The Prevalence of Pharmacogenomics Variants and Their Clinical Relevance Among the Pakistani Population

Evolutionary Bioinformatics
Volume 18: 1–14

DOI: 10.1177/11769343221095834
SAGE

Abdul Rafay Khan[1], Sayed Hajan Shah[1], Sadia Ajaz[2,*], Sadaf Firasat[1], Aiysha Abid[1] and Ali Raza[1]

[1]Center for Human Genetics and Molecular Medicine, Sindh Institute of Urology and Transplantation, Karachi, Pakistan. [2]Dr. Panjwani Center for Molecular Medicine and Drug Research, International Center for Chemical and Biological Sciences, University of Karachi, Karachi, Pakistan.

ABSTRACT

BACKGROUND: Pharmacogenomics (PGx), forming the basis of precision medicine, has revolutionized traditional medical practice. Currently, drug responses such as drug efficacy, drug dosage, and drug adverse reactions can be anticipated based on the genetic makeup of the patients. The pharmacogenomic data of Pakistani populations are limited. This study investigates the frequencies of pharmacogenetic variants and their clinical relevance among ethnic groups in Pakistan.

METHODS: The Pharmacogenomics Knowledge Base (PharmGKB) database was used to extract pharmacogenetic variants that are involved in medical conditions with high (1A + 1B) to moderate (2A + 2B) clinical evidence. Subsequently, the allele frequencies of these variants were searched among multiethnic groups of Pakistan (Balochi, Brahui, Burusho, Hazara, Kalash, Pashtun, Punjabi, and Sindhi) using the 1000 Genomes Project (1KGP) and ALlele FREquency Database (ALFRED). Furthermore, the published Pharmacogenomics literature on the Pakistani population was reviewed in PubMed and Google Scholar.

RESULTS: Our search retrieved (n = 29) pharmacogenetic genes and their (n = 44) variants with high to moderate evidence of clinical association. These pharmacogenetic variants correspond to drug-metabolizing enzymes (n = 22), drug-metabolizing transporters (n = 8), and PGx gene regulators, etc. (n = 14). We found 5 pharmacogenetic variants present at >50% among 8 ethnic groups of Pakistan. These pharmacogenetic variants include *CYP2B6* (rs2279345, C; 70%-86%), *CYP3A5* (rs776746, C; 64%-88%), *FLT3* (rs1933437, T; 54%-74%), *CETP* (rs1532624, A; 50%-70%), and *DPP6* (rs6977820, C; 61%-86%) genes that are involved in drug response for acquired immune deficiency syndrome, transplantation, cancer, heart disease, and mental health therapy, respectively.

CONCLUSIONS: This study highlights the frequency of important clinical pharmacogenetic variants (1A, 1B, 2A, and 2B) among multi-ethnic Pakistani populations. The high prevalence (>50%) of single nucleotide pharmacogenetic variants may contribute to the drug response/diseases outcome. These PGx data could be used as pharmacogenetic markers in the selection of appropriate therapeutic regimens for specific ethnic groups of Pakistan.

KEYWORDS: Pharmacogenomics, allele frequency, drug-metabolizing enzymes, drug-transporters, pharmacogenetic variants, Pakistan

RECEIVED: October 16, 2021. **ACCEPTED:** April 4, 2022.

TYPE: Bioinformatics Resources for Understanding the Epitranscriptome and General Omics-Original Research

FUNDING: The author(s) received no financial support for the research, authorship, and/or publication of this article.

DECLARATION OF CONFLICTING INTERESTS: The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

CORRESPONDING AUTHOR: Ali Raza, Center for Human Genetics and Molecular Medicine, Sindh Institute of Urology and Transplantation, Karachi, 74200, Pakistan.
Email: raza.a005@gmail.com

Introduction

The traditional medical practice has been revolutionized by precision medicine. Pharmacogenomics (PGx) plays a central role in the selection of targeted therapies that underpin precision medicine. The premise is that a patient's response to available treatment options for a particular disease depends upon genomic variations.[1] Pharmacogenetics and PGx are generally used interchangeably in the literature. However, pharmacogenetics focuses on variations in a single gene, while PGx includes variants at the genomic level. The DNA variations in genes encoding drug-metabolizing enzymes and drug transporters explain the differences in pharmacokinetics and pharmacodynamics in response to therapeutic regimens between individuals. The frequencies of such pharmacogenetic variants differ across the continents of the world.[2] Many efforts have been focused on incorporating PGx information for the therapeutic selection of several drugs. It might be available in low-income countries, in the near future. Pharmacogenomic consortiums play a critical role in the generation of such information. These include the Consortium for the Implementation of Clinical Pharmacogenetics Implementation Consortium (CPIC), the Canadian Pharmacogenomics Network for Drug Safety (CPNDS), and other pharmacogenomics initiatives.[3-5] To

*Sadia Ajaz is now affiliated to Dr. Panjwani Department of Human Genetics and Molecular Biology, University of Health Sciences, Lahore, Pakistan.

date, 133 clinical guidelines are available for the selection of the right dose of drug.[6] However, in terms of developing countries, genetic data are missing or scarce. To benefit from the pharmacogenomic revolution in these countries, it is essential to explore the available PGx data of their populations and apply it in clinical phenotypes appropriately.

The present study focuses on the collection of PGx data with high to moderate clinical evidence(s) for therapeutic purposes in the Pakistani population. Pakistan is the sixth most populous country in the world with an estimated population of over 207.74 million.[7] The population is diverse and divided into at least 18 major ethnic groups.[8]

In the present study, the Pharmacogenomics Knowledge Base (PharmGKB) database was explored for pharmacogenetic variants, which are correlated with clinical phenotypes in Pakistan. The frequencies of the variants were investigated in 8 ethnic groups (Balochi, Brahui, Burusho, Hazara, Kalash, Pashtun/Pathan, Punjabi, and Sindhi) included in the ALlele FREquency Database (ALFRED) and the 1000 Genomes Project (1KGP).[9,10] The purpose is to identify and quantify population-specific pharmacogenetic variants for their translation into clinical practices.

Methods

The general methodology of our study is shown in Figure 1.

Selection of clinical pharmacogenetic variants

A retrospective data analysis was performed for the pharmacogenetic variants associated with clinical outcomes using PharmGKB. The database lists pharmacogenetic variants with clinically annotated datasets. On the basis of clinical relevance, the contribution of variants is classified as high, moderate, or low. The present study was restricted to variants with a high to moderate level of clinical association.

Among the high association variants, level 1A represents the variant-drug combinations, which are currently used in clinical settings as recommended by CPIC and other registered professional societies and endorsed by the PGx guidelines. High association level 1B variants show significant variant-drug combinations with evidence available from more than a single cohort. The moderate association level 2A and 2B variants are the functionally important variants and the variants for which evidence for association is available from certain studies, respectively.[6,11]

Pharmacogenetic variants (1A + 1B and 2A + 2B) are known for their involvement in the drug-response of certain clinical phenotypes such as heart disease, cancer, etc. The prevalence of these clinically important PGx variants was investigated among ethnic groups of Pakistan.

Data extraction of allele frequencies on the basis of ethnicity for PGx-variants

The ALFRED and 1KGP data were accessed to determine the allele frequencies of selected pharmacogenetic variants (1A + 1B and 2A + 2B). ALFRED database provides the allele frequencies of pharmacogenetic variants in different ethnic groups throughout the world. In the case of Pakistan, data for 7 ethnic groups are available at ALFRED database: Balochi (n = 50), Brahui (n = 50), Burusho (n = 50), Hazara (n = 48), Kalash (n = 50), Pashtun (n = 46), and Sindhi (n = 50). The 1KGP has information on allele and genotype frequencies for a single ethnic group from Pakistan, Punjabis in Lahore (PJL; n = 158).

The allele frequencies of selected pharmacogenetic variants were manually extracted from ALFRED database. For 1KGP database, the Ensembl Rest API Package for R was used for data extraction.[12]

The pharmacogenetic variants data of Pakistani population/ethnic groups were also searched in the PubMed database, last accessed on August 19, 2019. The keywords used for searching include genotype, allele, frequency*, Pakistan*. We only selected studies of the human species without language limitations. We only included published articles that contain the allelic and genotypic frequency of pharmacogenomics variants in the Pakistani population.

