DMD Fast Forward. Published on July 17, 2012 as doi: 10.1124/dmd.112.046953

Human Liver Methionine Cycle: *MAT1A* and *GNMT* Gene Resequencing, Functional Genomics and Hepatic Genotype-Phenotype Correlation

Yuan Ji, Kendra K.S. Nordgren, Yubo Chai, Scott J. Hebbring, Gregory D. Jenkins, Ryan P. Abo, Yi Peng, Linda L. Pelleymounter, Irene Moon, Bruce W. Eckloff, Xiaoshan Chai, Jianping Zhang, Brooke L. Fridley, Vivien C. Yee, Eric D. Wieben and Richard M. Weinshilboum

From the Division of Clinical Pharmacology, Department of Molecular Pharmacology and Experimental Therapeutics (Y.J., K.K.S.N., Y.C., S.H., R.A., L.L.P., I.M., X.C., J.Z., R.M.W.), Division of Biomedical Statistical and Informatics, Department of Health Sciences Research (G.D.J., B.L.F.), Department of Biochemistry and Molecular Biology (B.W.E., E.D.W.), Mayo Clinic, Rochester, MN; Department of Biochemistry, Case Western Reserve University, Cleveland, OH (Y.P. and V.C.Y.); College of Pharmacy, Jinan University, Guangzhou, PR China (J.Z.)

Running title: *MAT1A* and *GNMT* Sequence Variation and Functional Genomics

Address correspondence and reprint requests to Richard M. Weinshilboum, MD, Department of Molecular Pharmacology and Experimental Therapeutics, Mayo Clinic, 200 First Street SW, Rochester, MN 55905. Telephone: (507) 284-2246, Fax: (507) 284-4455, E-mail: weinshilboum.richard@mayo.edu

Text pages	24
Number of Tables	4
Number of Figures	6
References	46
Words in **Abstract**	247
Words in **Introduction**	507
Words in **Discussion**	873

Abbreviations: AdoMet, S-adenosylmethionine; AdoHcy, S-adenosylhomocysteine; FR, flanking region; ORF, open reading frame; WT, wild type; UTR, untranslated region; EA, European-American; AA, African-American; HCA, Han Chinese-American; ns, nonsynonymous; BHMT, betaine homocysteine methyltransferase; SHMT, serine hydroxymethyltransferase

DMD Fast Forward. Published on July 17, 2012 as DOI: 10.1124/dmd.112.046953
This article has not been copyedited and formatted. The final version may differ from this version.

ABSTRACT

The "Methionine Cycle" plays a critical role in the regulation of concentrations of S-adenosylmethionine (AdoMet), the major biological methyl donor. We set out to study sequence variation in genes encoding the enzyme that synthesizes AdoMet in liver, methionine adenosyltransferase 1A (MAT1A), and the major hepatic AdoMet utilizing enzyme, glycine N-methyltransferase (GNMT), as well as functional implications of that variation. We resequenced *MAT1A* and *GNMT* using DNA from 288 subjects of three ethnicities, followed by functional genomic and genotype-phenotype correlation studies performed with 268 hepatic biopsy samples. We identified 44 and 42 polymorphisms in *MAT1A* and *GNMT*, respectively. Quantitative Western blot analyses for the human liver samples showed large individual variation in MAT1A and GNMT protein expression. Genotype-phenotype correlation identified two genotyped single nucleotide polymorphisms (SNPs), rs9471976 (corrected $p = 3.9 \times 10^{-10}$) and rs11752813 (corrected $p = 1.8 \times 10^{-5}$), and 42 imputed SNPs surrounding *GNMT* that were significantly associated with hepatic GNMT protein levels (corrected p-values < 0.01). Reporter gene studies showed that variant alleles for both genotyped SNPs resulted in decreased transcriptional activity. Correlation analyses among hepatic protein levels for Methionine Cycle enzymes showed significant correlations between GNMT and MAT1A ($p = 1.5 \times 10^{-3}$), and between GNMT and betaine homocysteine methyltransferase ($p = 1.6 \times 10^{-7}$). Our discovery of SNPs that are highly associated with hepatic GNMT protein expression as well as the "coordinate regulation" of Methionine Cycle enzyme protein levels provide novel insight into the regulation of this important human liver biochemical pathway.

Introduction

AdoMet, the methyl donor for most biological methylation reactions (Cantoni, 1951a; Cantoni, 1951b), is synthesized from methionine and ATP by methionine adenosyltransferase (MAT) (Fontecave et al., 2004). It is then utilized as a methyl donor for reactions catalyzed by methyltransferase (MT) enzymes to produce methylated compounds and S-adenosylhomocysteine (AdoHcy) (Clarke et al., 2003). AdoHcy is subsequently converted to homocysteine, which can either be remethylated to form methionine, completing the "Methionine Cycle", or converted to cysteine and glutathione (GSH) by the transsulfuration pathway (Fig. 1). The "Methionine and Folate Cycles" shown in Figure 1 have been implicated in the pathophysiology of diseases as diverse as cancer (Weinstein et al., 2006; Kasperzyk et al., 2009; Maruti et al., 2009; Stevens et al., 2010), cardiovascular disease (Arnesen et al., 1995; Frosst et al., 1995; Nygard et al., 1995; Clarke et al., 2003) and psychiatric illness (Smythies et al., 1997; Main et al., 2010).

The human liver expresses several unique Methionine Cycle enzymes (Fig. 1). Specifically, *MAT2A* encodes the enzyme that catalyzes AdoMet biosynthesis in non-hepatic tissues and in fetal liver, but *MAT1A* is expressed only in adult liver and encodes the enzyme that catalyzes AdoMet synthesis in the liver (Gil et al., 1996; Mato et al., 1997). Once formed, AdoMet is utilized by MT enzymes to generate methylated compounds, with AdoHcy as a reaction product. The tetrameric enzyme glycine N-methyltransferase (GNMT, EC 2.1.1.20), is highly expressed in liver (1-3% of total soluble protein) and poorly expressed in other tissues except the prostate and pancreas (Kerr, 1972; Heady and Kerr, 1973). GNMT catalyzes the methylation of glycine to form sarcosine and AdoHcy (Blumenstein and Williams, 1960). The biological role of sarcosine is not well understood, but a recent metabolomic study identified

sarcosine as a metabolic biomarker for the progression of prostate cancer (Sreekumar et al., 2009). There is also evidence that the reaction catalyzed by GNMT is a major factor regulating hepatic AdoMet concentrations (Mudd and Poole, 1975; Mudd et al., 1980; Balaghi et al., 1993). In addition to MAT1A and GNMT, betaine homocysteine methyltransferase (BHMT) is another Methionine Cycle enzyme that is expressed primarily in the liver (Li et al., 2008). BHMT catalyzes the hepatic remethylation of homocysteine to form methionine by transferring a methyl group from betaine to homocysteine (Skiba et al., 1982) (Fig. 1).

Rare mutations in *MAT1A* have been associated with persistent hypermethioninemia without elevation in either circulating homocysteine or tyrosine concentrations (Chamberlin et al., 2000). *GNMT* mutations have also been linked to hypermethioninemia (Mudd et al., 2001; Luka et al., 2002; Augoustides-Savvopoulou et al., 2003). These reports raise the possibility that common DNA sequence variation in the *MAT1A* and *GNMT* genes might modulate hepatic AdoMet concentrations and, as a result, hepatic methylation (Fig. 1). In the present study, we set out to systematically resequence both genes, followed by functional genomic studies and hepatic genotype-phenotype correlation analyses in an attempt to identify functional variants in *MAT1A* and *GNMT* that might contribute to variation in the regulation of AdoMet metabolism as well as Methionine and Folate Cycle function.

Materials and Methods

DNA Samples and Gene Resequencing. DNA samples for gene resequencing were obtained from the Coriell Cell Repository (Camden, NJ). Specifically, "Human Variation Panel" samples from 96 European-American (EA), 96 African-American (AA), and 96 Han Chinese-American (HCA) subjects (sample-sets HD100CAU, HD100AA, and HD100CHI, respectively)

were used in the resequencing studies. These samples had been collected, anonymized and deposited by the National Institute of General Medical Sciences. All subjects had provided written informed consent for the use of their DNA for research purposes. Our studies were reviewed and approved by the Mayo Clinic Institutional Review Board. Details of the DNA sequencing methods have been described previously (Ji et al., 2007). Briefly, all *MAT1A* exons, intron-exon splice junctions and ~1 kb of both 5'- and 3'-flanking regions (FRs) were amplified using the PCR, and the amplicons were sequenced on both strands in the Mayo Clinic Molecular Biology Core Facility using dye terminator sequencing chemistry. Because *GNMT* is a much smaller than *MAT1A*, the entire *GNMT* genes as well as approximately 1 kb of its 5'- and 3'-FRs were sequenced. Accession numbers for the reference sequences used were NM_000429.2. for *MAT1A* and NM_018960.4 for *GNMT*. Sequence chromatograms were analyzed using Mutation Surveyor®. Sequences of primers used to perform the PCR amplifications and gene resequencing studies are listed in Supplemental Table 1.

Functional Characterization of *MAT1A* Nonsynonymous (ns) SNPs. Mammalian expression constructs were created for wild-type MAT1A by subcloning the open reading frame (ORF) of *MAT1A* from the Origene clone SC119881 (Origene, Rockville, MD) into pcDNA^{TM}3.1D/V5-His-TOPO (Invitrogen, Carlsbad, CA) in frame with the V5-His tag, and site-directed mutagenesis was used to create variant allozyme expression constructs. These expression constructs were used to transfect COS-1 cells to obtain recombinant MAT1A allozymes for use in quantitative Western blot analyses and for the assay of allozyme enzyme activity. Bacterial expression constructs were also created for MAT1A wild type (WT) and variant allozymes and were expressed in BL21 *E.coli*. This bacterially expressed MAT1A was used to perform substrate kinetic experiments, as described previously (Wang et al., 2003).

Structural analysis of MAT1A allozymes was performed by using the 2.1 Å resolution crystal structure of human MAT1A bound to AdoMet (PDB accession code 2OBV). Visualization and analysis of the MAT1A structures and the computational "mutation" of Glu238Lys was carried out using the graphics program COOT (Purcell et al., 2007). Additional details with regard to the MAT1A structural analysis are described (see Supplemental Data).

MAT1A and GNMT Western Blot Analyses of Human Hepatic Biopsy Samples. Two hundred and sixty-eight human adult liver surgical biopsy samples were obtained at the Mayo Clinic in Rochester, MN. These samples were from Caucasian women undergoing medically indicated surgery. Tissue samples from only one sex were used to eliminate the possibility of confusion as a result of sex-dependent differences in enzyme protein expression. Characteristics of the patients from whom the biopsies were obtained have been described previously (Feng et al., 2009; Zhang et al., 2009; Nordgren et al., 2011). Use of these anonymized surgical biopsy samples was reviewed and approved by the Mayo Clinic IRB. Cytosol extracted from the hepatic tissue was stored at -80ºC prior to use.

Quantitative Western blot analyses were performed for both MAT1A and GNMT using the 268 liver cytosol preparations. A rabbit polyclonal antibody generated by Cocalico Biologicals, Inc. (Reamstown, PA), against MAT1A amino acids 208-228and a commercial mouse GNMT polyclonal antibody (Sigma-Aldrich, St. Louis, MO) were used to perform these studies. Purified His-tagged protein standards for MAT1A (50 ng) as well as a pooled sample of hepatic high-speed supernatant as a standard for GNMT were loaded on each gel and were stained for either MAT1A or GNMT protein. Levels of endogenous β-actin were also assayed for each gel and were used as a loading control. Specifically, each of the 268 liver cytosol samples was loaded in triplicate, and values for MAT1A and GNMT immunoreactive protein for the triplicate

DMD Fast Forward. Published on July 17, 2012 as DOI: 10.1124/dmd.112.046953
This article has not been copyedited and formatted. The final version may differ from this version.

samples were calculated based on the relative intensity of protein bands on the gel when compared with an appropriate protein standard assayed on the same gel.

