

Genotype Distribution of Dihydrofolatereductase Variants and their Role in Disease Susceptibility to Acute Lymphoblastic Leukemia in Indian Population: An Experimental and Computational Analysis

Sunitha Kodidela^{1*}, Suresh Chandra Pradhan¹, Jayaraman Muthukumar², Biswajit Dubashi³, Teresa Santos-Silva² and Debdatta Basu⁴

¹Department of Pharmacology, Jawaharlal Institute of Postgraduate Medical Education and Research (JIPMER), Puducherry, India

²UCIBIO@REQUIMTE, Department of Chemistry, Faculty of Science and Technology, New University of Lisbon, 2829-516 Caparica, Portugal

³Department of Medical Oncology, Jawaharlal Institute of Postgraduate Medical Education and Research (JIPMER), Puducherry, India

⁴Department of Pathology, Jawaharlal Institute of Postgraduate Medical Education and Research (JIPMER), Puducherry, India

*Corresponding author: Sunitha Kodidela, Department of Pharmacology, Jawaharlal Institute of Postgraduate Medical Education and Research (JIPMER), Puducherry, India, Tel: +917639656035; E-mail: kodidela.sunitha@gmail.com

Received date: February 26, 2016; Accepted date: March 25, 2016; Published date: March 28, 2016

Copyright: © 2016 Kodidela S, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

Objectives: To establish the allele and genotype frequencies of dihydrofolate reductase (DHFR) -317A>G and -680C>A variants in Indian population, to find the association of these variants with the risk of acute lymphoblastic leukemia (ALL) and to analyze the effects of non-synonymous SNPs (nsSNPs) and 3'untranslated region (3'UTR) variants of *DHFR* gene on its structure and function using *in-silico* tools.

Methods: A total of 235 unrelated healthy volunteers (controls) and 127 ALL patients (cases) were recruited for the study. DNA was extracted from peripheral leucocytes. Genotyping of DHFR polymorphisms was done by real-time PCR. We investigated the deleterious effect of nsSNPs and variants in 3'UTR of DHFR gene through computational platforms.

Results: In the present study, the frequency of DHFR -317G and -680 A alleles was found to be 33.3% and 59.8% respectively. The studied DHFR variants (rs408626 and rs442767) did not confer a significant risk for ALL. *In-silico* analysis revealed that three nsSNPs potentially affect the structure, function and activity of the DHFR protein. Four microRNA binding sites were found to be highly affected due to 3'UTR SNPs. Further, docking simulation suggested that the order of binding affinity of methotrexate towards native and all three mutant forms of DHFR is D153V (rs121913223)>native>G18R (rs61736208)>D187Y (rs200904105).

Conclusion: This is the first study to report the normative genotype distribution of DHFR variants in Indian population and also to report that DHFR (-317A>G and -680C>A) variants do not confer a potential risk for development of ALL. Inter-ethnic differences exist in the distribution of DHFR variant genotypes, and this can lead to variability in therapeutic response to DHFR substrates. Protein sequence analysis revealed rs200904105 influences the phosphorylation process (post-translational modification) of *DHFR* and docking simulation suggested methotrexate to have a higher affinity towards rs121913223 mutant form. Therefore, studies are warranted to explore the clinical impact of these variants in the Indian population.

Keywords: Acute lymphoblastic leukemia; Dihydrofolate reductase; Genotype; Methotrexate; Docking; Polymorphism

Introduction

Acute lymphoblastic leukemia (ALL), is a hematological malignancy characterized by an uncontrolled proliferation of lymphoblasts. Although it affects all age groups, it is the most frequent form of childhood cancers [1,2]. Estimated numbers of new cases of ALL in the United States in 2016 is 6,590 and out of which predicted deaths are 1,430 [3]. In India, the lymphoid leukemias are expected to be 18,449 by the year 2020 [4]. Though the causes of ALL are unknown, adverse gene-environment interactions are likely to be involved in the risk of developing ALL [5,6]. Leukemia commonly arises as a result of DNA translocations, different types of mutations in genes regulating blood cell development or homeostasis [7] and folate deficiency [8,9].

Folate pathway has two components such as methylation reactions and nucleotide synthesis. Polymorphisms in genes involved in methylation pathway were not found to influence the risk of ALL in Indian population [10-14]. Thus, exploring the variants in DNA synthesis pathway enzymes such as dihydrofolate reductase (DHFR) and thymidylate synthase (TYMS) may provide insights in to the susceptibility to ALL.

Dihydrofolate reductase (DHFR) is a crucial enzyme in the folate pathway which converts dihydrofolic acid (DHF) to tetrahydrofolic acid (THF). THF is essential for the synthesis of amino acids and nucleic acids, required for cell growth and proliferation. Impairment of folate pathway results, in an uncontrolled proliferation of cells leading to various cancers [5,15]. *DHFR* gene is located on chromosome 5 and has seven transcripts. The major transcript of *DHFR* gene was found to have various mutations such as three stop gained variants, twenty missense variants, seven splice region variants, eleven synonymous

variants, two coding sequence variants, ninety-one 5' prime UTR variants, seventy three 3' prime UTR variants, 1211 intron variants, 232 upstream gene variants and 210 downstream gene variants. Among these, non-synonymous SNPs and SNP in the regulatory region, play a significant role in gene function and are often reported to play a role in the development of diseases in humans [16-19].