Principal component analysis

Based on the genotypic frequency of the pharmacogenetic variants, the position of the Pakistani population was investigated among the world's populations. Therefore, the principal component analysis (PCA) was constructed using the 1KGP browser. Pakistani population was compared with the Asia (South and East Asians), Africa and Europe (Caucasians), and American (Hispanics) populations (list of populations (n = 26); Supplemental Table a). We used PGx genotypic data of the population for the PCA plot, available in 1KGP. Pakistani population; Punjabi, Lahore (PJL) data is only available in 1KGP were used for analysis.

The general steps of the principal component analysis are provided (Supplemental Figure a). Briefly, genotype data from all available world populations were downloaded in the Variant Call Format (VCF) and the Pedigree (PED) file formats from the 1KGP Server.[13] Genotype data were sorted according to the clinical association of pharmacogenetic variants (evidence of high, medium, and low clinical association). We only selected pharmacogenetic variants with high to moderate clinical association using Genome Analysis Toolkits (GATK) Select Variants commands.

The VCF files were converted into Binary Counterpart File (BCF) format using BCF Tools version 1.7. The BCF files were converted to PLINK format (bim, bed, and fam) using Plink 1.9. Finally, the PCA was generated by eigenvalues and eigenvector matrix files using Plink 1.9 (—pca), and the PCA plot was visualized using R-studio.

Results

Clinical pharmacogenetic variants

In the search for pharmacogenetic variants with a known high to moderate effect on an individual's response to drugs, a total

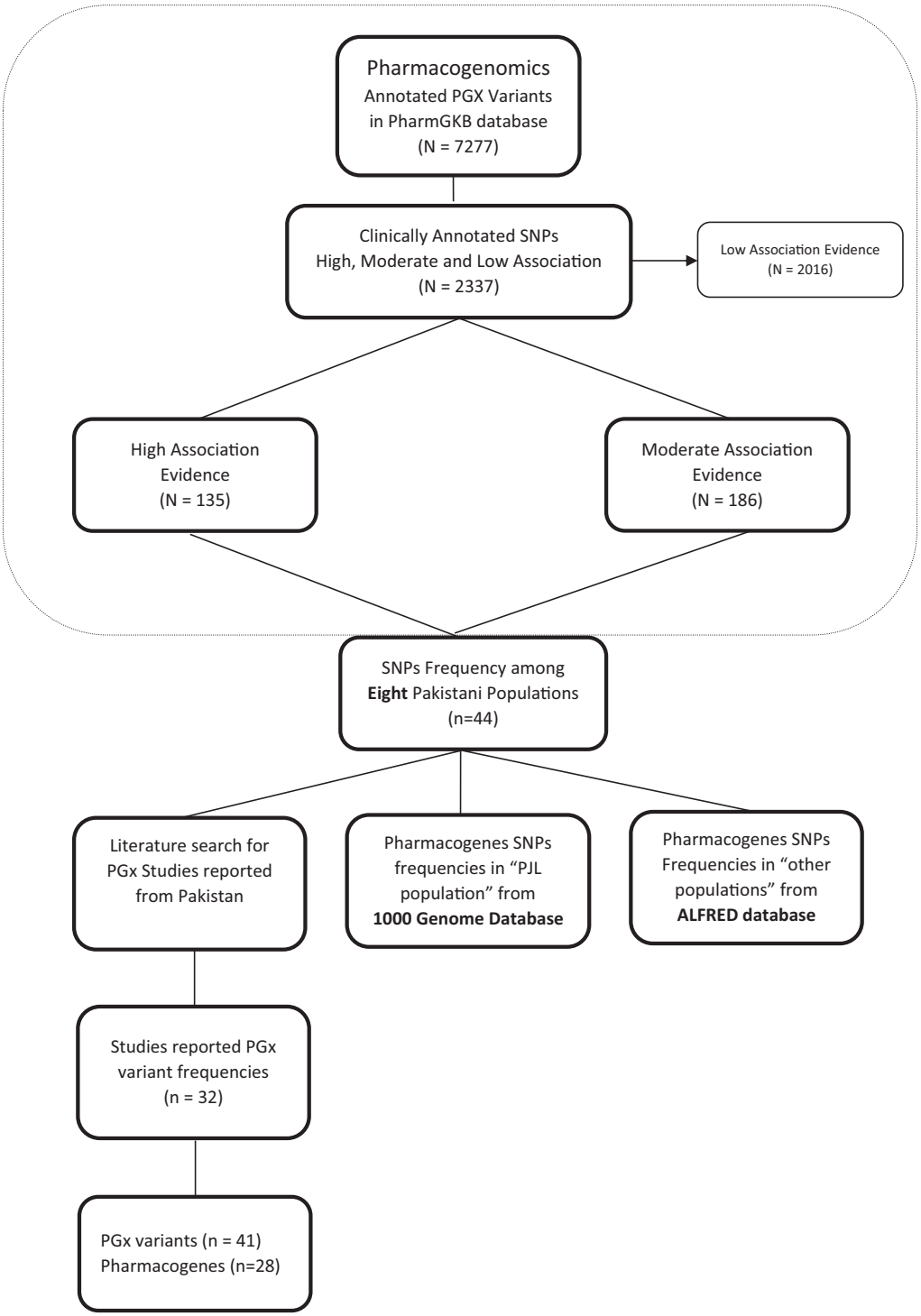

Figure 1. The overall design of the pharmacogenomics study.

of 321 variants were found. In the present study, we found 44 variants of PGx present in all available ethnic groups of the Pakistani population. These PGx variants belong to drug-metabolizing enzymes (n = 22), drug-metabolizing transporters (n = 8), and others (n = 14; includes transcriptional factors, binding proteins, regulatory variants, intron variants, etc.). The frequency data show the frequency of PGx variants/non-wild alleles among ethnic groups of Pakistan (Tables 1-3).

The 44 PGx variants are involved in the drug response of 12 clinical phenotypes (Table 4). These include drugs used for heart disease and associated factors (n = 13), cancers (n = 9), atherosclerosis (n = 4), mental health therapy (n = 5), drug addiction therapy (n = 4), antiretroviral therapy (n = 2), and pain management therapy (n = 2). Other variants (n = 5) were associated with diabetes type 2, asthma, hepatitis C, rheumatoid arthritis, and organ transplantation. The allele

frequencies of these PGx variants involved in the pharmacokinetics of drugs among Pakistani populations are given (Supplemental Table b).

Importantly, the frequency of 5 pharmacogenetic variants was found to be greater than >50% among the 8 ethnic groups of Pakistan. These include *CYP2B6* (rs2279345, C; 70%-86%), *CYP3A5* (rs776746, C; 64%-88%), *FLT3* (rs1933437, T; 54%-74%), *CETP* (rs1532624, A; 50%-70%) and an intronic variant of *DPP6* (rs6977820, C; 70%-86%) genes, which are involved in drug responses in acquired immune deficiency syndrome, organ transplantation, cancer, heart disease, and psychiatric therapy, respectively.

Published pharmacogenetic reports from Pakistani populations

We found a total of 753 articles using keywords. By selecting studies that contain frequencies of pharmacogenetic variants. Finally, 32 articles were retrieved after removing duplicate studies. These studies were classified as pharmacogenomics (n = 7) and genotype-phenotype association studies (n = 25). These published studies reported PGx variants (n = 41) of 28 genes in the Pakistani population. Most of the pharmacogenetic data (total studies, n = 21) are hospital patient data. However, very limited research institutes (total studies, n = 11) reported PGx data.

The total number of pharmacogenetic studies that reported allele frequencies from Pakistan are given in Table 5. The *CYP* genes are mostly studied. It includes *CYP2D6* (5 variants) and 2 variants of each *CYP2C9, CYP1A1, CYP2C6, CYPOR, SCNIA, VKORC1, IL-28B, MTHFR*, and *XRCC*. Interestingly, we found only 6 PGx variants associated with disease/drug response, which are clinically classified as high-moderate evidence. It includes *CYP2C9* (rs1057910, C; 14%), *CYP2D6* (rs3892097, T; 4.5%), *MTHFR* (rs1801133, T; 15.7%), *GTSP1* (rs1695, G; 25.5%), *OPRM1* (rs1799971, G; 14.5%), and *IFNL3* (rs8099917, G; 22%). The majority of PGx studies were conducted on cancer (n = 6). It includes breast cancers, bladder cancer, esophageal cancer, and renal cancers. Other PGx studies were include metabolic syndrome, diabetes, heart disease, and hepatitis.

