Published in final edited form as:

Clin Pharmacol Ther. 2017 March ; 101(3): 373–381. doi:10.1002/cpt.463.

A Genome-wide Approach Validates that Thiopurine Methyltransferase Activity is a Monogenic Pharmacogenomic Trait

Chengcheng Liu¹, Wenjian Yang¹, Deqing Pei², Cheng Cheng², Colton Smith¹, Wendy Landier³, Lindsey Hageman³, Yanjun Chen³, Jun J. Yang¹, Kristine R. Crews¹, Nancy Kornegay¹, Seth E. Karol⁴, F. Lennie Wong⁵, Sima Jeha⁴, John T. Sandlund⁴, Raul R. Ribeiro⁴, Jeffrey E. Rubnitz⁴, Monika L. Metzger⁴, Ching-Hon Pui⁴, William E. Evans¹, Smita Bhatia³, and Mary V. Relling¹

¹Department of Pharmaceutical Sciences, St. Jude Children's Research Hospital, Memphis, TN

²Department of Biostatistics, St. Jude Children's Research Hospital, Memphis, TN

³School of Medicine, University of Alabama at Birmingham, Birmingham, AL

⁴Department of Oncology, St. Jude Children's Research Hospital, Memphis, TN

⁵Department of Population Sciences, City of Hope, Duarte, CA

Abstract

We performed a genome-wide association study of primary erythrocyte TPMT activity in children with leukemia (n = 1026). Adjusting for age and ancestry, TPMT was the only gene that reached genome-wide significance (top hit rs1142345 or 719A>G, $P = 8.6 \times 10^{-61}$). Additional genetic variants (besides the 3 SNPs rs1800462, rs1800460 and rs1142345 defining TPMT clinical genotype) did not significantly improve classification accuracy for TPMT phenotype. Clinical mercaptopurine tolerability in 839 patients was related to TPMT clinical genotype ($P = 2.4 \times 10^{-11}$). Using 177 lymphoblastoid cell lines (LCLs), there were 251 SNPs ranked higher than the top TPMT SNP (rs1142345 $P = 6.8 \times 10^{-5}$), showing the limitation of LCLs for pharmacogenomic discovery. In a GWAS, TPMT activity in patients behaves as a monogenic trait, further bolstering the utility of TPMT genetic testing in the clinic.

Keywords

pharmacogenomics; single nucleotide polymorphism; personalized medicine; toxicity; clinical trials

Correspondence: Mary V. Relling, St. Jude Children's Research Hospital, 262 Danny Thomas Place, Memphis, TN 38105-2794, USA. Phone: 901-595-2348; Fax: 901-595-8869; mary.relling@stjude.org.

Authorship Contribution

M.V.R., C.L., W.Y., D.P., C.C., C.S., W.L., L.H., Y.C., J.J.Y., K.R.C., N.K., S.E.K., F.L.W., S.J., J.T.S., R.R., J.R., M.L.M., C.H.P., W.E.E., and S.B. wrote the manuscript; M.V.R. designed the research; M.V.R., C.L., W.Y., D.P., C.C., C.S., Y.C., J.J.Y., K.R.C., S.E.K., F.L.W., W.E.E., and S.B. analyzed the data.

Conflict of Interests disclosure: MVR and WEE: Prometheus Labs - Royalties from licensing TPMT genotyping.

Introduction

Thiopurine S-methyltransferase (TPMT) genomic variation has been associated with toxicities from mercaptopurine, a critical component in the treatment of autoimmune diseases, acute lymphoblastic leukemia (ALL) and lymphomas.(1, 2) Dosage individualization based on TPMT genotype or erythrocyte enzymatic activity (phenotype) is recommended for patients receiving thiopurines.(3) Without individualizing thiopurine doses based on TPMT, toxicity due to TPMT defects resulted in treatment interruptions,(2, 4) which have been associated with worse outcomes.(5–7)

Although TPMT activity has been strongly related to TPMT genotype, it may be influenced by other genes. A genome-wide study using HapMap CEU lymphoblastoid cell lines (LCLs) showed that a TPMT haplotype predicted TPMT activity, but there were also 96 genes ranked higher than TPMT.(8) Variation in erythrocyte TPMT activity within those who are wild-type for TPMT variants raises the question of whether there are other genetic modifiers of TPMT activity.(9–11) Moreover, racial and ethnic differences exist: patients of African ancestry have lower TPMT activity(12–14) and lower levels of methylated metabolites(15) than other populations, and Asian patients were relatively intolerant to thiopurines,(16) despite their low frequency of TPMT variant alleles.(17) However, a non-TPMT genetic basis for racial variation in TPMT activity is not known. Current strategies for clinical TPMT genetic testing test for three variants: rs1800462 (238G>C), rs1800460 (460G>A) and rs1142345 (719A>G). Here we studied two ethnically diverse cohorts of children with ALL using agnostic GWAS approaches to evaluate the contribution of genetic variation in TPMT and other genes to TPMT activity and effects on thiopurine tolerance, and to determine whether current TPMT genetic testing adequately classifies patients for their TPMT phenotype.

Results

Among the 1026 patients who had TPMT clinical genotype (Table 1), TPMT activity displayed normal distributions within the 938 patients whose TPMT clinical genotype was wild-type and the 87 heterozygotes who carried a known no-function variant of *2, *3A or *3C (Figure 1A). Heterozygous patients showed lower TPMT activity (median 10.4 U/mL erythrocytes, range 6.3–17.0 U/mL) than wild-type patients (median 19.8 U/mL, range 7.1–33.1 U/mL, $P = 1.2 \times 10^{-73}$). The one homozygous deficient patient (who carried *3A and *3C) had negligible activity of 0.5 U/mL. The percentage of TPMT heterozygotes tended to be higher in blacks (13%) than in whites (9%, Chi-square $P = 0.19$; Figure 1B). Overall, blacks ($n = 138$) also had lower TPMT activity than non-blacks ($n = 825$; $P = 7.7 \times 10^{-8}$). In a multivariate model, TPMT clinical genotype was the strongest determinant of erythrocyte TPMT activity after adjusting for clinical characteristics and genetic ancestry (meta-analysis $P = 2.3 \times 10^{-67}$; Table 2). African ancestry (meta- $P = 1.4 \times 10^{-7}$) and older age (meta- $P = 6.6 \times 10^{-4}$) were also associated with lower TPMT activity in both SJ and COG cohorts after adjusting for TPMT clinical genotype.