Functional polymorphisms in the *DHFR* gene can influence its expression or activity, thereby affecting the risk of folate-dependent diseases [19-22]. Polymorphisms in *DHFR* gene are also relevant in the context of sensitivity and resistance to antifolates [19,23,24] and anti-bacterial drugs, as they also affect the prognosis of such patients treated with these drugs. High activity of DHFR enzyme counteracts the action of anti-folate drugs which results in the development of resistance to chemotherapy. On the other hand, reduction of DHFR enzymatic activity decreases the THF levels that in turn can affect the DNA synthesis and methylation reactions in the body. The risk of cancer varies according to the geographical area, probably due to differences in the genetic makeup, diet and the environmental factors.

The admixture of various ethnic groups in India resulted in the heterogeneous population [25,26]. As the difference in the distribution of alleles could confound genotype-phenotype association studies and are the primary reasons for the conflicting results, it is important to explore the influence of DHFR variants in each population.

Thus, establishing the normal genotype frequency would provide the reference values for the distribution of these SNPs in a population, so that it can be compared with those of people with the risk of cancer. Sorting of deleterious SNPs is essential for narrowing down the number of mutations to be screened in genetic association studies and for a better understanding of the functional and structural aspects of protein before conducting an experimental study. To the best of our knowledge, the distribution of DHFR variants and their role in the risk of ALL has not been explored in Indian population.

Therefore, our objectives were, to establish the allele and genotype frequencies of *DHFR* -317A>G and -680 C>A variants in Indian population, to study the association of *DHFR* variants with the risk of ALL, and the study was also extended to perform computational analysis of nsSNPs and SNPs in the 3'-UTR of *DHFR* gene to identify possible deleterious mutations.

Materials and Methods

Our study was approved by the Institute ethics committee of JIPMER, Puducherry, India. Written informed consent was obtained from all study participants, and in the case of children, consent was obtained from their legally accepted representatives.

Study population

A case-control study consisting of 235 unrelated healthy volunteers and 127 patients diagnosed with ALL of either gender were recruited for the study. All the participants reported three consecutive generations residing in South India and spoke one of the Dravidian languages (Tamil, Telugu, Malayalam and Kannada) as their mother tongue. The mean age (\pm SD) of the cases and controls were 13.9 (\pm 13.37) and 24.5 (\pm 5.6) years respectively. There were 83 males and 44 females in the patient group and 156 males and 79 females in the control group. We could not recruit age-matched healthy children due to ethical issues that resulted in difference in the means of age between cases and controls. The details of sample collection and DNA

extraction were described previously [27]. Taqman assays for DHFR-317A>G (rs408626) and -680C>A (rs442767) genotyping were procured from Applied Biosystems (ABI, Foster City, CA, USA). Genotyping was done according to manufacturer instructions using Real Time-PCR (Applied Biosystems-7300). For quality control, genotyping was done in duplicates in 30 % of the randomly selected samples and found to be in 100% concordance.

Statistical analysis

The observed genotype frequencies were tested for Hardy-Weinberg equilibrium (HWE) using the Chi-square test. Fisher's exact test was used to test the differences in genotypes between patients with ALL and healthy volunteers (controls). Comparison between genotype and allele frequencies of SIs with 1000 genome project data was done using Fisher's exact test. Statistical analysis of data was performed using GraphPad Instat 3 (GraphPad Software Inc., San Diego, CA, USA) and SPSS software (version 16). The level of statistical significance was set at $p < 0.05$.

In-silico analysis

Flow chart of methodology used for In-silico analysis is given in Figure 1. Retrieval of SNP data and Prediction of functional consequences of missense mutations SNP data of *DHFR* gene were obtained from several databases namely the NCBI dbSNP [28], 1000 genomes [29] and the Ensembl genome browser [30]. We retrieved the protein sequence of Human *DHFR* from the UniProt database (P00374 (DYR_HUMAN) [31]. Minor allele frequencies of DHFR variants in other populations were obtained from the 1000 genome browser [29].

Functional effects of nsSNP were predicted using various bioinformatic tools (Figure 2). For simplicity, the nsSNPs that were found to cause disease due to protein function damage are labeled as deleterious while neutral variants are depicted as tolerated. The nsSNPs that were found to be deleterious by at least five programs were considered for further protein sequence analysis and molecular docking simulation. In addition to nsSNPs, variants affecting the 3'UTR function of DHFR were also identified using miRNASNP tool.

Protein sequence analysis

The protein sequence analysis was performed to explore the impact of considered deleterious substitutions at the sequence level for both native and three mutated forms of DHFR by several proteomics tools which are available at ExPasy (<http://www.expasy.org/tools>) suite.

Docking simulation

The three-dimensional structure of human DHFR was retrieved from Protein Data Bank (PDB, www.rcsb.org) and its accession code 4M6J. COOT [32] was used to generate the models of the three mutants based on the structure of native DHFR. The three mutants were subjected to energy minimization step using Modrefiner [33] and validated by Rampage web servers. Subsequently, the anti-cancer drug methotrexate was retrieved from the PubChem database (accession number CID 126941). The docking calculation begins with the preparation of the protein models and ligand molecules by adding polar hydrogens, assigning of charges, merging non-polar hydrogens and the files were saved in PDBQT format using MGLTools. Once the proteins and ligand files were prepared, the blind docking simulation was performed between proteins and drug molecule to explore all

possible interacting sites of the target protein. For this, a grid parameter file (grid box size: 105 Å × 124 Å × 105 Å, grid spacing: 0.375 Å and grid center: 6.390 Å × 7.300 Å × -18.160 Å) and docking parameter file (100 docking trials, 150 population size, 2500000 maximum number of energy evaluations, 27000 maximum number of generations, 0.02 mutation rate and 0.8 cross-over rate) were taken into account using Autogrid and Autodock programs [34]. Once the docking simulation was over, the best protein – drug complexes were selected based on estimated free energy of binding (ΔH) and inhibition constant (K_i) and finally exported into PDB format. The LigPlus [35] and PyMOL (<https://www.pymol.org>) programs were used for analyzing and comparing the docking results of native and three mutant forms of DHFR.

