A Genomewide Exploration Suggests a New Candidate Gene at Chromosome 11q23 as the Major Determinant of Plasma Homocysteine Levels: Results from the GAIT Project

Juan Carlos Souto[1,*] Francisco Blanco-Vaca,[2,*] José Manuel Soria,[1] Alfonso Buil,[1] Laura Almasy,[3] Jordi Ordoñez-Llanos,[2] Jesús Mª Martín-Campos,[2] Mark Lathrop,[4] William Stone,[1,3] John Blangero,[3] and Jordi Fontcuberta[1]

[1]Unitat de Trombosi i Hemostàsia, Departament d' Hematologia, and [2]Servei de Bioquimica, Hospital de la Santa Creu i Sant Pau, Barcelona; [3]Department of Genetics, Southwest Foundation for Biomedical Research, San Antonio; and [4]Centre National de Genotypage, Evry, France

Homocysteine (Hcy) plasma level is an independent risk marker for venous thrombosis, myocardial infarction, stroke, congestive heart failure, osteoporotic fractures, and Alzheimer disease. Hcy levels are determined by the interaction of genetic and environmental factors. The genetic basis is still poorly understood, since only the *MTHFR* 677 C→T polymorphism has been consistently associated with plasma Hcy levels. We conducted a genomewide linkage scan for genes affecting variation in plasma Hcy levels in 398 subjects from 21 extended Spanish families. A variance-components linkage method was used to analyze the data. The strongest linkage signal (LOD score of 3.01; genomewide $P = .035$) was found on chromosome 11q23, near marker *D11S908*, where a candidate gene involved in the metabolism of Hcy (the nicotinamide N-methyltransferase gene [*NNMT*]) is mapped. Haplotype analyses of 10 single-nucleotide polymorphisms within this gene found one haplotype associated with plasma Hcy levels ($P = .0003$). Our results, to our knowledge, represent the first genomic scan for quantitative variation in Hcy plasma levels. They strongly suggest that the *NNMT* gene could be a major genetic determinant of plasma Hcy levels in Spanish families. Since this gene encodes an enzyme involved in Hcy synthesis, this finding would be consistent with known biochemical pathways. These data could be relevant in determining the relationships between Hcy level, cardiovascular disease, osteoporosis, and Alzheimer disease.

Introduction

Multiple case-control and prospective studies have established an association between moderate hyperhomocysteinemia and increased risk of ischemic heart disease and stroke (Welch and Loscalzo 1998; Cattaneo 1999; Homocysteine Studies Collaboration 2002; Wald et al. 2002), congestive heart failure (Vasan et al. 2003), and deep vein thrombosis (Cattaneo 1999). Furthermore, recent data from the Framingham Study indicate that plasma homocysteine (Hcy) level is also a dose-related risk marker for the development of dementia and Alzheimer disease (Seshadri et al. 2002). There have been additional reports on the association between hyperhomocysteinemia and Alzheimer disease (Morris 2003), mental deterioration (Morris et al. 2001), and vascular dementia (McIlroy et al. 2002). These associations seems to be independent of age, sex, apolipoprotein E gene (*APOE*) genotype, plasma vitamin levels, and other putative risk factors (Seshadri et al. 2002). Furthermore, elevated Hcy concentration has been associated with risk of osteoporotic fractures in elderly persons (McLean et al. 2004; van Meurs et al. 2004)

The plasma Hcy level is a quantitative phenotype that is influenced by genetic and environmental factors. The main environmental determinants are sex, age, renal function, and vitamin intake (Brattsström et al. 1994). However, the genetic component is still poorly understood. Currently, the most widely accepted polymorphism associated with Hcy variability is *MTHFR* 677 C→T (Frosst et al. 1995). Other common polymorphisms within genes for enzymes involved in Hcy metabolism have been described. These polymorphisms, along with *MTHFR* 677 C→T, account for <10% of the variance in plasma Hcy level (Kluijtmans et al. 2003). A recent family study, which involved a segregation analysis, suggested the existence of a major gene that regulates Hcy plasma level, even after adjusting for effects of *MTHFR* 677 C→T (Jee et al. 2002).

As part of our Genetic Analysis of Idiopathic Thrombophilia (GAIT) Project, we observed that the herita-

Received December 28, 2004; accepted for publication March 21, 2005; electronically published April 22, 2005.

Address for correspondence and reprints: Dr. Juan Carlos Souto, Unitat de Trombosi i Hemostàsia, Hospital de la Santa Creu i Sant Pau, Sant Antoni Mª Claret 167, 08025 Barcelona, Spain. E-mail: jsouto@hsp.santpau.es

* These authors contributed equally to this study.

0002-9297/2005/7606-0003$15.00

bility of Hcy plasma levels (± SD) was 0.25 ± 0.08, indicating that ~25% of the phenotypic variation in this trait is due to the additive effect of genes (Souto et al. 2000*b*). Additional results from the GAIT Project (Souto et al. 2000*a*) have demonstrated that Hcy levels are genetically correlated with thrombotic risk (ρ_G = 0.65). This implies that some genes influence simultaneously the plasma levels of Hcy and the risk of thromboembolic disease.

In an effort to localize the genes influencing variation in plasma Hcy levels, we conducted a genomewide linkage scan of DNA from individuals in the GAIT sample.

Material and Methods

Study Population

The GAIT Project includes 21 extended families, 12 of whom were ascertained through a proband with thrombophilia, and 9 of whom were obtained randomly from the general population. A total of 398 individuals were examined, with a mean of 19 individuals examined per family. The composition of the families and the collection of lifestyle, medical, and family history data have been detailed elsewhere (Souto et al. 2000*b*). The study was performed in accordance with the Declaration of Helsinki of 1975; all adult patients provided informed consent and parents of minors provided informed consent for their children. All procedures were approved by the institutional review board of the Hospital de la Santa Creu i Sant Pau, Barcelona, Spain.