Clustering of Pakistani population-based on PGx variants by principal component analysis

The genotypic frequencies of the pharmacogenetic variants (n = 44) of Pakistan (PJL) from 1KGP were compared with the populations of Africans, Asians (South and East Asians), Europe (Caucasians), and Americas (Hispanics). The PCA results of pharmacogenetic variants show that the Pakistani population lies among South Asian, Hispanic, and Caucasian. Interestingly, both Asian populations (East and South Asian) show differences in terms of genotype frequencies of pharmacogenetic variants. African populations are very distinct from the rest of the world population, as shown in the PCA plot (Supplemental Figure b).

Discussion

This study links clinically important genes with high to moderate associations (1A, 1B, 2A, and 2B) listed in pharmacogenetics databases with their frequencies reported among ethnic groups of Pakistan. We reported 44 PGx variants which correspond to drug-metabolizing enzymes, drug transporters, and others such as gene regulators, etc. Importantly, the frequency of 5 pharmacogenetic variants was found to be greater than >50% among the 8 ethnic groups of Pakistan. These include *CYP2B6* (rs2279345, C; 70%-86%), *CYP3A5* (rs776746, C; 64%-88%), *FLT3* (rs1933437, T; 54%-74%), *CETP* (rs1532624, A; 50%-70%), and an intronic variant of *DPP6* (rs6977820, C; 61%-86%), which are involved in drug responses in acquired immune deficiency syndrome, organ transplantation, cancer, heart disease, and psychiatric therapy, respectively.[14-18] These highly prevalent single nucleotide variants (SNVs) contribute to drug metabolisms that can affect the ultimate outcome of the disease.

In addition to 5 SNVs, certain variants have a high-frequency in specific ethnic groups. For example, Kalash population has high-frequency of *HAS3* (rs2232228, G; 48%), *ERCC1* (rs3212986, T; 42%), and *SEMA3C* (rs7779029, C; 34%). The Hazara population also showed the highest variant frequency in *NQO1* (rs1800566, A; 48%), *MTHFR* (rs1801133, A; 32%), and *CYP2D6* (rs3892097, T; 16%). Similarly, Pashtun population has *VKORC1* (rs7294, T; 70%) and *GSTP1* (rs1695, G; 33%). The Sindhi and Burusho populations showed a high-frequency of variants for *NT5C2* (rs11598702, C; 34%). The Punjabi population showed the highest frequency in *VKORC1* (rs7294, T; 69%). Interestingly, the majority of the high-frequency variants were observed in the Brahui population at *CBR3* (rs1056892, A; 64%), *KIF6* (rs20455, G; 46%), GATM (rs1719247, T; 40%), *CYP4F2* (rs2108622, T; 40%), *CYP2C9* (rs4917639, C; 26%), *CYP2C9* (rs1057910, C; 12%), and *SLC28A3* (rs885004, A; 14%). Overall, the highly frequent pharmacogenetic variants were found mostly in Brahui, Kalash, and Hazara populations. On the other hand, Burusho population displays the lowest frequency of variants among all ethnic groups *CYP4F2* (rs2108622, T; 26%), *CYP2C9* (rs4917639, C; 8%), *CYP2C9* (rs1057910, C; 6%), and *GP1BA* (rs6065, T; 4%). Thus, the presence of either high-frequency variants or retained ancestral alleles may have been due to natural selection based on their geographical location, environmental, and other factors. This PGx knowledge may be used to anticipate which ethnic groups are likely to respond to specific therapeutic drugs.

Recently, a significant increase in alleles of pharmacogenetic variants was reported in healthy Pakistani populations of *CYP450 1A2, 2B6, 2C19, 3A5, ALDH3A1, GSTM1, ABCB1*, and *ABCC2*.[19] This study shows a significant difference in the prevalence of variants of drug-metabolizing enzymes and drug transporters compared to other ethnic groups. However, the

Table 1. Frequency data of pharmacogenomic variants* among ethnic groups of Pakistani.

S. NO.	GENE*	SNP ID	VARIANT	BALOCHI (N=50)	BRAHUI (N=50)	BURUSHO (N=50)	HAZARA (N=48)	KALASH (N=50)	PASHTUN (N=46)	PUNJABI (N=158)	SINDHI (N=50)
Frequencies of PGx variants of drug-metabolizing enzymes (DME)											
1	*ATIC*	rs4673993 (T>A/C/G)	C	0.48	0.46	0.34	0.35	0.48	0.33	0.45	0.44
2	*CBR3*	rs1056892 (G>A)	A	0.44	0.64	0.54	0.42	0.56	0.37	0.47	0.48
3	*COL22A1*	rs6988229 (C>T)	T	0.14	0.16	0.30	0.08	0.22	0.07	0.09	0.12
4	*COMT*	rs4680 (G>A)	A	0.54	0.42	0.48	0.52	0.54	0.41	0.52	0.42
5	*CYP2B6*	rs2279345 (T>A/C/G)	C	0.72	0.78	0.72	0.77	0.86	0.70	0.77	0.86
6	*CYP2C9*	rs4917639 (A>C/T)	C	0.1	0.26	0.08	0.23	0.22	0.17	0.17	0.22
7	*CYP2C9*	rs1057910 (A>C/G)	C	0.08	0.12	0.06	0.12	0.06	0.11	0.10	0.10
8	*CYP2D6*	rs3892097 (C>T)	T	0.07	0.05	0.07	0.16	0.09	0.10	0.08	0.11
9	*CYP3A5*	rs776746 (T>C)	C	0.8	0.88	0.78	0.75	0.76	0.87	0.64	0.78
10	*CYP4F2*	rs2108622 (C>G/T)	T	0.32	0.40	0.26	0.29	0.36	0.30	0.39	0.32
11	*EPHX1*	rs1051740 (T>C)	C	0.3	0.24	0.24	0.42	0.42	0.35	0.39	0.46
12	*EPHX1*	rs2234922 (A>G/T)	G	0.24	0.34	0.12	0.19	0.06	0.24	0.28	0.32
13	*ERCC1*	rs3212986 (C>A/G/T)	T	0.26	0.30	0.30	0.25	0.42	0.33	0.31	0.34
14	*FLT3*	rs1933437 (G>A/T)	T	0.6	0.74	0.54	0.62	0.62	0.74	0.63	0.56
15	*GSTP1*	rs1695 (A>G/T)	G	0.18	0.20	0.26	0.21	0.14	0.33	0.29	0.26
16	*HAS3*	rs2232228 (A>C/G)	G	0.3	0.38	0.32	0.37	0.48	0.35	0.33	0.18
17	*MTHFR*	rs1801133 (G>A/C)	A	0.10	0.12	0.26	0.32	0.26	0.18	0.17	0.21
18	*NEDD4L*	rs4149601 (G>A)	A	0.26	0.20	0.34	0.17	0.14	0.09	0.19	0.26
19	*NQO1*	rs1800566 (G>A)	A	0.16	0.42	0.32	0.48	0.22	0.39	0.30	0.30
20	*NT5C2*	rs11598702 (T>C/G)	C	0.18	0.18	0.34	0.25	0.28	0.17	0.27	0.34
21	*PRKCA*	rs16960228 (G>A)	A	0.06	0.12	0.02	0.13	0.00	0.04	0.01	0.04
22	*VKORC1*	rs7294 (C>T)	T	0.52	0.48	0.62	0.21	0.30	0.70	0.69	0.52

*Frequency data means variants/non-wild alleles reported in dbSNP database.

Table 2. Frequencies of PGx variants of drug metabolizing transporters (DMT).