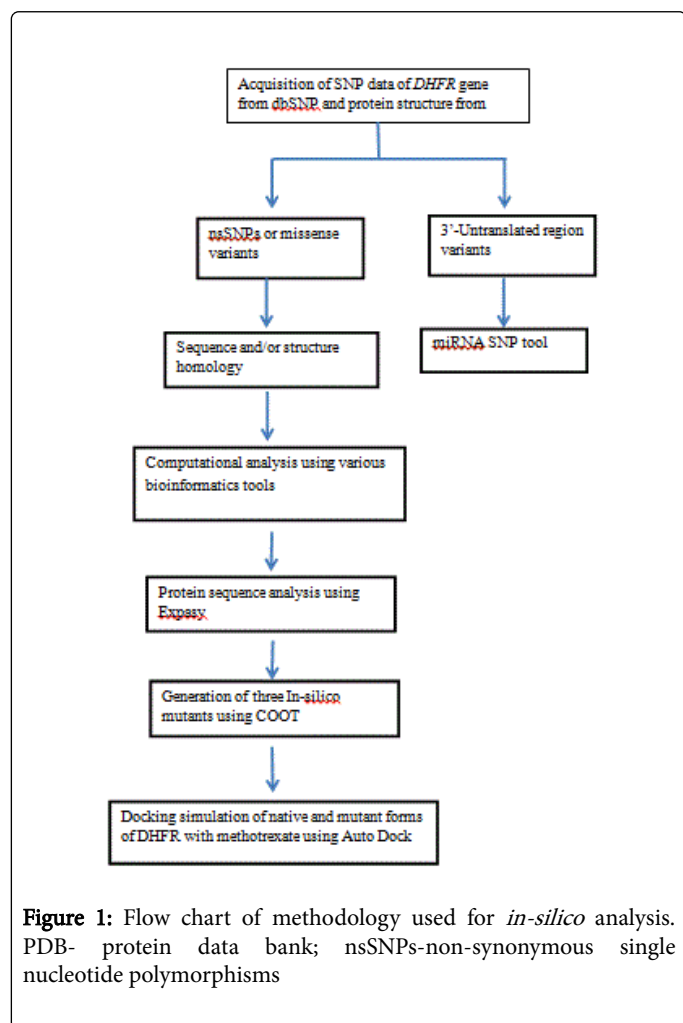


Figure 1: Flow chart of methodology used for *in-silico* analysis. PDB- protein data bank; nsSNPs-non-synonymous single nucleotide polymorphisms

Results

Association of DHFR -317 A>G and -680C>A variants with the risk of ALL

The genotypic and allelic distribution of DHFR -317 A>G variant in cases and controls is shown in Table 1. The genotype frequencies of 317A>G polymorphism was concordant with HWE ($p>0.05$) in both cases and controls whereas -680C>A polymorphism was discordant. There was no significant difference in the genotype and allele

frequencies of DHFR polymorphisms between patients and controls (Table 1).

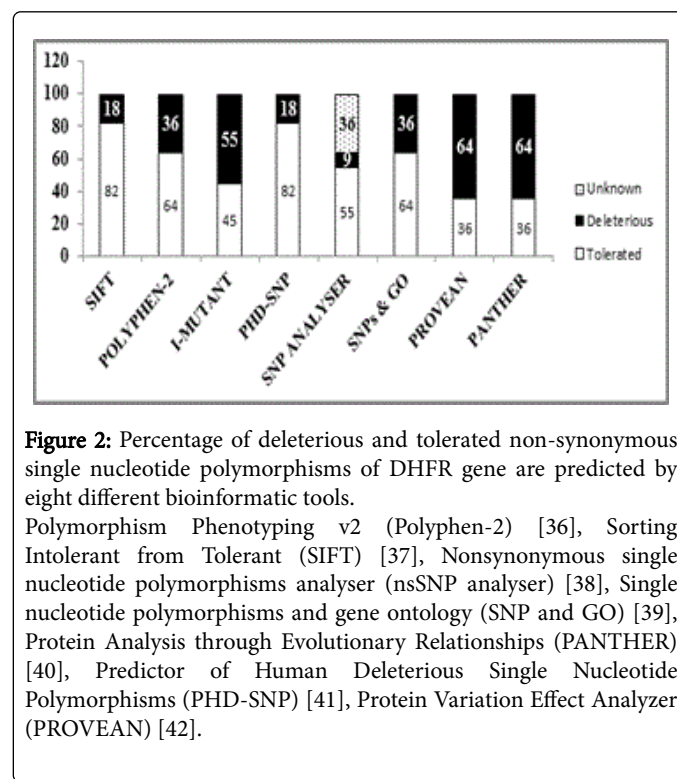


Figure 2: Percentage of deleterious and tolerated non-synonymous single nucleotide polymorphisms of DHFR gene are predicted by eight different bioinformatic tools.

