DOI: http://dx.doi.org/10.47829/JCMI.2023.62301

Journal of Clinical and Medical Images

Research Article

ISSN: 2640-9615 | Volume 6

In Silico Approach to Predict Indian Pharmacogenomics Variants Based on Analysis of Linkage Disequilibrium (LD) and Expression Quantitative Trait Locus (eQTL)

Kumar P*, Khanam A, Satyam S, Ghime D and Seth S

Predictive Diagnostic Lab, Technoculture Research Pvt. Ltd, Hathiyar, Bodhgaya, Bihar 824231, India

***Corresponding author:**

Permendra Kumar,
Predictive Diagnostic Lab, Technoculture Research Pvt. Ltd, Hathiyar, Bodhgaya, Bihar 824231, India,
orcid.org/0000-0002-7480-8569;
E-mail- permendra@technocultureresearch.org and permendrak@gmail.com

Received: 20 Dec 2022
Accepted: 17 Jan 2023
Published: 27 Jan 2023
J Short Name: JCMI

Citation:

Kumar P, In Silico Approach to Predict Indian Pharmacogenomics Variants Based on Analysis of Linkage Disequilibrium (LD) and Expression Quantitative Trait Locus (eQTL). J Clin Med Img. 2023; V6(23): 1-22

Keywords:

Linkage Disequilibrium; SNP; GTEx; Top LD; eQTL

1. Abstract

The drug response can vary based on ethnicity as it arises from differences in genetic makeup among the individuals and is associated with drug action, metabolism and action. The efficacy of drugs can be determined by genetic variation and may also lead to adverse effects. The drug responsive variants (Single nucleotide polymorphisms, SNPs) are obtained from pharmacogenomics studies based on the Indian population. We have used GTEx, Top LD, and LDlink computational tools to predict the LD and eQTL score. Furthermore, the Indian allelic frequency is predicted using the IndiGenome database. A total of 11 SNPs (rs1979277, rs3758149, rs1799853, rs9923231, rs2244500, rs4244285, rs10509681, rs1057910, rs749292, rs4775936 and rs700518) has been predicted to have significant LD and eQTL score to associate with the different drugs across the Indian population. The presented data helps in identification of drug target sites in a more appropriate way in the Indian population. Moreover, our findings highlight the importance of ethnicity based drug response and ultimately reduce and avoid the adverse effect of drugs and ensure effective drug treatment.

2. Introduction

Pharmacogenomics is an emerging field for the development of personalized medicine. It involves the study of the influence of genes and how it responds to a particular medicine in an individual [1,2]. It has been estimated the cost of $ 528 billion USD for the morbidity and mortality associated with unoptimized drugs in 2016 [3]. The adverse effect of drugs causes one of the leading preventable deaths as per the centers for disease control and prevention [4]. The variability in response of drugs among the individuals is primarily linked with its absorption, distribution, metabolism and elimination [1,2]. The variants in the antigen gene of leukocyte and drug target [5] can regulate the drug efficacy and further results in the adverse effect of the drug and leads to hospitalization along with mortality among children and adults [6-9]. The genotyping knowledge of drug response related loci in patients may reduce the adverse effect of drugs. For instance, the genotyping based dose of warfarin drug in cancer patients minimises the thromboembolism and internal bleeding associated with use of it [10]. Majority of the genetic variants are shared across the population having minor allele frequency > 0.05 [11] and a small fraction make the difference for the metabolic phenotypes in a population [12]. There are also reports of differences in genetic variants within and outside the ethnic group for a distribution and response of a drug [13-15]. The drug rosuvastatin is usually used to prevent the complications associated with cardiovascular condition as well as in treatment of

abnormal levels of lipid in blood. It has been observed the ethnic difference in pharmacokinetics of this drug. The systemic exposure in average to rosuvastatin drug has been found to be 2.3 fold higher in Chinese ethnicity than that of Caucasian however Asian Indian and Malays have the intermediate value [16].

There are genetic variations among the populations which make the drugs and markers used in pharmacogenotyping inappropriate for different populations. One of example is allele HLA-b*58:01 that is linked with allopurinol induced severe cutaneous adverse reaction and rs9263726 can be used as surrogate biomarkers for Japanese population however not in Han Chinese and Australian populations [17-18]. Similarly population based treatment for various diseases has been reported [19]. Therefore, the population specific genetic structure exploration has huge application in research related to medical as well as population genetic research. It also ensures the efficacy of drugs through the development of pharmacogenetic tests [20-22].

SNPs have been discovered as expression quantitative trait loci (eQTLs), highly related with gene expression, according to recent findings gathered from big initiatives like the Genotype-Tissue Expression (GTEx) and TopLD database [23]. eQTLs are genetic variants that affect how much a gene is actually expressed. These eQTLs may exert their effects through cis-regulatory (i.e., local regulation) or trans-regulatory (i.e., distant regulation) mechanisms. In transcriptional regulation, eQTL can affect functional gene transcript expression by DNA binding, mRNA splicing, and noncoding RNA production, which in turn imposes changes of downstream phenotypes. An important function of the eQTL regulatory mechanism is to link the influence of upstream genetic variation to the expression of traits further along the genetic pathway [24]. Disease etiology research has benefited from the use of eQTLs as a proxy for quantification of gene expression. It is interesting to note that eQTLs have been demonstrated to be enriched in polymorphisms related with complex variables in GWAS, such as risk of numerous malignancies and other diseases [25].

It is widely known that there are significant population-specific differences in the allelic frequency of the pharmacogenomic indicators connected to therapeutic response and unfavourable drug reactions. Several population-scale genome sequencing studies have been carried out to identify ethnic differences in the distribution of SNPs across the world's human populations, including South Asian tribes. India, the second most populous country in the world, has a noticeable genetic diversity because of its own culture, social structures, and biological tendencies. Hence, this study can inform future pharmacogenomic validation research on the Indian population, giving physicians a new angle to consider when making decisions, and obtaining improved drug response outcomes for the target population [26].

3. Material and Methods

3.1. Selection of SNPs Associated with the Drug Response

We select variants and reference (rs) numbers of alleles on the basis of previous pharmacogenomics studies in the Indian population using electronics databases like PubMed (https://pubmed.ncbi.nlm.nih.gov/) and dbSNPs for the variants reference number. 72 SNPs that are involved in the PGx studies are selected. And these SNPs are listed in (Table 1) and we performed further studies on these SNPs.

Table 1: Selection of pharmacogenomics variants with special reference to Indian population.

SNP(rsID)	Gene	Chromosome Position	Associated-Drug	Drugs-used For	Reference Paper
rs202242769	CYP21A2	chr6: 32040723	NA	change in hormone levels in the patients with congenital adrenal hyperplasia	[26]
rs854560	PON1	chr7: 95316772	Clopidogrel	Coronary artery disease, susceptibility to Microvascular complications of diabetes 5	[26]
rs1051740	EPHX1	chr1:225831932	carbamazepine	Epilepsy	[26]
rs1933437	FL3	chr13:28050157	Sunitinib	Anticancer Drug	[26]
rs2227291	ATP7A	chrX:78013005	Docetaxel	Breast Cancer, Used as Anti Cancer drug in chemotherapy	[26]
rs3918290	DPYD	chr1:97450058	Fluoropyrimidine	Cancer	[26]
rs1799971	OPRM1	chr6:154039662	Opioids	Pain Relief	[26]
rs4646	CYP19A1	chr15:51210647	Aromatase Inhibitors	Postmenopausal women with breast cancer	[30]
rs10046	CYP19A1	chr15: 51210789	Aromatase Inhibitors	Postmenopausal women with breast cancer	
rs700519	CYP19A1	chr15:51215771	Aromatase Inhibitors	postmenopausal women with breast cancer	
rs700518	CYP19A1	chr15: 51236915	Aromatase Inhibitors	postmenopausal women with breast cancer	
rs727479	CYP19A1	chr15:51242350	Aromatase Inhibitors	postmenopausal women with breast cancer	
rs4775936	CYP19A1	chr15:51243825	Aromatase Inhibitors	postmenopausal women with breast cancer	
rs10459592	CYP19A1	chr15:51243944	Aromatase Inhibitors	postmenopausal women with breast cancer	
rs749292	CYP19A1	chr15:51266534	Aromatase Inhibitors	postmenopausal women with breast cancer	
rs6493497	CYP19A1	chr15:51338638	Aromatase Inhibitors	postmenopausal women with breast cancer	
rs7176005	CYP19A1	chr15:51339082	Aromatase Inhibitors	postmenopausal women with breast cancer	

