



Pharmacogenetic determinants of thiopurines in an Indian cohort

Shaik Mohammad Naushad¹ · Mekala Janaki Ramaiah² · Vijay Kumar Kutala³ · Tajamul Hussain^{4,5} · Salman A. Alrokayan^{5,6}

Received: 4 July 2020 / Revised: 25 August 2020 / Accepted: 4 September 2020 / Published online: 15 September 2020
© Maj Institute of Pharmacology Polish Academy of Sciences 2020

Abstract

Background Several genetic variants of thiopurine metabolic pathway are associated with 6-thiopurine-mediated leucopenia. A population-based evaluation of these variants lays the foundation for Pharmacogenetic-guided thiopurine therapy.

Methods A total of 2000 subjects were screened for the pharmacogenetic determinants using the inifinium global screening array (GSA). The functional relevance of these variants was deduced using SNAP2, SIFT, Provean, Mutalyzer, Mutation Taster, Phyre2, SwissDock, AGGRESCAN, and CUPSAT.

Results The minor allele frequencies of NUDT15*3, NUDT15*5, TPMT*3C, TPMT*3B variant alleles were 6.78%, 0.11%, 1.98% and 0.69%, respectively. TPMT*3A genotype was observed in 0.35% subjects. No gender-based differences were observed in the incidence of these variants. Data from studies of the Indian population showed that 92.86% subjects heterozygous for NUDT15*3 and 60% subjects heterozygous for TPMT*3C exhibit thiopurine-mediated hematological toxicity. NUDT15 variants have no impact on the binding of ‘dGTP’ to the NUDT protein. NUDT15*3 variant increases aggregation ‘hot spot’ region and induces unfavourable torsion in the protein. NUDT15*5 destabilizes the protein and impairs Mg/Mn binding. TPMT*3A, TPMT*3B and TPMT*3C variants lower binding affinity to 6-mercaptopurine compared to the wild protein. TPMT*3C variant destabilizes the TPMT protein in the thermal experiment. Compared to the data of European and African/African American populations, NUDT15*3 frequency is higher and TPMT*3C frequency is lower in our population.

Conclusions TPMT variants were less frequent in Indian population, while NUDT15*3 is more frequent compared to European and African/African American populations. NUDT15*3 increases aggregation ‘hot spot’ and induces unfavourable torsion in the protein. NUDT15*5 and TPMT*3C destabilize the respective proteins. TPMT*3A, TPMT*3B and TPMT*3C are associated with a lower binding affinity towards 6-mercaptopurine.

✉ Shaik Mohammad Naushad
naushad@sandor.co.in

¹ Department of Biochemical Genetics and Pharmacogenomics, Sandor Speciality Diagnostics Pvt Ltd, Banjara Hills, Road No 3, Hyderabad 500034, India

² School of Chemical and Biotechnology, SASTRA University, Thanjavur, India

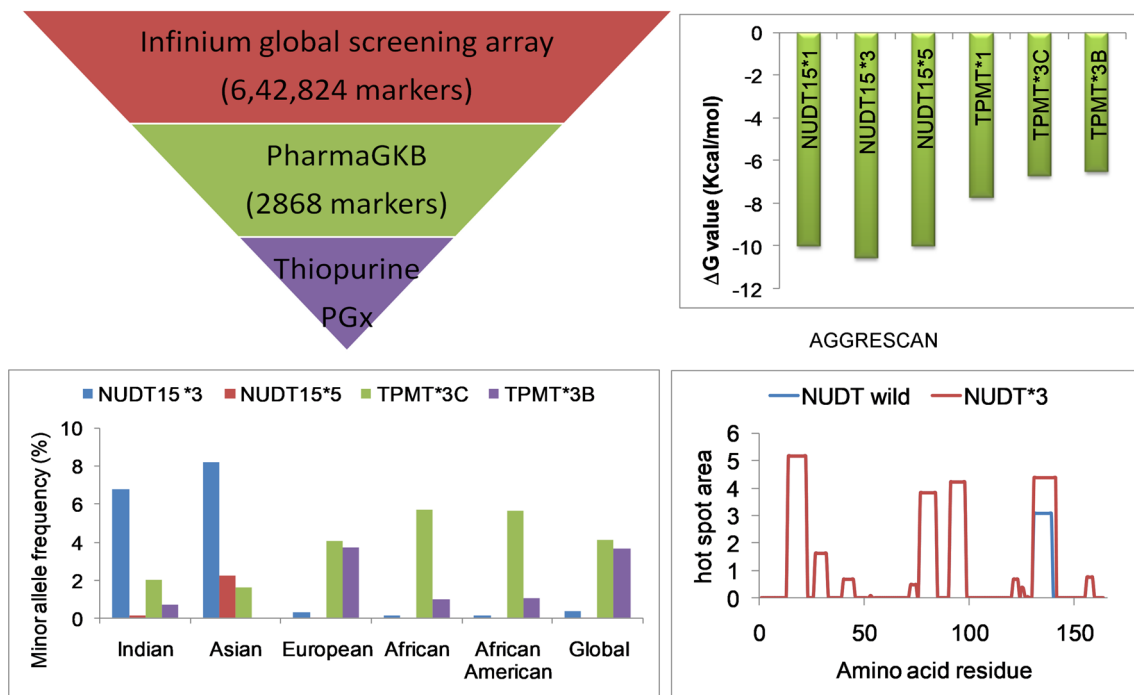
³ Department of Clinical Pharmacology and Therapeutics, Nizam’s Institute of Medical Sciences, Hyderabad, India

⁴ Center of Excellence in Biotechnology Research, College of Science, King Saud University, Riyadh 11451, Saudi Arabia

⁵ Biochemistry Department, College of Science, Research Chair for Biomedical Applications of Nanomaterials, King Saud University, Riyadh 11451, Saudi Arabia