Polymorphism Phenotyping v2 (Polyphen-2) [36], Sorting Intolerant from Tolerant (SIFT) [37], Nonsynonymous single nucleotide polymorphisms analyser (nsSNP analyser) [38], Single nucleotide polymorphisms and gene ontology (SNP and GO) [39], Protein Analysis through Evolutionary Relationships (PANTHER) [40], Predictor of Human Deleterious Single Nucleotide Polymorphisms (PHD-SNP) [41], Protein Variation Effect Analyzer (PROVEAN) [42].

DHFR Allele/genotype	Patients with ALL n(%)	Healthy volunteers n(%)	RR (95%CI)	p-value
rs408626 (-317A>G) genotype	N=127	N=235		
AA	55(43.3)	108(46)		
AG	54(42.5)	94(40)	0.92(0.68 to 1.25)	0.63
GG	18(14.2)	33(14)	0.95 (0.62 to 1.46)	0.86
AA vs AG+GG	55 (43) vs. 72(57)	108(46) vs. 127(54)	1.07(0.80 to 1.42)	0.65
Alleles				
A	164(64.6)	310(66)		
G	90(35.4)	160(34)	0.96(0.78 to 1.18)	0.74
rs442767(-680C>A) genotypes	N=123	N=235		
CC	0(0)	0(0)		
CA	94(76.4)	189(80)		
AA	29(23.6)	46(20)	0.85(0.61 to 1.19)	0.41
Alleles				

C	94(38.2)	189(40)		
A	152(61.7)	281(60)	0.94(0.76 to 1.16)	0.62

Table 1: Genotype and allele frequencies of DHFR variants in patients with acute lymphoblastic leukemia (ALL) and healthy individuals. DHFR: Dihydrofolate Reductase; RR: Relative Risk; CI: Confidence Interval

Comparison of DHFR variants in the study population with that of 1000 genome project data

The allele frequencies of DHFR variants (-317A>G and -680C>A) in the study population were found to be significantly different from other populations except SAS (Table 2).

DHFR	SI	AFR	AMR	EAS	EUR	BEB	GIH	ITU	PJL	STU	Mexican [19]	Chinese [21]
rs408626 (-317A>G)	235											
A	66.7	43.8	49	39.8	55	62.8	73.3	72.1	70.8	70.1	57	35.4
G	33.3	56.2*	51*	60.2*	45*	37.2	26.7	27.9	29.2	29.9	43*	64.6*
OFH	40	51.9	47.8	47.4	50.3	58.1	35.9	42.2	43.8	46.1	42	47.4
rs442767 (-680C>A)	-235											
C	40.2	96.8	60.1	41.5	68.1	72.1	79.1	79.4	77.6	77		
A	59.8	3.2*	39.9*	58.5	31.9*	27.9*	20.9*	20.6*	22.4*	23*		
OFH	80.4	5.7	42.9	48	43.9	46.5	35.9	37.3	32.3	36.3		

Table 2: Comparison of allele and genotype frequencies of DHFR variants (rs408626 and rs442767) in South Indians with the 1000 genome project phase-3 super populations and other study reports. DHFR: Dihydrofolate Reductase; OFH: Observed Frequency of Heterozygosity; SI: South India; AFR: African; AMR: American; EAS: East Asian; EUR: European; BEB: Bengali in Bangladesh; GIH: Gujarati Indian in Houston, ITU: Indian Telugu in the UK; PJL: Punjabi in Lahore, Pakistan; STU: Sri Lankan Tamil in the UK. South Asian sub-populations are highlighted in bold. * indicates P<0.05 compared to SI. (N) indicates number of patients in the respective population

Seven SNPs were observed to be deleterious, and 4 were found to be neutral by PROVEAN server. All eight tools suggested rs200232379 to be neutral. Three nsSNPs rs61736208 (G18R), rs121913223 (D153V), and rs200904105 (D187Y) were found to be deleterious or disease causing by five programs. Prediction of the influence of SNPs in 3'UTR identified that three polymorphisms rs76026972, rs11550953, rs56039509 led to the loss of miRNA binding sites such as hsa-miR-28-3p, hsa-miR-552, hsa-miR-3673 and hsa-miR-3911 respectively in DHFR gene (Figure 2).

In-silico analysis

A total of 20 nsSNPs were found to be non-synonymous in DHFR transcript-1. All SNPs were analyzed using eight different prediction tools (Figure 2).

SIFT analysis predicted that 2nsSNPs were deleterious to DHFR function, whereas 9 SNPs were tolerated. According to Polyphen-2 results, 3SNPs were probably damaging, 1 was possibly damaging, and 7 SNPs were benign. I-mutant predicted 6 SNPs cause disease and 5 SNPs were neutral. Two SNPs were found to cause disease and 9 were neutral by PHD-SNP. SNP analyzer predicted that 1 SNP causes disease, 6 SNPs were neutral, and the effect of 4 SNPs was unknown. According to SNP and GO program results, 4 SNPs were identified to cause disease and 7 were identified as neutral SNPs.