Seven coding variants (alleles TPMT*2, *3A, *3C, *8, *12, *24 and *X) were identified among 306 SJ and 578 COG patients who had Affymetrix 6.0 and Exome Array data (Table

S1). Genotypes of the three no-function variants (*2, *3A and *3C) based on Exome Array were 100% concordant with the clinical genotypes determined using PCR. Similar to previous reports,(18, 19) TPMT diplotypes varied substantially by race (Table 3). TPMT*3A was most prevalent in white (34/360, 9.5%) and Hispanic patients (17/242, 7.0%), but was completely absent in blacks (0/131). In contrast, TPMT*3C was the predominant variant allele in black patients (16/131, 12.2%), but was rare in whites (1/360, 0.3%) and Hispanics (2/242, 0.8%). No Asian patients (n = 55) carried any variant alleles. Patients heterozygous for the TPMT*24 allele (537G>T) had activity that was intermediate (n = 10; median 14.2 U/mL, range 8.0 – 23.9 U/mL) between those of TPMT clinical heterozygotes ($P = 3.9 \times 10^{-8}$) and *1/*1 wild-types ($P = 0.002$). TPMT*8 (644G>A), a variant that has been associated with African ancestry,(18) was found in 5 African-Americans, 1 Caucasian and 2 patients in “other” race group. Patients who carried a *8 allele tended to have TPMT activity (median 14.7 U/mL, range 12.1–21.1 U/mL; Figure 2) that was lower than those with TPMT *1/*1 (median 18.6 U/mL, range 7.1–33.1 U/mL; $P = 0.15$) but higher than those with heterozygote clinical genotype including *2, *3A or *3C ($P = 3.9 \times 10^{-8}$). The TPMT*X allele (677G>A) was found in one Hispanic patient with an intermediate activity of 12.8 U/mL.

The GWAS for TPMT activity was performed in two cohorts of 306 SJ and 578 COG patients separately (Figures 3A–B), followed by a meta-analysis that combined the results from the two cohorts (Figure 3C). After adjusting for age and ancestry, we observed a strong association between TPMT activity and polymorphisms in the TPMT region. There were 94 typed or imputed SNPs (coding and non-coding) with a meta-analysis P value that reached genome-wide significance ($P < 5 \times 10^{-8}$), all within 27kb of the TPMT gene (Table S2). Five of the 94 SNPs were interrogated on the arrays, including two missense variants rs1142345 (719A>G; $P = 8.6 \times 10^{-61}$) and rs1800460 (460G>A; $P = 2.0 \times 10^{-44}$), and three non-coding variants rs11964408 ($P = 4.0 \times 10^{-50}$), rs16880220 ($P = 4.2 \times 10^{-22}$) and rs6908777 ($P = 2.0 \times 10^{-18}$). The five SNPs were in the same LD block (Figure S1). Using a linear model, the three clinical no-function alleles of TPMT (*2, *3A and *3C) explained 24.9% of variability of TPMT activity. When we add to the model the coding variants for the alleles *8, *12, *24 and *X, 25.6% of the variability was explained (an increase of only 0.7%). The 92 SNPs that reached genome-wide significance (the 94 SNPs in Table S2, excluding the clinical SNPs of rs1142345 and rs1800460) accounted for another 0.9% in TPMT activity variability beyond that explained by the clinically tested alleles of *2, *3A and *3C.

To estimate the possible contribution of additional rare variants (beyond those interrogated on arrays) to TPMT activity, we took two approaches. First, we sequenced the TPMT locus in a subset (n = 296) of SJ patients. We detected 121 SNPs in TPMT and its adjacent region, and found 10 SNPs (all of them intronic and with common MAF) were associated with TPMT activity ($P < 0.05$) after adjusting for clinical genotype (Table S3), with the highest ranked SNP rs115104049 (upstream of TPMT) having $P = 0.0003$. The 10 variants explained 3.8% of variability of the TPMT activity in this subset of SJ patients, compared with 34% of variability explained by *2, *3A and *3C in the SJ subset. Secondly, we performed a simulation based on TPMT variants that have been deposited in the Exome Aggregation Consortium (ExAC) dataset (Table S4).(20) Among the 89 variants identified in

>60,000 individuals, 41 were predicted to be damaging; after excluding the three common variants, the 38 variants are very rare (MAF < 0.05%). Assuming these rare damaging TPMT variants have similar effect sizes as do TPMT *2/*3A/*3C, we estimate that 1.2% (95% CI: 0.78%–1.65%) of variation in TPMT activity could possibly be explained by additional rare TPMT variants, indicating a likely minimal contribution of other SNPs in the locus.

To identify whether there were non-TPMT genetic determinants, we conducted a second GWAS using TPMT clinical genotype as a covariate, and after adjusting for age, ancestry and TPMT clinical genotype, we observed a locus in the ADCY2 gene associated with TPMT activity in 306 SJ patients (Table S5 and Figure S2A). The top hit ($P = 3.3 \times 10^{-8}$) consisted of four imputed SNPs, all of which are in complete LD ($r^2 = 1$). Among the genotyped SNPs, an intronic SNP rs2892635 in ADCY2 ranked the highest ($P = 2.5 \times 10^{-7}$). This SNP was associated with higher TPMT activity across all race groups in SJ (Figure S3). However, none of the top ranked variants, including rs2892635, were replicated in the COG cohort at the $P < 0.05$ level. Among the 578 COG patients, the top non-TPMT hit associated with TPMT activity was a missense SNP rs142134087 in ARHGEF11 ($P = 2.0 \times 10^{-8}$; Table S6 and Figure S2B). This SNP was rare in the COG cohort (MAF = 0.2%) and absent in the SJ cohort; two individuals who were heterozygous for rs142134087 had high TPMT activity (with a mean of 31.9 U/mL). We also identified a SNP in a purine metabolism gene, FHIT (rs75341016, $P = 3.7 \times 10^{-7}$) associated with decreased TPMT activity. However, the top ranked non-TPMT SNPs for the COG cohort did not reach $P < 0.05$ in the SJ cohort. In the meta-analysis of non-TPMT SNPs, no SNP reached genome-wide significance (Table S7 and Figure S2C); rs1345390 in FTO was the top-ranked genotyped SNP associated with increased TPMT activity in SJ ($P = 4.5 \times 10^{-3}$) and COG patients ($P = 7.9 \times 10^{-5}$), and in the meta-analysis ($P = 1.6 \times 10^{-6}$).

We also assessed whether TPMT was the strongest genetic determinant of TPMT activity in HapMap LCLs. TPMT activity in the 88 YRI cell lines (median 16.7 U/mL, range 0.2–46.6 U/mL) was lower than that in the 89 CEPH cell lines (median 26.0 U/mL, range 9.1–98.0 U/mL; $P = 9.9 \times 10^{-14}$), consistent with the observation that African ancestry was associated with lower activity in patients (Table 2 and Figure 1B). Analyzing ~4 million SNPs, no SNP reached genome-wide significance. Unlike for primary erythrocyte data in patients, the top ranked SNPs for cell lines did not include TPMT SNPs. The highest ranked SNP within TPMT was rs1142345 (719A>G), which only ranked 252 ($P = 6.8 \times 10^{-5}$) among all SNPs (Table S8). Among the 94 SNPs that were genome-wide significant in patients (Table S2), 10 were also genotyped in LCLs, and 5 were associated with TPMT activity at the $P < 0.05$ level.