Plasma Hcy Determination

EDTA blood samples were obtained by venipuncture after a 12-h fast. Measures to prevent the flow of Hcy from erythrocytes to plasma were strictly implemented. Total Hcy plasma concentration was measured by use of high-performance liquid chromatography (Millipore) and fluorescence detection (Kontron Instruments). The analytical characteristics of this method have been analyzed and compared to a widely used commercial immunoassay (Blanco-Vaca et al. 2000).

DNA Extraction and Genotype Analysis

DNA extraction was performed in accordance with a standard protocol (Souto et al. 2000*b*). Subjects were genotyped for an autosomal genomewide scan with 363 highly informative DNA markers. Microsatellites consisted primarily of the ABI Prism genotyping set MD-10 (Applied Biosystems). PCR products were analyzed on PE 310, PE 377, and PE 3700 automated sequencers and were genotyped using PE Genotyper software (Applied Biosystems). The average heterozygosity of the microsatellite markers was 0.79, and the average distance between markers was 9.5 cM. The 677 C→T polymorphism in the *MTHFR* gene was genotyped using PCR with primers described elsewhere (Frosst et al. 1995) and with minor modifications to the reaction conditions. Marker maps for multipoint analyses were obtained from ABI Prism and from the Marshfield Center for Medical Genetics.

Genetic Linkage Analysis

Levels of Hcy in the GAIT sample exhibited a kurtosis of 29.5. For this reason, we used a natural logarithm to transform the Hcy values to achieve a normal distribution, and the new logarithmically transformed Hcy (lnHcy) levels had a kurtosis of only 0.48. This level of kurtosis did not affect the distribution of LOD scores (Blangero et al. 2000). The standard nominal *P* values for LOD scores were appropriate for the lnHcy linkage tests. The calculations also incorporated a simultaneous correction for covariates like sex, age, and BMI. Estimation of heritability (h^2) and the covariate effects of lnHcy were done using methods described elsewhere (Souto et al. 2000*b*). All analyses included an ascertainment correction, achieved by conditioning the likelihood of these pedigrees on the likelihoods of their respective probands (Boehnke and Lange 1984)

Standard multipoint variance-components linkage methods were used for the genome scan (Almasy and Blangero 1998). These methods have been implemented in the program SOLAR (available at the Southwest Foundation for Biomedical Research Web site). The statistical analysis is based on specification of the expected genetic covariances between arbitrary relatives as a function of the identity-by-descent (IBD) relationships at a given marker locus. To estimate the IBD matrix in these complex pedigrees, we used a Markov chain–Monte Carlo method as implemented in SIMWALK2 (Sobel and Lange 1996), which estimates the IBD probabilities at each cM. Bivariate linkage analyses with the use of a mixed discrete-/continuous-trait multivariate model were conducted with a modified version of SOLAR (Williams et al. 1999)

SNP Identification and Haplotype Analysis

We sequenced the nicotinamide N-methyltransferase gene (*NNMT* [UniGene accession number Hs.364345]) in 42 unrelated individuals from the GAIT Project using standard sequencing methods on an ABI 3100 automated sequencer (Applied Biosystems). The use of this number of individuals provided a >99% probability of detecting any polymorphism that had a frequency ≥0.05. To enrich the sequencing sample for functional variants, we chose 22 individuals from those in the higher tail of the Hcy distribution and 20 from those in the lower tail. We sequenced ~1.5 kb of the proximal promoter of *NNMT,* including exon 1, and we then con-

