MS, Montpetit A, Chagnon F, Ferretti V, Leboeuf M, Olivier JF, Phillips MS, Roumy S, Sallee C, Verner A, Hudson TJ, Kwok PY, Cai D, Koboldt DC, Miller RD, Pawlikowska L, Taillon-Miller P, Xiao M, Tsui LC, Mak W, Song YQ, Tam PK, Nakamura Y, Kawaguchi T, Kitamoto T, Morizono T, Nagashima A, Ohnishi Y, Sekine A, Tanaka T, Tsunoda T, Deloukas P, Bird CP, Delgado M, Dermitzakis ET, Gwilliam R, Hunt S, Morrison J, Powell D, Stranger BE, Whittaker P, Bentley DR, Daly MJ, de Bakker PI, Barrett J, Chretien YR, Maller J, McCarroll S, Patterson N, Pe'er I, Price A, Purcell S, Richter DJ, Sabeti P, Saxena R, Schaffner SF, Sham PC, Varilly P, Altshuler D, Stein LD, Krishnan L, Smith AV, Tello-Ruiz MK, Thorisson GA, Chakravarti A, Chen PE, Cutler DJ, Kashuk CS, Lin S, Abecasis GR, Guan W, Li Y, Munro HM, Qin ZS, Thomas DJ, McVean G, Auton A, Bottolo L, Cardin N, Eyheramendy S, Freeman C, Marchini J, Myers S, Spencer C, Stephens M, Donnelly P, Cardon LR, Clarke G, Evans DM, Morris AP, Weir BS, Tsunoda T, Mullikin JC, Sherry ST, Feolo M, Skol A, Zhang H, Zeng C, Zhao H, Matsuda I, Fukushima Y, Macer DR, Suda E, Rotimi CN, Adebamowo CA, Ajayi I, Aniagwu T, Marshall PA, Nkwodimmah C, Royal CD, Leppert MF, Dixon M, Peiffer A, Qiu R, Kent A, Kato K, Niikawa N, Adewole IF, Knoppers BM, Foster MW, Clayton EW, Watkin J, Gibbs RA, Belmont JW, Muzny D, Nazareth L, Sodergren E, Weinstock GM, Wheeler DA, Yakub I, Gabriel SB, Onofrio RC, Richter DJ, Ziaugra L, Birren BW, Daly MJ, Altshuler D, Wilson RK, Fulton LL, Rogers J, Burton J, Carter NP, Clee CM, Griffiths M, Jones MC, McLay K, Plumb RW, Ross MT, Sims SK, Willey DL, Chen Z, Han H, Kang L, Godbout M, Wallenburg JC, L'Archeveque P, Bellemare G,

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Saeki K, Wang H, An D, Fu H, Li Q, Wang Z, Wang R, Holden AL, Brooks LD, McEwen JE, Guyer MS, Wang VO, Peterson JL, Shi M, Spiegel J, Sung LM, Zacharia LF, Collins FS, Kennedy K, Jamieson R, Stewart J. A second generation human haplotype map of over 3.1 million SNPs. Nature 2007;449:851–861. [PubMed: 17943122]