rs1042713	ADRB2	chr5:148826877	Salbutamol	Asthma	[47]
rs1042714	ADRB2	chr5:148826910	Salbutamol	Asthma	
rs1128503	ABCB1	chr7: 87550285	NACT	Breast Cancer	[31]
rs242941	CRHR1	chr17: 45815154	Inhaled corticosteroid	Asthma	[48]
rs1799853	CYP2C9	chr10: 94942290	Cyclophosphamide	Breast Cancer,ovaries, and lymph system, and nerves (mainly in children).	[32]
rs1057910	CYP2C9	chr10: 94981296			
rs10509681	CYP2C8	chr10: 95038992			
rs4244285	CYP2C19	chr10:94781859			
rs1800566	NQO1	chr16: 69711242	5-FU, epirubicin/ adriamycin/ methotrexate, and cyclophosphamide	Breast Cancer	[37]
rs1801131	MTHFR	chr17: 69928988			
rs1801133	MTHFR	chr1:11796321			
rs1042522	P53	chr17: 7676154	Anthracycline		[49]
rs1065852	CYP2D6	chr22:42130692	Tamoxifen		[39]
rs1800460	TPMT	chr6:18138997	6-mercaptopurine	Acute Lymphoblastic Leukemia	[34]
rs1142345	TPMT	chr6:18130687	6-mercaptopurine		
rs61886492	GCPII	chr11:49164722	6-mercaptopurine		
rs1051266	RFC1	chr21:45537880	6-mercaptopurine		
rs3892097	CYP2D6	chr22: 42128945	Cisplatine	Head And Neck Cancer	[50]
rs1065852	CYP2D6	chr22: 42130692			
rs1695	GSTP1	chr11: 67585218	carboplatin	Lung Cancer	[51]
rs2072671	CDA	chr1: 20589208	Gemcitabine	Cancer	[39]
rs4451422	FPGS	chr9:127814318	Methotrexate(Mtx)	Rheumatoid Arthritis	[40]
rs2244500	TYMS	chr18:661005	Methotrexate(Mtx)		
rs3786362	TYMS	Chr18: 662247	Methotrexate(Mtx)		
rs2072671	TYMS	Chr18: 661647	Methotrexate(Mtx)		
rs1061235	HLA-A	chr6:29945521	Carbamazepine	Epilepsy	[52]
rs2395148	HLA-B	chr6:32353777	Carbamazepine	Epilepsy	[53]
rs2032582	ABCB1	chr7:87531302	Antipsychotics Drugs	schizophrenia	[45]
rs1045642	ABCB1	chr7:87509329	Antipsychotics Drugs		
rs265967	DRD1	chr5:175429787	Antipsychotics Drugs		
rs182137906	GLRA4	chrX: 103724259	Antipsychotics Drugs		
rs10934254	DRD3	chr3:114122787	Antipsychotics Drugs		
rs878567	HTR1A	chr5: 63960164	Antipsychotics Drugs		
rs1176744	HTR3B	chr11:113932306	Antipsychotics Drugs		
rs622342	SLC22A1	chr6: 160151834	Metformin	Type2Dibatese	[46]
rs11212617	ATM	chr11: 08412434	Metformin		[54]
rs10509681	CYP2C8	chr10:95038992	Pioglitazone, Repaglinide		[55]

rs1799931	NAT 2	chr8: 18400860	Isoniazid	Tuberculosis	[56]
rs1799930	NAT2	chr8: 18400593	Isoniazid	Tuberculosis	
rs1799929	NAT2	chr8: 18400484	Isoniazid	Tuberculosis	
rs9923231	VKORC1	chr16: 31096368	Acenocoumarol, Warfarin	Prevention and treatment of harmful blood clots	[33]
rs1057910	CYP2C9	chr10: 94981296	Acenocoumarol, Warfarin		
rs1799853	CYP2C9	chr10:94942290	Acenocoumarol, Warfarin		
rs776746	CYP3A5	chr7:99672916	Clopidogrel	prevent heart attacks and strokes in persons with heart disease	
rs4149056	SLCO1B1	chr12:21178615	Simvastatin	Treat high cholesterol and reduce risk of heart disease	[43]
rs2740574	CYP3A4	chr7:99784473	Atorvastatin		
rs3808607	CYP7A1	chr8: 58500365	Atorvastatin		
rs10757278	CDKN2B	chr9:22124478	NA	coronary artery disease	[57]
rs10757274	CDKN2B	chr9: 22096056	NA		
rs1333048	CDKN2B	chr9:22125348	NA		
rs2383206	CDKN2B	chr9: 22115027	NA		
rs3758149	GGH	chr8: 63039169	Methotrexate	Rheumatoid Arthritis	[58]
rs1979277	SHMT1	chr17:18328782			
rs34489327	TS	NA			

3.2. Linkage Disequilibrium and eQTL Analysis

Numbers of publicly available tools are present which allow rapid exploration of Linkage Disequilibrium (LD) between markers. We used Top-LD for LD and GTEx for eQTL analysis. The TOP-LD is an online tool to explore LD and that is based on deep whole genome sequencing (WGS) data from the TOPMed Program. GTEx tool is used to predict the complex patterns of genetic variation and gene regulation across diverse human tissue types.

3.3. TOP-LD

The TOP-LD is an online tool (http://topld.genetics.unc.edu/) to explore LD and that is based on deep WGS data from TOPMed Program. This program has the majority of the variants in TOP-LDs than 1000 Genomes so the low frequency or rare allele information is also included in TOP-LD [27]. This tool gives the LD proxy variants with the minor allele frequency <1%. We used selected SNPs as input and calculated LD and allelic frequency of these SNPs.

3.4. GTex

The GTEx database provides the information of how genes are expressed and identifies susceptibility to diseases inherited in human genes [28]. GTex portal (https://gtexportal.org/home/) has a number of tools i.e, 'Browser', 'Expression', 'QTL', 'Single Cell', 'eGTex', and 'Biobank'. The GTex 'QTL' tool has four modules and with the help of these modules we explore and compute eQTL and sQTL. We used Locus Browser (Variant-centric) of the QTL module to analysis of eQTL.

3.5. Allele Frequency of Selected SNPs in Indian Population

Publicly available IndiGenome (http://clingen.igib.res.in/indigen/) database contains the information on variants alleles frequency, allele number, allele count, number of homozygous and heterozygous alleles [29]. With the help of IndiGenome, we computed the allelic frequency of variants which are associated with the drug response and compared with the allelic frequency in the South Asians (SAS) population of 1000 Genomes.

4. Results

4.1. Linkage Disequilibrium and Allelic Frequency of Selected SNPs

A total of 72 SNPs were reported in high LD ($r^2 \geq 0.8$) see (Table 2.) We selected only those SNPs which are in perfect LD i.e $r^2=1$. Perfect LD means the two SNPs are not only separated by recombination but also have the same allelic frequency. A total of 33 SNPs out of 72 SNPs were found in perfect LD but some of them appeared in repetition so after filtration, 17 SNPs were considered in perfect LD. Genetic polymorphism of CYP19A1 gene could be involved in the severity of aromatase inhibitor adverse effects [30]. Variants rs10046, rs700518, rs4775936, rs10459592, rs749292, rs6493497, rs7176005 of CYP19A1 gene (on chromosome 15) were found in high LD. The SNPs rs1128503 in the ABCB1 gene can influence the various drug responses. The variant rs1128503 of ABCB1 gene is associated with the chemotherapy response [31]. Furthermore it is also associated with the plasma level of docetaxel [31]. The SNPs rs1799853 and rs1057910 of CYP2C9 gene are associated with the neoadjuvant chemotherapy (cyclophosphamide based NACT) [32] along with glimepiride and glipizide (Type 2 Diabetes). Moreover it also affects the Warfarin dose. The variants rs4244285 of CYP2C19 is associated with the NACT in

cancer patients and influencing the anti-tuberculosis drug [32]. It is found 40% in the patients of cardiovascular diseases in north India [33]. The variant rs1051266 of RFC1 gene is associated with the better tumour response in MTx taking patients [34]. The variant rs9923231 of VKORC1 is one of the factors for the acenocoumarol and Warfarin dose prediction [33].