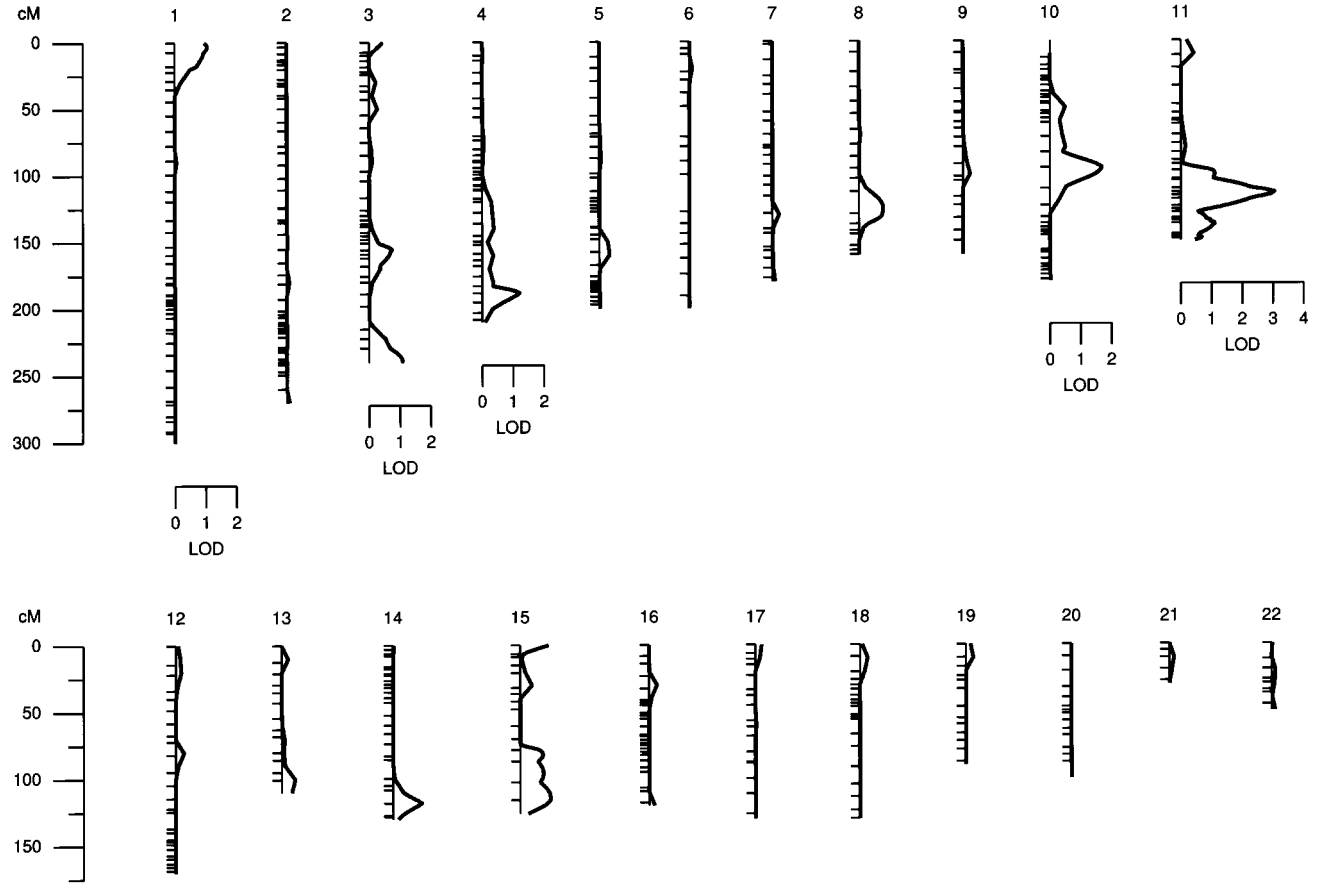

Figure 1 Linkage results from the autosomal multipoint genome scan. LOD score scales are shown for chromosomes 1, 3, 4, 10, and 11—the only chromosomes on which the maximum LOD score was >1. Hatch marks along the length of the chromosomes indicate the positions of the DNA markers.

tinued with the coding regions of exons 2, 3, 4, and 5. The latter included the 3′ UTR, with potential regulatory sequences.

Then, we genotyped the newly identified SNPs in our whole GAIT sample, as well as all previously reported *NNMT* SNPs (from the Celera database, available at the Applied Biosystems, myScience, Web site), by sequencing with the use of the SNP Genotyping Assays-on-Demand products from Applied Biosystems, in accordance with the recommendations of the manufacturer. The haplotypes with the SNPs detected in the *NNMT* gene were generated with the use of the "best" option in the MERLIN program (Abecasis et al. 2002). The association of each SNP or haplotype with Hcy plasma levels was tested by use of a quantitative-trait association analysis with the measured genotype approach, which involved testing for genotype-specific differences in the mean values of traits while allowing for the nonindependence among family members (Almasy and Blangero 2004). These analyses were performed using SOLAR.

Results

In the GAIT sample, the mean age-corrected Hcy plasma level (± SD) was 7.8 ± 3.1 μmol/liter. The range of Hcy values was 2.7–46.6 μmol/liter. After natural-logarithm transformation, lnHcy values showed a mean (± SD) of 2.05 ± 0.36. The significant regression coefficients for the covariates influencing lnHcy were 0.006 for age, −0.12 for female sex, and 0.004 for BMI. Together, they accounted for 18% of the variability of lnHcy levels. The heritability of the lnHcy phenotype (± SD) was 0.34 ± 0.08. This shows that genes play an important role in determining plasma levels (34% of phenotypic variability). This result is consistent with the previous estimate of $h^2 = 0.25 \pm 0.08$ for crude values of Hcy (Souto et al. 2000*b*).

Effects of the MTHFR *677 C→T Polymorphism*

The 677 C→T variant was present in 20 GAIT families. There were 174 heterozygotes and 59 homozygotes

Table 1

Genes Involved in Hcy Metabolism, with Their Chromosomal Locations and the Maximum LOD Score Obtained at Each Locus

Gene	Abbreviation	Location	LOD
Nicotinamide N-methyltransferase	*NNMT*	11q23	3.01
Methionine adenosyltransferase α I	*MATαI*	10q22	1.68
Methionine adenosyltransferase α II	*MATαII*	2p11	<1
Methionine adenosyltransferase β	*MATβ*	5q34	<1
5,10-Methylenetetrahydrofolate reductase	*MTHFR*	1p36	1.03
Other SAM-dependent methyltransferases[a]	…	…	<1
5-Methyltetrahydrofolate-homocysteine methyltransferase	*MTR*	1q43	<1
Methionine synthase reductase	*MTRR*	5p15	<1
Betaine-homocysteine methyltransferase	*BHMT*	5q14	<1
S-adenosylhomocysteine hydrolase	*AHCY*	20q11	<1
Cystathionine β-synthase	*CBS*	21q22	<1
Dihydrofolate reductase	*DHFR*	5q14	<1
Cystathionine γ-lyase	*CTH*	1p31	<1
Serine hydroxymethyltransferase 1	*SHMT1*	17p11	<1
Serine hydroxymethyltransferase 2	*SHMT2*	11q12	<1
Thymidylate synthase	*TS*	18p11	<1
Methylenetetrahydrofolate dehydrogenase 1	*MTHFD1*	14q23	<1
Methylenetetrahydrofolate dehydrogenase 2	*MTHFD2*	2p13	<1
S-adenosylmethionine decarboxylase	*AMD1*	6q21	<1
Choline dehydrogenase	*CHDH*	3p21	<1
Paraoxonase	*PON1*	7q21	<1
Folate receptor	*FOLR1*	11q13	<1
Folate carrier 19 member 1	*SLC19A1*	21q22	<1
Folate carrier 19 member 2	*SLC19A2*	1q24	<1
Folate carrier 19 member 3	*SLC19A3*	2q36	<1
Transcobalamin 1	*TCN1*	11q12	<1
Transcobalamin 2	*TCN2*	22q12	<1
Intrinsic factor	*GIF*	11q12	<1

[a] Apart from NNMT, there exist 38 other human SAM-dependent methyltransferases (Clarke and Banfield 2001); none of their respective gene loci showed a LOD score >1, and none is located on 11q23 or 10q22, where we obtained the highest LOD scores in our analysis.

for the T allele, and they exhibited higher mean Hcy levels than noncarriers (Hcy levels of 7.8 μmol/liter in CT heterozygotes, 9.5 μmol/liter in TT homozygotes, and 7.1 μmol/liter in CC noncarriers).