⁶ Biochemistry Department, College of Science, King Saud University, Riyadh 11451, Saudi Arabia

Graphic abstract



Keyword Thiopurine · NUDT15 · TPMT · Global screening array · Bioinformatics analysis

Abbreviations

ALFA	Aggregate allele and genotype frequencies computed from non-sensitive dbGaP studies
ALL	Acute lymphoblastic leukemia
AZT	Azathioprine
DNA-TG	DNA-incorporated thioguanine
GSA	Global Screening array
IBD	Inflammatory bowel disease
MAF	Minor allele frequency
MP	Mercaptopurine
NUDT15	Nudix hydrolase 15
PCNA	Proliferating cell nuclear antigen
SAM	S-Adenosyl methionine
SIFT	Sorting intolerant from tolerant
TGN	Thioguanine
TPMT	Thiopurine S-methyltransferase

Introduction

Nudix hydrolase 15 (NUDT15) catalyzes the hydrolysis of nucleoside triphosphates (dGTP, dTTP, dCTP), their oxidized forms (8-oxo-dGTP) and thiopurine pro-drug derivatives (thio-dGTP and 6-thio-GTP) [1]. NUDT15 may have a role in the synthesis of DNA and in the cell cycle progression

by stabilizing the proliferating cell nuclear antigen (PCNA) [2]. NUDT15 mutations result in poor metabolism of thiopurines and hence are associated with thiopurine induced early onset leukopenia [3]. Thiopurine S-methyltransferase (TPMT) catalyzes the S-methylation of thiopurines using S-adenosyl-L-methionine (SAM) as the methyl donor. TPMT deficiency is well documented to be associated with thiopurine-mediated haematological toxicity [4].

A Korean study on 978 patients with Crohn's disease revealed a strong association of rs116855232 (NUDT15*3, c.415 C>T, p.R139C) with thiopurine-induced early leukopenia with a sensitivity of 89.4% and specificity of 93.2% [3]. In the same cohort, TPMT variants showed lesser sensitivity (12.1%) and higher specificity (97.6%) [3]. A study on 160 Japanese patients with inflammatory bowel disease (IBD) has shown the association of NUDT15*3 variant with azathioprine induced early and late leukopenia and severe hair loss [5]. This variant was the strongest predictor of the interval between the initiation to discontinuation of azathioprine therapy [5]. The association of this variant with azathioprine (AZT) induced leukopenia was evident even in Chinese patients with autoimmune disorders [6].

The ratio of DNA-incorporated thioguanine (DNA-TG) in white blood cells to thioguanine (TGN) in red blood cells was reported to be higher in NUDT15 deficiency in a study

of Japanese children with acute lymphoblastic leukemia (ALL) [7]. A genome-wide association study identified that rs1142345 (TPMT c.719 A>G) and rs116855232 were associated with 6-mercaptopurine (6-MP) dose intensity [7]. Patients with NUDT15 TT-genotype were found to be highly sensitive to 6-MP requiring only 8.3% of recommended dose [8]. A significant loss in the catalytic efficiency (V_{\max}/K_m) was observed in the presence of rs116855232 (23.6 ± 0.9) and rs186364861 (27.7 ± 1.1) NUDT15 variant alleles when compared to wild genotype (107.9 ± 0.2) [9]. These loss-of-function alleles were shown to have an association with the thiopurine intolerance across three cohorts [9]. A recent study from India also demonstrated the association of rs116855232 with haematological toxicity in children with ALL [10]. Recently, we have reported a classification and regression model for predicting the 6-MP mediated haematological toxicity based on TPMT and ITPA genetic variants [11]. This model has projected that TPMT*12 and TPMT*3C are the key determinants of toxicity [11].

The current study was aimed to identify Indian population-specific pharmacogenetic determinants of thiopurines by examining the minor allele frequency of 2000 genomes using Infinium global screening array (GSA). Given wide ethnic and geographical variations in the distribution of these polymorphisms, identifying clinically actionable variants in a given population is essential to optimize pharmacogenetic testing to improve the efficacy and safety of 6-MP.

Materials and methods

Recruitment of subjects

A total of 2000 adults (1215 men and 785 women) in the age group of 18–72 year were recruited for this study from July 2012 to December 2019. These subjects were healthy and had no clinical implications at the time of recruitment. All the subjects consented for this study. The study protocol was approved by the Institutional Ethical Committee of Sandor Proteomic Pvt Ltd, Hyderabad (EC/SRP/008/2012, dated 26.6.12). Whole blood samples were collected in EDTA from these subjects.

Genotyping

Genomic DNA was isolated from these blood samples using a commercial kit (QIAamp DNA mini kit, Qiagen, USA) as per the manufacturer's protocol. A total of 200 ng of genomic DNA was used to genotype 6,42,824 markers using the Infinium GSA v 1.0 wherein genome amplification, fragmentation, and hybridization were performed as per the manufacturer's instructions. The samples were subjected to enzymatic base extension and fluorescent staining. The bead

chips were imaged using the iScan system and the analysis and genotype calling was performed using the Genome Studio V2011.1 software. The pharmacogenetic markers were filtered based on the PharmaGKB clinical annotations of drug-specific variants with specific emphasis on variants of level 1 and level 2 significance; and SNP missing call rate < 0.03.

Bioinformatics analysis

The data from the "Aggregate allele and genotype frequencies computed from non-sensitive dbGaP studies (ALFA)" project was used to deduce minor allele frequencies European, African, African American, Asian, and global populations. The impact of each mutation on the protein sequence was assessed using the mutalyzer software. Other effects such as splicing, binding properties, etc. were elucidated using the mutation taster database. Sorting Intolerant from Tolerant (SIFT) and Proven databases scores were used to assess whether the amino acid substitution is damaging/deleterious or tolerant. CUPSAT module was used to assess whether the identified genetic variants affect the protein stability or torsion under thermal and denaturant experimental conditions.