To partially address the striking difference in the ranking of TPMT SNPs for primary data in patients' erythrocytes compared to the cell line data, we performed 20 random permutations in the SJ and COG cohorts separately, using the same number of patients ($n = 177$) and the same model used for the GWAS of the HapMap LCLs. Among 169521 coding SNPs interrogated in the permutations, TPMT rs1142345 remained as the highest ranked with P values ranging from 3.8×10^{-7} to 6.2×10^{-13} (Figure S4). The top ranked SNPs associated

with TPMT activity in the permutation tests were similar to those from the patient GWAS in the two separate cohorts (data not shown).

We evaluated the classification performance of TPMT clinical genotype (based on 3 SNPs) to predict phenotype (Figure S5–S6). The threshold value for TPMT activity to divide patients with intermediate vs normal TPMT activity can vary. Using a threshold of 14 U/mL, the sensitivity and specificity of clinical genotype in predicting “intermediate-activity phenotype” was 85% and 92% in the SJ cohort, and 98% and 84% in the COG cohort, respectively. Using an optimal threshold to distinguish TPMT phenotypes determined by the ROC curve (15 U/mL for SJ and 13 U/mL for COG), the sensitivity and specificity were improved to 90% and 91% for SJ, and 94% and 91% for COG, respectively (Figure S5). The predictive capacity of clinical genotype based on the 3 SNPs, as evaluated by the AUC of the ROC curves, was greater than 0.95 in both cohorts. Adding other TPMT variants such as the *8 and *24 alleles, or the three genome-wide significant SNPs (rs11964408, rs6908777 and rs16880220) did not improve the classification performance of clinical TPMT genotype to assess phenotype (Figure S6).

Because TPMT activity relates to mercaptopurine toxicity, we studied the association between the top ranked TPMT SNPs and mercaptopurine tolerance. The median dose intensities in TPMT heterozygotes who carried one *2, *3A or *3C allele was 63%, 59% and 72%, respectively, which were lower than in those with *1/*1 genotype (median = 86%, $P = 2.4 \times 10^{-11}$ after adjusting for the protocol; Figure 4). The homozygous deficient patient tolerated only 6% of protocol-defined mercaptopurine therapy. Using a threshold of 80% of protocol dose (Figure S7) to distinguish high from low dose intensity, the negative predictive value of TPMT heterozygous genotype is 61%, i.e. 39% patients who had “wild-type” TPMT genetic test still missed >20% of their thiopurine doses. There was no statistically significant difference between the dose intensity in *1/*1 patients and those who carried *8, *12, *24 or *X variant alleles. After adjusting for protocol, TPMT clinical genotype and clinical features, we did not observe a strong association between dose intensity and other genetic factors for TPMT activity, including rs1345390 (FTO, $P = 0.34$), rs2892635 (ADCY2, $P = 0.09$), and rs142134087 (ARHGEF11, $P = 0.79$). As we reported,(21) the NUDT15 variant rs116855232 was associated with thiopurine dose intensity in the two cohorts (SJ $P = 0.030$ and COG $P = 7.3 \times 10^{-11}$), but it was not associated with TPMT activity ($P = 0.9$). Addition of this variant did not affect the association between TPMT clinical genotype and dose intensity ($P = 1.7 \times 10^{-13}$).

Discussion

Thiopurines are commonly used for autoimmune diseases and for therapy of ALL. TPMT genotype is an indicator for the starting dose of thiopurines,(22) and has been advocated to prevent toxicities in those with TPMT defects without compromising treatment outcome.(4, 23) However, some studies suggested that TPMT genotype might not be a perfect predictor for the phenotype with a concordance of 53–100% in heterozygotes.(24) Also, not all TPMT wild-type patients tolerated full dose of mercaptopurine; the average dose intensity for wild-types ranged between 80–85%.(21, 25)

Herein, we confirmed that TPMT activity is a monogenic pharmacogenomic trait in two independent ALL cohorts with diverse racial backgrounds using a genome-wide agnostic approach. The GWAS for TPMT activity (Figure 3) revealed a strong association in the TPMT region: all 94 genome-wide significant SNPs were in or adjacent to TPMT (Table S2). This is most likely because they are in LD with no-function variants in TPMT (Figure S1).

Heterozygosity at the three commonly tested no-function variants (*2, *3A and *3C) was associated with lower TPMT activity and lower dose intensity than those with *1/*1 genotype (Figures 2 and 4). Patients with *1/*8 (very common in African Americans) or *1/*24 genotypes had intermediate TPMT activity but tolerated normal dose intensity comparing to those with *1/*1, partially supporting the current categorization of these patients as TPMT wild-types. TPMT*X (677G>A at rs139392616) was described by Udaka et al(26) by sequencing the TPMT gene in 288 individuals; however its phenotype is not clear. Herein, we observed one Hispanic patient (MAF = 0.2% in Hispanics) who carried TPMT*X, and this patient had intermediate TPMT activity of 12.8 U/mL. Of 11,566 Latino subjects in the ExAC dataset, three individuals were heterozygous for this variant, yielding an MAF of 0.03%.

TPMT clinical genotype based on only 3 SNPs had high predictive capacity for TPMT phenotype (Figure S5). Although TPMT*8, TPMT*24 and three intronic variants detectable by arrays were associated with intermediate or low TPMT activity in the entire cohort (Figure 2 and Table S1), we were unable to significantly improve the test performance by considering these additional variants (Figure S6). As expected, sequencing does detect additional variants not interrogated by arrays: in the relatively small subset of 296 SJ patients, these additional variants explained 3.8% variability in TPMT activity, compared to the 34% of variability accounted for by clinical genotyping (based on 3 SNPs). To extrapolate the possible contribution of rare variants to TPMT activity based on sequencing data from a much larger sample size (the ~ 60,000 individuals in ExAC), we estimated that additional rare TPMT variants only explain an additional 1.2% of variation in TPMT activity beyond that explained by the 3 clinically genotyped variants. Unlike some other drug metabolism/disposition genes, no-function TPMT variants are relatively common and rare no-function alleles are less common,(27, 28) confirmed by lower population fixation index (FST) values.(29) Thus, genotyping of the three no-function SNPs (rs1800462, rs1800460 and rs1142345) is still the most reliable method to determine TPMT phenotype, and rare variants may be less of a problem for TPMT than for other actionable pharmacogenes.