Nominal evidence of linkage was found between lnHcy levels and the *MTHFR* locus in a two-point linkage analysis (LOD score of 0.66; $P = .04$). Additionally, the 677 C→T polymorphism was strongly associated with lnHcy levels ($P < 1.6 \times 10^{-5}$).

Genomewide Linkage Analysis

Multipoint linkage analyses were performed across all 22 autosomes. The results of the genome scan for a QTL influencing lnHcy levels are shown in figure 1. One region, on chromosome 11q23, showed strong evidence of linkage (LOD score of 3.01; nominal $P = .0001$; genomewide $P = .035$). This peak was the highest LOD score found and occurred between markers *D11S908* and *D11S4142*. This suggests that there is a gene in this region that influences plasma levels of Hcy. The one-unit support interval surrounding this peak has a chromosomal location in the range of 112–118 cM from the p terminus. This region contains the gene *NNMT,* which is involved in Hcy metabolism and is thus a candidate gene to explain the linkage signal.

Another chromosomal region produced a LOD score of 1.68 (nominal $P = .003$), suggesting linkage. The region was on chromosome 10q22 (fig. 1 and table 1), corresponding to marker *D10S537*. Such a LOD score would be expected to occur by chance approximately once every genome scan and must be considered suggestive until supported by other evidence. Again, there is a gene within this region that is involved in Hcy metabolism and is a candidate to explain the observed linkage signal: the methionine adenosyltransferase α I gene (*MATαI*).

The maximum LOD score found in the region of the *MTHFR* gene on chromosome 1p36.3 was 1.03 (nominal $P = .015$) (fig. 1 and table 1). These multipoint results show only a marginal increase from the LOD score of 0.66 obtained in the two-point linkage analysis with only the 677 C→T polymorphism as a marker.

Because our previous studies have suggested that Hcy levels are genetically correlated ($\rho_G = 0.65$) with the sus-

Table 2

Genotyped SNPs within *NNMT* and Results from the Association Analysis with Hcy Plasma Levels

SNP	Celera ID[a]	dbSNP ID	Location	Position[b] (bp)	Allele Frequencies[c]	*P*[d]
1	New[e]	*rs35077908*	Promoter	113,632,968	81/19	.38
2	c1058033	*rs505978*	Intron 1	113,634,267	76/24	.64
3	c1058034	*rs683271*	Intron 1	113,634,355	77/23	.65
4	***c2134722***	***rs694539***	Intron 1	113,638,629	85/15	***.017***
5	c2134725	*rs11214921*	Intron 2	113,642,022	19/81	.16
6	c2134727	*rs10891641*	Intron 2	113,644,291	21/79	.81
7	c10096025	*rs2852432*	Intron 2	113,669,214	27/73	.49
8	c10096016	*rs2852447*	Intron 4	113,684,911	75/25	.29
9	New[e]	*rs4646337*	3′ UTR	113,688,519	79/21	.98
10	c10096011	*rs11569688*	3′ UTR	113,688,653	18/81	.64

NOTE.—The SNP with a significant *P* value ($P < .05$) is shown in bold italics.

[a] Celera database available at the Applied Biosystems, myScience, Web site.

[b] The SNP position from the p terminus of chromosome 11.

[c] Estimated allele frequencies in the Spanish population, given as percentages (allele 1/allele 2), with family relationships taken into account.

[d] *P* value of association with Hcy levels. Values are not corrected for the number of markers tested.

[e] "New" indicates that the SNP was found in the present study and was not previously reported.

ceptibility to thrombosis (Souto et al. 2000*a*), we performed a bivariate linkage analysis. Bivariate analyses of related phenotypes have been shown to increase the power to detect linkage (Soria et al. 2002, 2003). In the combined analyses of lnHcy levels and thrombosis, the three LOD scores on chromosomes 11, 10, and 1 remained at their previous levels, with no significant QTL effects on susceptibility, providing no evidence that these QTLs influence thrombosis risk. Nevertheless, these results do not allow the possibility to be excluded.

Screening for SNPs in NNMT *and Haplotype Association Analysis*

We identified two novel SNPs. One (SNP 1 in table 2) is in the promoter region, and the other (SNP 9 in table 2) is in the 3′ UTR. These 2 identified SNPs, together with 10 SNPs previously reported in 2003 (Celera database), were genotyped for the whole GAIT sample. Of these 12 SNPs, 2 were excluded from the analysis because they correlated (i.e., showed perfect linkage disequilibrium) with other SNPs. The details of the 10 SNPs included in our analysis are given in table 2.