Using the crystal structure of *Homo sapiens* NUDT15 in complex with 6-thio-GMP (PDB ID: 5LPG) as the template, the three dimensional of NUDT15*3 and NUDT15*5 proteins were deduced through homology modeling with Phyre2 module. SwissDock was used to dock dGTP on to the wild and mutant proteins of NUDT15. The ΔG values were calculated to assess the binding affinity. AGGRESCAN module was used to identify the aggregation-prone areas in NUDT15.

Using the crystal structure of *Homo sapiens* thiopurine S-methyltransferase (PDB ID: 2BZG) as the template, the three dimensional of TPMT*3B and TPMT*3C proteins were deduced through homology modeling with Phyre2 module. SwissDock was used to dock 6-MP on to the wild and mutant proteins of TPMT. The ΔG values were calculated to assess the binding affinity. AGGRESCAN module was used to identify the aggregation-prone areas in TPMT.

Statistical analysis

The minor allele frequencies were calculated by dividing the number of variant alleles with the total number of (variant and wild) alleles. Fisher exact test was carried out to compare the minor allele frequencies of Indian population (baseline) with other populations by computing the data in 2×2 contingency table. The data is depicted in the form of odds ratio and 95% confidence interval. Logistic regression analysis was carried out to explore gender differences in the incidence of pharmacogenetic markers. Principal component

analysis was used to assess the influence of bioinformatics variables in differentiating level 1A and level 2B variants of thiopurines. Unit variance scaling is applied to rows; singular value decomposition (SVD) with imputation is used to calculate principal components.

Results

Filtering the markers based on pharmacological relevance

Out of the 6,42,824 genetic markers, 2868 were PharmGKB markers, out of which 33 are of high priority with level 1 significance with stringent CPIC guidelines. There are 110 markers with moderate relevance with level 2 significance in PharmGKB. From markers of level 3 significance, 68 were having borderline evidence of association. Out of these 211 markers, only 67 markers were present in our population (Fig. 1). From these 67 pharmacogenetic markers, only four markers were found to be relevant in the pharmacogenetics of 6-MP.

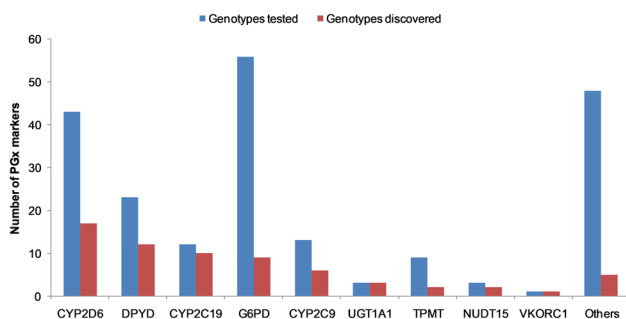


Fig. 1 The pharmacogenetic determinants of GSA. This illustrates the distribution of 211 pharmacogenetic (PGx) markers (level 1: 33, level 2: 110, level 3: 66) in the total panel with documented evidence as per PharmaGkb database. Further, the distribution of PGx markers identified in Indian population were depicted gene-wise in a comparative bar diagram

Distribution of pharmacogenetic determinants of 6-MP in our population

As shown in Table 1, NUDT15*3 (rs116855232) is the most prevalent allele with a minor allele frequency (MAF) of 6.778%. NUDT15*5 (rs186364861) is the least prevalent allele with a MAF of 0.1063%. TPMT*3C (rs1142345) and TPMT*3A (c.460G>A) has MAFs of 1.979% and 0.6937%, respectively. The cumulative variant allele frequency of NUDT15 and TPMT loci was 9.557%. In TPMT, seven subjects harboured *3A (*3B and *3C together) while 70 have *3C allele and 29 have *3B allele alone. All the variants are in accordance with Hardy–Weinberg equilibrium. Logistic regression analysis revealed no significant gender-based differences in the distribution of these variants (Table 2).

Bioinformatics analysis of NUDT15 and TPMT variants

As shown in Table 3, NUDT15*3 was predicted to be ‘tolerated’ and ‘neutral’ as per SIFT and Provean scores. However, this variant affects splicing (gcggtgtttaaaa/GAAC). The nudix hydrolase domain (amino acid residues 9–145) and the region that interacts with proliferating cell nuclear antigen (PCNA) (near residues 76–164) were predicted to be lost due to this variation. NUDT15*5 was predicted to be ‘damaging’ by SIFT and ‘neutral’ by Provean. This variant was found to affect the splicing (AGTC/gtgg). In addition to the nudix hydrolase domain, PCNA interacting domain, there is a loss of metal-binding domain 63 and 67 residues. TPMT*3C and TPMT*3A variants were predicted to be ‘damaging’ and ‘deleterious’ as per SIFT and Provean scores. TPMT*3C was predicted to affect splicing (gaaa/AGTTATAT). The strand between 236–244 residues was predicted to be lost due to this variant. SNAP2 scores of all the four variants are > 50 (potentially damaging).