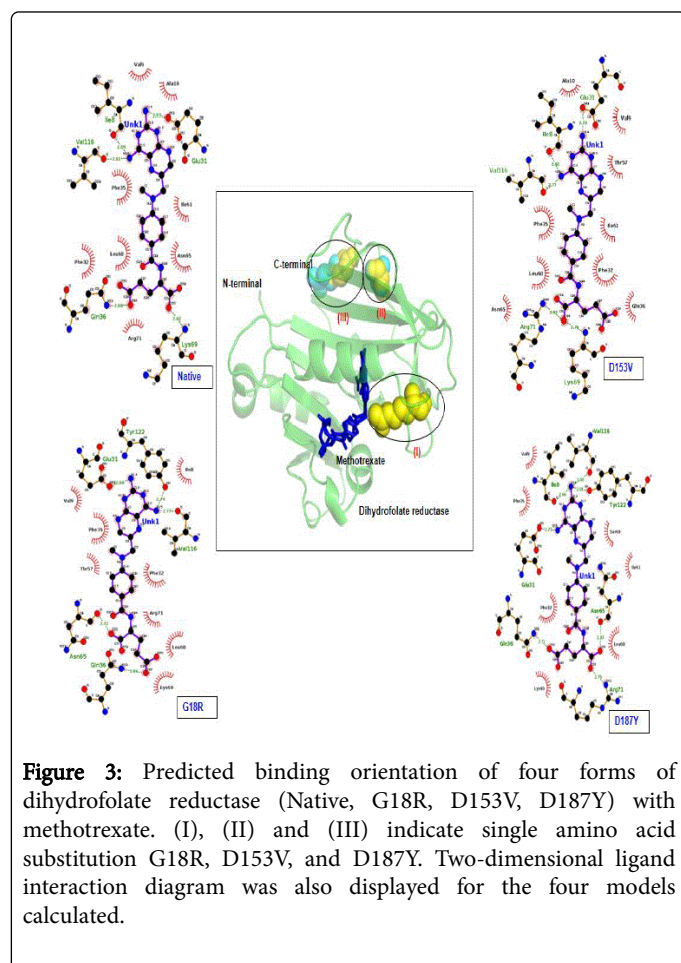
Protein sequence analysis and docking simulation

Protein sequence analysis of native and mutant forms of DHFR revealed, D187Y (rs200904105) substitution in DHFR gene can be phosphorylated whereas others couldn't be phosphorylated. In comparison with native and other mutant forms, methotrexate is showing a higher binding affinity towards D153V (rs121913223) of DHFR with the ΔH of -8.88 kcal/mol and K_i of 307.33 nM, respectively. Detailed docking results of methotrexate with all three mutant forms along with native is given in Table 3 and Figure 3.

Protein DHFR	Estimated free energy of binding (kcal/mol)	Estimated inhibition constant	Hydrogen bonds (Distance are mentioned in brackets with Å)	Hydrophobic interactions
Native	-8.36	749.72 nM	Ile8(2.89), Val116(2.81), Glu31(2.83), Lys69(2.82), Gln36(2.80)	Val9, Ala10, Phe35, Leu68, Phe32, Arg71, Asn65, Ile61
G18R (rs61736208)	-7.95	1.49 μ M	Tyr12 (2.79), Val116(2.77), Glu31(2.80), Asn65(2.72), Gln36(3.06)	Ile8, Phe32, Arg71, Leu68, Lys69, Thr57, Phe35, Val9
D153V (rs121913223)	-8.88	307.33 nM	Glu31(2.72), Ile8(2.86), Val116(2.77), Arg71(3.03), Lys69(2.76)	Ala10, Val9, Thr57, Ile61, Phe32, Gln36, Asn65, Leu68, Phe35

D187Y (rs200904105)	-7.86	1.74 μ M	Ile8(2.66), Val116(3.08), Tyr122(2.81), Glu31(2.73), Gln36(2.71), Arg71(2.75), Asn65(2.82)	Val9,Phe35,Ser60,Ile61,Leu68,Lys69 , Phe32
------------------------	-------	--------------	--	--

Table 3: Molecular docking results of DHFR (Native, G18R, D153V, and D187Y) with methotrexate. DHFR: Dihydrofolate Reductase. Common interacting residues are highlighted in bold



Discussion

To the best of our knowledge, this is the first study from India to provide insight into the allele and genotype frequency distribution of DHFR (-317A>G and -680C>A) polymorphisms. Our study is also the first study to report that the presence of studied DHFR variant alleles does not confer a potential risk for development of ALL. We have also performed a computational analysis and found that rs200904105 influences the phosphorylation process (post-translational modification) of DHFR and methotrexate has a higher binding affinity towards a rs121913223 mutant form of DHFR enzyme.

DHFR-317 G allele frequency in patients with ALL and healthy volunteers were found to be 35.4 % and 33.3% respectively in our study. The percentage of DHFR-680A allele frequency in patients and healthy volunteers was found to be 61.7 and 59.8 respectively. DHFR -680CC genotype was not observed or absent in SIs (Table 2). The DHFR-317G and -680 A alleles were not found to be associated with

disease susceptibility to ALL. Other studies have examined the role of DHFR variants in patients with ALL and reported conflicting results. A study by Yang et al. [21] in Chinese population reported that DHFR317A>G mutant is not associated with the risk of ALL. In contrast to our study and Yang et al. [21], Gomez-Gomez et al. [19] reported -317 homozygous mutant genotypes (GG) to be a risk factor for the development of ALL in the Mexican population. (OR=2.89, 95%CI 1.03 to 4.88). The contradictory findings could be due to differences in population genetics or due to environmental factors.

The -317 'A' allele was found to be the major allele in SI, Mexican, EUR and SAS, but G allele was the major allele in other population (Table 2). The percentage of G and A allele were highest in EAS (69.2%) and GIH (73.3%) populations respectively. We found that G allele of -317A>G in SIs was significantly lower compared to AFR, AMR, EAS and EUR populations. Further, the frequency of G allele was similar to SAS (Table 2). This observation suggests that for studies on -317A>G loci, all SAS populations can be grouped together because they may not present significant differences amongst their genetic loci.