31. Farrall M, Green FR, Peden JF, Olsson PG, Clarke R, Hellenius ML, Rust S, Lagercrantz J, Franzosi MG, Schulte H, Carey A, Olsson G, Assmann G, Tognoni G, Collins R, Hamsten A, Watkins H. Genome-wide mapping of susceptibility to coronary artery disease identifies a novel replicated locus on chromosome 17. PLoS Genet 2006;2:e72. [PubMed: 16710446]
32. Broadbent HM, Peden JF, Lorkowski S, Goel A, Ongen H, Green F, Clarke R, Collins R, Franzosi MG, Tognoni G, Seedorf U, Rust S, Eriksson P, Hamsten A, Farrall M, Watkins H. Susceptibility to coronary artery disease and diabetes is encoded by distinct, tightly linked SNPs in the ANRIL locus on chromosome 9p. Hum Mol Genet 2008;17:806–814. [PubMed: 18048406]
33. Habib GM, Barrios R, Shi ZZ, Lieberman MW. Four distinct membrane-bound dipeptidase RNAs are differentially expressed and show discordant regulation with gamma-glutamyl transpeptidase. J Biol Chem 1996;271:16273–16280. [PubMed: 8663190]
34. Kozak EM, Tate SS. Glutathione-degrading enzymes of microvillus membranes. J Biol Chem 1982;257:6322–6327. [PubMed: 6122685]
35. Finkelstein JD. Inborn errors of sulfur-containing amino acid metabolism. J Nutr 2006;136:1750S–1754S. [PubMed: 16702350]
36. Mayatepek E, Badiou S, Bellet H, Lehmann WD. A patient with neurological symptoms and abnormal leukotriene metabolism: a new defect in leukotriene biosynthesis. Ann Neurol 2005;58:968–970. [PubMed: 16315285]
37. Scaglia F, Brunetti-Pierri N, Kleppe S, Marini J, Carter S, Garlick P, Jahoor F, O'Brien W, Lee B. Clinical consequences of urea cycle enzyme deficiencies and potential links to arginine and nitric oxide metabolism. J Nutr 2004;134:2775S–2782S. [PubMed: 15465784]
38. Summar ML, Hall L, Christman B, Barr F, Smith H, Kallianpur A, Brown N, Yadav M, Willis A, Eeds A, Cermak E, Summar S, Wilson A, Arvin M, Putnam A, Wills M, Cunningham G. Environmentally determined genetic expression: clinical correlates with molecular variants of carbamyl phosphate synthetase I. Mol Genet Metab 2004;81:S12–19. [PubMed: 15050969]
39. Summar ML, Gainer JV, Pretorius M, Malave H, Harris S, Hall LD, Weisberg A, Vaughan DE, Christman BW, Brown NJ. Relationship between carbamoyl-phosphate synthetase genotype and systemic vascular function. Hypertension 2004;43:186–191. [PubMed: 14718356]
40. Pearson DL, Dawling S, Walsh WF, Haines JL, Christman BW, Bazyk A, Scott N, Summar ML. Neonatal pulmonary hypertension--urea-cycle intermediates, nitric oxide production, and carbamoyl-phosphate synthetase function. N Engl J Med 2001;344:1832–1838. [PubMed: 11407344]
41. Summar ML, Scott N, Cummings E, Hutcheson H, Dawling S, Christman B. Analysis of 200 patients undegoing bone marrow transplant shows allelic disequilibrium between drug related toxicity and a common exonic polymorphism in the CPS1 gene and correlates with disruption of urea cycle intermediates. Am J Hum Genet 1999;65:A25.
42. Deacon R, Lumb M, Perry J, Chanarin I, Minty B, Halsey MJ, Nunn JF. Selective inactivation of vitamin B12 in rats by nitrous oxide. Lancet 1978;2:1023–1024. [PubMed: 82036]
43. Ng J, Frith R. Nanging. Lancet 2002;360:384. [PubMed: 12241780]
44. Waclawik AJ, Luzzio CC, Juhasz-Pocsine K, Hamilton V. Myeloneuropathy from nitrous oxide abuse: unusually high methylmalonic acid and homocysteine levels. Wmj 2003;102:43–45. [PubMed: 12967021]
45. Drummond JT, Matthews RG. Nitrous oxide degradation by cobalamin-dependent methionine synthase: characterization of the reactants and products in the inactivation reaction. Biochemistry 1994;33:3732–3741. [PubMed: 8142373]
46. Kondo H, Osborne ML, Kolhouse JF, Binder MJ, Podell ER, Utley CS, Abrams RS, Allen RH. Nitrous oxide has multiple deleterious effects on cobalamin metabolism and causes decreases in activities of both mammalian cobalamin-dependent enzymes in rats. J Clin Invest 1981;67:1270–1283. [PubMed: 6112240]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

47. Rask H, Olesen AS, Mortensen JZ, Freund LG. N2O and urine methylmalonic acid in man. Scand J Haematol 1983;31:45–48. [PubMed: 6135251]

48. Loscalzo J. Adverse effects of supplemental L-arginine in atherosclerosis: consequences of methylation stress in a complex catabolism? Arterioscler Thromb Vasc Biol 2003;23:3–5. [PubMed: 12524215]

49. Okada M, Cheeseman IM, Hori T, Okawa K, McLeod IX, Yates JR 3rd, Desai A, Fukagawa T. The CENP-H-I complex is required for the efficient incorporation of newly synthesized CENP-A into centromeres. Nat Cell Biol 2006;8:446–457. [PubMed: 16622420]

50. Murray, R.; Granner, D.; Mayes, P.; Rodwell, V. Harper's illustrated biochemistry. New York: Lange Medical Books/McGraw-Hill; 2003.