The ΔG values for dGTP binding with NUDT15*1, NUDT15*3 and NUDT15*5 were -9.98 kcal/mol, -10.55 kcal/mol, and -10.01 kcal/mol, respectively (Fig. 2). The AGGRESCAN data is suggestive of increased aggregation-prone area in NUDT15*3 in comparison to

Table 1 Distribution of pharmacogenetic determinants of 6-mercaptopurine in Indian population

Gene	rs number	Nucleotide change	Protein change	WW	WM	MM	MAF (%)	P_{HWE}
NUDT15 (NM_018283.4)*3	rs116855232	c.415C>T	R139C	1738	253	9	6.778	1.00
NUDT15 (NM_018283.4)*5	rs186364861	c.52G>A	V18I	1996	4	0	0.1063	1.00
TPMT*3C	rs1142345	c.719A>G	Y240C	1923	77	0	1.979	0.61
TPMT*3B	rs1800460	c.460G>A	A154T	1973	27	0	0.6937	0.85

The distribution of wild, heterozygous and homozygous mutant genotypes was tested for Hardy–Weinberg equilibrium (HWE) and ‘p’ values were calculated based on χ^2 -test for observed and expected frequencies. TPMT*3A: 7; TPMT*3C: 70; TPMT*3B: 20

WW wild, WM heterozygous, MM homozygous mutant

Table 2 Logistic regression analysis to explore gender differences in the distribution of thiopurine SNPs

Variant	Males			Females			Males vs. females
	WW	WM	MM	WW	WM	MM	OR (95% CI)
rs116855232	1057	153	5	681	100	4	0.96 (0.75–1.24)
rs186364861	1213	2	0	783	2	0	0.63 (0.09–4.51)
rs1142345	1172	43	0	751	34	0	0.83 (0.52–1.32)
rs1800460	1201	14	0	772	13	0	0.72 (0.33–1.55)

No gender difference was observed in the distribution of these genotypes

WW wild, WM heterozygous, MM homozygous mutant, OR odds ratio, CI confidence interval

NUDT15*1 and NUDT15*5 (Fig. 3). CUPSAT analysis showed that NUDT15*3 variant contributes to unfavorable torsion in the thermal and denaturant experimental methods, although there is no effect on the overall stability. NUDT15*5 variant has destabilizing effect on the protein in thermal and denaturant experimental methods. TPMT*3C variant has destabilizing effect on the protein in thermal experiment. TPMT*3B variant has no influence on the protein stability or torsion (ΔG : 2.9 kcal/mol). (Table 3).

The ΔG values for 6-MP binding with TPMT*1, TPMT*3A, TPMT*3B, and TPMT*3C were -7.73 kcal/mol, -7.36 kcal/mol, -6.49 kcal/mol, and -6.71 kcal/mol, respectively (Fig. 4). The AGGRESCAN data showed no differences in the aggregation profiles of wild and mutant proteins of TPMT (Fig. 5). Principal component analysis was able to differentiate level 1A and level 2B variants using bioinformatics variables (Fig. 6).

Comparison with other populations

When Indian data of NUDT15*3 was compared with other populations, no significant differences were observed in South Asian population while East Asian population showed 1.44-folds higher prevalence compared to Indians. The European, African and African-American population showed significantly lower prevalence of NUDT15*3 allele. The NUDT15*5 prevalence of Indians was found to be similar to South Asians. However, East Asians showed 14.07-folds higher prevalence of this variant allele compared to Indians. The frequency of NUDT15*5 in European population was significantly lower than Indians. African and African American populations didn't exhibit this variant (Table 4).

The frequency of TPMT*3C variant in Indians was found to be similar to South Asians and East Asians. Europeans, Africans and African Americans showed 2.09-, 2.98- and 2.96-folds increase in TPMT*3C frequency, respectively compared to Indians. TPMT*3B variant was absent in East Asians while South Asians showed no significant difference when compared to Indians. European population showed 5.44-fold higher frequency of this variant while African and African American populations showed no significant differences in frequencies compared to Indians (Table 4).

Indian studies depicting association of these variants with thiopurine-mediated toxicity

There are three studies from India that examined the effect of TPMT and NUDT variants together on thiopurine-mediated toxicity. In a cohort of 101 patients with auto immune disorders, 7.9% were heterozygous for NUDT15*3 and 1.9% were homozygous mutant [12]. Following thiopurine therapy, heterozygous subjects developed leukopenia within 10 weeks of therapy while homozygous mutants had pancytopenia within two weeks of therapy [12]. In a cohort of 63 acute lymphoblastic leukemia patients, 3.17% were heterozygous for TPMT*3C while 9.52% were heterozygous for NUDT15*3 [13]. Subjects heterozygous for these variants exhibited neutropenia following 6-MP therapy [13]. Kodidela et al. observed one out of two heterozygous subjects of TPMT*3C (50%); 12 out of 14 heterozygous subjects of NUDT15*3 (85.71%) exhibiting severe haematological toxicity in acute lymphoblastic leukemia following treatment with 6-MP [10]. (Table 5) The data of all these three studies of the Indian population revealed that 26 out of 28 NUDT15*3 heterozygous subjects (92.86%) and 3 out of 5 TPMT*3C subjects (60%) exhibiting thiopurine-mediated hematological toxicity.

Discussion

The current study is the largest population-based study to explore the pharmacogenetic determinants of thiopurines from a set of 2868 PharmaGkb markers using GSA. Two variants of NUDT15 (*3 and *5) and two variants of TPMT (*3C and *3A) were identified to be clinically relevant determinants with documented evidence of their association with thiopurine-induced hematological toxicity in IBD, autoimmune disorders and ALL. The cumulative frequency of these allelic variants in our population is 9.557%. The MAFs are in the following order: NUDT15*3 > TPMT*3C > TPMT*3A > NUDT15*5. The comparison of population data revealed highest incidence of NUDT15*3, NUDT15*5 variants in East Asians; TPMT*3C in African/African Americans; and TPMT*3B in Europeans. A combination

Table 3 Bioinformatics analysis to deduce the functional relevance of the identified mutations

Variant	Class	SNAP2 score	SIFT score	Provean score	Thermal stability ΔG Kcal/mol	Thermal torsion (unfavorable)	Denaturant stability ΔG Kcal/mol	Denaturant torsion (unfavorable)	Splicing	Increased aggregation
NUDT15*3	1A	57	0.161	-2.02	1.11	1	3.79	1	1	1
NUDT15*5	2B	67	0.011	-0.85	-2.81	0	-1.06	0	1	0
TPMT*3C	1A	75	0.028	-5.15	-0.85	0	3.9	0	1	0
TPMT*3B	1A	85	0.023	-2.53	2.9	0	1.37	0	0	0