Variability in Drug response for affecting drug absorption, distribution, metabolism, and excretion (ADME) is well documented at individual and population level [35]. India is a country with a huge variety of culture, social backgrounds, environmental and so is a treasure for genetic diversity. We used IndiGenome database in which 1029 individuals of different ethnicities has been participated [36]. We calculated the allelic frequency of selected SNPs. There are several significant differences seen in the selected SNPs frequency in the Indian population and SAS population (1000 Genome). CYP19A1 variants rs10046 and rs749292 that are associated with the aromatase inhibitors in breast cancer patients [30], are found in high frequency in IndiGenome 0.3132 and 0.3660 compared to SAS population in 1000 Genome 0.287 and 0.313 respectively. The variant rs1128503 in ABCB1 gene is present in low frequency in Indian population (0.3891) as compared to SAS population in 1000 Genome (0.413). The MTHFR variant rs1801133 is associated with the toxicity of acute lymphoblastic leukaemia [37] that is present in higher frequency in the Indian population (IndiGenome= 0.1449) than the SAS population in 1000 Genome (0.119). The CYP2D6 variant rs1065852 is associated with increased risk of recurrence of breast cancer when patients are treated with the tamoxifen [38] is high frequency in the Indian population (0.1929) compared to SAS population in 1000 Genome (0.119). Moreover, the variant rs1142345 in TPMT gene reduced the activity of TPMT [34] is found high in the Indian population i.e 0.226 compared to SAS population in 1000 Genome i.e 0.017. The GCPII variant rs61886492 is associated with the 6-mercaptopurine mediated toxicity [34] is higher in the Indian population (IndiGenome=0.0343) than the SAS population in 1000 Genome (0.031). The variant rs1051266 of RFC1 gene is associated with increased toxicity of 6-Mercaptopurine [34], having high frequency found in the Indian population (IndiGenome = 0.6042) compared to SAS population (1000 Genome = 0.407. The FLT3 variant rs1933437 is associated with the toxicity of anticancer drug sunitinib [26] and is found as high frequency in the Indian population i.e, IndiGenome=0.675 compared to SAS population of 1000 Genome (0.344). The variant rs1799853 of CYP2C9 is associated with the grade 2–4 leucopenia in cyclophosphamide based NACT [32] found low frequency in Indian population rather than SAS population (IndiGenome = 0.0307, SAS (1000 Genome = 0.035). The Variant rs3918290 in DPYD is associated with fluoropyrimidine-based chemotherapy drug [26] is found to decreased prevalence in Indian population than SAS population (IndiGenome = 0.00049, SAS (1000 Genome = 0.008). The CDA variant rs2072671 is associated with cytarabine (Ara-C) induced-cytotoxicity after studying 100 adult patients with de novo acute myeloid leukemia [39] is low frequency i.e. 0.1885 compared to SAS population which is 0.23. The Variant rs2244500 in TYMS is found with 0.557 which is said to be high frequency in the Indian population compared to the 0.423 SAS 1000 Genome population and is associated with the poor response of methotrexate in people with arthritis [40]. The variant rs1045642 of ABCB1 is less prevalent in the Indian population (0.3659) compared to the SAS population (0.425). This variant of ABCB1 gene is associated with several kinds of diseases like adverse events to Mtx associated [41], good response to anti-epileptic drugs [42], associated with the LDL-cholesterol reduction in response to atorvastatin therapy [43]. It is also associated with the warfarin dose [35]. The rs265967 variant in DRD1 is high in indian population frequency (0.5141) compared to SAS population (0.412), DRD3 variant rs10934254 found to have high frequency in indian population i.e 0.5649 compared to SAS population (0.464), genetic polymorphism of HTR1A (rs878567) is high in indian population i.e, 0.7527 compared to 1000 Genome (SAS - 0.429) and HTR3B (rs1176744) is high in India population compared to 1000 Genome SAS population i.e 0.5879 and 0.405 respectively. These variants are associated with the unfavourable response to antipsychotic drugs [45]. The SNP rs622342 (SLC22A1) is found with high frequency in the Indian population i.e 0.4092 compared to other SAS populations (0.252). The SLC22A1 gene with variants rs622342 is associated with metformin response to diabetes [46]. The genetic variant rs11212617 of ATM gene which the metabolic target for metformin [46] is found to be less prevalent in the Indian population compared to the SAS population in 1000 Genomes i.e 0.0328 and 0.372 respectively. The variant rs10509681 in CYP2C8 gene is present at a high frequency (0.0753) compared to the SAS population in 1000 Genomes (0.03).

Table 2: LD and allelic frequency of SNPs in IndiGenome and SAS (1000 Genome)

Sl. No.	SNPs	LD-region	R-square (>=0.8) calculated by TOP-LD	Allele Frequency	
				IndiGenome	SAS(1000 Genome)
1	rs1799971	154014538-154039662	0.83	0.4395	0.4435
2	rs4646	51210647- 51206511	0.939	0.6194	0.6053

3	rs10046	51210789-51201371	1	0.3132	0.287
4	rs700519	51211945- 51226876	1	0.240	0.237
5	rs700518	51228009- 51241333	1	0.2870	0.291
6	rs727479	51224903- 51245609	0.988	0.7417	0.734
7	rs4775936	51241333- 51246526	1	0.3018	0.307
8	rs10459592	51237152- 51246613	1	0.4399	0.456
9	rs749292	51259149- 51271043	1	0.366	0.313
10	rs6493497	51326105- 51354540	1	0.2835	0.272
11	rs7176005	5126105- 51354540	1	2865	0.272
12	rs1042713	NA	NA	0.4688	0.446
13	rs1042714	148826465-148826812	1	0.8091	0.807
14	rs1128503	87528267- 87571770	1	0.3891	0.413
15	rs242941	45816121- 45815234	0.808	0.7769	0.785
16	rs1799853	94865198- 94991513	1	0.0307	0.035
17	rs1057910	94659408- 94997876	1	0.1093	0.109
18	rs10509681	94999426- 95091520	1	0.0328	0.03
19	rs4244285	94529595- 94912562	1	0.3678	0.358
20	rs1800566	69558546- 69789178	0.981	0.3503	0.358
21	rs1801131	11838976	0.825	0.4068	0.417
22	rs1801133	NA	NA	0.1449	0.119
23	rs1800460	NA	NA	0.0039	0.004
24	rs1142345	18102797- 18161001	1	0.0226	0.017
26	rs1051266	45528110- 45540065	1	0.6042	0.407
27	rs3892097	42040318- 42129809	0.808	0.1094	0.11
28	rs1065852	42092159- 42148767	0.987	0.1929	0.165
29	rs1933437	28038749- 28066426	0.972	0.675	0.344
30	rs2227291	77823273- 78069120	1	0.3607	0.352
31	rs3918290	NA	NA	0.0049	0.008
32	rs1695	67586499- 67588531	0.806	0.2790	0.294
33	rs2072671	20587026- 20590713	1	0.1885	0.23
34	rs4451422	127787292-127817814	0.991	0.396	0.397

35	rs2244500	659829-678947	1	0.557	0.429
36	rs3786362	NA	NA	0.1235	0.106
37	rs2847153	661647- 677931	0.991	0.3734	0.387
38	rs1045642	87509195- 87512181	0.967	0.3659	0.425
39	rs1051740	NA	NA	0.3659	0.377
40	rs1061235	87528267- 87535576	0.857	0.1312	NA
41	rs2395148	87509195- 87512181	0.967	0.4092	0.062
42	rs2032582	87528267- 87535576	0.857	0.3696	0.358
43	rs265967	114122778-114123597	0.992	0.5141	0.412
44	rs182137906	NA	NA	0.1136	0.098
45	rs10934254	114120388-114123597	0.992	0.5649	0.464
46	rs878567	63924648- 63988082	1	0.7527	0.429
47	rs1176744	113904553-113918577	0.894	0.5879	0.405
48	rs622342	160149051-160153503	0.979	0.4092	0.252
49	rs11212617	108175229-108418440	1	0.0328	0.373
50	rs10509681	94999426- 95091520	1	0.0753	0.03
51	rs1799931	18400860- 18417979	1	0.0753	0.069
52	rs1799930	18388623- 18424194	0.991	0.3563	0.36
53	rs1799929	18392142- 18426755	0.927	0.3052	0.321
54	rs9923231	31014320- 31119800	1	0.1893	0.145
55	rs1057910	99604750- 99728686	1	0.1093	0.109
56	rs1799853	94889882- 94991513	1	0.0307	0.035
57	rs776746	99604750- 99728686	1	0.0307	0.331
58	rs854560	95296302- 95318697	1	0.1877	0.184
59	rs4149056	21178615- 21227696	0.888	0.0513	0.043
60	rs2740574	99774507- 99806611	0.945	0.9721	0.041
61	rs3808607	58497880- 58502816	0.992	0.5351	0.448
62	rs10757278	22096056- 22125504	1	0.4995	0.497
63	rs10757274	22092925- 22125504	0.992	0.5019	0.483
64	rs1333048	22092925- 22125504	0.943	0.5161	0.484
65	rs2383206	22103184- 22125504	0.992	0.5204	0.48