Tag SNP Selection and Genotyping of Hepatic Biopsy DNA. Genotype data from a variety of sources were utilized to select tag SNPs for genes encoding proteins in the Methionine and Folate Cycles (see Fig. 1), including *MAT1A* and *GNMT,* for use in genotyping DNA samples from the same 268 human liver biopsy samples that we had phenotyped for level of hepatic MAT1A and GNMT protein. Specifically, a total of 768 tag SNPs for Methionine and Folate Cycle genes, including 42 for MAT1A and 5 for GNMT, were selected using SNP genotype data from 168 unrelated HapMap Caucasian subjects (http://www.hapmap.org, data Rel 27/phase II + III) as well as data from our resequencing studies for these genes (Shield et al., 2004; Martin et al., 2006; Feng et al., 2009; Nordgren et al.) and SNP data obtained by using Illumina 550K and 510S genome-wide BeadChips and Coriell Institute-generated Affymetrix 6.0 genome-wide SNP genotypes for DNA from the 96 EA subjects included in the Coriell Institute "Human Variation Panel" (Camden, NJ), the same cell lines from which we obtained the DNA for our gene resequencing experiments. The SNPs were selected to tag across the genes and to include approximately 20 kb of flanking sequence, with $r^2 \geq 0.8$ and a minor allele frequency (MAF) ≥ 0.025. LDselect (Carlson et al., 2004) was used to perform tagging, and genotyping of DNA from the 268 hepatic biopsy samples was performed in the Mayo Genotyping Shared Resource utilizing Illumina GoldenGate technology (Illumina, San Diego, CA). Of the 42 *MAT1A* tag SNPs, 4 were excluded because of call rates < 94%, leaving 38 *MAT1A* SNPs and 5 *GNMT* SNPs for use in the genotype-phenotype correlation analysis.

Genotype-Phenotype Correlation Analysis and "1000 Genomes" Imputation. In addition to SNPs genotyped using the DNA samples, genotypes of untyped SNPs across both

MAT1A and *GNMT* were imputed using "1000 Genomes" and HapMap (phase 2 release 22) data as reference sets. Tag SNP genotypes in the 268 hepatic DNA samples were the genetic background upon which imputation was performed. The software package MaCH 1.0 (Li et al., 2006) was used to perform imputation. To increase the likelihood of detecting functionally important variants at a distance from the genes, variants within 200 kb on either side of the two genes were included in the imputation. Imputation quality estimates were determined by masking 10% of the genotypes at random and imputing the masked genotypes to compare actual and imputed masked genotypes. Estimated allelic dosage values for the imputed genotypes were then used, in addition to the genotyped SNPs, to perform the association analyses. Associations between MAT1A or GNMT protein levels and *MAT1A* or *GNMT* genotypes for both imputed and typed SNPs were performed by using Spearman rank correlations and were tested vs. zero using a Wald test. Genotype-phenotype correlations for *MAT1A* SNPs or *GNMT* SNPs with $\log_2$-transformed protein levels for the two genes were calculated using PLINK (Purcell et al., 2007). Pairwise Linkage Disequilibrium (LD) was determined by using SNAP (Johnson et al., 2008). Mapping of transcription factors for SNPs with low p-values for the association analysis was performed by using the Encyclopedia of DNA Elements (ENCODE) data on the UCSC genome browser website (http://genome.ucsc.edu/) (Boyle et al.).

GNMT Reporter Gene and qRT-PCR Assays. Reporter gene assays were used to functionally characterize SNPs with low p-values for the genotype-phenotype correlation analyses. Specifically, DNA sequences (200~300 bp) harboring the SNP loci selected for study were cloned into a pGL3-promoter luciferase reporter vector that contained an SV40 promoter upstream of the luciferase gene (Promega Corporation, Fitchburg, WI). One microgram of each reporter gene construct was then co-transfected into the human hepatocellular carcinoma HepG2

cell and LNCaP human prostate cancer cell lines (American Type Culture Collection, ATCC, Manassas, VA), with 20 ng of the pRL-TK renilla luciferase vector as a control for transfection efficiency, followed by dual-luciferase assay performed 24 hrs after transfection (Promega). Two independent transfection studies were performed for each reporter gene construct, with triplicate independent transfections for each construct in each experiment. Values for relative activity were expressed as a percentage of the pGL3-promoter activity for vectors without an insert. Comparisons were made between pGL3 reporter gene constructs containing WT and variant nucleotides at the SNP loci. DNA samples used to amplify SNP regions were selected from the Coriell Institute "Human Variation Panel" DNA samples because the genotypes of these samples were known based both on our gene resequencing studies and the GWAS genotyping data available for these samples. Sequences of primers that were used to amplify genomic regions containing the SNP selected for study during the reporter gene assay are listed in Supplemental Table 2.

Finally, qRT-PCR was used to determine the level of GNMT expression in a series of cell lines. Specifically, total RNA was isolated from LNCaP, HEK293T, PC-3, DU145 and HepG2 cells as well as frozen liver tissues by using Quick-RNA Mini Prep kit (Zymo Research, Irvine, CA). Real-time PCR was performed with the power SYBR Green RNA-to-C_T 1-step Kit (Applied Biosystems, Carlsbad, California) using GAPDH as a control gene. The PCR conditions were 1 cycle of 30 minutes at 48°C, 10 minutes at 95°C, 40 cycles of 15 seconds at 95°C, 1 minute at 60°C, followed by 1 cycle of 15 seconds at 95°C, 1 minute at 60°C, and 15 seconds at 95°C. Primers for performing qRT-PCR for GNMT and GAPDH were purchased from QIAGEN (Valencia, California; QT00026285 for GNMT and QT01192646 for GAPDH).

DMD Fast Forward. Published on July 17, 2012 as DOI: 10.1124/dmd.112.046953
This article has not been copyedited and formatted. The final version may differ from this version.

Results

***MAT1A* and *GNMT* Resequencing.** Sanger sequencing was used to resequence the exons, splice junctions, and ~1000 bp of both the 5'- and 3'- FRs of *MAT1A* as well as the entire *GNMT* gene, including ~1000 bp of 5'- and 3'-FR, using 288 DNA samples, 96 each from AA, EA, and HCA subjects (Tables 1 and 2 and Fig. 2). *MAT1A* resequencing identified 44 polymorphisms, 27 in EA, 36 in AA and 20 in HCA, including one nsSNP (G712>A, Glu238Lys). This SNP was present, heterozygous, in DNA from 3 HCA subjects. Eleven of these 44 polymorphisms were novel, i.e., they had not been deposited in dbSNP or assigned an rsID number. *GNMT* resequencing identified 42 polymorphisms, 22 in EA, 32 in AA and 19 in HCA samples. Unlike the situation with *MAT1A,* we did not identify any nsSNPs during *GNMT* resequencing. Twenty seven of the 42 *GNMT* polymorphisms were novel. All polymorphisms observed during our *MAT1A* and *GNMT* resequencing were in Hardy-Weinberg equilibrium ($p > 0.05$).

MAT1A Allozyme Functional Characterization. Our gene resequencing effort identified one *MAT1A* nsSNP (G712>A, Glu238Lys) with a minor allele frequency (MAF) of 0.016 in DNA from HCA subjects. Mammalian and bacterial expression constructs were created for both WT and the Lys238 variant to generate recombinant MAT1A allozymes as well as bacterially purified protein that could both be assayed for MAT1A allozyme activity and used to perform substrate kinetic experiments. Alterations in amino acid sequence as a result of genetic polymorphisms can have functional consequences either because of changes in protein quantity or altered substrate kinetics. However, the MAT1A Lys238 variant allozyme did not differ significantly from the WT protein in terms of either (see the MAT1A allozyme apparent *Km* values, enzyme activities and relative protein quantities listed in Table 3). In addition, structural analysis of the human MAT1A allozymes showed that the WT Glu238 residue was surface-

exposed with a side chain exposed to solvent. Substitution of Glu238 with the hydrophilic lysine residue present in the variant enzyme could be easily accommodated by WT protein folding. Finally, Glu238Lys is distant from both the dimer and tetramer interfaces and, therefore, would be unlikely to interfere with the formation of either oligomeric form of the protein. These structural predictions all supported our observations of a lack of functional implications for this variant allozymes (see structural analysis in the Supplemental Data).

Hepatic Genotype-Phenotype Correlation Analysis. To determine whether DNA sequence variation in the *MAT1A* and *GNMT* genes might play a role in variation in the expression of these proteins in the human hepatic tissue where both enzymes are predominately expressed, we assayed levels of MAT1A and GNMT protein expression in 268 adult human liver biopsy samples by performing quantitative Western blot analyses. Both enzymes showed large individual variation, with a 100- and 1000-fold variation in MAT1A and GNMT protein expression, respectively (Fig. 3). The data in Figure 3 show a Gaussian frequency distribution for levels of MAT1A protein (A) but a skewed distribution of GNMT protein levels, with many samples that displayed low levels and two "outlier" points with very high levels of protein (B). The figure also shows representative Western blots for both enzymes (Fig. 3C).

As the next step in our analysis, we genotyped 42 tag SNPs for *MAT1A* and 5 tag SNPs for *GNMT* using DNA samples from the same 268 subjects from whom the liver biopsy samples used to perform the Western blot analyses had been obtained. Overall call rate of these 42 SNPs was over 98.0%, but call rates were < 94% for 4 SNPs, so those SNPs were excluded from the subsequent analysis. In addition, using the genotyped SNPs as a scaffold, we performed imputation using "1000 Genomes" data across both genes out to 200 kb from both the 5'- and 3'-ends of the genes (1000 Genomes Project Consortium, 2010). Imputation identified an

additional 150 *MAT1A* and 61 additional *GNMT* SNPs with MaCH "Rsq" values (an estimate of the squared correlation between imputed and true genotypes) of over 0.3. Only SNPs for EA subjects were imputed because the liver biopsy samples had been obtained entirely from EA subjects.

Hepatic GNMT protein levels were significantly correlated with age ($r = -0.17$, $p = 0.006$), but that was not the case for MAT1A ($r = 0.06$, $p = 0.30$). Therefore, the GNMT protein levels used in the association analysis were adjusted for age. Two of the genotyped *GNMT* tag SNPs, rs9471976 and rs11752813, were significantly associated with GNMT protein levels, with p-values of 6.4×10^{-12} and 2.88×10^{-7}, ($p = 3.9 \times 10^{-10}$ and $p = 2.5 \times 10^{-7}$ after correction for multiple comparisons), respectively (Fig. 4A). The correlation of GNMT protein expression with genotype for these two SNPs is shown graphically in Figure 4C. Similar analyses performed with imputed *GNMT* SNPs identified 38 additional markers within 200 kb on either side of the gene that were significantly associated with GNMT protein levels ($p < 10^{-4}$), with rs9471974 being the most significant imputed SNP ($p = 3.6 \times 10^{-10}$). No *MAT1A* SNPs were significantly associated with MAT1A protein levels (minimum $p = 0.03$, $p = 1.5$ after correction for multiple comparisons, Fig. 4B). A list of all SNPs with p-values $< 10^{-4}$ for association with GNMT protein levels is summarized in Supplemental Table 3.