S. NO.	SYMBOLS	SNP ID	VARIANT	BALOCHI (N=50)	BRAHUI (N=50)	BURUSHO (N=50)	HAZARA (N=48)	KALASH (N=50)	PASHTUN (N=46)	PUNJABI (N=158)	SINDHI (N=50)
1	*CHRNA3*	rs578776 (G>A/C)	C	0.74	0.64	0.62	0.35	0.56	0.57	0.47	0.48
2	*CHRNA3*	rs1051730 (G>A)	A	0.48	0.38	0.28	0.21	0.26	0.3	0.21	0.26
3	*DRD2*	rs1799978 (T>C)	C	0.10	0.10	0.00	0.08	0.08	0.07	0.12	0.06
4	*FCGR2A*	rs1801274 (A>C/G)	C	0.40	0.42	0.32	0.35	0.54	0.41	0.39	0.34
5	*GP1BA*	rs6065 (C>G/T)	T	0.06	0.10	0.04	0.12	0.10	0.07	0.07	0.04
6	*GRIK4*	rs1954787 (T>C)	C	0.48	0.58	0.50	0.67	0.50	0.67	0.60	0.46
7	*OPRD1*	rs678849 (C>G/T)	T	0.54	0.46	0.44	0.69	0.60	0.65	0.66	0.68
8	*OPRM1*	rs1799971 (A>G)	G	0.08	0.28	0.30	0.21	0.28	0.24	0.37	0.24

Table 3. Frequencies of PGx variants (at gene regulating regions) of drug-metabolizing enzymes/transporters.

S. NO.	SYMBOLS	SNP ID	FUNCTION	VARIANT	BALOCHI (N=50)	BRAHUI (N=50)	BURUSHO (N=50)	HAZARA (N=48)	KALASH (N=50)	PASHTUN (N=46)	PUNJABI (N=158)	SINDHI (N=50)
1	*ABCG2*	rs2231142 (G>C/T)	Transfer protein	T	0.04	0.06	0.08	0.13	0.00	0.11	0.10	0.06
2	*CALU*	rs339097 (A>G)	Binding protein	G	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00
3	*CCHCR1*	rs746647 (A>G)	Regulator protein	G	0.18	0.26	0.18	0.33	0.30	0.28	0.20	0.18
4	*CETP*	rs1532624 (C>A)	Transfer protein	A	0.52	0.54	0.50	0.67	0.64	0.70	0.54	0.54
5	*IFNL3/IL-28B*	rs8099917 (T>G)	Cytokine	G	0.28	0.14	0.12	0.10	0.22	0.13	0.19	0.18
6	*KIF6*	rs20455 (A>G)	Motor protein	G	0.22	0.46	0.40	0.46	0.28	0.20	0.45	0.44
7	*SEMA3C*	rs7779029 (T>C)	Secreted glycoprotein	C	0.10	0.10	0.10	0.13	0.34	0.13	0.09	0.08
8	*SLC28A3*	rs885004 (G>A)	Nucleoside transporters	A	0.06	0.14	0.10	0.04	0.04	0.11	0.07	0.10
9	*TCF7L2*	rs7903146 (C>G/T)	Transcription factor	T	0.24	0.50	0.36	0.17	0.28	0.39	0.25	0.34
10	*YEATS4*	rs7297610 (C>T)	Transcription factor	T	0.04	0.04	0.00	0.00	0.08	0.02	0.02	0.02
11	*Near FAM119A (METTL21A) and CREB1*	rs2952768 (T>A/C/G)	Intergenic variant	C	0.62	0.48	0.42	0.54	0.28	0.35	0.48	0.56
12	*GATM*	rs1346268 (T>A/C/G)	Intron variant	C	0.38	0.36	0.30	0.54	0.22	0.24	0.33	0.34
13	*DPP6*	rs6977820 (T>A/C)	Intron variant	C	0.78	0.80	0.74	0.90	0.82	0.61	0.68	0.67
14	*GATM*	rs1719247 (C>G/T)	Regulatory region	T	0.38	0.40	0.38	0.54	0.22	0.28	0.39	0.36

Table 4. The effect of pharmacogenomic variants in drugs metabolism for clinical phenotypes.

S. NO.	GENE	SNP ID	DRUGS	DISEASE	PHENOTYPE	REFERENCES
1	*CYP2B6*	rs2279345 (T>A/C/G)	Efavirenz	Acquire immune-deficiency syndrome	Variant C showed lower efavirenz plasma levels as compared to patients with ancestral T allele	Sukasem et al. (2012)
2	*CCHCR1*	rs746647 (A>G)	Nevirapine		Variant A are at increased risk of adverse drug reaction with nevirapine	Chantarangsu et al. (2011)
3	*CHRNA3*	rs578776 (G>A/C)	Nicotine	Addiction	Variant A may have a decreased risk for tobacco addiction.	Winterer et al. (2010)
4	*CHRNA3*	rs1051730 (G>A)			Variant A may have an increased risk for nicotine dependency, decreased lung function.	
5	*OPRD1*	rs678849 (C>G/T)	Buprenorphine		Variant T may have an increased response to buprenorphine.	Crist et al. (2013)
6	*OPRM1*	rs1799971 (A>G)	Naloxone		Variant G may have increased peak cortisol response	Hernandez-Avila et al. (2007)
7	*COL22A1*	rs6988229 (C>T)	Salbutamol	Asthma	Variant T increase bronchodilator (FEV1) response salbutamol	Duan et al. (2014)
8	*NEDD4L*	rs4149601 (G>A)	Diuretics, Hydrochlorothiazide	Atherosclerosis	Variant A may have poorer response as compared with patients.	McDonough et al. (2013)
9	*PRKCA*	rs16960228 (G>A)	Hydrochlorothiazide		Variant A may have increased reduction of diastolic blood pressure.	Turner et al. (2013)
10	*ABCG2*	rs2231142 (G>C/T)	Rosuvastatin		Variant C/T may have higher plasma concentrations of rosuvastatin and better response to treatment	Tomlinson et al. (2010)
11	*YEATS4*	rs7297610 (C>T)	Hydrochlorothiazide		Variant T may have a decreased the drug response	Duarte et al. (2013)
12	*CYP2D6*	rs3892097 (C>T)	Tamoxifen	Cancer	Variant T may have increase the risk of relapse in patients suffering from breast neoplasms.	Hertz et al. (2015)
13	*ERCC1*	rs3212986 (C>A/G/T)	Platinum regimens		Variant A decreased the risk for nephrotoxicity.	Tzvetkov et al. (2011)
14	*FLT3*	rs1933437 (G>A/T)	Sunitinib		Variant A increase the risk of leukopenia, thrombocytopenia, and neutropenia.	Kim et al. (2013)
15	*GSTP1*	rs1695 (A>G/T)	Cyclophosphamide and Epirubicin		Variant G has decreased drug response and also increase the severity of toxicity.	Oliveira et al. (2010)
16	*MTHFR*	rs1801133 (G>A/C)	Methotrexate		Variant A has poor response to treatment with increased risk of toxicity and has greater risk of folate deficiency.	Zgheib et al. (2014)
17	*NQO1*	rs1800566 (G>A)	Platinum compounds, Anthracyclines, and related compounds		Variant A may have worse outcome (overall survival and progression-free survival).	Fagerholm et al. (2008)
18	*NT5C2*	rs11598702 (T>C/G)	Gemcitabine		Variant C/G may have increased clearance of gemcitabine.	Mitra et al. (2012)
19	*FCGR2A*	rs1801274 (A>C/G)	Trastuzumab		Variant C/G may have decreased drug response.	Tamura et al. (2010)
20	*SEMA3C*	rs7779029 (T>C)	Irinotecan		Variant C may have increased severity of Neutropenia.	Han et al. (2013)
21	*TCF7L2*	rs7903146 (C>G/T)	Sulfonylureas	Diabetes type 2	Variant G/T showed poor results in maintaining hemoglobin A1c (HbA1c) and fasting plasma glucose levels with sulfonamides	Schroner et al. (2011)

(Continued)

Table 4. (Continued)