66	rs202242769	NA	NA	0.2273	0.035
67	rs3758149	63035084- 63102649	1	0.2860	0.285
68	rs1979277	18259902- 18357439	1	0.1563	0.148
69	rs34489327	NA	NA	NA	NA

4.2. Expression Quantitative Trait Loci (eQTL) of selected SNPs

Expression quantitative trait loci are the locus that explains a fraction of the genetic variance of a gene expression phenotype. It is very common to find variability in gene expression in different tissues due to the population including ethnic diversity, geographical diversity. Genotype tissue expression is a publicly available database to study tissue specific gene expression and regulation. It also provides open access to data including gene expression, QTL, histology and information. eQTL have identified important functions for non-coding SNP across the genome-some of which have been identified during response phenotypes. All linked SNPs ($r^2 > 0.8$) are considered as part of a single locus. Performing eQTL analysis of variants or SNPs which are in perfect LD. Positive normal enrichment score (NES values) of GTEx indicated that these SNPs are highly expressed in a particular tissue and it was shown as a red circle in the GTEx database. Small P-value of SNPs in GTEx indicated that variants have been identified as eQTL for a tissue. We explore the eQTL of all selected PGx and see how these SNPs are expressed in different tissues (see: Table 4). A total of 21 SNPs out of 72 SNPs not found to be expressed in the GTEx database. On the basis of the NES score (only positive NES score) and p-value (≤ 9), we filtered SNPs and listed them in (Table 3). Moreover, in this table we only listed those SNPs that are associated with particular drug responses expressed in that specific tissue [47-58].

Table 3: This table depicts the list of SNPs found significant on the basis of GTEx NES score and the tissue site of its expression.

SNPs	cis-eQTL	Tissue
rs1799971	NES=0.627 p-value=6.86	Brain-cerebellum
rs4646	p-value=4.70 NES=0.869	Cells-cultured fibroblasts
rs700518	p-value=8.31 NES= 0.101	Cells-cultured fibroblasts
rs4775936	p-value= 7.13 NES=0.0947	Cells-cultured fibroblasts
rs749292	p-value=3.45 NES=0.111	Adipose-subcutaneous
rs1057910	p-value=5.89 NES=0.505	Adipose-visceral
	p-value=5.04 NES=0.502	Breast
rs10509681	p-value=6.60 NES=0.349	Thyroid
rs4244285	p-value=5.0 NES=0.428	Skin-sun exposed
	p-value=11.6 NES=0.321	Stomach
	p-value=5 NES=0.428	Liver
rs1800566	p-value=3.86 NES= 0.67	Colon-transverse
rs1801131	p-value=5.6e-8 NES= 0.23	Breast - Mammary Tissue
rs1801133	P-value=0.000017 NES=-0.15	Adipose - Subcutaneous
rs1042522	p-value=8.21 NES=0.315	Thyroid
	p-value=6.09 NES=0.407	Pancreas
	p-value=4.14 NES=0.101	Cells-cultured fibroblasts
	p-value=5.47 NES=0.277	Colon-transverse
	p-value=4.89 NES=0.45	Brain-cerebellar hemisphere
	p-value=4.53 NES=0.159	Breast
	p-value=5.57 NES=0.165	Adipose-visceral
	p-value=6.20 NES=0.293	Esophagus-muscularis
rs1065852	p-value=8.5e-26 NES=0.65	Adipose - Subcutaneous
rs61886492	p-value=0.00055 NES= 0.17	Nerve - Tibial
	p-value=0.0000044 NES= 0.33	Lung
	p-value=0.0000013 NES= 0.23	Testis
	p-value=0.000064 NES=0.74	Brain - Hypothalamus
	p-value=0.000037 NES=0.76	Brain - Cortex
rs1695	p-value=-0.129 NES= 8.80	Testis
	p-value= -0.195 NES= 4.91	Pituitary

rs2072671	p-value=8.81 NES=0.449	Adrenal gland
	p-value=7.82 NES=0.257	Artery-aorta
	p-value=6.04 NES=0.298	Brain-cortex
	p-value=6.49 NES=0.337	Brain-caudate
	p-value=6.54 NES=0.363	Brain-putamen
	p-value=5.01 NES=0.236	Colon-sigmoid
	p-value=5.94 NES=0.223	Esophagus-Gastrophageal junction
	p-value=8.21 NES=0.276	Esophagus-muscularis
	p-value=8.25 NES=0.291	Heart-left ventricle
	p-value=6.19 NES=0.171	Lung
	p-value=7.35 NES=0.416	Pancreas
	p-value=4.28 NES=0.255	Pituitary
	p-value=6.46 NES=0.321	Prostate
	p-value=6.52 NES=0.204	Spleen
	p-value=8.03 NES=0.366	Testis
rs2244500	p-value=15.0 ES= 0.451	Esophagus-muscularis
	p-value=10.4 NES=0.429	Esophagus-Gastrophageal junction
rs2847153	p-value=8.27 NES=0.48	Esophagus-Gastrophageal junction
	p-value=12.3 NES=0.496	Esophagus-muscularis
rs2395148	p-value=5.17 NES=-0.419	Adipose-Subcutaneous
rs10509681	p-value=4.58 NES=0.307	Thyroid
	p-value=6.60 NES=0.349	Testis
rs9923231	p-value=7.79 NES=0.143	Heart-atrial appendages
	p-value=12.5 NES=0.187	Heart-left ventricle
rs1057910	p-value=12.0 NES=0.712	Adipose-subcutaneous
	p-value=5.89 NES=0.505	Adipose-visceral
	p-value=4.09 NES=0.565	Esophagus-Gastroesophageal junction
	p-value=5.84 NES=0.502	Breast
rs1799853	p-value=4.89 NES=0.281	Thyroid
rs3808607	p-value=-0.515 NES=10.5	Spleen
	p-value=-0.224 NES=4.71	Thyroid
rs3758149	p-value=4.03 NES=0.182	Heart
rs1979277	p-value=1.1E-29 NES=0.21	Lungs
	p-value=1.1E-21 NES=0.4	Heart
	p-value=0.0003 NES=0.097	Skin
	p-value=0.000019 NES=0.3	Brain -Caudate
	p-value=1.1E-14 NES=0.36	Nerve-Tibial

Table 4: eQTL analysis of all selected SNPs through GTEx server.

SNPs	**cis-eQTL**	**Tissue**	**LD_Gtex**
rs1799971	NES=0.627 p-value=6.86	Brain-cerebellum	1
rs4646	p-value=4.70 NES=0.869	Cells-cultured fibroblasts	1
	p-value=4.70 NES=0.0782	Whole blood	
rs10046	p-value=4.46 NES=0.109	Whole Blood	1
rs700519	p-value=4.42 NES= -0.256	Nerve-tibial	1
rs700518	p-value=8.31 NES= 0.101	Cells-cultured fibroblasts	1
	p-value=10.1 NES= 0.129	Whole blood	
	p-value=5.62 NES= 0.151	Skin-sun exposed	

rs727479	p-value=9.10 NES= 0.205	Skin-sun exposed	1
rs4775936	p-value= 7.13 NES=0.0947	Cells-cultured fibroblasts	1
	p-value= 8.35 NES= -0.118	Whole blood	
	p-value=5.0 NES= 0.143	Skin-sun exposed	
rs10459592	p-value=6.38 NES= 0.162	Skin-sun exposed	1
rs749292	p-value=3.45 NES=0.111	Adipose-subcutaneous	1
	p-value=4.79 NES=0.137	Skin-sun exposed	
	p-value=7.83 NES=0.112	Whole blood	
rs6493497	p-value=14.1 NES=0.421	Adipose-subcutaneous	1
	p-value=28.9 NES=0.699	Adipose-visceral	
	p-value=21.1 NES=0.406	Breast	
	p-value=4.30 NES=0.3	Esophagus-mucosa	
	p-value=60.8 NES=0.88	Muscle-skeletal	
rs7176005	p-value=14.3 NES=0.412	Adipose-subcutaneous	1
	p-value=31.9 NES=0.7	Adipose-visceral	
	p-value= 19.9 NES=0.377	Breast	
	p-value= 4.29 NES=0.287	Esophagus-mucosa	
	p-value=58.6 NES=0.856	Muscle-skeletal	
	p-value=3.56 NES=0.181	Skin-sun exposed	
rs1042713	NA	NA	NA
rs1042714	NA	NA	NA
rs1128503	p-value=29.9 NES= -0.704	Heart-atrial appendage	1
	p-value=3.94 NES=0.2	Brain-cerebellum	
rs242941	p-value=4.34 NES=0.329	Brain-cerebellum	1
rs1799853	p-value=4.89	Thyroid	1