The impact of each mutation on the protein function was elucidated by SNAP2, SIFT and Provean scores. CUPSAT module was used to determine the thermal/denaturant stability based on the amino acid atom potential and torsion angles

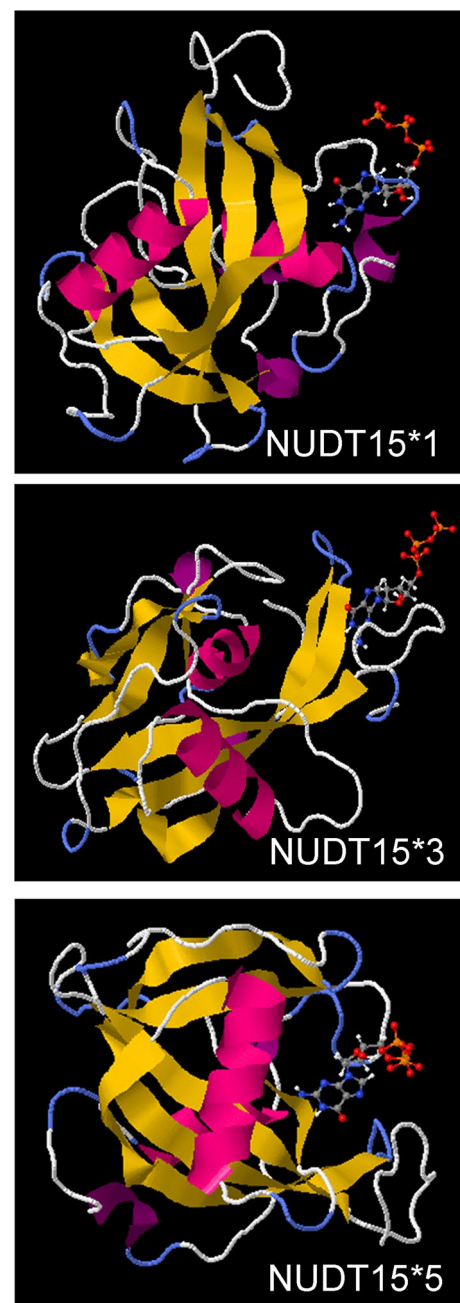


Fig. 2 Depicting the binding of dGTP with wild and mutant NUDT15 proteins. The three dimensional structures of wild and mutant NUDT15 proteins were docked with dGTP to determine the binding affinity of dGTP to NUDT15. The ΔG values for dGTP binding with NUDT15*1, NUDT15*3 and NUDT15*5 were -9.98 kcal/mol, -10.55 kcal/mol, and -10.01 kcal/mol, respectively

of bioinformatics tools were used to predict the functional relevance of the identified variants, which were able to differentiate level 1A and level 2B variants in the principal component analysis. The data from studies of the Indian population showed that 92.86% subjects heterozygous for

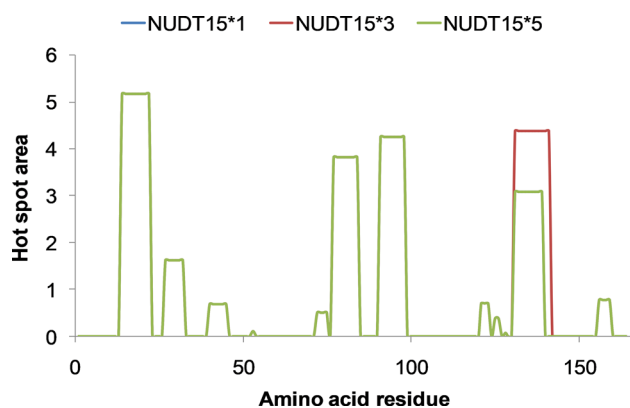


Fig. 3 AGGRESKAN of NUDT15 variants. Aggrescan module was used to predict “hot spots” of aggregation in NUDT15 based on aminoacid aggregation propensity value. The aggrescan profiles of NUDT15*1 (wild), NUDT15*3 and NUDT15*5 were overlapped to examine changes in aggrescan profile. NUDT15*3 variant showed an increase in aggregation-prone area compared to NUDT15*1 and NUDT15*5

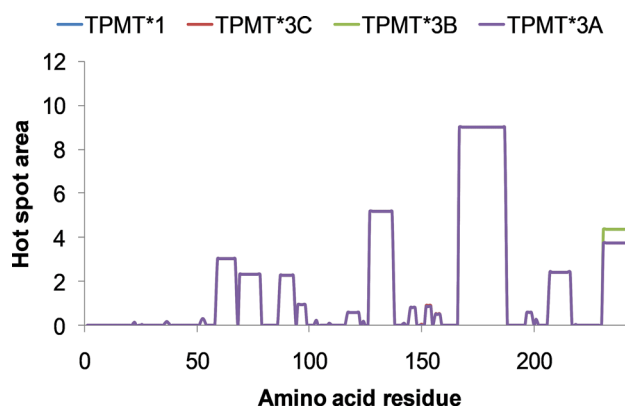


Fig. 5 AGGRESKAN of TPMT variants. Aggrescan module was used to predict “hot spots” of aggregation in TPMT based on aminoacid aggregation propensity value. The aggrescan profiles of TPMT*1, TPMT*3A, TPMT*3B and TPMT*3C were overlapped to examine changes in aggrescan profile. There is slight increase in aggregation-prone area in TPMT*3B compared to wild and other variant proteins, but it was not statistically significant