S. NO.	GENE	SNP ID	DRUGS	DISEASE	PHENOTYPE	REFERENCES
22	*CBR3*	rs1056892 (G>A)	Anthracyclines and related substances	Heart disease	Variant A may have decreased risk of cardiac damage after anthracycline exposure.	Blanco et al. (2012)
23	*CYP2C9*	rs4917639 (A>C/T)	Warfarin		Variant C/ T may require decreased dose	Cooper et al. (2008)
24	*CYP2C9*	rs1057910 (A>C/G)	Warfarin		Variant C/G may require a decreased dose.	
25	*CYP4F2*	rs2108622 (C>G/T)	Warfarin		Variant T may require a higher dose.	Kurnik et al. (2012)
26	*HAS3*	rs2232228 (A>C/G)	Anthracyclines and related substances		Variant C/G may have decreased cardiomyopathy risk when exposed to high-dose.	Wang et al. (2014)
27	*VKORC1*	rs7294 (C>T)	Warfarin		Variant T may require a higher dose.	Cen et al. (2010)
28	*GP1BA*	rs6065 (C>G/T)	Aspirin		Variant T may have a decreased risk for aspirin resistance.	Matsubara et al. (2008)
29	*CALU*	rs339097 (A>G)	Warfarin		Variant G may require a higher maintenance dose.	Voora et al. (2010)
30	*CETP*	rs1532624 (C>A)	HMG CoA reductase inhibitors		Variant A may have decreased response.	De Keyser et al. (2011)
31	*KIF6*	rs20455 (A>G)	Pravastatin		Variant G may have a higher risk of coronary disease and may be more likely to benefit from pravastatin.	Li et al. (2011)
32	*SLC28A3*	rs885004 (G>A)	Anthracyclines and related substances		Variant A may have decreased likelihood of cardiotoxicity.	Visscher et al. (2012)
33	*GATM*	rs1346268 (T>A/C/G)	Statin, Simvastatin		Variant C may be less likely to experience myopathy.	Mangravite et al. (2013)
34	*GATM*	rs1719247 (C>G/T)	Statin, Simvastatin		Variant G/T may be less likely to experience myopathy.	
35	*IFNL3*	rs8099917 (T>G)	PEG-interferon alfa and Ribavirin	Hepatitis C	variant G may have decreased response (lower sustained viral response) to PEG-interferon alfa and ribavirin therapy	Riva et al. (2012)
36	*EPHX1*	rs1051740 (T>C)	Carbamazepine	Mental health	Variant C may have higher metabolism of carbamazepine and may require an increased dose	Nakajima et al. (2005)
37	*EPHX1*	rs2234922 (A>G/T)	Carbamazepine		Variant G/T may require an increased dose of carbamazepine	Puranik et al. (2013)
38	*DRD2*	rs1799978 (T>C)	Risperidone		Variant C may be less likely to have improvement in symptoms	Xing et al. (2007)
39	*GRIK4*	rs1954787 (T>C)	Antidepressants		Variant C may be more likely to respond to antidepressant treatment	Pu et al. (2013)
40	*DPP6*	rs6977820 (T>A/C)	Antipsychotics		Variant A/C may have decreased likelihood of side effect of antipsychotic.	Tanaka et al. (2013)
41	*COMT*	rs4680 (G>A)	Opioid	Pain management	Variant A results an increased response to opioid as compared to patients with the G allele. T	Rakvåg et al. (2008)
42	*Near FAM119A (METTL21A) and CREB1.*	rs2952768 (T>A/C/G)	Opioids		Variant C may have increased opioid analgesic requirements after surgery as compared to patients with T allele	Nishizawa et al. (2014)
43	*ATIC*	rs4673993 (T>A/C/G)	Methotrexate	Rheumatoid arthritis	Patients with variant (C) treated with methotrexate showed a better response as compared to patients with ancestral T allele	Iannaccone et al. (2010, 2011)
44	*CYP3A5*	rs776746 (T>C)	Tacrolimus	Transplantation	Transplant patients with variant C allele have showed reduce tacrolimus metabolism, resulting in increased plasma drug levels.	Niioka et al. (2012)

Table 5. Frequencies of pharmacogenomics variants reported in published literature from Pakistani.

S. NO.	GENE	SNPS ID/HAPLOTYPE	PGX VARIANTS FREQ (%)	REFERENCES
1	*ABCB1*	rs2032582 (A>T)	T (61.5)	Farhat et al. (2015). *JCPSP*, 25(7), 486-490.
		rs1128503 (A>G)	G (38)	Farhat et al. (2014). *Ann. Pak. Inst. Med. Sci.* 10. 3-6.
2	*ALDH2*	rs671 (G>A)	A (32.5)	Saleem et al. (2018). *Lipids in Health and Disease*. 17. 10.1186/s12944-018-0874-6.
3	*APOA5*	rs662799 (G>C)	C (33.1)	Fiaz et al. (2019). *JPMA*. 69(3), 301-305.
4	*ApoE*	E2 allele (rs7412-T, rs429358-T)	T (3.4-11)	Mehboob ali et al. (2015). *Pakistan Journal of Zoology*. 47. 263-268.
		E3 allele (rs7412-C, rs429358-T)	(71-85)	
		E4 allele (rs7412-C, rs429358-C)	C (11-18)	
5	*ARMS2*	rs10490924 (G>T)	T (31)	Ayub et al. (2019). *Ann Hum Genet*. 2019; 83: 285-290. https://doi.org/10.1111/ahg.12311
6	*CFH*	rs1061170 (C>T)	T (60.5)	
7	*CXCL12*	rs1801157 (C>T)	T (38.5)	Khalid, S., & Hanif, R. (2017). *PeerJ*, 5, e3822. https://doi.org/10.7717/peerj.3822
8	*CYP1A1*	rs4646903 (A>T)	T (18)	Zakiullah et al. (2014). *APJCP*, 15(16), 6715-6720.
9	*CYP1B1*	rs1056836 (G>C)	C (16)	Sheikh et al. (2014). *Molecular vision, 20*, 991-1001.
10	*CYP2C9*	rs1057910 (A>C)	C (14)	Yasmeen et al. (2015). *J Thromb Thrombolysis* 40, 218-224.
		rs1799853 (C>T)	T (12)	
11	*CYP2D6**	rs1065852 (G>A)	A (7)	Nazir et al. (2016). *The Journal of the Pakistan Medical Association, 66*(12), 1554-1558.
		rs3892097 (C>T)	T (4.5)	Anwarullah et al. (2017). https://doi.org/10.1186/s41021-017-0078-8
		rs16947 (G>A)	A (38)	Ahmed et al. (2018). *Genes, 9*(10), 514.
		rs1135840 (C>G)	G (35)	
		rs28371725 (C>T)	T (20)	
12	*CYPOR*	rs1057868 (C>T)	T (32.5)	
		rs41301394 (C>T)	T (32)	
13	*FTO*	rs9939609 (T>A)	A (39.5)	Fawwad et al. (2016). *Diabetes & metabolic syndrome, 10*(1), 43-47.
14	*GSTM1*	Null/non-functional	(38-46)	Abid et al. (2016). *Urologic oncology, 34*(9), 419.e1-419.e12.
15	*GSTT1*	Null/non-functional	(11-23)	

(Continued)

Table 5. (Continued)

S. NO.	GENE	SNPS ID/HAPLOTYPE	PGX VARIANTS FREQ (%)	REFERENCES
16	*GSTP1*	rs1695 (A>G)	G (25.5)	Ali et al. (2017). *Familial Cancer* 16, 577-594 (2017).
17	*HLA-DRB1*			Fawwad et al. (2019). *Diabetes research and clinical practice*, 149, 9-17.
18	*IL-28B (IFNL3)*	rs12979860 (C>T)	T (27.5)	Aziz et al. (2015). *International journal of infectious diseases*. 30, 91-97.
		rs8099917 (T>G)	T (22)	
19	*IL-6*	rs1800795 (C>G)	G (65.5)	Saleem et al. (2018).*Lipids in Health and Disease*. 17. 10.1186/s12944-018-0874-6.
20	*ITGB3*	rs5918 (T>C)	C (73.5)	
21	*PON1*	rs662 (T>A,C,G)	C (42.2)	
22	*MTHFR*	rs1801133 (G>A)	A (16)	Ullah et al. (2019). Personalized *medicine, 16*(1), 35-49.
		rs1801131(T>G)	G (25-54)	
23	*OPRM1*	rs1799971(A>G)	G (14.5)	Ahmed et al. (2018). *Analysis. J Mol Neurosci* 65, 472-479.
24	*SCN1A*	rs2298771(A>G)	T (46.5)	Nazish et al. (2018). *Therapeutics and clinical risk management, 14*, 2305-2313.
25	*SCN2A*	rs17183814(G>A)	A (46.5)	
26	*VKORC1*	rs9923231 (C>A)	A (21)	Qayyum et al. (2018). *Clinical and applied thrombosis*/hemostasis *24*(2), 323-329.
		rs9934438 (G>A)	A (50.5)	
27	*XPD*	rs13181 (T>G)	G (4.5)	Hameed et al. (2016). *Pakistan journal of pharmaceutical sciences, 29*(3), 869-876.
28	*XRCC1*	rs25487(T>C)	C (63)	
	XRCC1	rs1799782 (G>A)	A (5)	

Highlighted PGx variants are associated with high-moderated disease/drug response.

study had a very low sample size from the main ethnic groups such as Punjabi (n = 8; 5.2%), Pashtun (n = 5; 3.2%), Sindhi (n = 10; 6.5%), and Balochi (not available).