	NES=0.281		
rs1057910	p-value=12.0 NES=0.712	Adipose-subcutaneous	1
	p-value=5.89 NES=0.505	Adipose-visceral	
	p-value=4.09 NES=0.565	Esophagus-Gastroesophageal junction	
	p-value=5.04 NES=0.502	Breast	
rs10509681	p-value=4.508 NES=0.307	Testis	1
	p-value=6.60 NES=0.349	Thyroid	
rs4244285	p-value=20.0 NES=0.525	Esophagus-mucosa	1
	p-value=5.0 NES=0.428	Skin-sun exposed	
	p-value=11.6 NES=0.321	Stomach	
	p-value=5 NES=0.428	Liver	
rs1800566	p-value=3.86 NES= 0.67	Colon-transverse	1
rs1801131	p-value=5.6e-8 NES= 0.23	Breast - Mammary Tissue	NA
rs1801133	P-value=0.000017 NES=-0.15	Adipose - Subcutaneous	NA
rs1042522	p-value=8.21 NES=0.315	Thyroid	1
	p-value=6.09 NES=0.407	Pancreas	
	p-value=4.14 NES=0.101	Cells-cultured fibroblasts	
	p-value=5.47 NES=0.277	Colon-transverse	
	p-value=4.89 NES=0.45	Brain-cerebellar hemisphere	
	p-value=4.53 NES=0.159	Breast	
	p-value=5.57 NES=0.165	Adipose-visceral	
	p-value=6.20 NES=0.293	Esophagus-muscularis	
rs1065852	p-value=8.5e-26 NES=0.65	Adipose - Subcutaneous	NA
rs1800460	p-value=5.08 NES= -0.396	Testis	1
	p-value=9.07 NES= -0.365	Muscle-skeletal	

	p-value=3.87 NES= -0.334	Skin-not sun exposed	
	p-value=4.16 NES= -0.495	Esophagus-muscularis	
	p-value=5.19 NES= -0.749	Colon-sigmoid	
rs1142345	p-value=5.17 NES= -0.418	Artery-aorta	1
	p-value=6.72 NES= -0.47	Artery-tibial	
	p-value=6.41 NES= -0.433	Esophagus-mucosa	
	p-value=4.71 NES= -0.381	Esophagus-muscularis	
	p-value=4.71 NES= -0.682	Skin-not sun exposed	
	p-value=15.5 NES= -0.711	Skin-sun exposed	
	p-value=6.33 NES= -0.359	Thyroid	
rs61886492	p-value=0.00055 NES= 0.17	Nerve - Tibial	
	p-value=0.0000044 NES= 0.33	Lung	
	p-value=0.0000013 NES= 0.23	Testis	
	p-value=0.000064 NES=0.74	Brain - Hypothalamus	
	p-value=0.000037 NES=0.76	Brain - Cortex	
rs1051266	NA	NA	NA
rs3892097	NA	NA	NA
rs1933437	p-value=15.09 NES= -0.489	Pancreas	1
	p-value=6.45 NES= -0.191	Adipose-subcutaneous	
	p-value=4.68 NES= -0.241	Artery-coronary	
	p-value=6.44 NES= -0.37	Brain-frontal cortex	
	p-value=5.72 NES= -0.4	Brain-Anterior singulate cortex	
	p-value=16.2 NES= -0.526	Brain- cortex	
	p-value=5.38 NES= -0.406	Brain-caudate	
	p-value=6.50 NES= -0.358	Brain-hypothalamus	

	p-value=5.46 NES= -0.352	Brain-substantia nigara	
	p-value=10.9 NES= -0.538	Brain-cerebellar hemisphere	
	p-value=5.37 NES= -0.173	Nerve-tibial	
	p-value=7.94 NES= -0.199	Skin-not sun exposed	
	p-value=16.8 NES= -0.278	Skin-sun exposed	
	p-value=17.1 NES= -0.665	Brain-cerebellum	
rs2227291	NA	NA	NA
rs3918290	NA	NA	NA
rs1695	p-value=-0.129 NES= 8.80	Testis	1
	p-value= -0.195 NES= 4.91	Pituitary	
	p-value= 8.87 NES=-0.133	Muscle Skeletal	
	p-value=10.8 NES= -0.131	Lung	
	p-value=8.30 NES= -0.177	Heart	
	p-value=10.9 NES=-0.154	Adipose-Subcutaneous	
	p-value= 9.32 NES= -0.1999	Adipose-Visceral	
	p-value= 7.57 NES= -0.193	Breast	
	p-value=5.42 NES=-.0219	Brain	
rs2072671	p-value=19.0 NES=0.464	Adipose-subcutaneous	1
	p-value=14.9 NES=0.341	Adipose-visceral	
	p-value=8.81 NES=0.449	Adrenal gland	
	p-value=7.82 NES=0.257	Artery-aorta	
	p-value=15.6 NES=0.365	Artery-tibial	
	p-value=12.1 NES=0.368	Breast	
	p-value=6.04 NES=0.298	Brain-cortex	
	p-value=6.49	Brain-caudate	

	NES=0.337		
	p-value=6.54 NES=0.363	Brain-putamen	
	p-value=5.01 NES=0.236	Colon-sigmoid	
	p-value=11.6 NES=0.24	Colon-transverse	
	p-value=5.94 NES=0.223	Esophagus-Gastrophageal junction	
	p-value=8.21 NES=0.276	Esophagus-muscularis	
	p-value=23.4 NES=0.347	Cells-cultured fibroblasts	
	p-value=23.1 NES=0.562	Heart-atrial appendage	
	p-value=8.25 NES=0.291	Heart-left ventricle	
	p-value=17.2 NES=0.403	Muscle-skeletal	
	p-value=12.0 NES=0.329	Nerve-tibial	
	p-value=6.19 NES=0.171	Lung	
	p-value=7.35 NES=0.416	Pancreas	
	p-value=4.28 NES=0.255	Pituitary	
	p-value=6.46 NES=0.321	Prostate	
	p-value=5.09 NES= -0.352	Skin-sun exposed	
	p-value=14.9 NES= -0.215	Skin- not sun exposed	
	p-value=10.4 NES=0.37	Small intestine-terminal illeum	
	p-value=6.52 NES=0.204	Spleen	
	p-value=27.4 NES=0.165	Whole blood	
	p-value=8.03 NES=0.366	Testis	
	p-value=30.6 NES=0.473	Thyroid	
rs4451422	p-value=5.76	Stomach	1
	p-value=8.27 NES= -0.138	Skin(exposed to sun)	
	p-value=8.42 NES= -0.151	Skin(not exposed to sun)	

	p-value=6.59 NES=-0.261	Pancreas	
	p-value=16.7 NES=-0.224	Nerve Taibil	
	p-value=7.46 NES= -0.195	Cell cultured Fibroblasts	
	p-value=5.87 NES= -0.16	Colon Transverse	
rs2244500	p-value=15.0 NES= 0.451	Esophagus-muscularis	1
	p-value=5.74 NES=-0.198	Esophagus-mucosa	
	p-value=10.4 NES=0.429	Esophagus-Gastrophageal junction	
rs3786362	NA	NA	NA
rs2847153	p-value=8.27 NES=0.48	Esophagus-Gastrophageal junction	1
	p-value=12.3 NES=0.496	Esophagus-muscularis	
rs2395148	p-value=5.17 NES=-0.419	Adipose-Subcutaneous	1
rs1051740	NA	NA	1
rs1061235	NA	NA	1
rs1045642	p-value=3.65 NES=0.155	Artery-aorta	1
	p-value=6.85 NES=0.195	Artery-tibial	
	p-value=3.65 NES=0.119	Nerve-tibial	
rs2032582	p-value=3.53 NES=0.145	Artery-tibial	1
rs265967	NA	NA	NA
rs182137906	NA	NA	NA
rs10934254	p-value=5.27 NES=0.0931	Lung	1
	p-value=4.15 NES=0.176	Pituitary	
rs878567	p-value=4.64 NES= -0.136	Adipose- visceral	1
	p-value=6.08 NES= -0.173	Breast	
	p-value=4.69 NES= -0.276	Brain-putamen	
	p-value=5.39 NES= -0.219	Esophagus-gastrophageal junction	
	p-value=3.91 NES= -0.136	Heart-atrial appendage	
	p-value=4.11	Prostate	