The measured-genotype association analysis revealed significant association between one of the SNPs (Celera accession number c2134722) and Hcy levels (table 2), supporting the presence of a QTL in the region of *NNMT*. In addition, in the GAIT sample, the 10 SNPs in the *NNMT* gene generated 31 haplotypes (table 3). In the measured-genotype association analysis, one of the haplotypes (Hap 18) was strongly associated with Hcy levels ($P = .00001$), even after correction for multiple testing (corrected $P = .00031$). Two other haplotypes (Hap 22 and Hap 31) showed an association with Hcy levels ($P = .01$ and .04, respectively), but the statistical significance was lost when the results were corrected for multiple testing. All carriers of Hap 18 ($n = 12$) were heterozygous and had a mean Hcy level ($\pm$ SD) of 13.2 $\pm$ 2.9 μmol/liter.

Discussion

A major challenge for medicine in the postgenomic era is to identify genetic variants that affect the risk of developing complex diseases such as cardiovascular or Alzheimer disease. It is usually extremely difficult to find these QTLs when the study focuses directly on disease status (i.e., disease or no disease). However, intermediate risk factors or risk markers can be analyzed by a genetic search. According to recent reports of epidemiological data, the Hcy plasma level is an independent intermediate risk marker for thromboembolic disease (both arterial and venous) (Welch and Loscalzo 1998; Cattaneo 1999; Homocysteine Studies Collaboration 2002; McIlroy et al. 2002; Wald et al. 2002), congestive heart failure (Vasan et al. 2003), Alzheimer disease (Morris et al. 2001; McIlroy et al. 2002; Seshadri et al. 2002; Morris 2003), and osteoporotic fractures (McLean et al. 2004; van Meurs et al. 2004).

In humans, the sole source of Hcy is the demethylation of methionine through the intervention of different methyltransferases. This pathway has several steps (fig. 2). First, methionine and ATP originate S-adeno-

Table 3

Haplotypes Observed within the *NNMT* Gene in the GAIT Sample, in Order of Frequency

Haplotype[a]	Hap ID	Frequency[b]	Hcy Level[c] (μmol/liter)	*P*[d]
1111222112	Hap 7	318	7.60	.69
2111222221	Hap 30	84	7.51	.99
1222111112	***Hap 22***	73	***6.94***	***.01***
2111222112	Hap 28	58	7.22	.87
1111221112	Hap 4	56	7.92	.89
1221222112	Hap 20	37	7.67	.59
1111222221	Hap 9	30	7.55	.69
1111221221	Hap 6	16	8.30	.09
1221211112	***Hap 18***	12	***13.18***	***.00001***
1222111212	Hap 24	9	7.97	.58
1111221212	Hap 5	8	8.55	.48
1221112112	Hap 16	8	7.64	.77
2111221212	Hap 27	8	9.80	.07
1112112112	Hap 13	7	6.66	.95
1111222212	Hap 8	5	6.42	.68
1112111221	Hap 12	5	8.00	.40
1111211222	Hap 3	4	6.22	.65
1112111212	Hap 11	3	7.10	.85
1221111112	Hap 14	3	10.76	.22
1221111222	Hap 15	3	7.37	.52
1221112221	Hap 17	3	7.00	.75
1221222221	Hap 21	3	6.00	.42
2111222211	Hap 29	3	8.20	.68
2112111112	***Hap 31***	3	***5.10***	***.04***
1111122112	Hap 2	2	10.00	.39
1111222222	Hap 10	2	9.85	.57
1221211222	Hap 19	2	5.90	.85
1222112112	Hap 26	2	5.50	.20
1111111221	Hap 1	1	7.70	.78
1222111122	Hap 23	1	3.90	.38
1222111221	Hap 25	1	8.20	.86

NOTE.—The haplotypes with significant *P* values ($P < .05$) are shown in bold italics.

[a] The composition of the haplotypes follows the same order of SNPs as in table 2.

[b] Frequency is given as the number of chromosomes containing the haplotype. There were only four haplotypes with homozygous carriers: Hap 7, Hap 20, Hap 22, and Hap 28. No statistically significant differences in Hcy levels between heterozygous and homozygous carriers were observed.

[c] Crude mean Hcy levels among haplotype carriers, without accounting for the nonindependence of family members.

[d] *P* value of association with Hcy levels. Values are not corrected for the number of haplotypes tested.

sylmethionine (SAM) by means of the enzyme methionine adenosyltransferase. Then, SAM is converted to S-adenosylhomocysteine (SAH) in a reaction catalyzed by SAM-dependent methyltransferases. This reaction originates a methyl group that is used for many different biochemical pathways through, in humans, at least 39 distinct methyltransferases (Clarke and Banfield 2001). SAH is subsequently broken down to form adenosine and Hcy that is remethylated to form methionine. Alternatively, Hcy can be converted to cysteine through the transsulfuration pathway.

The majority of potential candidate genes, which encode enzymes or proteins involved in Hcy metabolism (fig. 2 and table 1), did not produce a linkage signal for Hcy levels. The results of our genome scan suggest that the major QTL influencing variation in Hcy levels is located on chromosome 11q23 (fig. 1). Interestingly, this region has been linked to the risk of venous thrombosis in a genome scan performed with a single large pedigree with thrombosis and protein C deficiency (Hasstedt et al. 2004). This genetic protein C deficiency appeared insufficient to explain completely the venous thrombosis risk in the family. For this reason, Hasstedt et al. (2004) undertook a genomewide search of regions potentially harboring other genetic factors. In the search, a peak of significance was found at 113 cM from the p terminus of chromosome 11, just within our support interval (112–118 cM) for the Hcy QTL. We also investigated the possible pleiotropic effect of this QTL on the risk of thrombosis. The linkage signal in the bivariate linkage analysis was not significantly increased, and no significant QTL effect on susceptibility was observed. Thus, we can not directly replicate the observation by Hasstedt et al. (2004) of a thrombosis QTL on chromosome 11. However, lack of statistical power could be the reason for the failure to confirm this finding. A biologically plausible candidate gene, *NNMT,* is located within this region. NNMT is an enzyme involved in the synthesis of SAH, the precursor of Hcy, and it catabolizes nicotinamide and other pyridine compounds in a reaction that uses the methyl group generated during the conversion of SAM to SAH (Aksoy et al. 1994) (fig. 2). NNMT is widely synthesized in most tissues, especially in the liver (Yan et al. 1999). An important individual variability in the liver activity of this enzyme has been described, suggesting a trimodal distribution (Smith et al. 1998; Yan et al. 1999). In the absence of functional polymorphisms in the coding region of *NNMT,* this distribution has been shown to depend on variation at mRNA and protein levels (Smith et al. 1998; Yan et al. 1999).