GNMT Reporter Gene Studies. *GNMT* resequencing had not identified any polymorphisms that altered the encoded amino acid sequence. Therefore, our functional genomic experiments for *GNMT* focused on SNPs that were significantly associated with hepatic GNMT protein levels. Specifically, reporter gene constructs were created for the two genotyped SNPs with the lowest p-values for association with GNMT protein levels (rs9471976 and rs11752813). The locations of these SNPs relative to the site of *GNMT* transcription initiation is

shown in Figure 5A. As a first step, we next surveyed a series of cell lines in an attempt to determine whether they might express GNMT. Specifically, we performed qRT-PCR with mRNA preparations from HepG2, HEK293T, SU86, PC-3, DU-145 and LNCaP cell lines as well as pooled human liver sample preparations. As shown in Figure 5B, when compared with the HepG2 cell preparations, only LNCaP showed a relatively high level of GNMT mRNA expression. Please note that the y-axis is a logarithmic scale. Therefore, we used LNCaP and HepG2 cells in our reporter gene studies – LNCap because it highly expressed GNMT and HepG2 because of their origin from hepatic tissue – the major organ known to highly express GNMT. Results from dual-luciferase assays performed with HepG2 cells showed that DNA sequences around both rs9471976 and rs11752813 could enhance pGL3-promoter activity by up to 6-fold (Fig. 5C). There were statistically significant differences in terms of their ability to drive transcription between WT and variant allele sequences for both SNPs in HepG2 and LNCap cells (Figs. 5C and 5D). The effect of variant SNP genotypes, in both cases, was consistent with the liver genotype-phenotype association results, i.e., the G allele in rs11752813 and the A allele in rs9471976 were both associated with lower levels of hepatic GNMT protein. However, since these two SNPs, rs9471976 and rs11752813, had been chosen for genotyping by using a SNP tagging strategy, they may not represent the functional SNPs, but rather may be in LD with other variants that have functional significance.

Coordinate Regulation of the Hepatic Folate-Methionine Cycle. Finally, in an attempt to begin to understand the overall regulation of the Methionine and Folate Cycles in human liver, we performed association analyses among levels of hepatic protein for other enzymes in these pathways that had also been assayed in the same 268 liver samples used to perform the present study. Specifically, we correlated BHMT and serine hydroxymethyltransferase 1 (SHMT1)

protein levels as well as the level of enzyme activity for another important methyltransferase enzyme, catechol O-methyltransferase (COMT) (Feng et al., 2009; Zhang et al., 2009; Nordgren et al., 2011) with levels of MAT1A and GNMT protein expression (Table 4). BHMT, like MAT1A and GNMT, is predominately expressed in hepatic tissue (Li et al., 2008). Since there was a significant correlation between age and COMT activity ($p = 0.02$) and since GNMT protein level was also correlated with age ($p = 0.006$) in these samples, data for the genotype-phenotype association analyses described earlier and for the hepatic Methionine and Folate Cycle correlation analysis described here were adjusted for the age of the subject from whom the liver biopsy has been obtained for GNMT protein and COMT activity levels. In the 268 liver cytosol preparations studied, there was a significant correlation between GNMT and BHMT protein levels ($r = 0.34$, $p = 1.60 \times 10^{-7}$ after correction for multiple comparison) and between GNMT and MAT1A protein levels ($r = 0.23$, corrected $p = 1.50 \times 10^{-3}$, see Fig. 6). Level of activity for another AdoMet-dependent methyltransferase, COMT, was also associated with expression level of protein for these same two enzymes, but less significantly ($p = 0.03$ and 0.04, respectively). These results raise the possibility of coordinate regulation of Methionine Cycle enzyme protein levels in human hepatic tissue. Our association analysis of *GNMT* SNPs with levels of hepatic protein expression also showed that the rs9471976 and rs11752813 *GNMT* SNPs were significantly associated with SHMT1 protein level ($p = 0.007$ and 0.02, respectively).

Discussion

AdoMet is the major methyl donor for biological methylation reactions, including those involved in DNA and histone methylation as well as the methyl conjugation of hormones, neurotransmitters and drugs (Mato et al., 1997). In the adult human liver, AdoMet is synthesized

from methionine and ATP by a reaction catalyzed by MAT1A (Kluijtmans et al., 2003), and a major process in its hepatic utilization is the GNMT-catalyzed formation of sarcosine and AdoHcy (Kluijtmans et al., 2003). These reactions represent critical steps in the hepatic "Methionine Cycle" (see Fig. 1). Because of the importance of methylation for a variety of cellular functions, intercellular AdoMet concentrations and the AdoMet/AdoHcy ratio have been reported to contribute to variation in biological processes ranging from the epigenetic regulation of gene expression to biogenic amine neurotransmitter biosynthesis and metabolism (Reynolds et al., 1984; Carney et al., 1987; Ulrey et al., 2005). In addition, AdoMet concentrations are a major factor determining levels of "downstream" metabolites in the Methionine Cycle such as AdoHcy, homocysteine and glutathione. For example, elevated plasma homocysteine levels have been associated with increased risk for cardiovascular disease and individual variation in homocysteine concentrations are thought to be influenced by genetic factors. However, to date, the C667T variant (rs1801133) in the methylene tetrahydrofolate reductase (MTHFR) gene is one of only a small number of common genetic polymorphisms that are known to be associated with variation in circulating homocysteine concentrations (Frosst et al., 1995). In an attempt to begin to dissect the possible role of genetic variation in genes encoding enzymes in the Methionine Cycle on a variety of cellular processes, in the present study we set out to determine common genetic variation in genes encoding the hepatic AdoMet synthesizing and degrading enzymes, MAT1A and GNMT. To do that, we took a systematic approach that began with gene resequencing, followed by imputation using "1000 Genomes Project" data, functional genomic experiments and genotype-phenotype correlation analyses performed with human hepatic biopsy samples. The gene resequencing studies identified common DNA sequence variation across all 9 *MAT1A* exons, exon-intron splice junctions, and across the entire *GNMT* gene, as well as ~1 kb

of 5'- and 3'-FRs for both genes (Fig. 2 and Supplemental Tables 1 and 2). Resequencing of 288 DNA samples identified one nsSNP in *MAT1A*, but this variant did not appear to have functional consequences (Table 3). Using "1000 Genomes" data that were released soon after we completed our gene resequencing studies, we were able to extend our ability to scan these genes for functional markers out to 200 kb on either side of both genes.

Genotype-phenotype association analyses were then performed using 268 adult human surgical hepatic biopsy samples since the liver is the organ in which both MAT1A and GNMT are predominantly expressed. We observed two *GNMT* genotyped SNPs, rs9471976 and rs11752813, with p-values for association with level of hepatic GNMT protein expression of 6.4×10^{-12} and 2.9×10^{-7}, respectively, and 38 imputed SNPs within ± 200 kb surrounding *GNMT* that were significantly associated with hepatic GNMT protein levels ($p < 10^{-4}$). Many of these SNPs were in LD and were located in a region within 10 kb 5'- to *GNMT*, suggesting a possible role for this region in the regulation of *GNMT* transcription. Results of reporter gene studies performed with both HepG2 and LNCaP cells suggested that sequences around the rs9471976 and rs11752813 SNPs could increase transcription up to 6-fold, and there were significant differences in reporter gene activity between the WT and variant alleles (Fig. 5). In contrast, none of the 42 genotyped or 150 imputed *MAT1A* SNPs were highly associated with hepatic MAT1A protein levels ($p > 0.03$) (see Fig. 4B). These observations indicate that genetic regulation of GNMT expression in the human liver might be influenced by SNPs in a region at the 5'-end of the *GNMT* gene. Whether these polymorphisms influence concentrations of hepatic AdoMet or other Methionine Cycle metabolites, eg, homocysteine, will have to be determined in the course of future experiments.

Finally, we examined the possibility of the coordinate regulation of hepatic Methionine and Folate Cycle protein expression by determining correlations among protein expression or the activities of enzymes in this pathway in the 268 liver biopsy samples that we had used to perform the Western blot analyses (Table 4). Among the 6 proteins included in that analysis, GNMT, MAT1A and BHMT are expressed predominately in adult human liver, while COMT and SHMT1 are ubiquitously expressed. In a previous analysis of Methionine and Folate Cycle mRNA expression in the "Human Variation Panel" lymphoblastoid cell lines that were used as a source of DNA for our gene resequencing studies (Hebbring et al., 2012), we observed strong correlations among mRNA expression levels for SHMT1, SHMT2, COMT, MAT2A and MAT2B (Hebbring et al., 2012). Unfortunately, MAT1A, GNMT and BHMT are not expressed in lymphoblastoid cells. The results of these two complementary association analyses for hepatic tissue and lymphoblastoid cells both raise the possibility of the "coordinate regulation" of the expression of genes encoding Methionine and Folate Cycle enzymes. In summary, the results of the present study, when joined with those of previous reports, indicate the existence of the coordinate regulation of Methionine Cycle enzymes, with possible functional consequences for methylation in the human liver.

Authors Contributions

Participated in research design: YJ, KKSN, SJH, YP, VCY, RMW

Conducted the research: YJ, KKSN, YC, SJH, YP, LLP, IM, XC, JZ, EDW

Performed data analysis: YJ, KKSN, SJH, GDJ, RPA, LF, VCY, RMW

Wrote or contributed to writing of the manuscript: YJ, KIKSN, RPA, YP, VCY, RMW

REFERENCES

1000 Genomes Project Consortium (2010) A map of human genome variation from population-scale sequencing. *Nature* **467:**1061-1073.

Arnesen E, Refsum H, Bonaa KH, Ueland PM, Forde OH and Nordrehaug JE (1995) Serum total homocysteine and coronary heart disease. *Int J Epidemiol* **24:**704-709.

Augoustides-Savvopoulou P, Luka Z, Karyda S, Stabler SP, Allen RH, Patsiaoura K, Wagner C and Mudd SH (2003) Glycine N -methyltransferase deficiency: a new patient with a novel mutation. *J Inherit Metab Dis* **26:**745-759.

Balaghi M, Horne DW and Wagner C (1993) Hepatic one-carbon metabolism in early folate deficiency in rats. *Biochem J* **291 (Pt 1):**145-149.

Blumenstein J and Williams GR (1960) The enzymic N-methylation of glycine. *Biochem Biophys Res Com* **3:**259-263.

Boyle AP, Song L, Lee BK, London D, Keefe D, Birney E, Iyer VR, Crawford GE and Furey TS (2011) High-resolution genome-wide in vivo footprinting of diverse transcription factors in human cells. *Genome Res* **21:**456-464.

Cantoni GL (1951a) Activation of methionine for transmethylation. *J Biol Chem* **189:**745-754.

Cantoni GL (1951b) Methylation of nicotinamide with soluble enzyme system from rat liver. *J Biol Chem* **189:**203-216.

Carlson CS, Eberle MA, Rieder MJ, Yi Q, Kruglyak L and Nickerson DA (2004) Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium. *Am J Hum Genet* **74:**106-120.

Carney MW, Toone BK and Reynolds EH (1987) S-adenosylmethionine and affective disorder. *Am J Med* **83:**104-106.

Chamberlin ME, Ubagai T, Mudd SH, Thomas J, Pao VY, Nguyen TK, Levy HL, Greene C, Freehauf C and Chou JY (2000) Methionine adenosyltransferase I/III deficiency: novel mutations and clinical variations. *Am J Hum Genet* **66:**347-355.

Clarke R, Lewington S and Landray M (2003) Homocysteine, renal function, and risk of cardiovascular disease. *Kidney Int Suppl***:**S131-133.

Feng Q, Keshtgarpour M, Pelleymounter LL, Moon I, Kalari KR, Eckloff BW, Wieben ED and Weinshilboum RM (2009) Human S-adenosylhomocysteine hydrolase: common gene sequence variation and functional genomic characterization. *J Neurochem* **110:**1806-1817.