The variant allele of rs442767 SNP namely A allele has also demonstrated significant variation in its distribution. The percentage of A allele was highest in CHS subpopulation (63.8%) of EAS (Supplementary Table 1). Allele A was absent in Luhya in Webuye, Kenya subpopulation of AFR. DHFR-680CC genotype (rs442767) was not found in SIs, and AA genotype is absent in most of the AFR subpopulations (Supplementary Table 1). The absence of an individual genotype in different populations resulted in discordant HWE. Population stratification as a cause of discordance is unlikely in the present study subjects because of the high prevalence of endogamy [25,43]. DHFR-680A allele frequency was significantly greater than AFR, AMR, EUR and SAS sub-populations. The allele frequencies of -680C>A were similar when compared to EAS sub-populations and Peruvian in Lima, Peru subpopulation of AMR (Supplementary Table 1).

These findings suggest inter-ethnic variability in the distribution of DHFR polymorphisms. Thus, susceptibility to the diseases and the treatment outcome may vary between populations and confound genotype-phenotype association studies. Hence, normative frequencies must be established in genetically distinct populations before exploring the influence of genotypic distributions on the phenotype.

In our study, three SNPs (rs61736208, rs121913223, and rs200904105) were found to be deleterious by most of the prediction tools. None of these SNPs are validated by 1000 genome project and Hap-Map population. Hence, there is a need to confirm the following SNPs the rs61736208, rs121913223, and rs200904105 in different populations. A Genome-wide study by Cario et al. [44] reported rs121913223 ASP153Val (A458T) to be pathogenic, causing megaloblastic anemia and cerebral folate deficiency. They also reported no difference in DHFR mRNA levels between wild-type and mutated cells, whereas protein expression is reduced in cells with the DHFR mutation. DHFR deficiency can lead to hematological and neurological

diseases. Therefore, the role of rs121913223 alone and in combination with rs61736208, rs200904105 need to be evaluated in ALL.

Protein sequence analysis of the native and three mutant forms of DHFR concluded that the mutant D187Y can influence the phosphorylation process (post-translational modification) of *DHFR*. Molecular docking studies were performed to understand the role of single amino-acid substitution on protein-drug interactions. The order of binding affinity of methotrexate towards all the three mutant forms along with native DHFR is D153V (rs121913223)>native>G18R (rs61736208)>D187Y (rs200904105). Though the substitution (G18R, D153V, and D187Y) is not directly involved in the drug binding, it may increase the drug binding ability in the case of D153V whereas it is decreased in other instances. Further experimental studies are necessary to validate the *in-silico* hypothesis. Prediction of the influence of SNPs in 3'UTR identified that three polymorphisms rs76026972, rs11550953, rs56039509, and rs56039509 led to the loss of miRNA binding sites such as hsa-miR-28-3p, hsa-miR-552, hsa-miR-3673 and hsa-miR-3911 respectively.

DHFR plays a crucial role in folate metabolism, and its variations may play a dual role. Polymorphisms associated with increased expression of DHFR may protect against cancer, due to the higher levels of 5,10-methylene-THF needed for thymidylate synthesis, whereas decreased expression may affect methylation reactions by changing 5-methyl-THF pool and, consequently, increase cancer risk. Both genomic DNA hypomethylation and gene-specific promoter CpG island hypermethylation are important epigenetic mechanisms of carcinogenesis [45-47]. Therefore, the functional role of studied DHFR (-317A>G and -680C>A) variants needs to be explored. Several enzymes are involved in folate pathway to regulate THF pools and DNA methylation. Influence of one enzyme variant may be abrogated by another variant [48]. Hence, studies that consider the role of gene-gene and folate-gene interactions are warranted.

Clinical relevance of studied DHFR variants

DHFR is an important target of MTX and mutations in the *DHFR* gene were found to influence drug response. To the best of our knowledge, only four studies have been reported regarding the clinical significance of *DHFR* variants worldwide. Interestingly, the results of all these studies were contradictory to each other. A study on patients with ALL of the European ethnicity reported that -317AA genotype carriers had lower event-free survival (EFS), and both the studied *DHFR* variants (-317A>G and -680C>A) were not associated with MTX toxicity [23]. In Mexican and SI patients with ALL, -317 GG genotype was found to be involved in the relapse and lower overall survival [19,27]. DHFR variants were found to have no influence on the response to methotrexate in the Spanish population, but the DHFR-680C>A variant was found to be associated with neutropenia [49]. Discordant results of *DHFR* variants could be majorly due to the difference in the distribution of DHFR variant alleles. As the difference in the distribution of alleles could confound genotype-phenotype association studies, it is important to explore the distribution of *DHFR* variants in each population to establish its role in predicting the therapeutic efficacy of *DHFR* substrates.

Our study provides normative frequency data of *DHFR* variants in South Indian population. Studied DHFR variants were not found to have an influence on susceptibility to ALL. Inter-ethnic differences exist in the distribution of *DHFR* variant genotypes, and this can lead to variability in therapeutic response to *DHFR* substrates. Computational analysis revealed three nSNPs (rs61736208,

rs121913223, and rs200904105) to be deleterious to *DHFR* function and three 3'UTR SNPs affect miRNA binding sites of *DHFR*. Protein sequence analysis revealed rs200904105 influences the phosphorylation process (post-translational modification) of DHFR. Docking simulation suggested that methotrexate has highest and lowest binding affinity towards rs121913223 and rs200904105 mutant forms of *DHFR* respectively. Therefore, studies are warranted to explore the clinical impact of these variants on the Indian population.