Cardiovascular diseases (CVDs) are the leading cause of death worldwide. The WHO estimated 17.9 million people died from CVDs, 85% of them due to heart attack and strokes.[20] Pakistani populations are at the highest risk of coronary heart disease. According to WHO estimates, the proportional mortality from CVD is 29% in Pakistan.[21] Atherosclerosis is the main cause of heart attacks, stroke, peripheral vascular diseases, etc. Hypertension and dyslipidemia/elevated cholesterol levels are the major contributing factors to atherosclerosis and heart disease.[22] Most variants of PGx (n = 13) are involved in heart disease and associated factors such as anticoagulants, 3-hydroxy-3-methylglutaryl coenzyme A (HMG-CoA) reductase inhibitors, and lipid-lowering drugs. It includes *VKORC1* (rs7294, T); *CYP4F2* (rs2108622, T); *CALU* (rs339097, G), *CYP2C9* (rs4917639, C/T and rs1057910, C/G) in warfarin pharmacokinetics.[23-26] The *CETP* variants (rs1532624, A) and *KIF6* (rs20455, G), are involved in HMG-CoA reductase inhibitors (cholesterol synthesis) and lipid-lowering drugs, respectively.[17,27] However, *GATM* (rs1346268, C and rs1719247, G/T), *HAS3* (rs2232228, C/G), *CBR3* (rs1056892, A), and *SLC28A3* (rs885004, A) with myopathy/cardiomyopathy are reported.[28-31]

The *VKORC1* allele variant (rs7294, T) was found to be the highest in Pashtun and Punjabi (70%) populations and the lowest in Hazara (21%) and Kalash (30%) populations. Allele frequencies for *CYP4F2* (rs2108622, T) and *CYP2C9* (rs4917639, C) vary from 26% to 40% and 8% to 26% among ethnic groups, respectively. Interestingly, the frequency of *CYP2C9* (rs1057910, C) was found to be low (6%-12%). However, *CALU* (rs339097, G) is not found in any of the Pakistani populations. The frequency of variants of *CETP* (rs1532624, A) is 50% to 70% in our population. The frequency of *KIF6* allele variants (rs20455, G) is 20% to 46% among ethnic groups. The variant frequencies of *GATM* (rs1346268, C) and (rs1719247, T) are 22% to 54%. The allele frequencies of variant *HAS3* (rs2232228, G), *CBR3* (rs1056892, A), and *SLC28A3* (rs885004, A) are 18% to 48%, 37% to 64% and 4% to 14%, respectively. The *GP1BA* (rs6065, T) have a decreased risk of aspirin resistance.[32] The frequency of *GP1BA* (rs6065, T) variants for aspirin metabolism is 4% to 12% in the Pakistani population.

Atherosclerosis

We found 4 variants of pharmacogenetics involved in atherosclerosis; *NEDD4L* (rs4149601, A), *ABCG2* (rs2231142, T), *YEATS4* (rs7297610, T), and *PRKCA* (rs16960228, A). These variants showed an association with drugs prescribed for blood pressure and cholesterol. Importantly, the variants *NEDD4L* (rs4149601, A) and *YEATS4* (rs7297610, T) showed a poor response to hydrochlorothiazide.[33,34] The frequency of *NEDD4L* (rs4149601, A) and *YEATS4* (rs7297610, T) is 9% to 34% and 0% to 8%, respectively. On the other hand, *PRKCA* (rs16960228, A) showed an increased reduction in diastolic blood pressure.[35] The *ABCG2* variants (rs2231142, C/T) showed increased plasma rosuvastatin concentrations and better response to treatment.[36] The frequency of *PRKCA* (rs16960228, A) and *ABCG2* (rs2231142, /T) is 0% to 13% in Pakistanis. The alternative allele frequencies of the drugs related to atherosclerosis showed the presence of very low variant frequencies *NEDD4L* (rs4149601, A) 9% to 34%, *PRKCA* (rs16960228, A), and *ABCG2* (rs2231142, T) 0% to 13% and *YEATS4* (rs7297610, T) 0% to 8% among Pakistani populations.

Cancer is the second leading cause of death (~9.6 million) worldwide. In Pakistan, the number of new cancer cases is 0.17 million and the number of deaths is 0.11 million, as reported by The Global Cancer Observatory.[37] Based on incidence and death by cancer type, breast (19%) and lip/oral (11%) cancers are more common, followed by lungs (5.6%), esophagus (4.6%), and leukemia (4%). We found 9 pharmacogenetic variants that are involved in cancer drug metabolism. These are *FLT3* (rs1933437, A), *FCGR2A* (rs1801274, C/G), *ERCC1* (rs3212986, T), *NQO1* (rs1800566, A), *GSTP1* (rs1695, G), *NT5C2* (rs11598702, C), *MTHFR* (rs1801133, A), *SEMA3C* (rs7779029, C), and *CYP2D6* (rs3892097, T).

The *FCGR2A* (rs1801274, variant C or G) showed a reduced/low drug response of Trastuzumab—anti-HER2.[38] The highest and lowest allele frequency of *FCGR2A* (rs1801274, C) was reported in Kalash (54%) and Burusho (32%) populations, respectively. The *NT5C2* variant (rs11598702, C) has increased clearance levels of gemcitabine drugs.[39] Sindhi and Burusho populations have the highest *NT5C2* (rs11598702, C) frequency of 34%, and the lowest frequency (17%) in Pashtun populations. The *ERCC1* (rs3212986, A) has been reported to reduce the risk of nephrotoxicity.[40] Other variants of pharmacogenetics *FLT3* (rs1933437, A), *NQO1* (rs1800566, A), *GSTP1* (rs1695, G), *MTHFR* (rs1801133, A), *SEMA3C* (rs7779029, C), and *CYP2D6* (rs3892097, T) cause adverse drug response such as increased risk of toxicity, leukopenia, thrombocytopenia, neutropenia, and relapses.[16,41-44] Allele frequencies ranges are as 54% to 74% for *FLT3* (rs1933437, A); 16% to 48% for *NQO1* (rs1800566, A); 14% to 33% for *GSTP1* (rs1695, G); 10% to 32% for *MTHFR* (rs1801133, A); 8% to 34% for *SEMA3C* (rs7779029, C), and 5% to 16% for *CYP2D6* (rs3892097, T) among ethnic groups. Interestingly, the Kalash populations showed the highest variant frequency of 54% for *FCGR2A* (rs1801274, C), 42% for *ERCC1* (rs3212986, T), and 34% for *SEMA3C* (rs7779029, C). The Hazara populations showed 48% *NQO1* (rs1800566, A), 32%*MTHFR* (rs1801133, A), and 16% *CYP2D6* (rs3892097 T,) respectively. The Pashtun population also showed the highest variant frequency of 74% for *FLT3* (rs1933437, T) and 33% for GST*P1* (rs1695, G).

Mental health

A total of 5 clinically significant pharmacogenetic variants were found. These are *DPP6* (rs6977820, A/C), *GRIK4* (rs1954787, C), *EPHX1* (rs1051740, C and rs2234922, G/T), and *DRD2* (rs1799978, C) which are associated with drug metabolism, response or recommended dosage to antipsychotics (neuroleptics or major tranquilizers), antidepressants, carbamazepine, and risperidone.[18,45-48] The range of alternative allele frequencies for psychotic disorders accounts for 70% to 86% for *DPP6* (rs6977820, C), 46% to 67% for *GRIK4* (rs1954787, C), 24% to 46% for *EPHX1* (rs1051740, C), and 6% to 34% (rs2234922, G/T) and 0% to 12% *DRD2* (rs1799978, C) among Pakistani ethnic groups.