Fontecave M, Atta M and Mulliez E (2004) S-adenosylmethionine: nothing goes to waste. *Trends Biochem Sci* **29:**243-249.

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

Gil B, Casado M, Pajares MA, Bosca L, Mato JM, Martin-Sanz P and Alvarez L (1996) Differential expression pattern of S-adenosylmethionine synthetase isoenzymes during rat liver development. *Hepatology* **24:**876-881.

Heady JE and Kerr SJ (1973) Purification and characterization of glycine N-methyltransferase. *J Biol Chem* **248:**69-72.

Hebbring SJ, Chai Y, Ji Y, Abo RP, Jenkins GD, Fridley B, Zhang J, Eckloff BW, Wieben ED and Weinshilboum RM (2012) Serine hydroxymethyltransferase 1 and 2: gene sequence variation and functional genomic characterization. *J Neurochem* **120:**881-890.

Ji Y, Moon I, Zlatkovic J, Salavaggione OE, Thomae BA, Eckloff BW, Wieben ED, Schaid DJ and Weinshilboum RM (2007) Human hydroxysteroid sulfotransferase SULT2B1 pharmacogenomics: gene sequence variation and functional genomics. *J Pharmacol Exp Ther* **322:**529-540.

Johnson AD, Handsaker RE, Pulit SL, Nizzari MM, O'Donnell CJ and de Bakker PI (2008) SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics* **24:**2938-2939.

Kasperzyk JL, Fall K, Mucci LA, Hakansson N, Wolk A, Johansson JE, Andersson SO and Andren O (2009) One-carbon metabolism-related nutrients and prostate cancer survival. *Am J Clin Nutr* **90:**561-569.

Kerr SJ (1972) Competing methyltransferase systems. *J Biol Chem* **247:**4248-4252.

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

Li F, Feng Q, Lee C, Wang S, Pelleymounter LL, Moon I, Eckloff BW, Wieben ED, Schaid DJ, Yee V and Weinshilboum RM (2008) Human betaine-homocysteine methyltransferase (BHMT) and BHMT2: Common gene sequence variation and functional characterization. *Mol Genet Metab* **94:**326-335.

Li Y, Ding J and Abecasis GR (2006) Mach 1.0: rapid haplotype reconstruction and missing genotype inference. Am Soc Hum Genet. **579**:416, New Orleans.

Luka Z, Cerone R, Phillips JA, 3rd, Mudd HS and Wagner C (2002) Mutations in human glycine N-methyltransferase give insights into its role in methionine metabolism. *Hum Genet* **110:**68-74.

Main PA, Angley MT, Thomas P, O'Doherty CE and Fenech M (2010) Folate and methionine metabolism in autism: a systematic review. *Am J Clin Nutr* **91:**1598-1620.

Martin YN, Salavaggione OE, Eckloff BW, Wieben ED, Schaid DJ and Weinshilboum RM (2006) Human methylenetetrahydrofolate reductase pharmacogenomics: gene resequencing and functional genomics. *Pharmacogenet Genomics* **16:**265-277.

Maruti SS, Ulrich CM and White E (2009) Folate and one-carbon metabolism nutrients from supplements and diet in relation to breast cancer risk. *Am J Clin Nutr* **89:**624-633.

Mato JM, Alvarez L, Ortiz P and Pajares MA (1997) S-adenosylmethionine synthesis: molecular mechanisms and clinical implications. *Pharmacol Ther* **73:**265-280.

Mudd SH, Cerone R, Schiaffino MC, Fantasia AR, Minniti G, Caruso U, Lorini R, Watkins D, Matiaszuk N, Rosenblatt DS, Schwahn B, Rozen R, LeGros L, Kotb M, Capdevila A, Luka Z, Finkelstein JD, Tangerman A, Stabler SP, Allen RH and Wagner C (2001) Glycine N-methyltransferase deficiency: a novel inborn error causing persistent isolated hypermethioninaemia. *J Inherit Metab Dis* **24:**448-464.

Mudd SH, Ebert MH and Scriver CR (1980) Labile methyl group balances in the human: the role of sarcosine. *Metabolism* **29:**707-720.

Mudd SH and Poole JR (1975) Labile methyl balances for normal humans on various dietary regimens. *Metabolism* **24:**721-735.

Nordgren KK, Peng Y, Pelleymounter LL, Moon I, Abo R, Feng Q, Eckloff B, Yee VC, Wieben E and Weinshilboum RM (2011) Methionine adenosyltransferase 2A/2B and methylation: gene sequence variation and functional genomics. *Drug Metab Dispos* **39:**2135-2147.

Nygard O, Vollset SE, Refsum H, Stensvold I, Tverdal A, Nordrehaug JE, Ueland M and Kvale G (1995) Total plasma homocysteine and cardiovascular risk profile. The Hordaland Homocysteine Study. *JAMA* **274:**1526-1533.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ and Sham PC (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81:**559-575.

Reynolds EH, Carney MW and Toone BK (1984) Methylation and mood. *Lancet* **2:**196-198.

Shield AJ, Thomae BA, Eckloff BW, Wieben ED and Weinshilboum RM (2004) Human catechol O-methyltransferase genetic variation: gene resequencing and functional characterization of variant allozymes. *Mol Psychiatry* **9:**151-160.

Skiba WE, Taylor MP, Wells MS, Mangum JH and Awad WM, Jr. (1982) Human hepatic methionine biosynthesis. Purification and characterization of betaine:homocysteine S-methyltransferase. *J Biol Chem* **257:**14944-14948.

Smythies JR, Gottfries CG and Regland B (1997) Disturbances of one-carbon metabolism in neuropsychiatric disorders: a review. *Biol Psychiatry* **41:**230-233.

Sreekumar A, Poisson LM, Rajendiran TM, Khan AP, Cao Q, Yu J, Laxman B, Mehra R, Lonigro RJ, Li Y, Nyati MK, Ahsan A, Kalyana-Sundaram S, Han B, Cao X, Byun J, Omenn GS, Ghosh D, Pennathur S, Alexander DC, Berger A, Shuster JR, Wei JT, Varambally S, Beecher C and Chinnaiyan AM (2009) Metabolomic profiles delineate potential role for sarcosine in prostate cancer progression. *Nature* **457:**910-914.

Stevens VL, McCullough ML, Sun J and Gapstur SM (2010) Folate and other one-carbon metabolism-related nutrients and risk of postmenopausal breast cancer in the Cancer Prevention Study II Nutrition Cohort. *Am J Clin Nutr* **91:**1708-1715.

Ulrey CL, Liu L, Andrews LG and Tollefsbol TO (2005) The impact of metabolism on DNA methylation. *Hum Mol Genet* **14** R139-147.

Wang SH, Kuo SC and Chen SC (2003) High-performance liquid chromatography determination of methionine adenosyltransferase activity using catechol-O-methyltransferase-coupled fluorometric detection. *Anal Biochem* **319:**13-20.

Weinstein SJ, Stolzenberg-Solomon R, Pietinen P, Taylor PR, Virtamo J and Albanes D (2006) Dietary factors of one-carbon metabolism and prostate cancer risk. *Am J Clin Nutr* **84:**929-935.

Zhang J, Ji Y, Moon I, Pelleymounter LL, Ezequel Salavaggione O, Wu Y, Jenkins GD, Batzler AJ, Schaid DJ and Weinshilboum RM (2009) Catechol O-methyltransferase pharmacogenomics: human liver genotype-phenotype correlation and proximal promoter studies. *Pharmacogenet Genomics* **19:**577-587.

Footnotes

YJ and KKSN contributed equally to the manuscript

This work was supported in part by National Institutes of Health grants [R01 GM28157, R01 CA132780, U19 GM61388 (The Pharmacogenomics Research Network), R21 GM86689], a KL2 Mentored Career Development Award [NCRR Grant KL2 RR024151], a Gerstner Family Mayo Career Development Award in Individualized Medicine and a PhRMA Foundation "Center of Excellence Award in Clinical Pharmacology."

FIGURE LEGENDS

Figure 1. Methionine and Folate Cycles in the human liver. MAT, methionine adenosyltransferase; GNMT, glycine N-methyltransferase; AHCY, adenosylhomocysteine hydrolase; BHMT, betaine homocysteine methyltransferase; MTHFR, methylenetetrahydrofolate reductase; SHMT, serine hydroxymethyltransferase; AdoMet, S-adenosylmethionine, AdoHcy, S-adenosylhomocysteine, Hcy, homocysteine.

Figure 2. Human *MAT1A* and *GNMT* polymorphisms observed during gene resequencing. The figure shows a schematic representation of the human (A) *MAT1A* and (B) *GNMT* gene structures, with arrows indicating the locations of polymorphisms observed during the resequencing studies. Black rectangles represent exons encoding the opening-reading frame, and open rectangles represent potions of exons encoding untranslated region sequences. The colors of arrows indicate minor allele frequencies. EA, European-American; AA, African-American; HCA, Han Chinese-American.

Figure 3. Human hepatic liver MAT1A and GNMT protein levels. (A) MAT1A protein level frequency distribution; (B) GNMT protein level frequency distribution; (C) Representative Western blot analysis gel with “a”, “b”, “c” and “d” representing individual samples assayed in triplicate.

Figure 4. Genotype-phenotype correlations for human hepatic MAT1A and GNMT. Negative log-transformed p-values for single SNP associations with human liver (A) GNMT protein levels and (B) MAT1A protein levels using both genotyped SNPs and SNPs imputed by using “1000

Genomes" data. Figures 4A and B show SNPs within 20 kb of the 3'- and 5'-ends of the genes and include only imputed SNPs with imputation quality score R-sq values > 0.3. Black circles represent genotyped tag SNPs and red triangles represent imputed SNPs. * – indicates the threshold for significance after Bonferroni correction for *GNMT*. The structures of both genes are also shown schematically. (C) The relationship between SNP genotypes for the *GNMT* rs9471976 and rs11752813 SNPs and hepatic GNMT protein is shown. Spearman's rank correlation coefficients (r) and p-values are also shown.

Figure 5. GNMT qRT-PCR in a series of cell lines as well as dual luciferase reporter gene assays for the *GNMT* rs9471976 and rs11752813 SNPs performed with HepG2 and LNCaP cells. (A) Schematic representation of the *GNMT* gene with the locations of rs9471976 and rs11752813. (B) GNMT mRNA measured by qRT-PCR in a series of cell lines expressed as a percentage of the value for LNCaP cells. (C) Luciferase reporter gene assays for HepG2 cells. (D) Luciferase reporter assays for LNCaP cells. Each bar for the reporter gene studies represents relative luciferase activity reported as a % of the pGL3-Promoter construct activity. Values represent mean ± SEM for 6 independent transfections.

Figure 6. Correlations between levels of MAT1A and BHMT protein in adult human liver biopsy samples plotted against log GNMT protein levels in the same samples adjusted for age. (A) GNMT vs. MAT1A; (B) GNMT vs. BHMT. Spearman correlation coefficients, r, as well as p-values for associations after correction for multiple comparisons are also shown. In both panels, the level of log GNMT protein has been corrected for patient age.

DMD Fast Forward. Published on July 17, 2012 as DOI: 10.1124/dmd.112.046953
This article has not been copyedited and formatted. The final version may differ from this version.