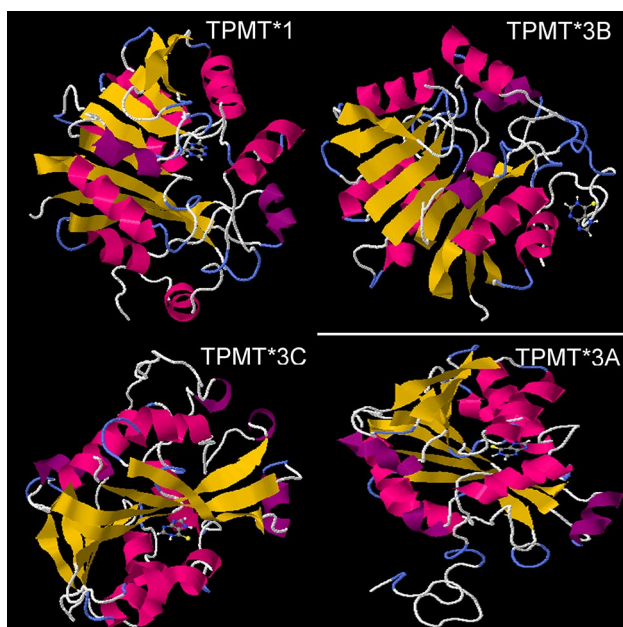


Fig. 4 Depicting the binding of 6-MP with wild and mutant TPMT proteins. The three dimensional structures of wild and mutant TPMT proteins were docked with 6-MP to determine the binding affinity changes of 6-MP towards TPMT. The ΔG values for 6-MP binding with TPMT*1, TPMT*3A, TPMT*3B, and TPMT*3C were -7.73 kcal/mol, -7.36 kcal/mol, -6.49 kcal/mol, and -6.71 kcal/mol, respectively

NUDT15*3 and 60% subjects heterozygous for TPMT*3C exhibit thiopurine-mediated hematological toxicity.

In vitro and cellular studies have suggested that 6-thio-dGTP and 6-thio-GTP are the favored substrates for NUDT15 [14]. The NUDT15*3 variation was not associated

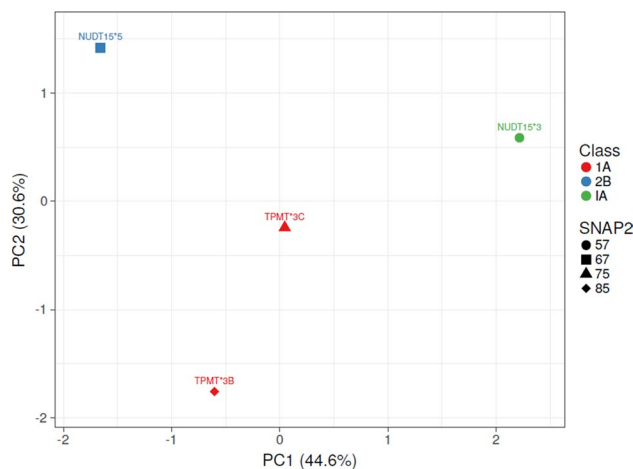


Fig. 6 Principal component analysis of bioinformatics variables in segregating thiopurine pharmacogenetic determinants. Unit variance scaling is applied to rows; SVD with imputation is used to calculate principal components. X and Y axis show principal component 1 and principal component 2 that explain 44.6% and 30.6% of the total variance, respectively. $N=4$ data points. Using Bioinformatics variables, level 1A (TPMT*3B, TPMT*3C, NUDT15*3) and level 2B (NUDT15*5) variants were clearly differentiated

with loss of enzyme activity, but it was shown to influence the protein stability negatively as a result of the loss of supportive intramolecular bonds that induce rapid proteasomal degradation in cells [14]. This study corroborated with our in silico analyses that depicted no differences in ΔG values of NUDT15*1, NUDT15*3 and NUDT15*5; and increased aggregation-prone area in NUDT15*3 mutant. SIFT and Provean scores for this variant showing “Tolerated/Neutral” corroborated with the above study. Our bioinformatics

Table 4 Comparison of allele frequencies with other populations

Population	NUDT15*3 rs116855232		NUDT15*5 rs186364861		TPMT*3C rs1142345		TPMT*3B rs1800460	
	M/W	OR (95%CI)	M/W	OR (95%CI)	W/M	OR (95%CI)	W/M	OR (95%CI)
Indian	271/3729	Reference	4/3996	Reference	79/3921	Reference	28/3972	Reference
South Asian	137/1819	1.04 (0.84–1.28)	2/1954	1.02 (0.19–5.59)	33/1923	0.85 (0.57–1.28)	8/1948	0.58 (0.27–1.28)
East Asian	192/1824	1.44 (1.12–1.41)*	28/1988	14.07 (4.93–40.17)*	44/1972	1.11 (0.76–1.61)	0/2016	–
European	470/158,406	0.04 (0.035–0.048)*	5/154,699	0.03 (0.01–0.12)*	7195/170,933	2.09 (1.67–2.61)*	9426/245,750	5.44 (3.75–7.90)*
African	2/1686	0.016 (0.004–0.07)*	0/1672	–	248/4128	2.98 (2.31–3.86)*	76/7808	1.38 (0.89–2.13)
African American	2/1654	0.017 (0.004–0.07)*	0/1640	–	238/3990	2.96 (2.29–3.83)*	76/7552	1.43 (0.92–2.21)

The data from the “Aggregate allele and genotype frequencies computed from non-sensitive dbGaP studies (ALFA)” project was used to deduce minor allele frequencies European, African, African Americans. South Asian and East Asian data was obtained from 1000 genome project. Fisher exact test was carried out using our data as baseline to compare the distribution with other populations

M mutant/variant allele, *W* wild allele, *OR* odds ratio, *CI* confidence interval

* $p < 0.0001$

Table 5 Indian studies showing association of NUDT15*3 and TPMT*3C with thiopurine mediated toxicity