We found no non-TPMT genes with a significant effect on TPMT activity in this GWAS. Given an alpha level of 1×10^{-7} , we had 80% power to detect any SNPs with MAF of at least 5% with an effect size of 40% of those of TPMT variants. There could be trans-SNPs with weaker effect sizes or lower MAF that could influence TPMT activity, but a larger sample size would be needed to discover those SNPs.

After adjusting for TPMT genotype, TPMT activity was associated with variants in ADCY2 among SJ patients and with ARHGEF11 among COG patients (Figure S4). However, the top variants that were identified in one cohort failed to replicate in the other cohort (Tables S5

and S6). The reasons for the conflicting results between two cohorts may include (1) the SJ protocol used a TPMT-guided dosing strategy while the COG protocol did not; (2) the top variant in the COG cohort (rs142134087 in ARHGEF11) was rare and absent in the SJ cohort; and (3) the variants have small effect sizes and require higher sample sizes to replicate. In any case, the effect of additional genetic factors on TPMT activity was not comparable to the effect of TPMT genotype.

Variants in FTO(30, 31) were modestly associated with TPMT activity (Table S7). Interestingly, over-expression of FTO upregulated the expression of TPMT in mice.(32) However, in this study the FTO variants only accounted for a minor change in TPMT activity: among the 896 patients wild-type for TPMT clinical genotype, those with two vs one vs zero C alleles at rs1345390 had a median activity of 19.5, 19.0, and 18.2 U/mL vs ($P = 7.3 \times 10^{-6}$ after adjusting for clinical features).

The HapMap project has been a useful resource in pharmacogenomic studies. Here we showed that TPMT variant rs1142345 was associated with TPMT activity in 177 HapMap LCLs but only ranked 252 among all HapMap SNPs (Table S8), while the permutation tests in patients showed that rs1142345 was always top-ranked (Figure S3). It should be noted that rs1142345 is in complete LD with the TPMT haplotype (named HAP1) that we previously reported as the highest ranked TPMT SNP associated with TPMT activity in HapMap CEU LCLs.(8)The difference between the results from LCLs and from patients may be attributed to different conditions in cultured cell lines and in vivo, particularly because LCLs are transformed by EBV infection, which may impact genetic regulation.(33) In the current study we accounted for different LCL growth rates,(8, 9) and for the time gap between transformation of LCLs from CEU and from YRI samples(34) by adjusting for genetic ancestries. However, gene expression in LCLs may be affected by EBV copy number.(35) Non-genetic factors may also have an impact on enzyme function in LCLs.(36) Moreover, TPMT activity may differ in lymphocytes vs erythrocytes.(37) We have reported previously that the expression of PACSIN2 and its cis-SNP (rs2413739) were related to TPMT activity in 30 HapMap CEU trios.(9) In the current study, rs2413739 was not associated with TPMT activity in patients ($P = 0.14$ and 0.54 in SJ and COG, respectively, after adjusting for TPMT clinical genotype). We compared our results with another study that used a pathway approach;(38) among the top ranked SNPs from that study, only the 2 SNPs in TPMT were replicated in our study of HapMap LCLs and in SJ and COG patients (Table S9). Thus, LCLs may serve as a valuable resource for validating GWAS results, but their use for discovery may be challenging.(8)

Race-specific differences in TPMT activity and mercaptopurine metabolism have been observed ,(12–17) and African ancestry has been related to lower TPMT activity(12–14) and lower levels of methylated metabolites.(15) Some studies reported a higher prevalence of TPMT variants in blacks(21) but other did not.(18, 39) In the current study, black patients had a lower TPMT activity compared to non-blacks ($P = 7.7 \times 10^{-8}$), and tended to have a higher frequency of no-function variants than whites ($P = 0.19$; Figure 1B).The addition of TPMT genotype to the multivariate model did not affect the association of African ancestry with low activity (Table 1), suggesting that other genetic factors may be involved; however, in a meta-analysis of African-ancestry patients only (SJ $n = 44$ and COG $n = 87$), all

genome-wide significant SNPs ($n = 21$) locate within TPMT, similar to the observation in all patients. After adjusting for TPMT genotype, we were unable to find additional strong genetic associations ($P < 10^{-6}$) with TPMT activity (data not shown). Given that black patients tend to have a high risk of relapse(40–42) and that TPMT activity might be associated with relapse risk,(23, 43, 44) genetic and non-genetic factors should be considered in dose individualization in those patients.

There have been conflicting reports on the relationship between age and TPMT activity. (45–49) Here we observed that older age was associated with lower TPMT activity after adjusting for TPMT genotype in both cohorts (Table 1).

In summary, we used a genome-wide agnostic approach to confirm that TPMT activity was a monogenic trait, and that current strategies for clinical TPMT genetic testing based on only three genomic variants adequately classify patients for their TPMT phenotype in a multiethnic cohort, thereby bolstering the use of TPMT genetic testing to assess TPMT status in the clinic.

Methods

Patients and Phenotypes

We studied 1026 children with ALL, including 306 treated on St. Jude Children's Research Hospital (SJ) Total Therapy XV (NCT00137111)(40) and 720 treated on the Children's Oncology Group (COG) clinical trial AALL03N1 (NCT00268528)(21) (Table 1). These patients were included in a previous study of mercaptopurine tolerability.(21) SJ Total XV was approved by the Institutional Review Boards (IRB) at SJ, and COG AALL03N1 was approved by the National Cancer Institute and the IRBs of participating institutions. Informed consent was obtained from patients, parents or guardians in accordance with the Declaration of Helsinki.

A preliminary TPMT clinical genotype was based on PCR tests for three known no-function variants: rs1800462, rs1800460 and rs1142345 (contributing to *2, *3A and *3C alleles) . (50) Considering the impact of anemia and blood transfusion on TPMT enzymatic activity, (50, 51) blood samples were collected after remission induction and 90 days after erythrocyte transfusions.