To test whether *NNMT* variability accounts for the observed linkage signal for Hcy plasma levels, we studied 10 SNPs within this gene. One of them showed significant association with Hcy levels ($P = .017$), although this result was no longer significant when corrected for the testing of 10 markers. Moreover, a haplotype built with the 10 SNPs within *NNMT* was more strongly associated with higher Hcy levels than were the remaining haplotypes ($P = .00031$, corrected for multiple testing). Although we have not yet identified the functional variants that are responsible for the observed effects, the significant associations of the intragenic SNP and haplotype, in conjunction with the linkage signal in this QTL, strongly support the hypothesis that genetic variants in *NNMT* are the main determinants of plasma

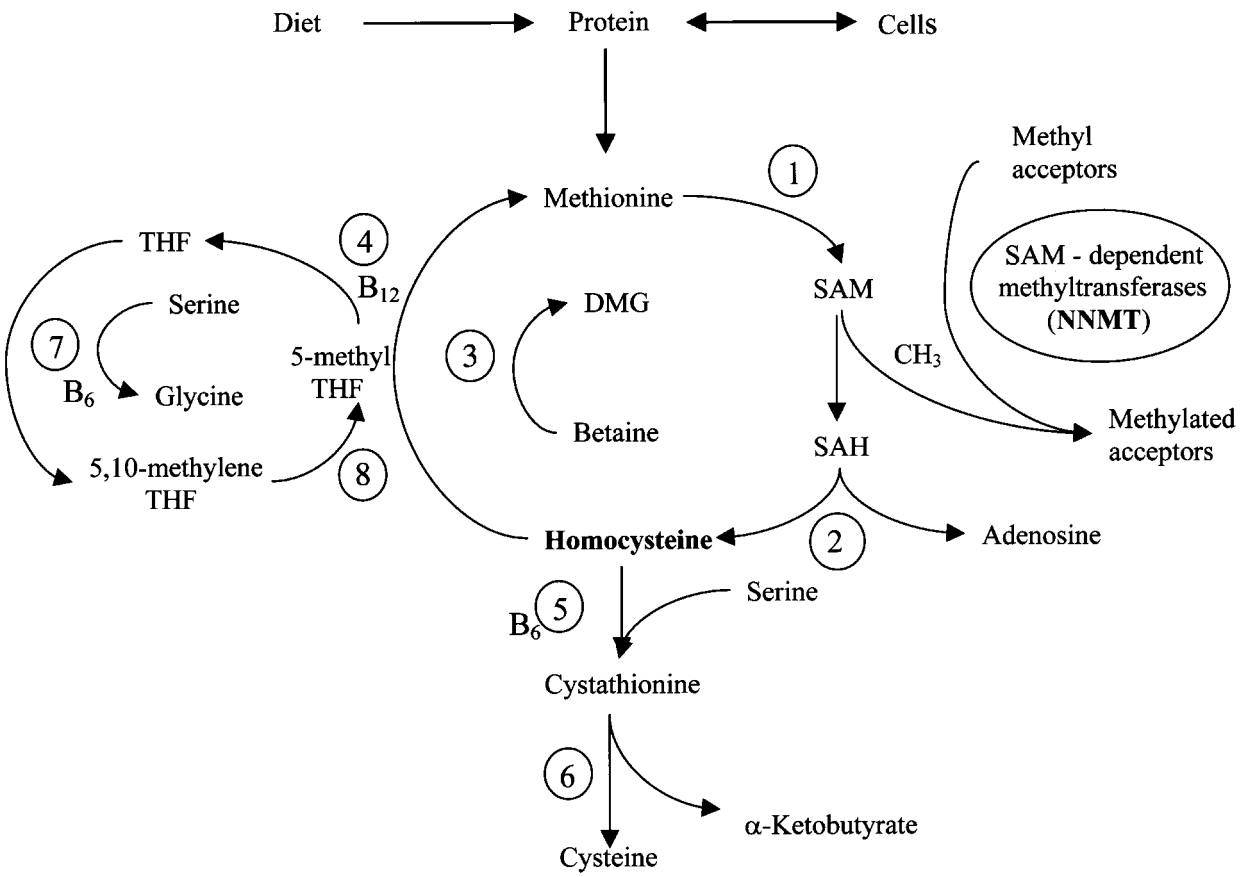

Figure 2 Metabolic pathways involved in Hcy regulation. Enzymes are numbered as follows: (1) methionine adenosyltransferase, (2) SAH hydrolase, (3) betaine-homocysteine methyltransferase, (4) 5-methyltetrahydrofolate-homocysteine methyltransferase, (5) cystathionine β-synthase, (6) cystathionine γ-lyase, (7) serine hydroxymethyltransferase, and (8) 5,10-methylenetetrahydrofolate reductase. B_6 = vitamin B_6; B_{12} = vitamin B_{12}; CH_3 = methyl group; DMG = dimethylglycine; THF = tetrahydrofolate. NNMT is one of at least 39 SAM-dependent methyltransferases described in mammals (Clarke and Banfield 2001). The gene encoding NNMT is located on chromosome 11q23, at the same QTL that has been linked to plasma Hcy levels.