51. Selhub J, Morris MS, Jacques PF. In vitamin B12 deficiency, higher serum folate is associated with increased total homocysteine and methylmalonic acid concentrations. Proc Natl Acad Sci U S A 2007;104:19995–20000. [PubMed: 18056804]

52. Geiszt M, Kopp JB, Varnai P, Leto TL. Identification of renox, an NAD(P)H oxidase in kidney. Proc Natl Acad Sci U S A 2000;97:8010–8014. [PubMed: 10869423]

53. Sharma K, Ramachandrarao S, Qiu G, Usui HK, Zhu Y, Dunn SR, Ouedraogo R, Hough K, McCue P, Chan L, Falkner B, Goldstein BJ. Adiponectin regulates albuminuria and podocyte function in mice. J Clin Invest 2008;118:1645–1656. [PubMed: 18431508]

54. McVean GA, Myers SR, Hunt S, Deloukas P, Bentley DR, Donnelly P. The fine-scale structure of recombination rate variation in the human genome. Science 2004;304:581–584. [PubMed: 15105499]

55. Winckler W, Myers SR, Richter DJ, Onofrio RC, McDonald GJ, Bontrop RE, McVean GA, Gabriel SB, Reich D, Donnelly P, Altshuler D. Comparison of fine-scale recombination rates in humans and chimpanzees. Science 2005;308:107–111. [PubMed: 15705809]