COG patients received a protocol-planned dose of 75 mg/m²/day, which was adjusted based on the degree of leukopenia and toxicities. SJ patients received 50–75 mg/m²/day, with those heterozygous for TPMT clinical genotype receiving a starting mercaptopurine dose of 50–60 mg/m²/d . Mercaptopurine dosage was also titrated based on TPMT activity and thiopurine metabolites. To assess tolerability, dose intensity for each patient was estimated as the total cumulative prescribed dose)/(cumulative protocol dose)].(21, 25)

Genotyping and GWAS

Germline DNA from blood or bone marrow obtained at remission from SJ ($n = 306$) and COG ($n = 578$) patients was typed using Affymetrix Genome-Wide Human SNP Array 6.0 (or GeneChip Human Mapping 500K) and by Illumina HumanExome Beadchip Array.(21)

For SNPs not interrogated on the arrays, imputation was performed using the 1000 Genomes Project (<http://www.1000genomes.org>) as the reference with MaCH-Admix software (University of North Carolina). Imputed SNPs with a minor allele frequency (MAF) < 1% were excluded. Genetic ancestries (European, African, Native American and Asian) were determined in 306 SJ and 657 COG patients using STRUCTURE,(52) and either treated as continuous variables or used to categorize patients as white (European ancestry > 90%), black (African ancestry > 70%), Hispanics (non-black patients with Native American ancestry > 10%), Asian (Asian ancestry > 90%), and “other” groups. We also identified additional variants by sequencing the TPMT gene in SJ patients (n = 296; Table S3), as described.(53)

Genotype calls were coded as 0, 1 or 2 for the number of B alleles (AA, AB or BB) assuming an additive genetic model. To identify covariates, we used multiple linear regression to evaluate the association between TPMT activity and variables including age (years) and ancestry (as % European, African, Asian and Native American ancestries)(52) as continuous variables. After excluding SNPs with a call rate < 95%, over 9 million SNPs (including 578233 typed SNPs) were interrogated in the initial GWAS for TPMT activity in the SJ and COG cohorts separately. A meta-analysis was performed by combining the GWAS results from SJ and COG using Stouffer’s Z-score method.(54)

To determine non-TPMT genetic covariates for TPMT activity, we conducted a second GWAS that included the covariates of ancestry and age as continuous variables, and TPMT clinical genotype coded as 0 (wild-type), 1 (heterozygous) or 2 (homozygous) for the number of *2, *3A, or *3C alleles. A meta-analysis was performed as described above.

TPMT activity GWAS in HapMap LCLs

We obtained 177 Epstein-Barr virus (EBV)-transformed LCLs, including 89 CEPH (forming 30 CEPH family trios) and 88 YRI (forming 30 YRI family trios) genotyped in the International HapMap project, from Coriell (Camden, New Jersey <http://www.ccr.coriell.org>). TPMT activity for each cell line was determined.(8, 9) Genotypes for ~4 million SNPs were obtained from release 28 of the International HapMap Project (<http://www.hapmap.org>). A GWAS for TPMT activity was performed using genetic ancestry as a continuous variable.

Permutation-based GWAS in patients

To compare GWAS results from patients to those from HapMap LCLs, it was necessary to adjust for the smaller sample size of the LCLs to achieve the same level of statistical power. For each permutation test, we randomly selected 177 patients from either SJ or COG cohorts (the same as the number of HapMap LCLs, n = 177) in 20 independent permuted GWASs. Within each permutation, the genotype-phenotype association of 169521 SNPs on the Illumina Exome Array was evaluated using the same method as in the HapMap GWAS.

Classification performance of TPMT genotype for phenotype

Sensitivity, specificity, positive predictive value and negative predictive value of TPMT genotypes for dichotomized TPMT phenotypes were calculated using the numbers of

patients with true positive, false positive, true negative and false negative results (Figures S5 and S6). The classification performance of TPMT genotypes was evaluated in terms of receiver-operator-characteristic (ROC) curves (Figures S5 and S6).

Statistics

The Mann-Whitney U test was used to compare continuous variables between groups and the Chi-square test was used to compare categorical variables. A multivariate general linear model was used to test covariates for their association with the dependent variables. Linkage disequilibrium (LD) was calculated using Haploview(55) (version 4.2).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This study was supported by grants GM 92666, CA 21765, CA 142665, CA 36401, P50 GM 115279, U10 CA98543 (COG Chair's grant), U10 CA98413 (COG Statistical Center), and U24 CA114766 (COG Specimen Banking) from the National Cancer Institute, and by the American Lebanese Syrian Associated Charities.

References

1. Relling MV, Pui CH, Cheng C, Evans WE. Thiopurine methyltransferase in acute lymphoblastic leukemia. *Blood*. 2006; 107(2):843–4. [PubMed: 16401827]
2. Teml A, Schaeffeler E, Herrlinger KR, Klotz U, Schwab M. Thiopurine treatment in inflammatory bowel disease: clinical pharmacology and implication of pharmacogenetically guided dosing. *Clinical pharmacokinetics*. 2007; 46(3):187–208. [PubMed: 17328579]
3. Marshall E. Preventing toxicity with a gene test. *Science*. 2003 Oct 24; 302(5645):588–90. [PubMed: 14576416]
4. Relling MV, Hancock ML, Rivera GK, Sandlund JT, Ribeiro RC, Krynetski EY, et al. Mercaptopurine therapy intolerance and heterozygosity at the thiopurine S-methyltransferase gene locus. *Journal of the National Cancer Institute*. 1999; 91:2001–8. [PubMed: 10580024]
5. McLeod HL, Coulthard S, Thomas AE, Pritchard SC, King DJ, Richards SM, et al. Analysis of thiopurine methyltransferase variant alleles in childhood acute lymphoblastic leukaemia. *British journal of haematology*. 1999 Jun; 105(3):696–700. Epub 1999/06/03. eng. [PubMed: 10354134]
6. Lennard L, Welch J, Lilleyman JS. Mercaptopurine in childhood leukaemia: the effects of dose escalation on thioguanine nucleotide metabolites. *British journal of clinical pharmacology*. 1996 Oct; 42(4):525–7. Epub 1996/10/01. eng. [PubMed: 8904630]
7. Bhatia S, Landier W, Hageman L, Chen Y, Kim H, Sun CL, et al. Systemic Exposure to Thiopurines and Risk of Relapse in Children With Acute Lymphoblastic Leukemia: A Children's Oncology Group Study. *JAMA oncology*. 2015 Jun 1; 1(3):287–95. [PubMed: 26181173]
8. Jones TS, Yang W, Evans WE, Relling MV. Using HapMap tools in pharmacogenomic discovery: the thiopurine methyltransferase polymorphism. *Clinical pharmacology and therapeutics*. 2007 May; 81(5):729–34. Epub 2007/03/03. eng. [PubMed: 17329987]
9. Stocco G, Yang W, Crews KR, Thierfelder WE, Decorti G, Londero M, et al. PACSIN2 polymorphism influences TPMT activity and mercaptopurine-related gastrointestinal toxicity. *Human molecular genetics*. 2012 Nov 1; 21(21):4793–804. Epub 2012/08/01. eng. [PubMed: 22846425]
10. Scheuermann TH, Keeler C, Hodsdon ME. Consequences of binding an S-adenosylmethionine analogue on the structure and dynamics of the thiopurine methyltransferase protein backbone. *Biochemistry*. 2004 Sep 28; 43(38):12198–209. [PubMed: 15379558]