Hcy variability in the GAIT sample. It is very probable that these variants will be located within noncoding regions, and they may influence the regulation of transcription. These regions were not included in our sequencing because of the huge size of the *NNMT* gene, which spans >50 kb. In support of the involvement of methyltransferases in plasma Hcy concentration, it is worth mentioning a recent report that demonstrated that mice deficient in another methyltransferase have a 50% decrease in plasma Hcy (Noga et al. 2003). Thus, we hypothesize a direct relationship between NNMT activity and plasma Hcy concentration.

The second-highest LOD score obtained in our genome scan is located on chromosome 10q22. This locus contains *MATαI,* a gene encoding another enzyme involved in Hcy formation (fig. 2). In the context of a genomewide exploration, a LOD score of 1.68 would be merely suggestive of linkage, but the presence of an obvious candidate gene reinforces the necessity of refining the search in this region. We genotyped three SNPs within *MATαI* that were reported in 2003 (Celera database), but none was significantly associated with Hcy levels (data not shown). Despite the attractiveness of *MATαI* as a candidate, the present study did not prove that it has an effect on Hcy levels. Nevertheless, an absence of association with some SNPs does not rule out implication of the gene in determining the phenotype. Other SNPs in the same gene that are not in linkage disequilibrium with those analyzed might be responsible for the linkage signal.

We have also examined the widely reported influence of the *MTHFR* 677 C→T genotype on Hcy levels. First, we used a two-point linkage analysis to assess the LOD score between this marker and the phenotype. We also explored this linkage in the genomewide scan. We confirmed previous findings of a direct effect of *MTHFR* on Hcy levels (Husemoen et al. 2003; Kluijtmans et al. 2003). However, our results indicate that the *MTHFR*

677 C→T polymorphism is not the major genetic determinant of the quantitative variation in Hcy plasma levels.

In conclusion, we found evidence that the *NNMT* gene, located in the region 11q23, influences plasma Hcy levels. It is probably the major genetic determinant of this phenotype in the Spanish population. It is possible that another gene on chromosome 10q22 also influences Hcy levels. Given the previous reports about the relevance of Hcy in certain diseases, further studies are warranted. These should include the exhaustive dissection (including exons, introns, and the 5′ and 3′ regulatory regions) of the *NNMT* gene. This task will require a high-throughput resequencing and genotyping effort, combined with advanced statistical genetic analyses, to determine the most likely functional variants.

Acknowledgments

This study was supported by National Institutes of Health (NIH) grant HL70751; Fondo de Investigación Sanitaria (FIS), Spanish Health Ministry, grants FIS PI020375 and FIS C03-08; grant SAF 2002-03449, partially supported by European Regional Development Funds (FEDER) (Spanish Ministry of Science and Technology); and grant RETICS C03/01 (Cardiovascular Network, FIS, Spanish Health Ministry). J.M.S. is supported by FIS 99/3048. A.B. is supported by FIS 01/A046. Statistical genetic analysis was supported by NIH grants MH59490 and HL70751. Fundació "La Caixa" and Fundació d'Investigació Sant Pau also contributed to the project. We are grateful to Rosa Arcelús, for homocysteine determinations, and Dr. Carlos Mª Díaz, for his help with *NNMT* sequencing. Finally, we are deeply grateful to all of the families who participated in the study.

Electronic-Database Information

The accession number and URLs for data presented herein are as follows:

Applied Biosystems, myScience, Web site, http://myscience.appliedbiosystems.com/ (for Celera database and *NNMT* SNPs)

dbSNP, http://www.ncbi.nlm.nih.gov/SNP/ (for *NNMT* SNPs)

Marshfield Center for Medical Genetics, http://research.marshfieldclinic.org/genetics/ (for the marker map used in multipoint linkage analysis)

Southwest Foundation for Biomedical Research, http://www.sfbr/org/ (for SOLAR program package)

UniGene, http://www.ncbi.nlm.nih.gov/UniGene/ (for *NNMT* [accession number Hs.364345] and other genes involved in Hcy metabolism)

References

Abecasis GR, Cherny SS, Cookson WO, Cardon LR (2002) Merlin-rapid analysis of dense genetic maps using sparse gene flow trees. Nat Genet 30:97–101

Aksoy S, Szumlanski CL, Weinshilboum RM (1994) Human liver nicotinamide N-methyltransferase: cDNA cloning, expression, and biochemical characterization. J Biol Chem 269:14835–14840

Almasy L, Blangero J (1998) Multipoint quantitative-trait linkage analysis in general pedigrees. Am J Hum Genet 62:1198–1211

——— (2004) Exploring positional candidate genes: linkage conditional on measured genotype. Behav Genet 34:173–177

Blanco-Vaca F, Arcelús R, González-Sastre F, Ordóñez-Llanos J, Queraltó-Compañó JM (2000) Comparison of the Abbott Imx and a high-performance liquid chromatography method for measuring total plasma homocysteine. Clin Chem Lab Med 38:327–329

Blangero J, Williams JT, Almasy L (2000) Robust LOD scores for variance component-based linkage analysis. Genet Epidemiol 19:S8–S14

Boehnke M, Lange K (1984) Ascertainment and goodness of fit of variance component models for pedigree data. Prog Clin Biol Res 147:173–192

Brattsström L, Lindgren A, Israelsson B, Andersson A, Hultberg B (1994) Homocysteine and cysteine: determinants of plasma levels in middle-aged and elderly subjects. J Intern Med 236:633–641

Cattaneo M (1999) Hyperhomocysteinemia, atherosclerosis and thrombosis. Thromb Haemost 81:165–176