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

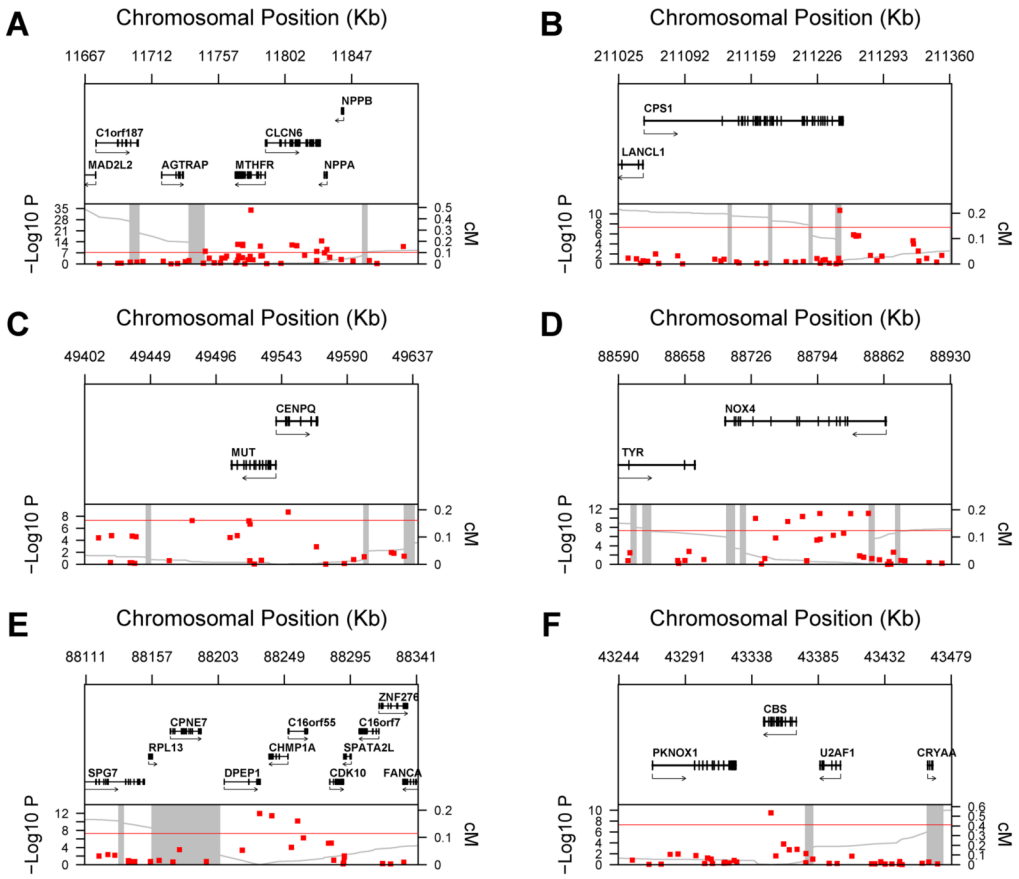

Figure 1. Genetic Context of Significant Associations

Genomic context for each of six loci with significant association with homocysteine concentration. (A) *MTHFR* locus (1p36.22); (B) *CPS1* locus (2q34); (C) *MUT* locus (6p12.3); (D) *NOX4* locus (11q14.3); (E) *DPEP1* locus (16q24.3); (F) *CBS* locus (21q22.3). Upper panel: Genes from RefSeq release 25. Only one isoform is shown when multiple splicing variants are known. Lower Panel: SNPs are shown according to their physical location and $-\log_{10}$ P-values for association with homocysteine (red dots). The red line represents the genome-wide significance threshold of 5×10^{-8}. Also shown is the genetic distance in cM from the lowest P-value SNP (light grey line) along with the position of recombination hotspots (light grey vertical bars). Recombination rates and hotspots are based on HapMap data, as described by McVean et al.[54] and Winckler et al.[55]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

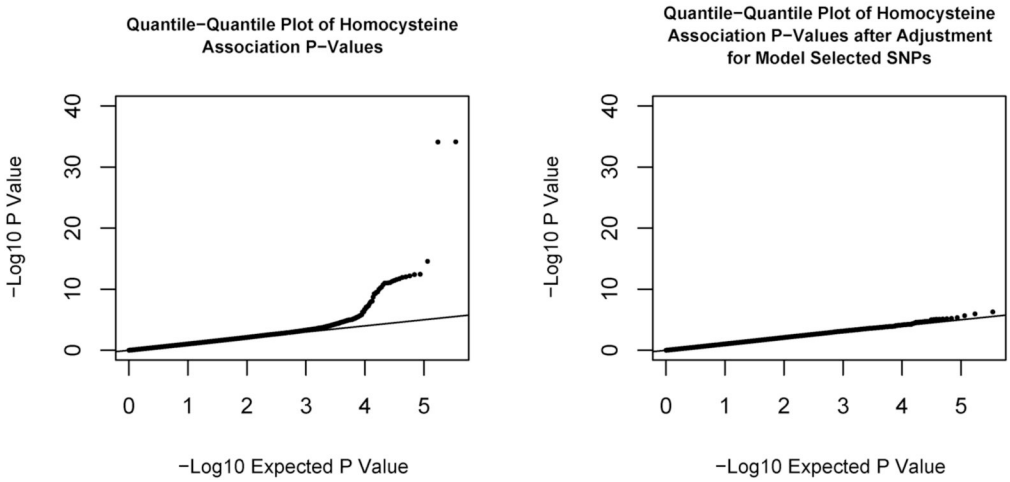

Figure 2. Quantile-Quantile plot of Homocysteine Association P-Values

The quantile-quantile plot of homocysteine association P-values is shown on the left. On the right, the same quantile-quantile plot is shown, but after adjusting homocysteine values for the 7 SNPs retained by the model selection algorithm (see text for details).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 1

Clinical characteristics of the samples used.

	WGHS (N=13,974)	PROCARDIS (N=840)
%Female Patients	100%	19.2%
Age (yrs.)	54.7 (7.0)	61.0 (7.0)
BMI (kg/m^2)	25.8 (4.7)	28.0 (4.2)
Menopause*	54.6%	NA
HRT†	44.9%	NA
Smoking	11.6%	25.0 %
Homocysteine (μmol/L)	11.4(4.7)	13.3(6.0)
Folate Intake (μg/day)	434 (227)	NA
Vitamin B_6 Intake (mg/day)	5.2 (17.4)	NA
Vitamin B_{12} Intake (μg/day)	8.6 (9.6)	NA

Results are given as mean (standard deviation), as appropriate.

*In women only.

†HRT: Hormone replacement therapy use, in women only.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

Table 2

SNPs with a P-Value Lower than 10^{-6} for Association with Homocysteine Concentration.