11. Fessing MY, Krynetski EY, Zambetti GP, Evans WE. Functional characterization of the human thiopurine S- methyltransferase (TPMT) gene promoter. *EurJBiochem*. 1998; 256(3):510–7.
12. McLeod HL, Lin JS, Scott EP, Pui CH, Evans WE. Thiopurine methyltransferase activity in American white subjects and black subjects. *Clinical pharmacology and therapeutics*. 1994; 55:15–20. [PubMed: 8299312]
13. Jones CD, Smart C, Titus A, Blyden G, Dorvil M, Nwadike N. Thiopurine methyltransferase activity in a sample population of black subjects in Florida. *Clinical pharmacology and therapeutics*. 1993; 53:348–53. [PubMed: 8453854]
14. Cooper SC, Ford LT, Berg JD, Lewis MJ. Ethnic variation of thiopurine S-methyltransferase activity: a large, prospective population study. *Pharmacogenomics*. 2008 Mar; 9(3):303–9. [PubMed: 18303966]
15. Bell BA, Brockway GN, Shuster JJ, Erdmann G, Sterikoff S, Bostrom B, et al. A comparison of red blood cell thiopurine metabolites in children with acute lymphoblastic leukemia who received oral mercaptopurine twice daily or once daily: a Pediatric Oncology Group study (now The Children's Oncology Group). *Pediatric blood & cancer*. 2004 Aug; 43(2):105–9. Epub 2004/07/06. eng. [PubMed: 15236274]
16. Takatsu N, Matsui T, Murakami Y, Ishihara H, Hisabe T, Nagahama T, et al. Adverse reactions to azathioprine cannot be predicted by thiopurine S-methyltransferase genotype in Japanese patients with inflammatory bowel disease. *Journal of gastroenterology and hepatology*. 2009 Jul; 24(7): 1258–64. Epub 2009/08/18. eng. [PubMed: 19682195]
17. Otterness D, Szumlanski C, Lennard L, Klemetsdal B, Aarbakke J, Park-Hah JO, et al. Human thiopurine methyltransferase pharmacogenetics: gene sequence polymorphisms. *Clinical pharmacology and therapeutics*. 1997; 62:60–73. [PubMed: 9246020]
18. Hon YY, Fessing MY, Pui CH, Relling MV, Krynetski EY, Evans WE. Polymorphism of the thiopurine S-methyltransferase gene in African Americans. *Human molecular genetics*. 1999; 8:371–6. [PubMed: 9931346]
19. McLeod H, Krynetski E, Relling MV, Evans WE. Genetic polymorphism of thiopurine methyltransferase and its clinical relevance for childhood acute lymphoblastic leukemia. *Leukemia*. 2000; 14:567–72. [PubMed: 10764140]
20. Exome Aggregation Consortium (ExAC) data set. Cambridge, MA: <http://exac.broadinstitute.org>Internet [accessed Jul 2016]
21. Yang JJ, Landier W, Yang W, Liu C, Hageman L, Cheng C, et al. Inherited NUDT15 variant is a genetic determinant of mercaptopurine intolerance in children with acute lymphoblastic leukemia. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2015 Apr 10; 33(11):1235–42. [PubMed: 25624441]
22. Schaeffeler E, Fischer C, Brockmeier D, Wernet D, Moerike K, Eichelbaum M, et al. Comprehensive analysis of thiopurine S-methyltransferase phenotype-genotype correlation in a large population of German-Caucasians and identification of novel TPMT variants. *Pharmacogenetics*. 2004 Jul; 14(7):407–17. [PubMed: 15226673]
23. Schmiegelow K, Forestier E, Kristinsson J, Soderhall S, Vettenranta K, Weinshilboum R, et al. Thiopurine methyltransferase activity is related to the risk of relapse of childhood acute lymphoblastic leukemia: results from the NOPHO ALL-92 study. *Leukemia*. 2009 Mar; 23(3): 557–64. Epub 2008/11/07. eng. [PubMed: 18987654]
24. Karas-Kuzelicki N, Mlinaric-Rascan I. Individualization of thiopurine therapy: thiopurine S-methyltransferase and beyond. *Pharmacogenomics*. 2009 Aug; 10(8):1309–22. [PubMed: 19663675]
25. Relling MV, Hancock ML, Boyett JM, Pui CH, Evans WE. Prognostic importance of 6-mercaptopurine dose intensity in acute lymphoblastic leukemia *Blood*. 1999; 93:2817–23.
26. Udaoka T, Torii C, Takahashi D, Mori T, Aramaki M, Kosaki R, et al. Comprehensive screening of the thiopurine methyltransferase polymorphisms by denaturing high-performance liquid chromatography. *Genetic testing*. 2005 Summer;9(2):85–92. [PubMed: 15943548]
27. Drogemoller BI, Wright GE, Warnich L. Considerations for rare variants in drug metabolism genes and the clinical implications. *Expert opinion on drug metabolism & toxicology*. 2014 Jun; 10(6): 873–84. [PubMed: 24673405]