Addiction

Treatment of substance abuse. We found 4 variants of pharmacogenetics associated with response to drug addiction therapy. These are *OPRD1* (rs678849, T), *CHRNA3* (rs578776, A and rs1051730, A), and *OPRM1* (rs1799971, G) for buprenorphine, nicotine, and naloxone, respectively.[49-51] The range of allele frequencies of *OPRD1* (rs678849, T), *CHRNA3* (rs578776 A), *CHRNA3* (rs1051730 A), and *OPRM1* (rs1799971 G) is 44% to 69%, 35% to 74%, 21% to 48%, and 8% to 37%, respectively.

Acquired immuno-deficiency syndrome (AIDs). The *CYP2B6* variant (rs2279345, C) and *CCHCR1* (rs746647, G) are 2 pharmacogenetic variants involved in drug response (efavirenz and nevirapine) to HIV treatment. The frequency of the *CYP2B6* variant allele (rs2279345, C) is quite high (70%–86%) among all ethnic groups. This may reflect that the Pakistani population (three-fourths) is at risk of low efavirenz levels, particularly (86%) Sindhi and Kalash populations. However, the allele frequency of the *CCHCR1* allele (rs746647, G) is 18% to 33%, the risk of adverse risk reaction with nevirapine.

Pain management. Pharmacovariants in *COMT* (rs4680, A) and (rs2952768, C) have shown an association with opioid drug metabolism.[52,53] The variant frequency *COMT* (rs4680, A) and (rs2952768, C) were found to be 41% to 54% and 28% to 56% in all ethnic groups. The highest frequency of variants of (rs4680, A) is found in the Kalash (54%) and Pashtun (41%) populations. The lowest frequency of variants of *COMT* (rs2952768, C) is found in the Kalash (28%) and Baloch (62%) populations.

Diabetes. A pharmacogenetic variant of *TCF7L2* (rs7903146, G/T) has shown association with glucose metabolism. Patients with ancestral allele C have been reported to be better able to maintain hemoglobin A1c (HbA1c) and fasting plasma glucose levels with sulfonamides compared to patients with the variant allele (T).[54] The frequency of allele variants (rs7903146, T) is 17% to 50% among the Pakistani population. The lowest frequency of variant alleles is reported in the Hazara population (17%) and the highest (50%) in Brahui population.

Asthma. Salbutamol/albuterol is the common drug used for asthma (bronchodilator). It reduced the response of the bronchodilator (FEV1) in patients with an ancestral C allele compared to patients with collagen (*COL22A1*) gene variant alleles (rs6988229, T).[55] In Pakistan, the frequency of *COL22A1* gene variant alleles (rs6988229, T) ranges from 7% to 30%, with the lowest (7%) in Pashtun population and the highest (30%) in Burusho population.

Rheumatoid arthritis. Methotrexate is a chemotherapy agent used to suppress the immune system. Studies have shown an association of the *ATIC* (5-aminoimidazole-4-carboxamide ribonucleotide formyl transferase/IMP cyclohydrolase) polymorphism (rs4673993, C) with rheumatic drug therapy.[56] Patients with rheumatoid arthritis with variant C treated with methotrexate showed a better response compared to patients with the ancestral T allele. The frequency of variants of *ATIC* is 33% to 48% in all ethnic groups of Pakistan. The highest frequency (48%) is found in Balochi and Kalash populations and the lowest (33%) in Pashtun population.

Hepatitis C. Peg-interferon alfa and ribavirin are the important drugs used in the treatment of liver disease (HBV and HCV). Interestingly, HCV genotypes and interferon-lambda SNPs contributed to the virological response to the drug. Single interferon-lambda SNPs (*INFL3* rs8099917, G) have shown an association with the sustained virologic response in genotype 1, 3, and 4 infected patients.[57] Patients with the variant allele (G) may have a decreased response (lower sustained viral response) to peg-interferon alfa and ribavirin therapy with HCV genotype 1 compared to patients with the ancestral (T) allele.[58] The frequency of *INFL3* (G) variants is 10% to 28% among ethnic groups. Baloch showed the highest frequency of variants (28%) and the lowest were found in Hazara (10%) compared to other ethnic populations.

Organ transplantation. The pharmacogenetic variant *CYP3A5* (rs776746) allele has shown evidence of a strong association with immunosuppressive drug metabolism (tacrolimus). Transplant patients with the variant allele *CYP3A5* (rs776746, C) have shown a reduction in tacrolimus metabolism, resulting in increased plasma drug levels.[59] Therefore, patients are at high risk for drug-related toxicity and need a low dose of tacrolimus. The overall frequency of *CYP3A5* (rs776746, C) was found to be much higher (64%-88%) in all ethnic groups, compared to the values reported in other South Asian populations. The Punjabis population has the lowest prevalence (64%) and Brahui has the highest prevalence (88%) of this variant. This indicates that two-thirds of

Pakistan's population would exhibit poor metabolizers of immunosuppressant drugs.

The genotypic data of 44 PGx variants are reported for Punjabi at 1KGP. The Punjabis constitute about half (45%) of the total Pakistani population. The Punjabi identity is primarily linguistic, and all those who speak this Indo-Aryan language as their first language are classified as Punjabis.[59] The Punjabis are further divided into many castes, tribes, and clans, residing in the most populous province of Pakistan, Punjab. Due to the large admixture found in this area, it often makes it difficult to classify indigenous Punjabi castes/tribes. Therefore, using the Punjabi population from Lahore in the principal component analysis as a reference for the Pakistani population may not truly represent the total pool of PGx in our country. The availability of other population pharmacogenetic data would better locate the position of Pakistani ethnic groups among the world population.

The principal component analysis plot shows that the Pakistani population is located among the South Asian population groups. Interestingly, the frequencies of pharmacogenetic variants in South Asia were distant from those in East Asia. The population cluster pattern also showed that African populations are distinct from the rest of the world's populations, indicating high ancestral alleles. Human migration is believed to have started from Africa to Europe and Asia. Thus, low variant frequencies are present in Africans followed by Europe and Asia. Recently, it has been reported that due to the presence of the variant allele *CYP3A5* (rs776746, C) in Asians, the tacrolimus concentration/dose (Co/D) ratio is significantly lower in the *CYP3A5* expresser group compared to the non-expresser in Asian and European populations at any posttransplant period.[60]

There are certain limitations in our study. First, we only selected PGx variants with high to moderate clinical evidence and their frequency among Pakistani populations (Punjabi, Sindhi, Pashtun, Balochi, Hazara, Kalash, Burusho, and Brahui). Low clinical evidence PGx variants were excluded from the study; primarily due to limited studies. Interestingly, it should be noted that most of the PGx-based studies are conducted on non-Asian populations. Further studies involving Asian populations may expand our knowledge about clinically significant variants.

To date, the PGx data for the Pakistani population are very limited. This may be due to the lack of trained manpower and financial constraints.[61] Additionally, there is no national database of pharmacogenetics available in Pakistan. Therefore, this study highlights clinically important pharmacogenetic variants and their frequencies in Pakistani populations and may provide targeted therapeutic drugs based on the genetic makeup of the patients. Ethnic specific PGx variants knowledge may help in selecting the right therapeutic drugs and its dosage for specific clinical phenotypes. This will impact on translational of PGx knowledge into clinical practice.

Author Contributions

AR and AK designed the study, AK, AR, SA, SHS, and SF search clinical data and analyze ethnic data. AK wrote the manuscript. SA, AA, and AR made substantial edits to the manuscript. All authors reviewed and approved the manuscript.

Disclaimer

It may please be noted that neither the manuscript nor any part of this manuscript is under consideration for publication elsewhere.

ORCID iDs

Abdul Rafay Khan https://orcid.org/0000-0001-6275-0956
Ali Raza https://orcid.org/0000-0001-9765-8818

Supplemental Material

Supplemental material for this article is available online.