SNP	Region	Position (kb)	Nearest Gene (Candidate Gene)	Function*	MAF†	HW‡	Beta§	P-Value§
rs12564559	**1p36.22**	**11748.1**	**AGTRAP (MTHFR)**	**intron**	**0.11**	**0.58**	**−0.034**	**9.4E-09**
rs3737967	**1p36.22**	**11770.0**	**MTHFR**	**CNS (R492H)**	**0.05**	**0.52**	**−0.063**	**4.1E-13**
rs2274976	**1p36.22**	**11773.5**	**MTHFR**	**CNS (R594Q)**	**0.05**	**0.35**	**−0.062**	**1.9E-12**
rs1801133	**1p36.22**	**11779.0**	**MTHFR**	**CNS (A222V)**	**0.33**	**0.02**	**0.048**	**8.1E-35**
rs17367504	**1p36.22**	**11785.4**	**MTHFR**	**intron**	**0.16**	**0.26**	**−0.035**	**5.4E-12**
rs2076003	**1p36.22**	**11806.7**	**CLCN6 (MTHFR)**	**intron**	**0.05**	**0.72**	**−0.062**	**9.4E-13**
rs7537765	**1p36.22**	**11809.9**	**CLCN6 (MTHFR)**	**intron**	**0.16**	**0.29**	**−0.035**	**2.6E-12**
rs198414	**1p36.22**	**11823.4**	**CLCN6 (MTHFR)**	**untranslated**	**0.14**	**0.97**	**−0.031**	**1.1E-08**
rs198358	**1p36.22**	**11826.7**	**CLCN6 (MTHFR)**	**-**	**0.24**	**0.41**	**−0.034**	**2.7E-15**
rs5065	1p36.22	11828.7	NPPA (MTHFR)	CNS (*152R)	0.15	0.74	−0.028	9.0E-08
rs5063	**1p36.22**	**11830.2**	**NPPA (MTHFR)**	**CNS (V32M)**	**0.05**	**0.46**	**−0.055**	**5.6E-10**
rs1999594	**1p36.22**	**11881.8**	**NPPB (MTHFR)**	**-**	**0.46**	**0.77**	**−0.025**	**8.5E-12**
rs2639453	**1p36.22**	**11905.1**	**KIAA2013 (MTHFR)**	**intron**	**0.22**	**0.34**	**0.032**	**6.7E-13**
rs6682554	**1p36.22**	**11909.2**	**KIAA2013 (MTHFR)**	**-**	**0.45**	**0.36**	**0.024**	**2.7E-10**
rs7550536	**1p36.22**	**11989.6**	**MFN2 (MTHFR)**	**intron**	**0.40**	**0.70**	**0.021**	**1.4E-08**
rs7422339	**2q34**	**211248.8**	**CPS1**	**CNS (T1406N)**	**0.31**	**0.92**	**0.027**	**1.9E-11**
rs2501976	6p12.3	49478.9	MUT	-	0.37	0.15	0.021	5.3E-08
rs6458690	6p12.3	49519.6	MUT	intron	0.37	0.27	0.021	5.9E-08
rs9473558	6p12.3	49520.4	MUT	CNS (R532H)	0.37	0.17	0.020	2.1E-07
rs4267943	**6p12.3**	**49547.8**	**CENPQ (MUT)**	**CNS (G63R)**	**0.36**	**5.6E-05**	**0.024**	**2.0E-09**
rs317191	**11q14.3**	**88730.5**	**NOX4**	**intron**	**0.07**	**0.64**	**−0.048**	**1.1E-10**
rs317150	**11q14.3**	**88763.4**	**NOX4**	**intron**	**0.08**	**0.63**	**−0.041**	**5.3E-10**
rs7929532	**11q14.3**	**88778.6**	**NOX4**	**intron**	**0.07**	**0.53**	**−0.048**	**4.4E-11**
rs9299894	**11q14.3**	**88796.3**	**NOX4**	**intron**	**0.07**	**0.14**	**−0.050**	**9.8E-12**
rs317148	11q14.3	88810.2	NOX4	intron	0.44	0.73	−0.019	5.1E-07
rs319016	11q14.3	88820.8	NOX4	intron	0.47	0.32	0.019	1.8E-07
rs10501705	**11q14.3**	**88827.9**	**NOX4**	**intron**	**0.07**	**0.27**	**−0.050**	**1.0E-11**
rs11018628	**11q14.3**	**88846.2**	**NOX4**	**intron**	**0.07**	**0.18**	**−0.050**	**9.6E-12**
rs2061821	15q25.3	83923.7	AKAP13	CNS (M452T)	0.38	0.71	0.019	5.5E-07