28. Gordon AS, Tabor HK, Johnson AD, Snively BM, Assimes TL, Auer PL, et al. Quantifying rare, deleterious variation in 12 human cytochrome P450 drug-metabolism genes in a large-scale exome dataset. *Human molecular genetics*. 2014 Apr 15; 23(8):1957–63. [PubMed: 24282029]
29. Brunham LR, Chan SL, Li R, Aminkeng F, Liu X, Saw WY, et al. Pharmacogenomic diversity in Singaporean populations and Europeans. *The pharmacogenomics journal*. 2014 Dec; 14(6):555–63. [PubMed: 24861855]
30. Frayling TM, Timpson NJ, Weedon MN, Zeggini E, Freathy RM, Lindgren CM, et al. A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science (New York, NY)*. 2007 May 11; 316(5826):889–94.
31. Keller L, Xu W, Wang HX, Winblad B, Fratiglioni L, Graff C. The obesity related gene, FTO, interacts with APOE, and is associated with Alzheimer's disease risk: a prospective cohort study. *Journal of Alzheimer's disease : JAD*. 2011; 23(3):461–9. [PubMed: 21098976]
32. Merkestein M, McTaggart JS, Lee S, Kramer HB, McMurray F, Lafond M, et al. Changes in gene expression associated with FTO overexpression in mice. *PloS one*. 2014; 9(5):e97162. [PubMed: 24842286]
33. Caliskan M, Cusanovich DA, Ober C, Gilad Y. The effects of EBV transformation on gene expression levels and methylation profiles. *Human molecular genetics*. 2011 Apr 15; 20(8):1643–52. [PubMed: 21289059]
34. International HapMap C. The International HapMap Project. *Nature*. 2003 Dec 18; 426(6968):789–96. [PubMed: 14685227]
35. Choy E, Yelensky R, Bonakdar S, Plenge RM, Saxena R, De Jager PL, et al. Genetic analysis of human traits in vitro: drug response and gene expression in lymphoblastoid cell lines. *PLoS genetics*. 2008 Nov.4(11):e1000287. Epub 2008/12/02. eng. [PubMed: 19043577]
36. Zhang W, Ratain MJ, Dolan ME. The HapMap Resource is Providing New Insights into Ourselves and its Application to Pharmacogenomics. *Bioinform Biol Insights*. 2008; 2:15–23. Epub 2008/04/09. Eng. [PubMed: 18392109]
37. Van Loon JA, Weinshilboum RM. Thiopurine methyltransferase biochemical genetics: human lymphocyte activity. *Biochemical genetics*. 1982; 20:637–58. [PubMed: 7138494]
38. Matimba A, Li F, Livshits A, Cartwright CS, Scully S, Fridley BL, et al. Thiopurine pharmacogenomics: association of SNPs with clinical response and functional validation of candidate genes. *Pharmacogenomics*. 2014 Mar; 15(4):433–47. [PubMed: 24624911]
39. Ameyaw MM, Collie-Duguid ES, Powrie RH, Ofori-Adjei D, McLeod HL. Thiopurine methyltransferase alleles in British and Ghanaian populations. *Human molecular genetics*. 1999 Feb; 8(2):367–70. Epub 1999/02/05. eng. [PubMed: 9931345]
40. Pui CH, Campana D, Pei D, Bowman WP, Sandlund JT, Kaste SC, et al. Treating childhood acute lymphoblastic leukemia without cranial irradiation. *N Engl J Med*. 2009 Jun 25; 360(26):2730–41. Epub 2009/06/26. eng. [PubMed: 19553647]
41. Kadan-Lottick NS, Ness KK, Bhatia S, Gurney JG. Survival variability by race and ethnicity in childhood acute lymphoblastic leukemia. *JAMA : the journal of the American Medical Association*. 2003; 290(15):2008–14. [PubMed: 14559954]
42. Pollock BH, DeBaun MR, Camitta BM, Shuster JJ, Ravindranath Y, Pullen DJ, et al. Racial differences in the survival of childhood B-precursor acute lymphoblastic leukemia: a Pediatric Oncology Group Study. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2000; 18(4):813–23. [PubMed: 10673523]
43. Lennard L, Lilleyman JS, Van Loon J, Weinshilboum RM. Genetic variation in response to 6-mercaptopurine for childhood acute lymphoblastic leukaemia. *Lancet*. 1990; 336:225–9. [PubMed: 1973780]
44. Schmiegelow K, Bjork O, Glomstein A, Gustafsson G, Keiding N, Kristinsson J, et al. Intensification of mercaptopurine/methotrexate maintenance chemotherapy may increase the risk of relapse for some children with acute lymphoblastic leukemia. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2003; 21(7):1332–9. [PubMed: 12663723]

45. Lennard L, Chew TS, Lilleyman JS. Human thiopurine methyltransferase activity varies with red blood cell age. *British journal of clinical pharmacology*. 2001 Nov; 52(5):539–46. Epub 2001/12/12. eng. [PubMed: 11736862]
46. Ganiere-Monteil C, Medard Y, Lejus C, Bruneau B, Pineau A, Fenneteau O, et al. Phenotype and genotype for thiopurine methyltransferase activity in the French Caucasian population: impact of age. *European journal of clinical pharmacology*. 2004 Apr; 60(2):89–96. Epub 2004/03/17. eng. [PubMed: 15022030]
47. McLeod HL, Krynetski EY, Wilimas JA, Evans WE. Higher activity of polymorphic thiopurine S-methyltransferase in erythrocytes from neonates compared to adults. *Pharmacogenetics*. 1995; 5:281–6. [PubMed: 8563768]
48. Pettersson B, Almer S, Albertioni F, Soderhall S, Peterson C. Differences between children and adults in thiopurine methyltransferase activity and metabolite formation during thiopurine therapy: possible role of concomitant methotrexate. *TherDrug Monit*. 2002; 24(3):351–8.
49. Serpe L, Calvo PL, Muntoni E, D'Antico S, Giaccone M, Avagnina A, et al. Thiopurine S-methyltransferase pharmacogenetics in a large-scale healthy Italian-Caucasian population: differences in enzyme activity. *Pharmacogenomics*. 2009 Nov; 10(11):1753–65. Epub 2009/11/07. eng. [PubMed: 19891552]
50. Yates CR, Krynetski EY, Loennechen T, Fessing MY, Tai HL, Pui CH, et al. Molecular diagnosis of thiopurine S-methyltransferase deficiency: genetic basis for azathioprine and mercaptopurine intolerance. *Annals of internal medicine*. 1997; 126:608–14. [PubMed: 9103127]
51. McLeod HL, Relling MV, Liu Q, Pui CH, Evans WE. Polymorphic thiopurine methyltransferase in erythrocytes is indicative of activity in leukemic blasts from children with acute lymphoblastic leukemia. *Blood*. 1995; 85:1897–902. [PubMed: 7703493]
52. Yang JJ, Cheng C, Devidas M, Cao X, Fan Y, Campana D, et al. Ancestry and pharmacogenomics of relapse in acute lymphoblastic leukemia. *Nature genetics*. 2011 Mar; 43(3):237–41. Epub 2011/02/08. eng. [PubMed: 21297632]
53. Liu C, Yang W, Devidas M, Cheng C, Pei D, Smith C, et al. Clinical and Genetic Risk Factors for Acute Pancreatitis in Patients With Acute Lymphoblastic Leukemia. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2016 Jun 20; 34(18):2133–40. [PubMed: 27114598]
54. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*. 2010 Sep 1; 26(17):2190–1. Epub 2010/07/10. eng. [PubMed: 20616382]
55. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*. 2005; 21(2):263–5. [PubMed: 15297300]

Study Highlights

What is the current knowledge on the topic?

Dose-limiting toxicity due to *TPMT* defects resulted in treatment interruptions. Genotypic and phenotypic assays of *TPMT* activity in patients are used to adjust doses of thiopurines, assuming *TPMT* genotyping has adequate sensitivity and specificity.

What question did this study address?

This pharmacogenomics study evaluated the contribution of genetic variation in *TPMT* and other genes to *TPMT* activity, and their effects on thiopurine tolerance in two ALL cohorts.

What this study adds to our knowledge?

TPMT activity behaves as a monogenic trait in both cohorts. Additional genetic variants (besides the three SNPs rs1800462, rs1800460 and rs1142345 defining *TPMT* clinical genotype) did not substantially improve classification accuracy for *TPMT* phenotype.

How this might change clinical pharmacology or translational science?