Human *MAT1A* Polymorphisms

Polymorphism Location	Nucleotide Sequence Change	Amino Acid Change	MAF			rs No.	NCBI36/hg18 (dbSNP 130)
			AA	EA	HCA		
5'FR (-1015)	G > A		0.000	0.005	0.000		82040174
5'FR (-823)	A > T		0.005	0.000	0.000		82039982
5'FR (-806)	G > C		0.005	0.000	0.000		82039965
5'FR (-686)	C > A		0.016	0.000	0.000		82039845
5'FR (-625)	C > T		0.016	0.000	0.000		82039784
5'FR (-424)	T > C		0.026	0.147	0.026	rs17677908	82039583
5'UTR (-153)	C > T		0.000	0.042	0.000	rs11595587	82039312
Intron 1 (111)	G > A		0.115	0.073	0.000	rs3862534	82038958
Intron 1 (-9)	C > G		0.469	0.188	0.005	rs10887721	82035334
Intron 2 (14)	G > A		0.005	0.031	0.000		82035234
Intron 2 (199)	del of TAAT		0.245	0.281	0.042	rs71482773	82035046
Intron 2 (-306)	A > C		0.146	0.005	0.000	rs28539197	82034080
Intron 2 (-211)	C > T		0.026	0.000	0.000		82033985
Intron 2 (-162)	G > A		0.005	0.000	0.000		82033936
Intron 3 (96)	T > C		0.068	0.276	0.057	rs2236569	82033556
Intron 3 (-104)	G > C		0.052	0.068	0.000		82030632
Intron 3 (-54)	T > C		0.026	0.099	0.005	rs71481597	82030582
Intron 4 (90)	C > T		0.068	0.286	0.021	rs2282367	82030326
Intron 4 (274)	G > A		0.026	0.099	0.005	rs71481596	82030142
Intron 4 (313)	C > T		0.052	0.073	0.000	rs41284066	82030103
Exon 5 (426)	C > T		0.068	0.286	0.021	rs1143694	82030032
Exon 5 (429)	C > T		0.010	0.000	0.000		82030029
Intron 5 (-47)	G > A		0.052	0.073	0.000		82026377
Exon 6 (712)	G > A	Glu238Lys	0.000	0.000	0.016		82026168
Intron 6 (-95)	C > T		0.000	0.000	0.016		82025030
Intron 6 (-85)	G > A		0.021	0.000	0.000		82025020
Exon 7 (870)	G > A		0.120	0.297	0.021	rs10788546	82024834
Exon 7 (882)	C > T		0.120	0.297	0.021	rs10887711	82024822
Exon 7 (885)	A > T		0.010	0.000	0.000	rs17851642	82024819
Intron 7 (44)	C > T		0.026	0.151	0.026	rs55855057	82024709
Intron 7 (68)	del of A		0.010	0.000	0.000		82024685
Intron 7 (75)	G > A		0.000	0.000	0.005		82024678
Intron 7 (98)	C > T		0.074	0.276	0.021	rs10788545	82024655
Intron 7 (274)	C > T		0.000	0.000	0.005		82024479
Exon 8 (1005)	C > T		0.016	0.000	0.000	rs61734474	82024336
Intron 8 (14)	T > C		0.260	0.469	0.443	rs2994388	82024242
Intron 8 (15)	G > A		0.000	0.005	0.000		82024241
Intron 8 (114)	G > A		0.021	0.021	0.000	rs72809554	82024142
Intron 8 (336)	G > A		0.026	0.000	0.000		82023920
Intron 8 (-44)	C > T		0.135	0.223	0.531	rs4933327	82023663
Exon 9 (1131)	T > C		0.260	0.468	0.438	rs2993763	82023574
3'UTR (1207)	C > A		0.000	0.005	0.000		82023498
3'UTR (1255)	C > T		0.245	0.234	0.026	rs7087728	82023450
3'UTR (1261)	G > T		0.026	0.000	0.000		82023444

Table 1. Human *MAT1A* genetic polymorphisms. Exons and untranslated regions (UTRs) are numbered relative to the A (nucleotide 1) in the ATG translation initiation codon in exon 1. Negative numbers were assigned to positions 5′ to that location, and positive numbers to positions 3′. Nucleotides located within introns are numbered based on their distance from the nearest splice junction, with distances from 3′-splice junctions assigned positive numbers, and distances from 5′ splice junctions assigned negative numbers. Exon sequences have been "boxed". Polymorphisms identified previously are noted by rs number. AA, African-American; EA, European-American; HCA, Han Chinese-American subjects.

DMD Fast Forward. Published on July 17, 2012 as DOI: 10.1124/dmd.112.046953
This article has not been copyedited and formatted. The final version may differ from this version.

Human *GNMT* Polymorphisms

Polymorphism Location	Nucleotide Sequence Change	Amino Acid Change	MAF AA	MAF EA	MAF HCA	rs No.	NCBI36/hg18 (dbSNP 130)
							43035794
5'FR (-690)	C > T		0.005	0.005	0.000		43035829
5'FR (-655)	C > T		0.010	0.000	0.000		43035899
5'FR (-585)	T > C		0.021	0.000	0.000		43035937
5'FR (-547)	C > T		0.000	0.010	0.000		43035948
5'FR (-536)	deletion of GTTACCGT		0.016	0.000	0.000		43035995
5'FR (-489)	C > G		0.443	0.297	0.057	rs11752813	43036439
5'FR (-45)	C > T		0.542	0.313	0.068	rs10948059	43036730
Intron 1 (41)	G > A		0.156	0.000	0.000	rs5031030	43036736
Intron 1 (47)	G > T		0.099	0.401	0.628	rs2296805	43036832
Intron 1 (143)	C > A		0.000	0.005	0.000		43036925
Intron 1 (236)	deletion of AC		0.000	0.000	0.005		43037339
Intron 1 (650)	G > A		0.089	0.000	0.000	rs58057801	43037414
Intron 1 (725)	C > T		0.005	0.016	0.000		43037511
Intron 1 (-417)	C > T		0.242	0.422	0.758		43037638
Intron 1 (-290)	C > G		0.005	0.000	0.000		43037806
Intron 1 (-122)	G > A		0.000	0.021	0.000		43037817
Intron 1 (-111)	A > G		0.266	0.422	0.745	rs7760250	43037821
Intron 1 (-107)	C > T		0.000	0.000	0.053		43037867
Intron 1 (-61)	C > T		0.016	0.000	0.000		43037873
Intron 1 (-55)	C > G		0.000	0.010	0.000		43038039
Exon 2 (318)	C > T		0.000	0.000	0.005		43038287
Intron 2 (232)	A > G		0.188	0.052	0.053	rs3800292	43038437
Intron 2 (-41)	A > G		0.052	0.000	0.000		43038746
Intron 3 (-42)	C > A		0.000	0.005	0.000		43038855
Exon 4 (519)	G > A		0.027	0.000	0.000		43038924
Exon 4 (588)	C > T		0.005	0.000	0.000		43039187
Intron 5 (22)	deletion of G		0.218	0.083	0.122		43039193
Intron 5 (28)	C > T		0.158	0.039	0.025	rs4987174	43039202
Intron 5 (37)	G > A		0.495	0.522	0.131	rs4987173	43039239
Intron 5 (74)	G > C		0.066	0.394	0.635	rs2296804	43039505
3'UTR (968)	G > A		0.006	0.000	0.000		43039614
3'FR (1070)	deletion of TTAT		0.089	0.395	0.621	rs5875822	43039708
3'FR (1174)	C > G		0.000	0.000	0.010		43039732
3'FR (1193)	G > A		0.042	0.026	0.146	rs736158	43039757
3'FR (1223)	G > A		0.000	0.000	0.005		43039839
3'FR (1305)	G > A		0.026	0.042	0.000	rs1051218	43039871
3'FR (1337)	insertion of T		0.026	0.000	0.000		43040026
3'FR (1492)	C > A		0.016	0.000	0.000		43040058
3'FR (1524)	G > A		0.026	0.000	0.000		43040178
3'FR (1664)	G > T		0.458	0.489	0.120	rs1129187	43040180
3'FR (1646)	T > C		0.271	0.416	0.760	rs1129186	43040220
3'FR (1686)	A > G		0.026	0.000	0.000		43035794

Table 2. Human *GNMT* genetic polymorphisms. Exons and untranslated regions (UTRs) are numbered relative to the A (nucleotide 1) in the ATG translation initiation codon located in exon 1. Negative numbers were assigned to positions 5′ to that location, and positive numbers to positions 3′. Nucleotides located within introns (IVSs) are numbered based on their distance from the nearest splice junction, with distances from 3′-splice junctions assigned positive numbers, and distances from 5′ splice junctions assigned negative numbers. Exon sequences have been "boxed". Polymorphisms identified previously are noted by rs number. AA, African-American; EA, European-American; HCA, Han Chinese-American subjects.

MAT1A Allozyme	***K*m (μM) Methionine**	***K*m (μM) ATP**	**Allozyme Activity % of WT**	**Protein Quantity % of WT**
WT	315 ± 35	938 ± 85	100 ± 2	100 ± 3
Lys238	327 ± 54	894 ± 76	102 ± 3	104 ± 4

Table 3. MAT1A allozyme functional genomics. Values are mean ± SEM for three independent determinations. The variant allozyme values did not differ significantly (p-values > 0.05) from those for the WT allozyme for either apparent *K*m values or enzyme activity (measured using bacterially expressed protein) or relative protein quantity after expression in COS-1 cells.

DMD Fast Forward. Published on July 17, 2012 as DOI: 10.1124/dmd.112.046953
This article has not been copyedited and formatted. The final version may differ from this version.

r_s value \ p-value	COMT*[1]	**BHMT**	SHMT1	**MAT1A**	**GNMT[1]**
COMT*[1]	---	0.03	0.03	0.04	0.02
BHMT	0.18	---	0.18	0.89	**1.60E-07**
SHMT1	0.18	0.15	---	1.00	0.17
MAT1A	-0.13	0.10	0.00	---	**1.50E-03**
GNMT[1]	0.19	**0.34**	0.15	**0.23**	---

Table 4. Spearman correlations and p-values for correlations of levels of protein expression measured in 268 hepatic biopsy samples. The three enzymes that are expressed primarily in liver, GNMT, MAT1A and BHMT, are bolded. * indicates that enzyme activity was used to perform the correlation analysis for COMT. [1] indicates that values for relative protein expression or enzyme activity were adjusted for age. The p-values listed in the table have been corrected for multiple comparisons.