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

SNP	Region	Position (kb)	Nearest Gene (Candidate Gene)	Function*	MAF†	HW‡	Beta§	P-Value§
rs1126464	**16q24.3**	**88231.9**	**DPEP1**	**CNS (E351Q)**	**0.24**	**0.93**	**−0.031**	**1.2E-12**
rs460879	**16q24.3**	**88240.4**	**CHMP1A (DPEP1)**	**intron**	**0.46**	**0.19**	**0.026**	**3.8E-12**
rs459920	**16q24.3**	**88258.3**	**C16orf55 (DPEP1)**	**intron**	**0.44**	**0.36**	**0.024**	**6.5E-11**
rs2377058	16q24.3	88262.3	C16orf55 (DPEP1)	intron	0.36	0.86	−0.019	5.6E-07
rs6586282	**21q22.3**	**43351.6**	**CBS**	**intron**	**0.18**	**0.37**	**−0.030**	**3.2E-10**

*CNS: Coding Non-Synonymous SNP

†MAF: Minor allele frequency

‡HW: Deviation from Hardy-Weinberg equilibrium P-value

§All analyses were performed using adjusted log-transformed homocysteine values (see text for details). SNPs with P-values lower than 5×10^{-8} are shown in bold.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

Table 3

Multiple Linear Regression Statistics of SNPs Retained by the Forward Model Selection Algorithm.

SNP	Chromosome	Position (kb)	Nearest Gene (Candidate Gene)	Function*	MAF†	HW‡	Beta§	P-Value§
rs1801133	1p36.22	11779.0	MTHFR	CNS	0.33	0.02	0.057	1.7E-44
rs17350396	1p36.22	11823.4	CLCN6 (MTHFR)	untranslated	0.17	0.82	0.036	4.2E-12
rs7422339	2q34	211248.8	CPS1	CNS	0.31	0.92	0.027	1.9E-11
rs4267943	6p12.3	49547.8	CENPQ (MUT)	CNS	0.36	0.00	0.024	2.0E-09
rs11018628	11q14.3	88846.2	NOX4	intron	0.07	0.18	−0.050	9.6E-12
rs1126464	16q24.3	88231.9	DPEP1	CNS	0.24	0.93	−0.031	1.2E-12
rs6586282	21q22.3	43351.6	CBS	intron	0.18	0.37	−0.030	3.2E-10

*CNS: Coding Non-Synonymous SNP

†MAF: Minor allele frequency

‡HW: Deviation from Hardy-Weinberg equilibrium P-value

§All analyses were performed using adjusted log-transformed homocysteine values (see text for details). Both rs1801133 and rs17350396 (MTHFR locus) were included in the same multivariable linear model.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 4

Haplotype Analysis of rs1801133 and rs17350396 (*MTHFR* Locus).

Haplotype		Frequency	Beta*	P-Value*
rs1801133	**rs17350396**			
G	G	0.17	0.013	1.1E-02
A	A	0.33	0.048	3.3E-35
G	A	0.50	−0.051	8.6E-43

Omnibus (2 df) P-Value = 5.5 E-45

*All analyses were performed using adjusted log-transformed homocysteine values

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 5

Partition of Homocysteine Variance According to Genetic and Clinical Variables.

Category	Variable	Variable R^2	Category R^2
Clinical Covariates	Age	0.0063	0.0871
	Body Mass Index	0.0020	
	Smoking Status	0.0226	
	HRT Status	0.0074	
	Folate Intake	0.0483	
	Vitamin B_6 Intake	0.0002	
	Vitamin B_{12} Intake	0.0003	
1p36.22 (MTHFR) Locus	rs1801133	0.0095	0.0126
	rs17350396	0.0030	
2q34 (CPS1) Locus	rs7422339	0.0027	0.0027
6p12.3 (MUT) Locus	rs4267943	0.0027	0.0027
11q14.3 (NOX4) Locus	rs11018622	0.0029	0.0029
16q24.3 (DPEP1) Locus	rs1126464	0.0029	0.0029
21q22.3 (CBS) Locus	rs6586282	0.0023	0.0023
TOTAL			0.1131