Compliance with Ethical Standards

All procedures performed in the present study involving human participants were by the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study.

Conflicts of Interest

Authors declare no conflicts of interest.

Acknowledgments

We thank the Jawaharlal Institute of Postgraduate Medical Education and Research (JIPMER), Puducherry, India, for providing an intramural fund to conduct the study. Muthukumaran J thanks, UCIBIO (UID/Multi/04378/2013), Fundação para a Ciência e a Tecnologia (FCT) (PTDC/QUI-BIQ/117799/2010), (SFRH/BPD/97719/2013) Portugal and Professor Maria João Romão from the protein crystallography group at FCT-UNL for her support to carry out *in silico* part of the research work. The authors also like to thank A. G. Santos and D.F.Jana at LAQV for providing access to the computational cluster QtRex from the Chemistry Department at FCT-UNL.

References

1. Pui CH, Robison LL, Look AT (2008) Acute lymphoblastic leukaemia. *Lancet* 371: 1030-1043.
2. Hunger SP, Lu X, Devidas M, Camitta BM, Gaynon PS, et al. (2012) Improved survival for children and adolescents with acute lymphoblastic leukemia between 1990 and 2005: a report from the children's oncology group. *J Clin Oncol* 30: 1663-1669.
3. Siegel RL, Miller KD, Jemal A (2016) Cancer statistics, 2016. *CA Cancer J Clin* 66: 7-30.
4. Takiar R, Nadayil D, Nandakumar A (2010) Projections of number of cancer cases in India (2010-2020) by cancer groups. *Asian Pac J Cancer Prev* 11: 1045-1049.
5. Lamm N, Maoz K, Bester AC, Im MM, Shewach DS, et al. (2015) Folate levels modulate oncogene-induced replication stress and tumorigenicity. *EMBO Mol Med* 7: 1138-1152.
6. Milne E, Greenop KR, Scott RJ, Haber M, Norris MD, et al. (2015) Folate pathway gene polymorphisms, maternal folic acid use, and risk of childhood acute lymphoblastic leukemia. *Cancer Epidemiol. Biomarkers Prev* 24: 48-56.
7. Küppers R, Klein U, Hansmann ML, Rajewsky K (1999) Cellular origin of human B-cell lymphomas. *N Engl J Med* 341: 1520-1529.
8. Blount BC, Mack MM, Wehr CM, MacGregor JT, Hiatt RA, et al. (1997) Folate deficiency causes uracil misincorporation into human DNA and chromosome breakage: implications for cancer and neuronal damage. *Proc Natl Acad Sci USA* 94: 3290-3295.