Methionine and Folate Cycles in Human Liver

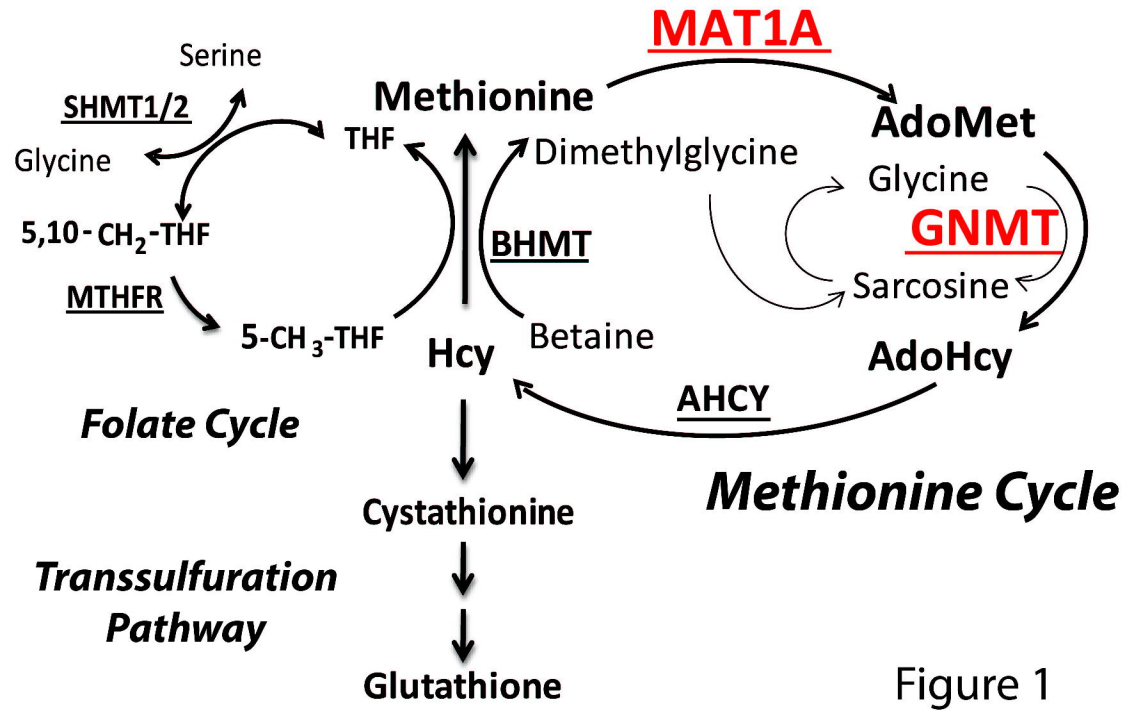


Figure 1

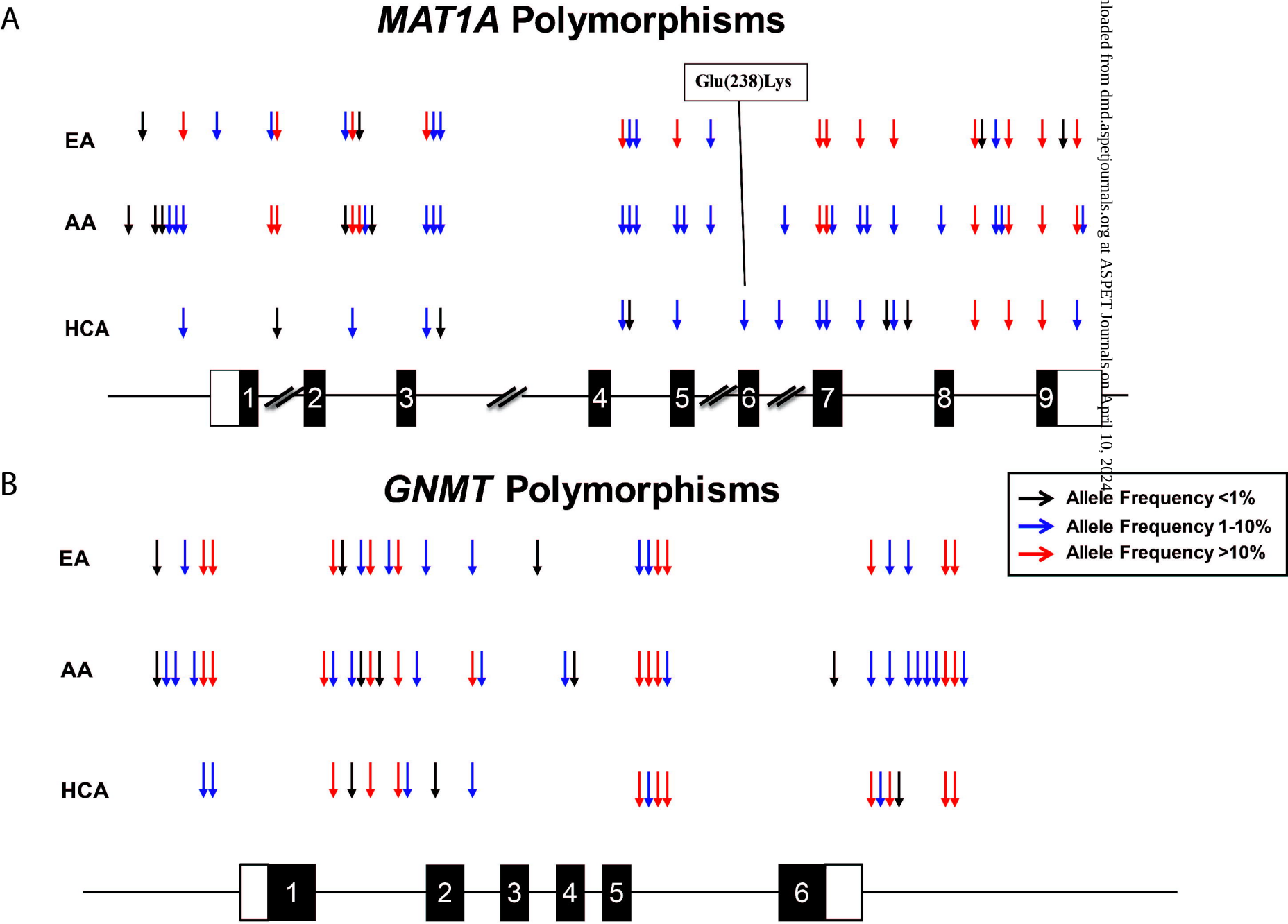


Figure 2

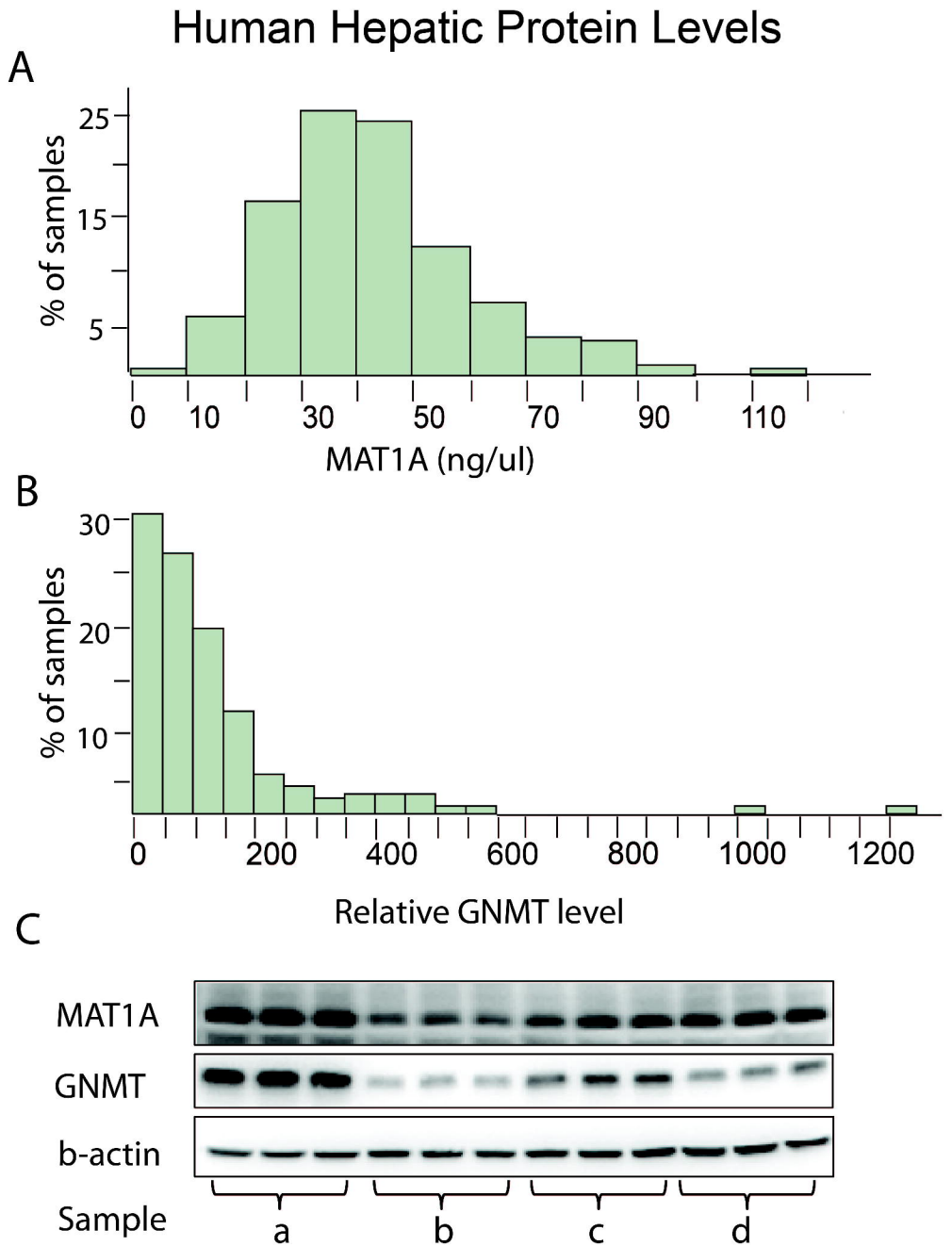


Figure 3

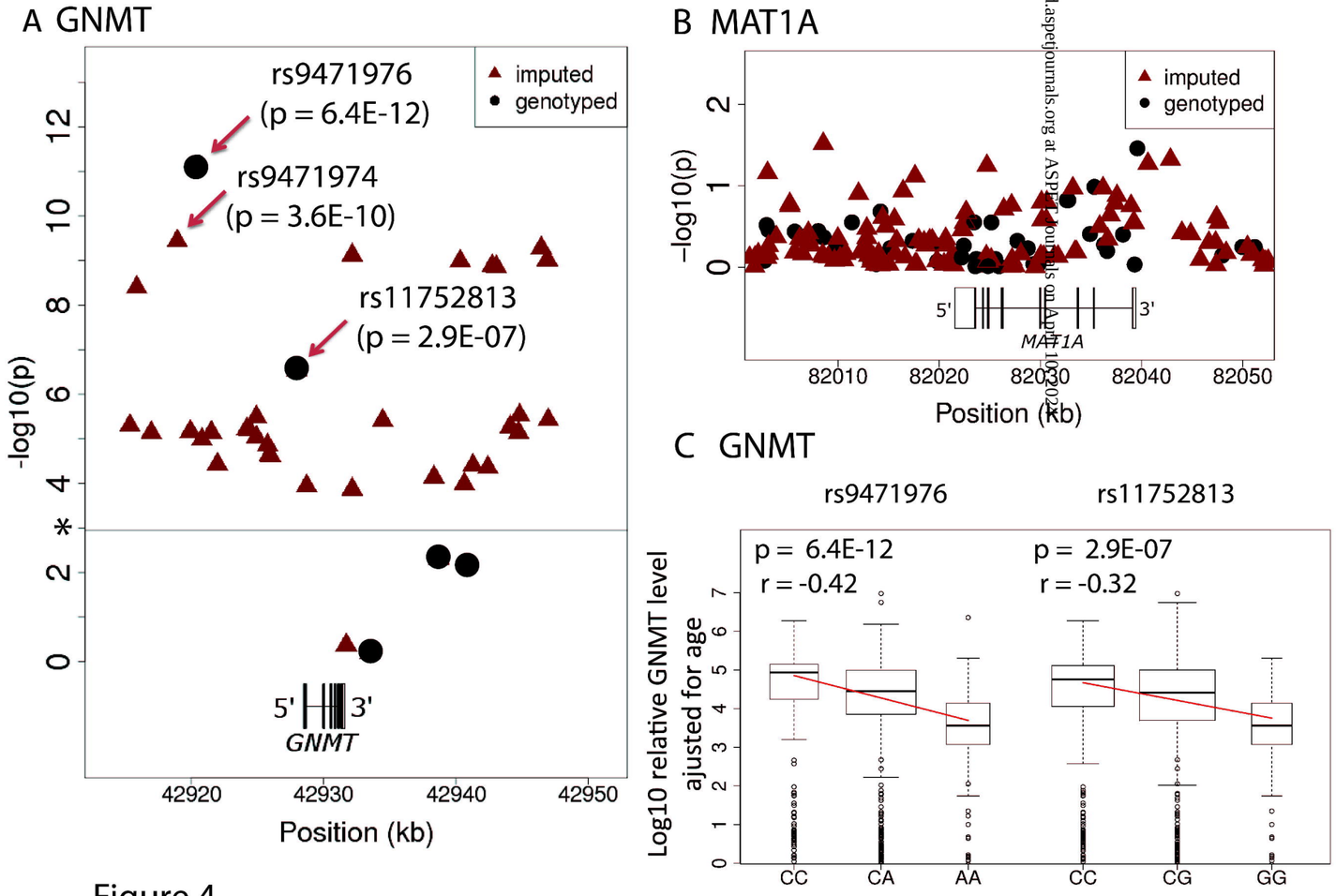


Figure 4

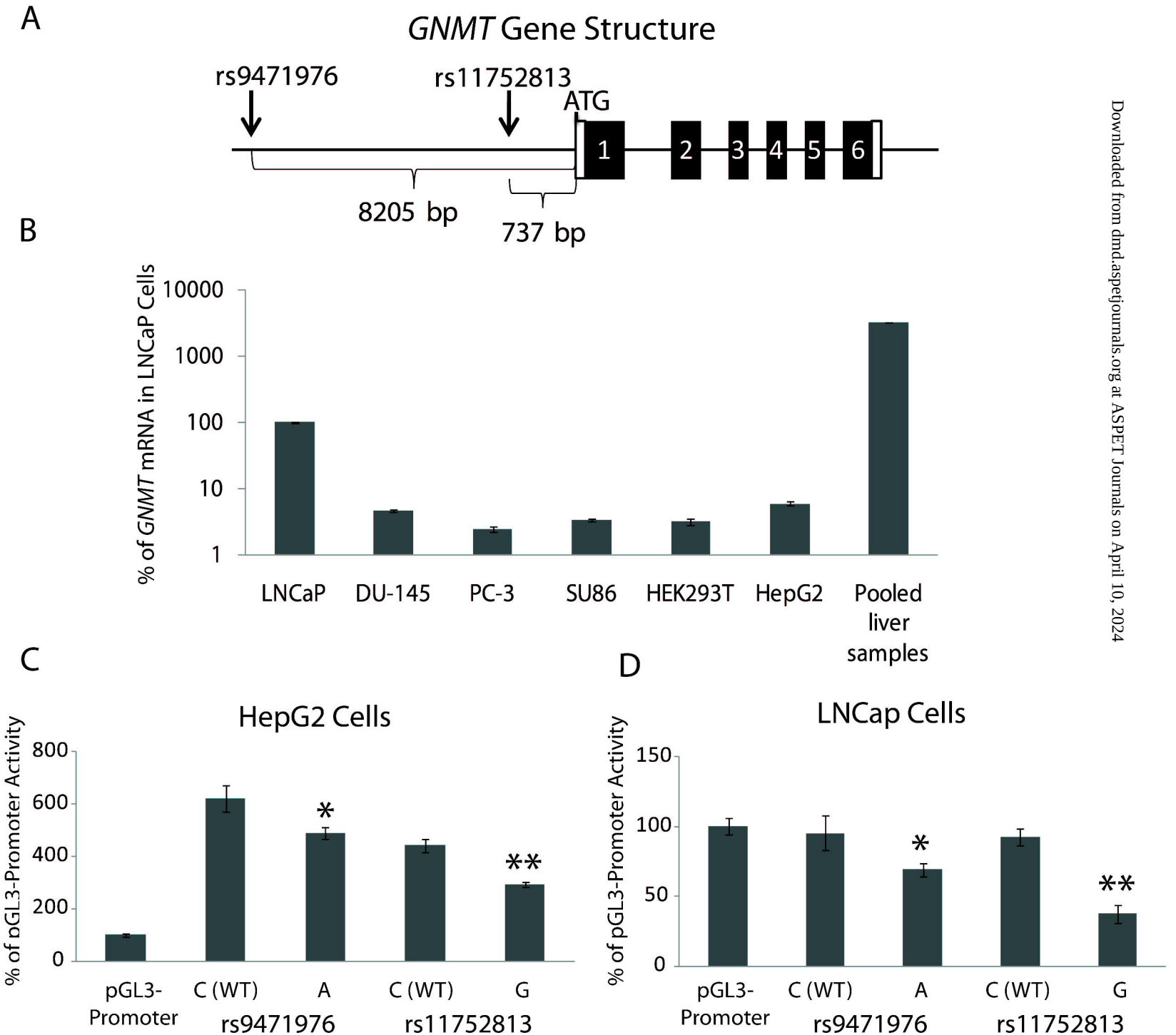
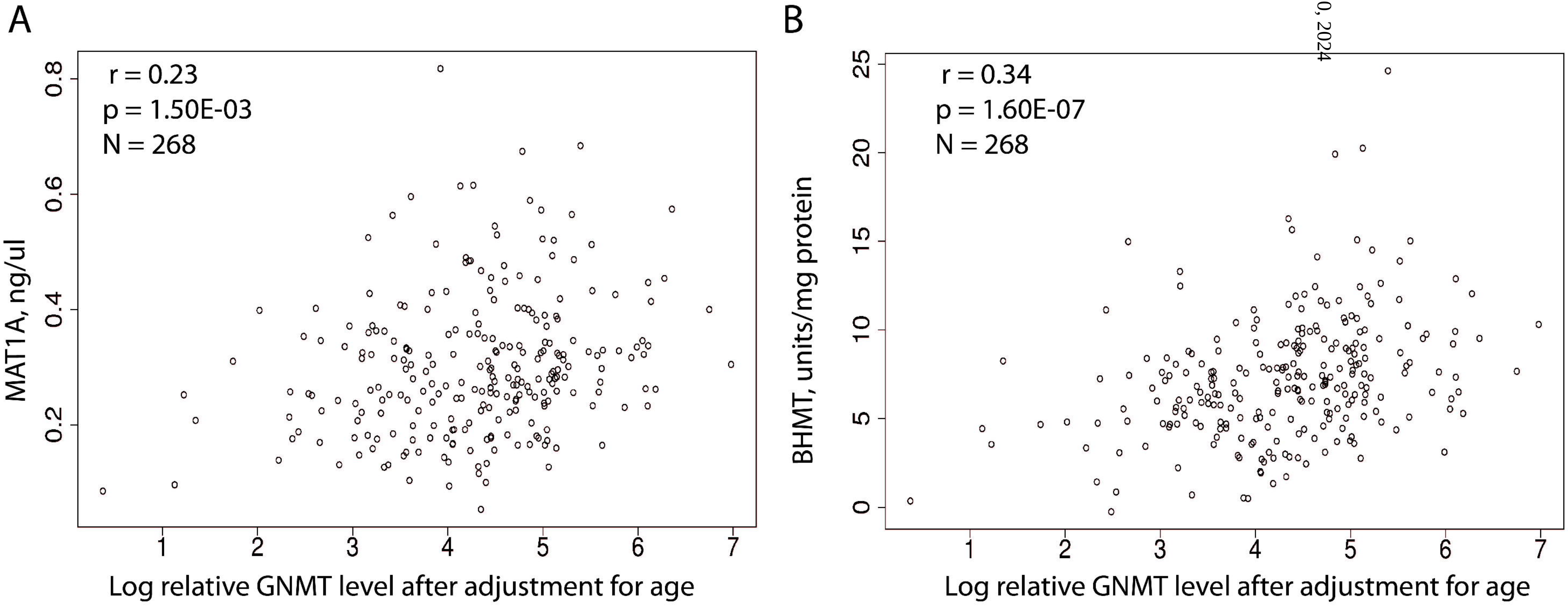


Figure 5

* p<0.05; ** p<0.005 compared to WT

Figure 6



SUPPLEMENTAL DATA

Human Liver Methionine Cycle: *MAT1A* and *GNMT* Gene Resequencing, Functional Genomics and Hepatic Genotype-Phenotype Correlation

Yuan Ji, Kendra K.S. Nordgren, Yubo Chai, Scott J. Hebbring, Gregory D. Jenkins, Ryan P. Abo, Yi Peng, Linda L. Pelleymounter, Irene Moon, Bruce W. Eckloff, Xiaoshan Chai, Jianping Zhang, Brooke L. Fridley, Vivien C. Yee, Eric D. Wieben and Richard M. Weinshilboum

Drug Metabolism and Disposition

METHODS

1. MAT1A mammalian expression

The open reading frame (ORF) of *MAT1A* was amplified from Origene clone SC119881 (Origene, Rockville, MD), and the amplicon was subcloned into pcDNA^{TM}3.1D/V5-His-TOPO (Invitrogen, Carlsbad, CA) in frame with the V5-His tag. The sequence of the insert was verified by DNA sequencing. Site-directed mutagenesis was used to create variant allozyme expression constructs. Sequences of the variant allozyme constructs were verified by DNA sequencing, and transient expression of the constructs in COS-1 cells was performed as described previously (1, 2).

2. MAT1A and S-COMT bacterial expression and purification