	NES= -0.159		
	p-value=5.97 NES= -0.242	Testis	
	p-value=6.46 NES= -0.15	Thyroid	
	p-value= NES= -0.183	Heart-left ventricle	
	p-value=6.60 NES= -0.17	Nerve-tibial	
	p-value=5.17 NES= -0.149	Esophagus-Muscularis	
	p-value=7.01 NES= -0.322	Brain-cerebellar hemisphere	
rs1176744	p-value=5.16 NES= -0.104	Muscle-skeletal	1
rs622342	NA	NA	1
rs11212617	p-value=5.35 NES= -0.134	Artery-tibial	1
	p-value=5.58 NES= -0.144	Cells-cultured fibroblasts	
	p-value=4.98 NES= -0.0978	Muscle-skeletal	
	p-value=3.76 NES= -0.112	Skin-not sun exposed	
	p-value=4.61 NES= -0.114	Skin-sun exposed	
	p-value=4.47 NES= -0.107	Adipose-subcutaneous	
rs10509681	p-value=4.58 NES=0.307	Thyroid	1
	p-value=6.60 NES=0.349	Testis	
rs1799931	NA	NA	NA
rs1799930	p-value=7.27 NES=0.412	Testis	1
rs1799929	NA	NA	NA
rs9923231	p-value=16.7 NES=0.584	Adrenal gland	1
	p-value=6.35 NES=0.262	Brain-cortex	
	p-value=20.2 NES=0.231	Cells-cultured fibroblsts	
	p-value=7.79 NES=0.143	Heart-atrial appendages	
	p-value=12.5 NES=0.187	Heart-left ventricle	
	p-value=5.06 NES=0.1	Lung	
	p-value=13.2	Muscle-skeletal	

	NES=0.199		
	p-value=4.71 NES=0.101	Nerve-tibial	
	p-value=6.48 NES=0.224	Pituitary	
	p-value=4.44 NES=0.172	Spleen	
	p-value=11.7 NES=0.303	Testis	
	p-value=7.64 NES=0.109	Whole blood	
rs1057910	p-value=12.0 NES=0.712	Adipose-subcutaneous	1
	p-value=5.89 NES=0.505	Adipose-visceral	
	p-value=4.09 NES=0.565	Esophagus-Gastroesophageal junction	
	p-value=5.84 NES=0.502	Breast	
rs1799853	p-value=4.89 NES=0.281	Thyroid	1
rs776746	NA	NA	NA
rs854560	p-value=4.89 NES= -0.623	Adrenal gland	1
	p-value= NES= -0.369	Brain-Hypothalamus	
	p-value= NES= -0.325	Colon-transverse	
	p-value= NES= -0.314	Liver	
	p-value= NES= -0.424	Ovary	
	p-value= NES= -0.447	Prostate	
	p-value= NES= -0.518	Testis	
rs4149056	NA	NA	NA
rs2740574	NA	NA	NA
rs3808607	p-value=4.42 NES=-0.369	Prostate	1
	p-value=-0.515 NES=10.5	Spleen	
	p-value=-0.224 NES=4.71	Thyroid	
rs10757278	NA	NA	NA
rs10757274	NA	NA	NA
rs1333048	NA	NA	NA
rs2383206	NA	NA	NA

rs202242769	NA	NA	NA
rs3758149	p-value=18.9 NES=0.3772	Lungs	1
	p-value=4.03 NES=0.182	Heart	
	p-value=35.2 NES=0.331	Skin(exposed to sun)	
	p-value=22.3 NES=0.274	Skin not exposed	
rs1979277	p-value=1.1E-29 NES=0.21	Lungs	
	p-value=1.1E-21 NES=0.4	Heart	
	p-value=0.0003 NES=0.097	Skin	
	p-value=0.000019 NES=0.3	Brain -Caudate	
	p-value=1.1E-14 NES=0.36	Nerve-Tibial	
rs34489327	NA	NA	NA

5. Discussion

Ethnic diversity plays a major role in the pharmacogenomics study. Several studies have been reported that the variant associated with the drug response in different types of diseases is not same in every population or not affected as with the same variants in the population. The clinical pharmacogenetics implementation consortium (CPIC) Guideline provides the information about the variants which are associated with the drug response including better response, less response and adverse drug response and suggest for genetic tasting of patients before giving the drugs like Warfarin, Abacavir, Phenytoin etc. But not all risk variants are frequently present in the different populations. Also, variants which are risk factors in one population are not considered as risk factors in other populations. So, in this study we focused on the pharmacogenomics aspect analyzed on the basis of the Indian population. We considered all the research work related to pharmacogenomics in the Indian population from available literature along with online databases.

Linkage Disequilibrium is the tendency of alleles at a locus to be inherited together. Also LD can provide insight into the history of populations including their matting pattern, geographical subpopulation structure, natural selection, gene conversion mutation, and changes in allele frequency overtime. Here we used TOPLD database for analysing and exploring the selected SNPs whether in LD or not. Almost all SNPs are present in LD. We selected all SNPs that are in perfect LD and with the proxy variants. Proxy variants or SNP is a SNP that is in LD with the actual causal variants. Two SNPs are proxy for each other by considering allele frequency. So, proxies SNPs may affect the drug response in individuals in absence of each other. We listed the proxy SNPs of these selected SNPs (Table 5). In case of some diseases or drug response or drug adverse reaction, variants which are previously reported the risk or good association or bad association or risk of adverse reaction may be not present in individuals or population but the response is reported. So, in this case looking for proxy SNPs gives the causal variants. Proxy SNPs for aromatase inhibitor drug target, multi-drug resistant gene 1 for the antiepileptic drug, CYP29*2 and CYP2C9*3 associated with the NACT would be used as causal variants in these diseases and drug response.

Furthermore, comparison of allelic frequency of these SNP in indian population and SAS population in 1000 Genome database also mentioned in Table 2. A total of 24 variants out of 72 variants were found in high prevalence in the Indian population as compared to SAS population. Moreover a total of 18 SNPs out of 72 SNPs found to be less frequent in the Indian population. The other remaining 30 SNPs have very few marginal differences in India and 1000 genome (SAS). A total of 11 SNPs (rs1979277, rs3758149, rs1799853, rs9923231, rs2244500, rs4244285, rs10509681, rs1057910, rs749292, rs4775936 and rs700518) has been predicted to have significant LD and eQTL score to associate with the different drugs across the Indian population. SHMT1 and GGH gene is the drug target for Mtx drug in Rheumatoid Arthritis patients and variants rs1979277(SHMT1) and rs3758149 (GGH) is found to be in perfect LD and associated with the MTx-adverse events found to be expressed in lungs, Heart, Skin, Brain-caudate and Nerve-Tibial tissues. The rs1799853 and rs1057910 variants of CYP2C9 is associated with cardiovascular diseases and also with the NACT treatment in breast cancer has been found to be

expressed in thyroid, adipose-subcutaneous adipose-visceral, esophagus-gastroesophageal junction and breast tissue. The variants rs9923231 of VKORC1 gene is associated with the acenocoumarol, warfarin dose determination is found to be perfect LD and also expressed in heart-atrial appendage and Heart-left ventricular tissue. The variant rs2244500 of TYMs gene is associated with the MTx response in rheumatoid arthritis patients expressed in esophagus-muscularis and esophagus-gastroesophageal junction and it is also found to be in perfect LD. The variant rs4244285 of CYP2C19 is associated with the cyclophosphamide based NACT in Breast Cancer,ovaries, and lymph system, and nerves (mainly in children) is found to be positively expressed in stomach liver and skin (sun exposed) tissues. The rs10509681 variant of CYP2C8 gene is associated with NACT treatment of breast cancer also with type2 diabetes is expressed in thyroid and testis tissues. The variants rs749292, rs4775936, and rs700518 of CYP19A1 which is associated the aromatase inhibitors in breast cancer has been expressed in adipose-subcutaneous, Cell cultured fibroblasts.

Table 5: Pair of SNP proxy ($r^2 = 1$) calculated by TOPLD.