9. Duthie SJ (1999) Folic acid deficiency and cancer: mechanisms of DNA instability. *Br Med Bull* 55: 578-592.
10. Bellampalli R, Phani NM, Bhat KG, Prasad K, Bhaskaranand N, et al. (2015) Significance of 5,10-methylenetetrahydrofolate reductase gene variants in acute lymphoblastic leukemia in Indian population: an experimental, computational and meta-analysis. *Leuk. Lymphoma* 56: 1450-1459.
11. Hussain SR, Naqvi H, Raza ST, Ahmed F, Babu SG, et al. (2012) Methylenetetrahydrofolate reductase C677T genetic polymorphisms and risk of leukaemia among the North Indian population. *Cancer Epidemiol* 36: e227-e231.
12. Sazawal S, Chaubey R, Kaur P, Chikkara S, Kumar B, et al. (2014) MTHFR Gene Polymorphisms and the Risk of Acute Lymphoblastic Leukemia in Adults and Children: A Case Control Study in India. *Indian J Hematol Blood Transfus* 30: 219-225.
13. Sadananda AMN, Chandoy S, Ramachandra N, Appaji L, Aruna KBS, et al. (2010) Methylenetetrahydrofolate reductase gene polymorphisms and risk of acute lymphoblastic leukemia in children. *Indian J Cancer* 47: 40-45.
14. Sood S, Das R, Trehan A, Ahluwalia J, Sachdeva MU, et al. (2010) Methylenetetrahydrofolate reductase gene polymorphisms: association with risk for pediatric acute lymphoblastic leukemia in north Indians. *Leuk Lymphoma* 51: 928-932.
15. Zhang D, Wen X, Wu W, Guo Y, Cui W (2015) Elevated homocysteine level and folate deficiency associated with increased overall risk of carcinogenesis: meta-analysis of 83 case-control studies involving 35,758 individuals. *PLoS ONE* 10: e0123423.
16. Ramensky V, Bork P, Sunyaev S (2002) Human non-synonymous SNPs: server and survey. *Nucleic Acids Res* 30: 3894-3900.
17. Yang SD, Cai YL, Jiang P, Li W, Tang JX (2014) Association of a miR-502-binding site single nucleotide polymorphism in the 3'-untranslated region of SET8 and the TP53 codon 72 polymorphism with cervical cancer in the Chinese population. *Asian Pac J Cancer Prev* 15: 6505-6510.
18. Spinella JF, Healy J, Saillour V, Richer C, Cassart P, et al. (2015) Whole-exome sequencing of a rare case of familial childhood acute lymphoblastic leukemia reveals putative predisposing mutations in Fanconi anemia genes. *BMC Cancer* 15: 539.
19. Gómez-Gómez Y, Organista-Nava J, Saavedra-Herrera MV, Rivera-Ramírez AB, Terán-Porcayo MA, et al. (2012) Survival and risk of relapse of acute lymphoblastic leukemia in a Mexican population is affected by dihydrofolate reductase gene polymorphisms. *Exp Ther Med* 3: 665-672.
20. Orjuela MA, Cabrera-Muñoz L, Paul L, Ramirez-Ortiz MA, Liu X, et al. (2012) Risk of retinoblastoma is associated with a maternal polymorphism in dihydrofolate reductase (DHFR) and prenatal folic acid intake. *Cancer* 118: 5912-5919.
21. Yang L, Liu L, Wang J, Qiu L, Mi Y, et al. (2011) Polymorphisms in folate-related genes: impact on risk of adult acute lymphoblastic leukemia rather than pediatric in Han Chinese. *Leuk Lymphoma* 52: 1770-1776.
22. Martinelli M, Girardi A, Cura F, Carinci F, Morselli PG, et al. (2014) Evidence of the involvement of the DHFR gene in nonsyndromic cleft lip with or without cleft palate. *Eur J Med Genet* 57: 1-4.
23. Dulucq S, St-Onge G, Gagné V, Ansari M, Sinnett D, et al. (2008) DNA variants in the dihydrofolate reductase gene and outcome in childhood ALL. *Blood* 111: 3692-3700.
24. Al-Shakfa F, Dulucq S, Brukner I, Milacic I, Ansari M, et al. (2009) DNA variants in region for noncoding interfering transcript of dihydrofolate reductase gene and outcome in childhood acute lymphoblastic leukemia. *Clin. Cancer Res* 15: 6931-6938.
25. Tamang R, Singh L, Thangaraj K (2012) Complex genetic origin of Indian populations and its implications. *J Biosci* 37: 911-919.
26. Aggarwal S, Phadke SR (2015) Medical genetics and genomic medicine in India: current status and opportunities ahead. *Mol Genet Genomic Med* 3: 160-171.
27. Kodidela S, Pradhan SC, Dubashi B, Basu D (2015) Influence of dihydrofolate reductase gene polymorphisms rs408626 (-317A>G) and rs442767 (-680C>A) on the outcome of methotrexate-based maintenance therapy in South Indian patients with acute lymphoblastic leukemia. *Eur J Clin Pharmacol* 71: 1349-1358.
28. dbSNP Short Genetic Variations.
29. Deorowicz S, Grabowski S (2011) Compression of DNA sequence reads in FASTQ format. *Bioinformatics* 27: 860-862.
30. Ensembl genome browser 81: Homo sapiens - Summary - Gene: DHFR (ENSG00000228716).
31. UniProt.
32. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66: 486-501.
33. Xu D, Zhang Y (2011) Improving the physical realism and structural accuracy of protein models by a two-step atomic-level energy minimization. *Biophys J* 101: 2525-2534.
34. Rollins DE, von Bahr C, Glaumann H, Moldéus P, Rane A (1979) Acetaminophen: potentially toxic metabolite formed by human fetal and adult liver microsomes and isolated fetal liver cells. *Science* 205: 1414-1416.
35. Laskowski RA, Swindells MB (2011) LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J Chem Inf Model* 51: 2778-2786.
36. PolyPhen-2: prediction of functional effects of human nsSNPs.
37. SIFT Human Protein.
38. nsSNPAnalyzer: predicting disease-associated nonsynonymous single nucleotide polymorphisms.
39. Capriotti E, Calabrese R, Fariselli P, Martelli PL, Altman RB, et al. (2013) WS-SNPs&GO: a web server for predicting the deleterious effect of human protein variants using functional annotation. *BMC Genomics* 14 Suppl 3: S6.
40. PANTHER - Gene List Analysis.
41. PhD-SNP: Predictor of human Deleterious Single Nucleotide Polymorphisms.
42. Choi Y, Chan AP (2015) Provean Web Server: a Tool to Predict the Functional Effect of Amino Acid Substitutions and Indels *Bioinformatics* 31: 2745-2747.
43. Sebro R, Lange C, Laird NM, Rogus JJ, Risch NJ (2012) Differentiating population stratification from genotyping error using family data. *Ann Hum Genet* 76: 42-52.
44. Cario H, Smith DEC, Blom H, Blau N, Bode H, et al. (2011) Dihydrofolate reductase deficiency due to a homozygous DHFR mutation causes megaloblastic anemia and cerebral folate deficiency leading to severe neurologic disease. *Am. J. Hum. Genet* 88: 226-231.
45. Ehrlich M, Lacey M (2013) DNA hypomethylation and hemimethylation in cancer. *Adv Exp Med Biol* 754: 31-56.
46. Kulis M, Esteller M (2010) DNA methylation and cancer. *Adv Genet* 70: 27-56.
47. Merlo A, Herman JG, Mao L, Lee DJ, Gabrielson E, et al. (1995) CpG island methylation is associated with transcriptional silencing of the tumour suppressor p16/CDKN2/MTS1 in human cancers. *Nat. Med* 1: 686-692.
48. Kamel AM, Moussa HS, Ebid GT, Bu RR, Bhatia KG (2007) Synergistic effect of methyltetrahydrofolate reductase (MTHFR) C677T and A1298C polymorphism as risk modifiers of pediatric acute lymphoblastic leukemia. *J Egypt Natl Canc Inst* 19: 96-105.
49. Salazar J, Altés A, del Rio E, Estella J, Rives S, et al. (2012) Methotrexate consolidation treatment according to pharmacogenetics of MTHFR ameliorates event-free survival in childhood acute lymphoblastic leukaemia. *Pharmacogenomics* 12: 379-385.