The *MAT1A* WT and variant constructs were cloned into the bacterial expression vector pET28a(+) (Novagen, Madison, WI) and transformed into BL21(DE3) *E. coli* to express MAT1A-His fusion proteins for use in substrate kinetic studies. The human catechol O-methyltransferase (COMT) gene open reading frame (ORF) for the soluble form of the enzyme (S-COMT) (3) was cloned into pET28a(+) and was also expressed in BL21 *E. coli*. After purification with a Ni-NTA column (Qiagen, Valencia, CA), the bacterially expressed S-COMT protein was used as a "reagent" in the MAT enzyme assay described subsequently.

The pET-28a-MAT1A and pET-28a-S-COMT vectors were transformed into *E. coli* BL21(DE3) cells that were grown to an OD_{600} of ~0.6 at 37°C before induction with 0.4 mM IPTG and incubation overnight at 18°C. Cells were harvested by centrifugation and resuspended in lysis buffer (50 mM Tris-HCl, pH 8.0, 0.5 M NaCl, 5 mM imidazole) with the addition of Complete protease inhibitor cocktail (Roche, Indianapolis, IN) before being disrupted using a M-110Y microfluidizer (Microfluidics, Newton, MA). After centrifugation, the supernatant was loaded onto Ni-NTA resin and was washed with 50 mM Tris-HCl, pH 8.0, containing 0.5 M NaCl, and 5 mM imidazole. The expressed protein was eluted with 50 mM Tris-HCl, pH 8.0, that contained 0.5 M NaCl and, sequentially, with increasing concentrations of imidazole from 50-500 mM. The combined eluted fractions were dialyzed against 50 mM Tris-HCl, pH 8.0, with 150 mM NaCl and 2 mM DTT. The WT and the Lys238 allozymes were subjected to a final purification step performed by gel filtration on a Superdex 200 column (GE Biosciences, Piscataway, NJ) equilibrated with 50 mM Tris-HCl, pH 8.0, with 150 mM NaCl, 5% glycerol, and 2 mM DTT. The void volume of the column was ~8.0 mL, and the calibration standards used were β-amylase (200 kDa, eluting at 11.8 mL), albumin (67 kDa, eluting at 13.3 mL), and ovalbumin (43 kDa, eluting at 14.5 mL).