SNPs	LD-proxy variants	Chromosome No.
rs700518	rs2414096	15
rs700518	rs7176330	15
rs700518	rs56097510	15
rs4775936	rs10851498	15
rs4775936	rs17523284	15
rs749292	rs12050767	15
rs749292	rs8029537	15
rs749292	rs28637352	15
rs6493497	rs1961177	15
rs6493497	rs1870049	15
rs6493497	rs12101686	15
rs6493497	rs11070843	15
rs7176005	rs1961177	15
rs7176005	rs1870049	15
rs7176005	rs12101686	15
rs7176005	rs11070843	15
rs1128503	rs10276036	7
rs1128503	rs4728702	7
rs1128503	rs10225464	7
rs1799853	rs56090603	10
rs1799853	rs4917636	10
rs1799853	rs9332100	10
rs1799853	rs9332101	10
rs1057910	rs74963911	10
rs1057910	rs111598382	10
rs1057910	rs111309918	10
rs1057910	rs148648466	10
rs10509681	rs72818673	10
rs10509681	rs112152869	10
rs10509681	rs113526640	10
rs10509681	rs72818678	10
rs4244285	rs12571533	10
rs4244285	rs58335703	10
rs4244285	rs4532967	10
rs4244285	rs4986894	10
rs1051266	rs4819130	21
rs1051266	rs9976727	21
rs2244500	rs2015944	18

rs2244500	rs1001761	18
rs2244500	rs2612095	18
rs9923231	rs12934418	16
rs9923231	rs10871454	16
rs9923231	rs12448321	16
rs9923231	rs9939417	16
rs3758149	rs11545076	8
rs3758149	rs11988534	8
rs3758149	rs12544045	8
rs3758149	rs35535527	8
rs3758149	rs12547126	8
rs3758149	rs34554414	8
rs1979277	rs4924847	17
rs1979277	rs921866	17
rs1979277	rs7222684	17
rs1979277	rs4925171	17
rs1979277	rs12952556	17
rs1979277	rs12952667	17
rs1979277	rs6502649	17

6. Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

7. Notes

The authors declare no competing financial interest.

8. Conflict of Interest

The authors declare no conflicts of interest for publishing this work.

9. Authors Contribution

Afreen Khanam and Permendra kumar have equal contribution in preparing this manuscript.

10. Acknowledgement

All authors are grateful to Technoculture Research Pvt. Ltd. Bodhgaya, Bihar (India), for providing the facility to carry out the presented work. The authors thank Mr. Vinay Tiwary (CMD, Technoculture Research Pvt. Ltd.) for his constant support and motivation during the research. The authors acknowledge Mr. Sarthak Tiwari (Advisor, Technoculture Research Pvt. Ltd.) for his technical help and manuscript preparation.

References

1. Brown AMK, Renaud Y, Ross C, Hansen M, Mongrain I, Valois D, et al. Development of a broad-based ADME panel for use in pharmacogenomic studies. Pharmacogenomics. 2014; 15: 1185–1195.
2. Grossman I. ADME pharmacogenetics: current practices and future outlook. Expert Opin Drug Metab Toxicol. 2009; 5: 449–462.
3. Watanabe JH, McInnis T, Hirsch JD. Cost of Prescription Drug–Related Morbidity and Mortality. Ann Pharmacother. 2018; 52: 829-837.
4. Sarkar U, Lopez A, Maselli JH, Gonzales R. Adverse Drug Events in U.S. Adult Ambulatory Medical Care. Health Serv Res. 2011; 46: 1517-1533.
5. Mukerjee G, Huston A, Kabakchiev B, Piquette-Miller M, Van Schaik R, Dorfman R, et al. User considerations in assessing pharmacogenomic tests and their clinical support tools. npj Genomic Med. 2018; 3: 26.
6. Impicciatore P, Choonara I, Clarkson A, Provasi D, Pandolfini C, Bonati M. Incidence of adverse drug reactions in paediatric in/out-patients: a systematic review and meta-analysis of prospective studies. Br J Clin Pharmacol. 2001; 52: 77–83.
7. Lazarou J, Pomeranz BH, Corey PN. Incidence of Adverse Drug Reactions in Hospitalized Patients. JAMA. 1998; 279: 1200.
8. Mouton JP, Mehta U, Parrish AG, Wilson DPK, Stewart A, Njuguna CW et al. Mortality from adverse drug reactions in adult medical inpatients at four hospitals in South Africa: a cross-sectional survey. Br J Clin Pharmacol. 2015; 80: 818–826.
9. Sanghera DK, Bejar C, Sapkota B, Wander GS, Ralhan S. Frequencies of poor metabolizer alleles of 12 pharmacogenomic actionable genes in Punjabi Sikhs of Indian Origin. Sci Rep. 2018; 8: 15742.
10. Goulding R, Dawes D, Price M, Wilkie S, Dawes M. Genotype-guided drug prescribing: a systematic review and meta-analysis of randomized control trials. Br J Clin Pharmacol. 2015; 80: 868–877.
11. Nebert DW, Menon AG. Pharmacogenomics, ethnicity, and susceptibility genes. Pharmacogenomics J. 2001; 1: 19-22.
12. Suarez-Kurtz G, Parra EJ. Population Diversity in Pharmacogenetics: A Latin American Perspective. 2018; 133-154.
13. Ortega VE, Meyers DA. Pharmacogenetics: Implications of race and ethnicity on defining genetic profiles for personalized medicine. J Allergy Clin Immunol. 2014; 133: 16-26.