Author	Disease	Thiopurine dose	Gene	WW	WM	MM	Association with toxicity
Shah et al. [12]	Autoimmune disorders	25 mg/day	TPMT*3C	100	1	0	No adverse drug reaction in WM
			NUDT15*3	91	8	2	Leukopenia in 10 weeks (WM); Pancytopenia in 2 weeks (MM)
Khera et al. [13]	Acute lymphoblastic leukemia	Mean cumulative dose: 10,927 vs. 12,533 mg/m ²	TPMT*3C	61	2	0	Risk of neutropenia pronounced in WM
			NUDT15*3	57	6	0	Risk of neutropenia pronounced in WM
Kodidela et al. [10]	Acute lymphoblastic leukemia	4100 ± 1973 mg/m ² (no toxicity) 3657 ± 1697 mg/m ² (toxicity)	TPMT*3C	69	2	0	One WM showed severe haematological toxicity
			NUDT15*3	57	14	0	Twelve WMs showed severe haematological toxicity

WW wild, WM heterozygous, MM homozygous mutant

analysis showed loss of Nudix hydrolase domain and loss of the region interacting with PCNA.

NUDT15 415T-allele carriers and subjects with TT-genotype were associated with 3.79- and 6.54-fold increased risk for thiopurine-induced leukocytopenia, respectively in a meta-analysis of 1138 patients [15]. This analysis further showed that NUDT15*3 variant is common in Asians and Hispanics, but rare in Europeans and Africans as demonstrated in this study also [14]. Another meta-analysis of 3538 patients projected that three variants of NUDT15 i.e., rs116855232 (*3), rs554405994, and rs186364861 (*5) as the clinically relevant predictors of thiopurine-induced

leucopenia with 8.44-, 4.34- and 2.74-folds risk, respectively [16].

Subjects heterozygous for the loss-of-function variants of TPMT (*2, *3A, and *3C) were reported to exhibit lower TPMT activity and lower dose intensity than the subjects with *1/*1 genotype [17]. Molecular dynamic simulations revealed that TPMT*3C (Y240C) is located in the beta-strand 9, which is distant from the active site [18]. The side-chain contacts between the 240 residue and helix alpha8 are lost due to mutation [18]. A154T polymorphism is within the co-substrate binding site and the larger threonine alters the packing of substrate binding residues at P68, L69, and Y166 thus increasing the solvent exposure of the polymorphic site

[18]. This rearrangement contributes to the lack of activity in the A154T mutant [18]. Among the different TPMT variants, *3A, *3B and *5 were reported to have undetectable activity in an in vitro functional study [19]. Our in silico study corroborates with these studies in demonstrating the lower binding affinity of 6-MP in TPMT*3B and TPMT*3C variants in comparison to TPMT*1 (wild protein). No differences were observed between the wild and mutant TPMT proteins in terms of aggregation hot spots.

Among the identified variants of the current study, TPMT*3A, TPMT*3B, TPMT*3C, and NUDT15 rs116855232 were considered as level 1A pharmacogenetic determinants while NUDT15*5 was considered as level 2B determinant. As per the 2018 update of Clinical Pharmacogenetics Implementation Consortium (CPIC) guidelines, subjects heterozygous for TPMT*3B and *3C mutations should be treated with a 30–80% of reduced dose of 6-MP (standard dose: 75 mg/m²/day or 1.5 mg/kg/day) and dose adjustments are recommended based on the degree of myelosuppression [20]. In homozygous mutants and subjects with *3A genotype, a tenfold reduction in the daily dose is recommended and dose adjustments to be done based on the degree of myelosuppression. Subjects heterozygous for NUDT15*3 allele were recommended to be treated with 30–80% of reduced dose of 6-MP with the dose adjustments based on the degree of myelosuppression [20]. Homozygous mutants of NUDT15*3 should be given 10 mg/m²/day and dose adjustments should be performed based on the degree of myelosuppression [20].

The major strengths of the current study are (i) population-based evaluation of NUDT15 and TPMT alleles that serves as the foundation for the future research on the implementation of pharmacogenetic testing in the clinical setting; (ii) application of different computational tools for a better understanding of the impact of these mutations on the protein level. Data from studies of the Indian population was used to demonstrate the association of TPMT and NUDT variants with thiopurine-mediated toxicity. The major limitation of the current study is lack of data corresponding to biochemical assays of TPMT activity or TGN-levels and emphasis was given only on the genetic aspects. Nevertheless, all the reported variants were well characterized earlier by several researchers that justify the clinical relevance. PharmaGkb clinical annotation showing level 1A for four variants level 2B for one variant substantiates it further.

To conclude, the cumulative MAF of level 1A alleles of 6-MP pharmacogenetic determinants (NUDT15*3, TPMT*3A, TPMT*3B, TPMT*3C) is 9.4507%. The MAF of level 2B allele (NUDT15*5) was 0.1063%. NUDT15*3 and NUDT15*5 variants showed no differences in the binding affinity towards dGTP. However, the substitution of arginine with cysteine in NUDT15*3 variant increased the aggregation-prone area. TPMT*3A, TPMT*3B and

TPMT*3C variants exhibited lower binding affinity to 6-MP than the wild TPMT (TPMT*1). The NUDT15 and TPMT variant alleles are well documented to increase hematological toxicity following thiopurine therapy in IBD, autoimmune disorders, and ALL, which is also further evident from the data from studies of Indian population. Bioinformatics variables were able to differentiate level 1A and level 2B alleles in the Principal Component Analysis. Pharmacogenetic testing for these markers followed by the implementation of CPIC guidelines in therapy may improve the efficacy and safety of thiopurine therapy.

Acknowledgements The study was funded by intramural Grants of Sandor Speciality Diagnostics Pvt Ltd, Hyderabad, India. The authors are grateful to the Deanship of Scientific Research, King Saud University for funding through Vice Deanship of Scientific Research Chairs.