This study confirms that current strategies for clinical *TPMT* genetic testing based on only three genomic variants adequately classify patients for their *TPMT* phenotype, thereby bolstering the use of *TPMT* genetic testing to assess *TPMT* status in the clinic.

Figure 1. TPMT activity by genotype and race

TPMT activity was measured in individual patients who were wild-type (WT; green), heterozygous (HET; red) or homozygous (HOM; yellow) for *TPMT*, defined by the presence of known no-function alleles *2, *3A or *3C. (A) Distribution of TPMT activity by clinical *TPMT* genotype (n = 1026). (B) Box plot of TPMT activity by genetically-determined race (n = 963). The number and percentage of heterozygous patients (based on clinical *TPMT* genotype) in each race group are shown. P value was determined using a general linear model after adjusting for protocol.

Figure 2. TPMT activity by diplotypes in the combined cohort (n = 884)

TPMT activity was measured in 306 SJ and 578 COG patients whose *TPMT* diplotypes were determined using Illumina HumanExome Beadchip Array. *TPMT* heterozygous (HET) was defined as the presence of known no-function alleles *2, *3A or *3C (shown in dashed box). P values between diplotypes were determined using a general linear model that included protocol as covariate.

Figure 3. Genome-wide association study featured *TPMT* as the strongest determinant of TPMT activity

Manhattan plots of the association between ~9 million typed or imputed SNPs and TPMT activity in (A) 306 SJ patients; (B) 578 COG patients; and (C) meta-analysis based on the combined results from SJ and COG patients using Stouffer's Z-score method. P values were determined using a general linear model after adjusting for genetic ancestry.

Figure 4. Dose intensity by diplotypes in the combined cohort (n = 839)

Overall dose intensity (i.e. prescribed dose/protocol dose) during maintenance therapy was estimated in 261 SJ patients and 578 COG patients whose *TPMT* diplotypes were determined using Illumina HumanExome Beadchip Array. *TPMT* heterozygous (HET) was defined as the presence of the known no-function alleles *2, *3A or *3C (shown in dashed box). P values between diplotypes were determined using a general linear model that included protocol as covariate.

Table 1

Patient characteristics by cohort.

Clinical characteristics	SJ (n = 306)	COG (n = 720)
Age		
< 10 years old	235 (76.8%)	545 (75.7%)
10 years old	71 (23.2%)	175 (24.3%)
Gender		
Male	171 (55.9%)	492 (68.3%)
Female	135 (44.1%)	228 (31.7%)
Race ^a		
White	206 (67.3%)	201 (30.6%)
Black	44 (14.4%)	94 (14.3%)
Hispanic	38 (12.4%)	213 (32.4%)
Asian	4 (1.3%)	53 (8.1%)
Other	14 (4.6%)	96 (14.6%)
<i>TPMT</i> clinical genotype ^b		
WT	266 (86.9%)	672 (93.3%)
HET	40 (13.1%)	47 (6.5%)
HOM	(0%)	1 (0.2%)

^aGenetic ancestries (European, African, Native American and Asian) were determined in 306 SJ patients and 657 COG patients using STRUCTURE,²² and either treated as a continuous variable or used to categorize patients as white (European ancestry > 90%), black (African ancestry > 70%), Hispanics (non-black patients with Native American ancestry > 10%), Asian (Asian ancestry > 90%), and “other” groups.

^b*TPMT* clinical genotype was determined using peripheral blood DNA at the beginning of remission induction, using PCR to detect the presence of three known no-function variants: rs1800462, rs1800460 and rs1142345 (contributing to *2, *3A and *3C alleles) and categorized as homozygous (HOM), heterozygous (HET) or wild-type (WT).

Table 2

Multivariate analysis of TPMT activity in both SJ and COG cohorts.

Variable	Meta-analysis		SJ (n = 306)		COG (n = 720)	
	P	Coefficient	P	Coefficient	P	Coefficient
Age (year)	6.6×10^{-4}	-0.16	8.0×10^{-3}	-0.066	0.030	
Male	0.98	-0.038	0.95	0.010	0.97	
African ancestry	1.4×10^{-7}	-2.72	4.7×10^{-3}	-2.49	3.7×10^{-6}	
Native American ancestry	0.15	-1.19	0.51	-0.89	0.17	
Asian ancestry	0.67	1.79	0.45	-0.08	0.89	
<i>TPMT</i> clinical genotype ^a	2.3×10^{-67}	-10.47	8.2×10^{-30}	-7.67	1.1×10^{-39}	

P values and coefficients are estimated using multivariate linear regression models. Regression coefficient indicates increase (positive) or decrease (negative) in TPMT activity (U/mL) with the presence of categorical variables (e.g. gender) or with certain levels of change in continuous variables (e.g. every one year increase in age). For example, patient with 100% African ancestry is estimated to be 2.72 U/mL lower than those with 100% European ancestry. Meta-analysis is performed on the results from SJ and COG patients using Stouffer's Z-score combined P-value method.

^aNumber of known no-function alleles (*2, *3A or *3C), i.e. homozygous deficiency = 2, heterozygous = 1.

Table 3

TPMT diplotypes across race groups in the combined SJ and COG cohort (n = 884).

Genotype	Diplotype	White (n = 360) n (%)	Black (n = 131) n (%)	Hispanic (n = 242) n (%)	Asian (n = 55) n (%)	Other (n = 96) n (%)
HOM	*3A/*3C	1 (0.3%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
HET	*1/*2	1 (0.3%)	0 (0%)	4 (1.7%)	0 (0%)	1 (1%)
	*1/*3A	33 (9.2%)	0 (0%)	17 (7.0%)	0 (0%)	2 (2.1%)
	*1/*3C	0 (0%)	16 (12.2%)	2 (0.8%)	0 (0%)	2 (2.1%)
WT	*1/*8	1 (0.3%)	5 (3.8%)	0 (0%)	0 (0%)	2 (2.1%)
	*1/*12	0 (0%)	1 (0.8%)	1 (0.4%)	0 (0%)	0 (0%)
	*1/*24	0 (0%)	7 (5.3%)	1 (0.4%)	0 (0%)	2 (2.1%)
	*1/*X	0 (0%)	0 (0%)	1 (0.4%)	0 (0%)	0 (0%)
	*1/*1	324 (90%)	102 (77.4%)	216 (89.3%)	55 (100%)	87 (90.6%)

TPMT diplotypes were determined using the Exome Array data, and categorized as homozygous variant (HOM), heterozygous (HET) or wild-type (WT). Race was categorized using genetic ancestries: white (European ancestry > 90%), black (African ancestry > 70%), Hispanics (non-black patients with Native American ancestry > 10%), Asian (Asian ancestry > 90%) and other.