14. Claudio-Campos K, Duconge J, Cadilla CL, Ruaño G. Pharmacogenetics of drug-metabolizing enzymes in US Hispanics. Drug Metab Pers Ther. 2015; 30.
15. McGraw J, Waller D. Cytochrome P450 variations in different ethnic populations. Expert Opin Drug Metab Toxicol. 2012; 8: 371-382.
16. Lee E, Ryan S, Birmingham B, Zalikowski J, March R, Ambrose H et al. Rosuvastatin pharmacokinetics and pharmacogenetics in white and Asian subjects residing in the same environment. Clin Pharmacol Ther. 2005; 78: 330-341.
17. Dou Y, Peng P, Cai C, Ye A, Kong L, Zhang R, et al. HLA-B*58:01 and rs9263726 have a linkage, but not absolute linkage disequilibrium in Han Chinese population. Drug Metab Pharmacokinet. 2018; 33: 228–231.
18. Maekawa K, Nishikawa J, Kaniwa N, Sugiyama E, Koizumi T, Kurose K, et al. Development of a Rapid and Inexpensive Assay for Detecting a Surrogate Genetic Polymorphism of HLA-B*58:01: A Partially Predictive but Useful Biomarker for Allopurinol-related Stevens-Johnson Syndrome/toxic Epidermal Necrolysis in Japanese. Drug Metab Pharmacokinet. 2012; 27: 447-450.
19. Hariprakash JM, Vellarikkal SK, Keechilat P, Verma A, Jayarajan R, Dixit V, et al. Pharmacogenetic landscape of DPYD variants in south Asian populations by integration of genome-scale data. Pharmacogenomics. 2018; 19: 227–241.
20. Bachtiar M, Lee CGL. Genetics of Population Differences in Drug Response. Curr Genet Med Rep. 2013; 1: 162-170.
21. Alhusain L, Hafez AM. Nonparametric approaches for population structure analysis. Hum Genomics. 2018; 12: 25.
22. Paschou P, Ziv E, Burchard EG, Choudhry S, Rodriguez-Cintron W, Mahoney MW et al. PCA-Correlated SNPs for Structure Identification in Worldwide Human Populations. PLoS Genet. 2007; 3: e160.
23. Westra H-J, Franke L. From genome to function by studying eQTLs. Biochim Biophys Acta - Mol Basis Dis. 2014; 1842: 1896–1902.
24. Yang HC, Lin CW, Chen CW, Chen JJ. Applying genome-wide gene-based expression quantitative trait locus mapping to study population ancestry and pharmacogenetics. BMC Genomics. 2014; 15: 319.
25. Macauda A, Piredda C, Clay-Gilmour AI, Sainz J, Buda G, Markiewicz M, et al. Expression quantitative trait loci of genes predicting outcome are associated with survival of multiple myeloma patients. Int J Cancer. 2021; 149: 327–336.
26. Sahana S, Bhoyar RC, Sivadas A, Jain A, Imran M, Rophina M, et al. Pharmacogenomic landscape of Indian population using whole genomes. Clin Transl Sci. 2022; 15: 866–877.
27. Huang L, Rosen JD, Sun Q, Chen J, Wheeler MM, Zhou Y, et al. TOP-LD: A tool to explore linkage disequilibrium with TOPMed whole-genome sequence data. Am J Hum Genet. 2022; 109: 1175–1181.
28. Stanfill AG, Cao X. Enhancing Research Through the Use of the Genotype-Tissue Expression (GTEx) Database. Biol Res Nurs. 2021; 23: 533–540.
29. Jain A, Bhoyar RC, Pandhare K, Mishra A, Sharma D, Imran M, et al. Indi Genomes: a comprehensive resource of genetic variants from over 1000 Indian genomes. Nucleic Acids Res. 2020.
30. Umamaheswaran G, Dkhar SA, Kalaivani S, Anjana R, Revathy M, Jaharamma M, et al. Haplotype structures and functional polymorphic variants of the drug target enzyme aromatase (CYP19A1) in South Indian population. Med Oncol. 2013; 30(3): 665.
31. Priyadarshini R, Raj GM, Sundaram R, Kayal S, Ramesh A, Shewade DG, et al. Association of plasma docetaxel levels with ABCB1 gene polymorphisms and tumour response in locally advanced breast cancer patients of South India on neo-adjuvant chemotherapy. Breast Cancer. 2021; 28: 414–423.
32. Tulsyan S, Agarwal G, Lal P, Mittal B. Significant role of CYP450 genetic variants in cyclophosphamide based breast cancer treatment outcomes: a multi-analytical strategy. Clin Chim Acta. 2014; 434: 21–28.
33. Shalia KK, Doshi SM, Parikh S, Pawar PP, Divekar SS, Varma SP, et al. Prevalence of VKORC1 and CYP2C9 gene polymorphisms in Indian population and its effect on warfarin response. J Assoc Physicians India. 2012; 60: 34–8.
34. Dorababu P, Naushad SM, Linga VG, Gundeti S, Nagesh N, Kutala VK, et al. Genetic variants of thiopurine and folate metabolic pathways determine 6-MP-mediated hematological toxicity in childhood ALL. Pharmacogenomics. 2012; 13: 1001–1008.
35. Sweeney GD. Variability in the human drug response. Thromb Res. 1983; 29: 3-15.
36. Jain A, Bhoyar RC, Pandhare K, Mishra A, Sharma D, Imran M, et al. IndiGenomes: a comprehensive resource of genetic variants from over 1000 Indian genomes. Nucleic Acids Res. 2020. 49(D1): D1225-D1232.
37. Chaturvedi P, Tulsyan S, Agarwal G, Lal P, Agrawal S, Mittal RD, et al. Relationship of MTHFR and NQO1 Pharmacogenetics and Chemotherapy Clinical Outcomes in Breast Cancer Patients. Biochem Genet. 2015; 53: 211–222.
38. Damodaran SE, Pradhan SC, Umamaheswaran G, Kadambari D, Reddy KS, Adithan C, et al. Genetic polymorphisms of CYP2D6 increase the risk for recurrence of breast cancer in patients receiving tamoxifen as an adjuvant therapy. Cancer Chemother Pharmacol. 2012; 70: 75–81.
39. Iyer SN, Ankala A, Singhal RS, Hegde MR. Determination of common genetic variants in cytidine deaminase (CDA) gene in Indian ethnic population. Gene. 2013; 524: 35–39.
40. Senapati S, Singh S, Das M, Kumar A, Gupta R, Kumar U, et al. Genome-wide analysis of methotrexate pharmacogenomics in rheumatoid arthritis shows multiple novel risk variants and leads for TYMS regulation. Pharmacogenet Genomics. 2014; 24: 211–219.
41. Muralidharan N, Antony PT, Jain VK, Mariaselvam CM, Negi VS. Multidrug resistance 1 (MDR1) 3435C>T gene polymorphism influences the clinical phenotype and methotrexate-induced adverse events in South Indian Tamil rheumatoid arthritis. Eur J Clin Pharmacol. 2015; 71: 959–965.

42. Shaheen U, Prasad DKV, Sharma V, Suryaprabha T, Ahuja YR, Jyothy A, et al. Significance of MDR1 gene polymorphism C3435T in predicting drug response in epilepsy. Epilepsy Res. 2014; 108: 251–256.

43. Kadam P, Ashavaid TF, Ponde CK, Rajani RM. Genetic determinants of lipid-lowering response to atorvastatin therapy in an Indian population. J Clin Pharm Ther. 2016; 41: 329–333.

44. Nahar R, Saxena R, Deb R, Parakh R, Shad S, Sethi PK, et al. CYP2C9, VKORC1, CYP4F2, ABCB1 and F5 variants: Influence on quality of long-term anticoagulation. Pharmacol Reports. 2014; 66: 243–249.

45. Gupta M, Moily NS, Kaur H, Jajodia A, Jain S, Kukreti R. Identifying a predictive model for response to atypical antipsychotic monotherapy treatment in south Indian schizophrenia patients. Genomics. 2013; 102: 131–135.

46. Umamaheswaran G, Praveen RG, Damodaran SE, Das AK, Adithan C. Influence of SLC22A1 rs622342 genetic polymorphism on metformin response in South Indian type 2 diabetes mellitus patients. Clin Exp Med. 2015; 15: 511–517.

47. Sahi PK, Shastri S, Lodha R, Gupta N, Pandey RM, Kabra SK, et al. ADRB2 polymorphism and salbutamol responsiveness in Northern Indian children with mild to moderate exacerbation of asthma. Indian Pediatr. 2016; 53: 211–215.

48. Awasthi S, Gupta S, Agarwal S, Sharma N. CRHR1 Gene SNPs and Response to Systemic Corticosteroids in Indian Asthmatic Children During Acute Exacerbation. Indian J Pediatr. 2015; 82: 781–786.

49. Gopinath S, Shewade DG, Dubashi B, Kadambari D, Ramesh A, Kesavan R, et al. Effect of p53 Codon 72 Polymorphism on Clinical Outcome in Locally Advanced Breast Cancer Patients Receiving Anthracycline Based Neoadjuvant Chemotherapy. J Young Pharm. 2017; 9: 114–119.

50. Shukla P, Pant M, Gupta D, Parmar D. CYP 2D6 polymorphism: A predictor of susceptibility and response to chemoradiotherapy in head and neck cancer. J Cancer Res Ther. 2012; 8: 40.

51. Tulsyan S, Chaturvedi P, Agarwal G, Lal P, Agrawal S, Mittal RD, et al. Pharmacogenetic Influence of GST Polymorphisms on Anthracycline-Based Chemotherapy Responses and Toxicity in Breast Cancer Patients: A Multi-Analytical Approach. Mol Diagn Ther. 2013; 17: 371–379.

52. Duraivel M, Arunkumar R, Ruckmani A, Arunkumar G, Sreekanth M, Bhanu K, et al. Study to assess the prevalence of HLA-A 3101 allele among Indian epileptic patients and its influence on safety and efficacy of antiepileptic therapy. Int J Basic Clin Pharmacol. 2015; 1.

53. Aggarwal R, Sharma M, Modi M, Garg V kumar, Salaria M. HLA-B∗1502 is associated with carbamazepine induced Stevens–Johnson syndrome in North Indian population. Hum Immunol. 2014; 75: 1120–1122.

54. Vilvanathan S, Gurusamy U, Mukta V, Das A, Chandrasekaran A. Allele and genotype frequency of a genetic variant in ataxia telangiectasia mutated gene affecting glycemic response to metformin in South Indian population. Indian J Endocrinol Metab. 2014; 18: 850.

55. Moulik NR, Kumar A, Agrawal S, Mahdi AA, Kumar A. Effect of folate status and methylenetetrahydrofolate reductase genotypes on the complications and outcome of high dose methotrexate chemotherapy in north Indian children with acute lymphoblastic leukemia. Indian J Med Paediatr Oncol. 2016; 37: 85–89.

56. Mishra S, Daschakraborty S, Shukla P, Kapoor P, Aggarwal R. N-acetyltransferase and cytochrome P450 2E1 gene polymorphisms and susceptibility to antituberculosis drug hepatotoxicity in an Indian population. Natl Med J India. 2013; 26: 260–5.

57. Manjula G, Pranavchand R, Kumuda I, Reddy BS, Reddy BM. The SNP rs7865618 of 9p21.3 locus emerges as the most promising marker of coronary artery disease in the southern Indian population. Sci Rep. 2020; 10: 21511.

58. Ghodke-Puranik Y, Puranik AS, Shintre P, Joshi K, Patwardhan B, Lamba J, et al. Folate metabolic pathway single nucleotide polymorphisms: a predictive pharmacogenetic marker of methotrexate response in Indian (Asian) patients with rheumatoid arthritis. Pharmacogenomics. 2015; 16: 2019–2034.