3. MAT enzyme activity assay

Recombinant allozymes were assayed for MAT1A activity using an *L*-[methyl-^{14}C]-methionine (specific activity 55 mCi/mmol, American Radiolabeled Chemicals Inc., St. Louis, MO) radiochemical COMT coupled assay (4, 5). This assay was a modification of the method described by Wang et al. (5). Specifically, the [^{14}C]-labeled methyl group of AdoMet was transferred enzymatically to 3,4-dihydroxybenzoic acid (DBA) in a reaction catalyzed by COMT to form [^{14}C]-methyl labeled radioactive 3-methoxy-4-hydroxybenzoic acid, as described previously (6). The reaction mixture for this step contained 5 mM Tris-HCl buffer (pH 7.4), 1 mM DBA, 4 mM ATP, 15 mM $MgCl_2$, 150 mM KCl, 5 mM DTT, 1 mM *L*-[methyl-^{14}C]-methionine (1 μCi/μmol), 20 μg purified COMT protein, and recombinant enzyme (in 5 mM potassium phosphate buffer, pH 7.4), in a final reaction volume of 248 μL. "Blanks" consisted of reactions that lacked ATP, DBA, or enzyme. The reaction mixture was incubated at 37°C for 30 min with shaking, and the reaction was stopped by the addition of 80 μL 1N HCl. 2.5 mL of toluene was added, and the mixture was vortexed for 10 sec. After centrifugation at 700xg for 10 min at room temperature, 1.5 mL of the organic layer was aspirated and added to 4 mL of Bio-Safe II liquid scintillation counting fluid. Radioactivity was then measured in a Beckman Coulter LS6500 liquid scintillation counter (Brea, CA). For substrate kinetic experiments, *L*-methionine concentrations varied from 8 to 1000 μM and concentrations of ATP varied from 8 to 5000 μM. Apparent K_m values were calculated using Prism 4 (GraphPad Software, La Jolla, CA). Average levels of allozyme activity and protein were compared using the two-sample *t*-test (Microsoft Excel, Redmond, WA).

4. β-Galactosidase enzyme activity assay

β-Galactosidase activity was measured spectrophotometrically using the Promega β-Galactosidase Assay System (Madison, WI), and levels of recombinant MAT enzyme activity and MAT1A protein for the COS-1 cell transfection experiments were corrected on the basis of the activity of the cotransfected β-galactosidase.

5. MAT1A structural analysis

A search of the Protein Data Bank identified several crystal structures for rat MAT1A, ranging in resolution from 2.7 to 3.5 Å, as well as one human MAT1A structure at a resolution of 2.1 Å (Structural Genomics Consortium, unpublished). Both rat MAT1A and human MAT1A crystallized as tetramers organized as dimers of dimers. Due to its higher resolution and direct relevance to this study, the 2.1 Å resolution crystal structure of human MAT1A bound to AdoMet, the reaction product (PDB accession code 2OBV), was used as a starting point for the structural analysis. The crystallized human MAT1A contained residues 16-395, together with 4 additional N-terminal residues that resulted from a subcloning artifact. The human MAT1A dimer was formed by extensive interactions between two monomers related by 2-fold (180°) rotational symmetry.

References

1. Ji Y, Moon I, Zlatkovic J, Salavaggione OE, Thomae BA, Eckloff BW, Wieben ED, Schaid DJ, Weinshilboum RM. Human hydroxysteroid sulfotransferase SULT2B1 pharmacogenomics: gene sequence variation and functional genomics. J Pharmacol Exp Ther 2007;322(2):529-40.
2. Ji Y, Salavaggione OE, Wang L, Adjei AA, Eckloff B, Wieben ED, Weinshilboum RM. Human phenylethanolamine N-methyltransferase pharmacogenomics: gene re-sequencing and functional genomics. J Neurochem 2005;95(6):1766-76.
3. Shield AJ, Thomae BA, Eckloff BW, Wieben ED, Weinshilboum RM. Human catechol O-methyltransferase genetic variation: gene resequencing and functional characterization of variant allozymes. Mol Psychiatry 2004;9(2):151-60.
4. Chou TC, Lombardini JB. A rapid assay procedure for ATP:L-methionine adenosyltransferase. Biochim Biophys Acta 1972;276(2):399-406.
5. Wang SH, Kuo SC, Chen SC. High-performance liquid chromatography determination of methionine adenosyltransferase activity using catechol-O-methyltransferase-coupled fluorometric detection. Anal Biochem 2003;319(1):13-20.
6. Raymond FA, Weinshilboum RM. Microassay of human erythrocyte catechol-O-methyltransferase: removal of inhibitory calcium ion with chelating resin. Clin Chim Acta 1975;58(2):185-94.

Forward or Reverse	Region amplified by the primer pair	PCR Primer Sequence
MAT1A		
F	Exon 1	CCTAACTTTTGCTTCCCACAGT
R		CTGAAGGGCTAGAAGAGGGAAT
F	Exons 2 & 3	TACGTGTCATCCAAAATTAGGC
R		CCTTATGCCAGAGTCTTTGACC
F	Exons 4 & 5	CTGGCTAGTTAGGGAACCCC
R		ACCATTCTTTGAATGCCCAG
F	Exon 6	TTTTTGAACGTAAGAATGTGTCAGA
R		TTCACCGACTTTACATTTAGTCCA
F	Exons 7, 8 & 9	GAGTTCTGGGCTAAGGGGTC
R		TGACAGGACAGGCTAAATGAGA
GNMT		
F	5'FR, Exons 1 & 2	CCCTGTTACATTTTGTGAGTTTAAATA
R		CTAGACCTGCATACCCCACTTGT
F	Exons 2, 3, 4 & 5	CGTGTGGCAGCTGTATATCG
R		AAGTGGTACCTCAAGCCAGGA
F	Exons 5, 6 & 3'UTR	CTATAAGGTGGGGCCCTCTG
R		TGGGAGACAAACCTAGTCCTG

Supplemental Table 1. Sequence of primers used in *MAT1A* and *GNMT* resequencing amplifications for the gene resequencing studies. The genomic region amplified by the primer pair is also listed.

SNP	Primer	Size of region amplified (bp)	Primer sequences
rs9471976	Forward	214	AGG ACA GGT ACC CCC TTT TTG GTT AGG CTG TC
	Reverse		AGG ACA ACG CGT TTT TAA ATG CTT GCG TGC AG
rs11752813	Forward	225	AGG ACA GGT ACC TGT CAC CAT GTC CCA GCT AA
	Reverse		AGG ACA ACG CGT CGA AGG AAG GCA TCA GCA TA

Supplemental Table 2. Sequences of oligonucleotides used to perform PCR reactions to amplify regions containing SNPs for luciferase reporter gene constructs. Sequences for restriction sites (Acc65I for forward primer and MIu1 for reverse primer) were added at the 5'-ends of each primer used in the reporter gene studies and are underlined in the table.

SNP rs no.	Spear man correlation	p-value (raw)	p-value (corrected)	R-sq (quality score)	Chr. location	Gene location	Distance to *GNMT*
rs9471976	-0.42	6.38E-12	3.89E-10	N/A	43028527	5'-	7951 bp
rs9471974	0.38	3.62E-10	2.21E-08	0.83	43026962	5'-	9516 bp
rs9462858	0.38	5.18E-10	3.16E-08	0.68	43054468	3'-	14872 bp
rs1129187	0.38	7.53E-10	4.59E-08	0.93	43040178	3'-	582 bp
rs9462859	0.37	9.56E-10	5.83E-08	0.68	43054921	3'-	15325 bp
rs9471983	0.37	1.03E-09	6.28E-08	0.72	43048343	3'-	8747 bp
rs9986447	0.37	1.31E-09	7.99E-08	0.69	43050757	3'-	11161 bp
rs9462857	0.37	1.39E-09	8.48E-08	0.70	43051076	3'-	11480 bp
rs6941212	0.36	4.02E-09	2.45E-07	0.74	43023898	5'-	12580 bp
rs11752813	-0.32	2.88E-07	1.76E-05	N/A	43035995	5'-	483 bp
rs11752813	0.32	2.88E-07	1.76E-05	1.00	43035995	5'-	483 bp
rs73432512	0.31	4.18E-07	2.55E-05	0.49	43030043	5'-	6435 bp
rs4714634	0.29	1.98E-06	1.21E-04	0.54	43009098	5'-	27380 bp
rs6907751	0.29	2.97E-06	1.81E-04	0.67	43052828	3'-	13232 bp
rs6920547	0.29	3.24E-06	1.98E-04	0.87	43032933	5'-	3545 bp
rs3763236	0.29	3.73E-06	2.28E-04	0.55	43010486	5'-	25992 bp
rs9462860	0.29	3.77E-06	2.30E-04	0.72	43054991	3'-	15395 bp
rs2274514	0.29	3.88E-06	2.37E-04	0.91	43042478	3'-	2882 bp
rs7770760	0.28	5.05E-06	3.08E-04	0.67	43023381	5'-	13097 bp
rs9471987	0.28	5.35E-06	3.26E-04	0.70	43052118	3'-	12522 bp
rs13216214	0.28	5.97E-06	3.64E-04	0.84	43032225	5'-	4253 bp
rs4714640	0.28	6.25E-06	3.81E-04	0.83	43032402	5'-	4076 bp
rs2104616	0.28	6.82E-06	4.16E-04	0.81	43027947	5'-	8531 bp
rs58497441	0.28	7.18E-06	4.38E-04	0.79	43029539	5'-	6939 bp
rs9471988	0.28	7.26E-06	4.43E-04	0.72	43052751	3'-	13155 bp
rs4714638	0.28	7.31E-06	4.46E-04	0.80	43024995	5'-	11483 bp
rs6940837	0.28	9.24E-06	5.64E-04	0.87	43032947	5'-	3531 bp
rs4469287	0.27	1.04E-05	6.34E-04	0.82	43028822	5'-	7656 bp
rs9296404	0.27	1.39E-05	8.48E-04	0.85	43033781	5'-	2697 bp
rs6458312	0.27	1.57E-05	9.58E-04	0.53	43012252	5'-	24226 bp
rs9462855	0.26	2.32E-05	1.42E-03	0.92	43033914	5'-	2564 bp
rs9462856	0.26	2.40E-05	1.46E-03	0.93	43034002	5'-	2476 bp
rs6458313	0.26	3.72E-05	2.27E-03	0.79	43029996	5'-	6482 bp
rs9471985	0.26	3.98E-05	2.43E-03	0.81	43049302	3'-	9706 bp
rs112538187	0.25	4.39E-05	2.68E-03	0.36	43034812	5'-	1666 bp
rs9296407	0.25	4.43E-05	2.70E-03	0.80	43050429	3'-	10833 bp
rs2234185	0.25	4.61E-05	2.81E-03	0.36	43005052	5'-	31426 bp
rs57295928	0.25	4.84E-05	2.95E-03	0.32	43031068	5'-	5410 bp
rs3293	0.25	7.25E-05	4.42E-03	0.81	43046355	3'-	6759 bp
rs7759302	0.24	9.61E-05	5.86E-03	0.37	43037230	Intron 1	
rs2395943	0.24	1.04E-04	6.34E-03	0.81	43048651	3'-	9055 bp
rs2296805	0.24	1.15E-04	7.02E-03	0.91	43036736	Intron 1	
rs1129186	0.24	1.35E-04	8.24E-03	0.90	43040180	3'-	584 bp
rs73432523	0.23	2.34E-04	1.43E-02	0.32	43032054	5'-	4424 bp

Supplemental Table 3. SNPs that were significant associated with hepatic GNMT protein level after multiple comparisons (corrected p-values > 0.05 or 5.00E-02) are listed. Chromosome (chr.) locations (NCBI/hg18, dbSNP 130), Spearman correlation coefficient values, imputation (R-sq) quality scores, as well as p-values for genotype vs. GNMT protein level are listed. Location and distance to *GNMT*, NM_018960.4 (43036478 – 43039596) of the SNPs are also listed. Only SNPs with a R-sq score for imputation > 0.3 are listed. Genotyped SNPs are highlighted with grey shading. Hepatic GNMT protein levels were adjusted for age.