Clarke S, Banfield K (2001) S-adenosylmethionine-dependent methyltransferases. In: Carmel R, Jacobsen DW (eds) Homocysteine in health and disease. Cambridge University Press, Cambridge, United Kingdom, pp 63–78

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

Hasstedt SJ, Scott BT, Callas PW, Vossen CY, Rosendaal FR, Long GL, Bovill EG (2004) Genome scan of venous thrombosis in a pedigree with protein C deficiency. J Thromb Haemost 2:868–873

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

Husemoen LLN, Thomsen TF, Fenger M, Jørgensen HL, Jørgensen T (2003) Contribution of thermolabile methylenetetrahydrofolate reductase variant to total plasma homocysteine levels in healthy men and women. Inter99 (2). Genet Epidemiol 24:322–330

Jee SH, Song KS, Shim WH, Kim HK, Suh I, Park JY, Won SY, Beaty TH (2002) Major gene evidence after MTHFR-segregation analysis of serum homocysteine in families of patients undergoing coronary arteriography. Hum Genet 111:128–135

Kluijtmans LA, Young IS, Boreham CA, Murray L, McMaster D, McNulty H, Strain JJ, McPartlin J, Scott JM, Whitehead AS (2003) Genetic and nutritional factors contributing to hyperhomocysteinemia in young adults. Blood 101:2483–2488

McIlroy SP, Dynan KB, Lawson JT, Patterson CC, Passmore AP (2002) Moderately elevated plasma homocysteine,

methylenetetrahydrofolate reductase genotype, and risk for stroke, vascular dementia, and Alzheimer disease in Northern Ireland. Stroke 33:2351–2356

McLean RR, Jacques PF, Selhub J, Tucker KL, Samelson EJ, Broe KE, Hannan MT, Cupples LA, Kiel DP (2004) Homocysteine as a predictive factor for hip fracture in older persons. N Engl J Med 350:2042–2049

Morris MS (2003) Homocysteine and Alzheimer disease. Lancet Neurol 2:425–428

Morris MS, Jacques PF, Rosenberg IH, Selhub J (2001) National Health and Nutrition Examination Survey. Hyperhomocysteinemia associated with poor recall in the third National Health and Nutrition Examination Survey. Am J Clin Nutr 73:927–933

Noga AA, Stead LM, Zhao Y, Brosnan ME, Brosnan JT, Vance DE (2003) Plasma homocysteine is regulated by phospholipid methylation. J Biol Chem 278:5952–5955

Seshadri S, Beiser A, Selhub J, Jacques PF, Rosenberg IH, D'Agostina RB, Wilson PWF, Wolf PA (2002) Plasma homocysteine as a risk factor for dementia and Alzheimer disease. N Engl J Med 346:476–483

Smith ML, Burnett D, Bennet P, Waring RH, Brown HM, Williams AC, Ramsden DB (1998) A direct correlation between nicotinamide N-methyltransferase activity and protein levels in human liver cytosol. Biochim Biophys Acta 1442:238–244

Sobel E, Lange K (1996) Descent graphs in pedigree analysis: applications to haplotyping, location scores, and marker sharing statistics. Am J Hum Genet 58:1323–1337

Soria JM, Almasy L, Souto JC, Bacq D, Buil A, Faure A, Martínez-Marchán E, Mateo J, Borrell M, Stone W, Lathrop M, Fontcuberta J, Blangero J (2002) A quantitative-trait locus in the human factor XII gene influences both plasma factor XII levels and susceptibility to thrombotic disease. Am J Hum Genet 70:567–574

Soria JM, Almasy L, Souto JC, Buil A, Martínez-Marchán E, Mateo J, Borrell M, Stone WH, Lathrop M, Fontcuberta J, Blangero J (2003) A new locus on chromosome 18 that influences normal variation in activated protein C resistance phenotype and factor VIII activity and its relation to thrombosis susceptibility. Blood 101:163–167

Souto JC, Almasy L, Borrell M, Blanco-Vaca F, Mateo J, Soria JM, Coll I, Felices R, Stone W, Fontcuberta J, Blangero J (2000*a*) Genetic susceptibility to thrombosis and its relationship with physiological risk factors: the GAIT study. Am J Hum Genet 67:1452–1459

Souto JC, Almasy L, Borrell M, Garí M, Martínez E, Mateo J, Stone WH, Blangero J, Fontcuberta J (2000*b*) Genetic determinants of hemostasis phenotypes in Spanish families. Circulation 101:1546–1551

van Meurs JBJ, Dhonukshe-Rutten RAM, Pluijm SMF, van der Klift M, de Jonge R, Lindemans J, de Groot LC, Hofman A, Witteman JC, van Leeuwen JP, Breteler MM, Lips P, Pols HA, Uitterlinden AG (2004) Homocysteine levels and the risk of osteoporotic fracture. N Engl J Med 350:2033–2041

Vasan RS, Beiser A, D'Agostino RB, Levy D, Selhub J, Jacques PF, Rosenberg IH, Wilson PWF (2003) Plasma homocysteine and risk for congestive heart failure in adults without prior myocardial infarction. JAMA 289:1251–1257

Wald DS, Law M, Morris JK (2002) Homocysteine and cardiovascular disease: evidence on causality from a meta-analysis. BMJ 325:1202–1208

Welch GN, Loscalzo J (1998) Homocysteine and atherothrombosis. N Engl J Med 338:1042–1050

Williams JT, Van Eerdewegh P, Almasy L, Blangero J (1999) Joint multipoint linkage analysis of multivariate qualitative and quantitative traits. I. Likelihood formulation and simulation results. Am J Hum Genet 65:1134–1147

Yan L, Otterness DM, Weinshilboum RM (1999) Human nicotinamide N-methyltransferase pharmacogenetics: gene sequence analysis and promoter characterization. Pharmacogenetics 9:307–316