Author contributions The study was conceived and designed by SMN and VKK. The recruitment of subjects, extraction of DNA, GSA analysis, and genotyping were performed by SMN, MJR, VKK, TA, and SAA. The data analysis and bioinformatics analysis were performed by SMN, MJR, and TA. The manuscript was drafted by SMN and MJR. The manuscript was revised for the critical content by TA and SAA. The final version of the manuscript was approved by all the authors.

Compliance with ethical standards

Conflict of interest All the authors hereby declare no conflicts of interest.

References

1. Carter M, Jemth AS, Hagenkort A, Page BD, Gustafsson R, Griese JJ, et al. Crystal structure, biochemical and cellular activities demonstrate separate functions of MTH1 and MTH2. *Nat Commun.* 2015;4(6):7871.
2. Yu Y, Cai JP, Tu B, Wu L, Zhao Y, Liu X, et al. Proliferating cell nuclear antigen is protected from degradation by forming a complex with MutT Homolog2. *J Biol Chem.* 2009;284(29):19310–20.
3. Yang SK, Hong M, Baek J, Choi H, Zhao W, Jung Y, et al. A common missense variant in NUDT15 confers susceptibility to thiopurine-induced leukopenia. *Nat Genet.* 2014;46(9):1017–20.
4. Evans WE, Hon YY, Bomgaars L, Coutre S, Holdsworth M, Janco R, et al. Preponderance of thiopurine S-methyltransferase deficiency and heterozygosity among patients intolerant to mercaptopurine or azathioprine. *J Clin Oncol.* 2001;19(8):2293–301.
5. Sato T, Takagawa T, Kakuta Y, Nishio A, Kawai M, Kamikozuru K, et al. NUDT15, FTO, and RUNX1 genetic variants and thiopurine intolerance among Japanese patients with inflammatory bowel diseases. *Intest Res.* 2017;15(3):328–37.
6. Fei X, Shu Q, Zhu H, Hua B, Wang S, Guo L, et al. NUDT15 R139C variants increase the risk of azathioprine-induced leukopenia in Chinese autoimmune patients. *Front Pharmacol.* 2018;9:460.
7. Moriyama T, Nishii R, Lin TN, Kihira K, Toyoda H, Jacob N, et al. The effects of inherited NUDT15 polymorphisms on thiopurine active metabolites in Japanese children with acute lymphoblastic leukemia. *Pharmacogenet Genomics.* 2017;27(6):236–9.
8. 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. *J Clin Oncol*. 2015;33(11):1235–42.
9. Moriyama T, Nishii R, Perez-Andreu V, Yang W, Klussmann FA, Zhao X, et al. NUDT15 polymorphisms alter thiopurine metabolism and hematopoietic toxicity. *Nat Genet*. 2016;48(4):367–73.
 10. Kodidela S, Dorababu P, Thakkar DN, Dubashi B, Sundaram R, Muralidharan N, et al. Association of NUDT15*3 and FPGS 2572C>T variants with the risk of early hematologic toxicity during 6-MP and low-dose methotrexate-based maintenance therapy in Indian Patients with acute lymphoblastic leukemia. *Genes (Basel)*. 2572C;11(6):E594.
 11. Naushad SM, Dorababu P, Rupasree Y, Pavani A, Raghunadharao D, Hussain T, et al. Classification and regression tree-based prediction of 6-mercaptopurine-induced leucopenia grades in children with acute lymphoblastic leukemia. *Cancer Chemother Pharmacol*. 2019;83(5):875–80.
 12. Shah SAV, Paradkar MU, Desai DC, Ashavaid TF. Preemptive NUDT15 genotyping: redefining the management of patients with thiopurine-induced toxicity. *Drug Metab Pers Ther*. 2018;33(1):57–60.
 13. Khera S, Trehan A, Bhatia P, Singh M, Bansal D, Varma N. Prevalence of TPMT, ITPA and NUDT 15 genetic polymorphisms and their relation to 6MP toxicity in north Indian children with acute lymphoblastic leukemia. *Cancer Chemother Pharmacol*. 2019;83(2):341–8.
 14. Valerie NC, Hagenkort A, Page BD, Masuyer G, Rehling D, Carter M, et al. NUDT15 Hydrolyzes 6-thio-deoxyGTP to mediate the anticancer efficacy of 6-thioguanine. *Cancer Res*. 2016;76(18):5501–11.
 15. Zhang AL, Yang J, Wang H, Lu JL, Tang S, Zhang XJ. Association of NUDT15 c.415C>T allele and thiopurine-induced leukocytopenia in Asians: a systematic review and meta-analysis. *Ir J Med Sci*. 2018;187(1):145–53.
 16. Cargin S, Genazzani AA, Canonico PL, Terrazzino S. Diagnostic accuracy of NUDT15 gene variants for thiopurine-induced leukopenia: a systematic review and meta-analysis. *Pharmacol Res*. 2018;135:102–11.
 17. Liu C, Yang W, Pei D, Cheng C, Smith C, Landier W, et al. Genomewide approach validates thiopurine methyltransferase activity is a monogenic pharmacogenomic trait. *Clin Pharmacol Ther*. 2017;101(3):373–81.
 18. Rutherford K, Daggett V. Four human thiopurine S-methyltransferase alleles severely affect protein structure and dynamics. *J Mol Biol*. 2008;379(4):803–14.
 19. Salavaggione OE, Wang L, Wiepert M, Yee VC, Weinshilboum RM. Thiopurine S-methyltransferase pharmacogenetics: variant allele functional and comparative genomics. *Pharmacogenet Genom*. 2005;15(11):801–15.
 20. Relling MV, Schwab M, Whirl-Carrillo M, Suarez-Kurtz G, Pui CH, Stein CM, et al. Clinical pharmacogenetics implementation consortium guideline for thiopurine dosing based on TPMT and NUDT15 genotypes: 2018 update. *Clin Pharmacol Ther*. 2019;105(5):1095–105.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.