REFERENCES

1. Relling MV, Klein TE. CPIC: clinical pharmacogenetics implementation consortium of the pharmacogenomics research Network. *Clin Pharmacol Ther*. 2011;89:464-467.
2. Lakiotaki K, Kanterakis A, Kartsaki E, Katsila T, Patrinos GP, Potamias G. Exploring public genomics data for population pharmacogenomics. *PLoS One*. 2017;12:e0182138.
3. The Clinical Pharmacogenetics Implementation Consortium (CPIC®). 2019. Accessed July 20, 2019. https://cpicpgx.org
4. The Canadian Pharmacogenomics Network for Drug Safety. 2019. Accessed July 20, 2019. http://cpnds.ubc.ca/
5. Luzum J, Pakyz R, Elsey A, et al. The Pharmacogenomics Research Network Translational Pharmacogenetics Program: outcomes and metrics of pharmacogenetic implementations across diverse healthcare systems. *Clin Pharmacol Ther*. 2017;102:502-510.
6. Whirl-Carrillo M, McDonagh EM, Hebert JM, et al. Pharmacogenomics knowledge for personalized medicine. *Clin Pharmacol Ther*. 2012;92:414-417.
7. Pakistan Bureau of Statistics, Government of Pakistan. Brief on Census-2017. Updated January 21, 2022. Accessed July 10, 2019. https://www.pbs.gov.pk/content/brief-census-2017
8. Raza A, Firasat S, Khaliq S, et al. HLA class I and II polymorphisms in the Gujjar population from Pakistan. *Immunol Invest*. 2013;42:691-700.
9. Gonzalez-Galarza FF, Christmas S, Middleton D, Jones AR. Allele frequency net: a database and online repository for immune gene frequencies in worldwide populations. *Nucleic Acids Res*. 2011;39:D913-D919.
10. The 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature*. 2015;526:68-74.
11. McDonagh EM, Whirl-Carrillo M, Garten Y, Altman RB, Klein TE. From pharmacogenomic knowledge acquisition to clinical applications: the PharmGKB as a clinical pharmacogenomic biomarker resource. *Biomark Med*. 2011;5:795-806.
12. Yates A, Beal K, Keenan S, et al. The ensembl REST API: ensembl data for any language. *Bioinformatics*. 2015;31:143-145.
13. The 1000 Genomes Browser. 2019. Accessed September 20, 2019. ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/.
14. Sukasem C, Cressey TR, Prapaithong P, et al. Pharmacogenetic markers of CYP2B6 associated with efavirenz plasma concentrations in HIV-1 infected Thai adults. *Br J Clin Pharmacol*. 2012;74:1005-1012.
15. Rojas L, Neumann I, Herrero MJ, et al. Effect of CYP3A5*3 on kidney transplant recipients treated with tacrolimus: a systematic review and meta-analysis of observational studies. *Pharmacogenomics J*. 2015;15:38-48.
16. Kim HR, Park HS, Kwon WS, et al. Pharmacogenetic determinants associated with sunitinib-induced toxicity and ethnic difference in Korean metastatic renal cell carcinoma patients. *Cancer Chemother Pharmacol*. 2013;72:825-835.
17. de Keyser CE, Eijgelsheim M, Hofman A, et al. Single nucleotide polymorphisms in genes that are associated with a modified response to statin therapy: the Rotterdam Study. *Pharmacogenomics J*. 2011;11:72-80.

18. Tanaka S, Syu A, Ishiguro H, et al. DPP6 as a candidate gene for neuroleptic-induced tardive dyskinesia. *Pharmacogenomics J*. 2013;13:27-34.
19. Afsar NA, Bruckmueller H, Werk AN, Nisar MK, Ahmad HR, Cascorbi I. Implications of genetic variation of common drug metabolizing enzymes and ABC transporters among the Pakistani population. *Sci Rep*. 2019;9:7323.
20. World Health Organization (WHO). Cardiovascular diseases (CVDs). 2019. Accessed September 20, 2019. https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)
21. World Health Organization. Diseases (NCD) Country Profiles. 2018. Accessed September 10, 2019. https://www.who.int/nmh/countries/2018/pak_en.pdf?ua=1
22. Liaquat A, Javed Q. Current trends of cardiovascular risk determinants in Pakistan. *Cureus*. 2018;10:e3409.
23. Voora D, Koboldt D, King C, et al. A polymorphism in the VKORC1-regulator calumenin predicts higher warfarin doses in African-Americans. *Clin Pharmacol Ther*. 2010;87:445-451.
24. Kurnik D, Qasim H, Sominsky S, et al. Effect of the VKORC1 D36Y variant on warfarin dose requirement and pharmacogenetic dose prediction. *Thromb Haemost*. 2012;108:781-788.
25. Cen HJ, Zeng WT, Leng XY, et al. CYP4F2 rs2108622: a minor significant genetic factor of warfarin dose in Han Chinese patients with mechanical heart valve replacement. *Br J Clin Pharmacol*. 2010;70:234-240.
26. Cooper GM, Johnson JA, Langaee TY, et al. A genome-wide scan for common genetic variants with a large influence on warfarin maintenance dose. *Blood*. 2008;112:1022-1027.
27. Li Y, Sabatine MS, Tong CH, et al. Genetic variants in the KIF6 region and coronary event reduction from statin therapy. *Hum Genet*. 2011;129:17-23.
28. Mangravite LM, Engelhardt BE, Medina MW, et al. A statin-dependent QTL for GATM expression is associated with statin-induced myopathy. *Nature*. 2013;502:377-380.
29. Blanco JG, Sun CL, Landier W, et al. Anthracycline-related cardiomyopathy after childhood cancer: role of polymorphisms in carbonyl reductase genes—a report from the children's Oncology Group. *J Clin Oncol*. 2012;30:1415-1421.
30. Visscher H, Ross CJ, Rassekh SR, et al. Pharmacogenomic prediction of anthracycline-induced cardiotoxicity in children. *J Clin Oncol*. 2012;30:1422-1428.
31. Wang X, Liu W, Sun CL, Armenian SH, Hakonarson H, Hageman L, et al. Hyaluronan synthase 3 variant and anthracycline-related cardiomyopathy: a report from the children's oncology group. *J Clin Oncol*. 2014;32:647-653.
32. Matsubara Y, Murata M, Watanabe G, Ikeda Y. Enhancing effect of the 145Met-allele of GPIb alpha on platelet sensitivity to aspirin under high-shear conditions. *Thromb Res*. 2008;123:331-335.
33. McDonough CW, Burbage SE, Duarte JD, et al. Association of variants in NEDD4L with blood pressure response and adverse cardiovascular outcomes in hypertensive patients treated with thiazide diuretics. *J Hypertens*. 2013;31:698-704.
34. Duarte JD, Turner ST, Tran B, et al. Association of chromosome 12 locus with antihypertensive response to hydrochlorothiazide may involve differential YEATS4 expression. *Pharmacogenomics J*. 2013;13:257-263.
35. Turner ST, Boerwinkle E, O'Connell JR, et al. Genomic association analysis of common variants influencing antihypertensive response to hydrochlorothiazide. *Hypertension*. 2013;62:391-397.
36. Tomlinson B, Hu M, Lee VWY, et al. ABCG2 polymorphism is associated with the low-density lipoprotein cholesterol response to rosuvastatin. *Clin Pharmacol Ther*. 2010;87:558-562.
37. The Global Cancer Observatory. Globocan Report. 2018. Accessed September 20, 2019. https://gco.iarc.fr/today/data/factsheets/populations/586-pakistan-fact-sheets.pdf
38. Tamura K, Shimizu C, Hojo T, et al. FcγR2A and 3A polymorphisms predict clinical outcome of trastuzumab in both neoadjuvant and metastatic settings in patients with HER2-positive breast cancer. *Ann Oncol*. 2011;22:1302-1307.
39. Mitra AK, Kirstein MN, Khatri A, et al. Pathway-based pharmacogenomics of gemcitabine pharmacokinetics in patients with solid tumors. *Pharmacogenomics*. 2012;13:1009-1021.
40. Tzvetkov MV, Behrens G, O'Brien VP, Hohloch K, Brockmöller J, Benöhr P. Pharmacogenetic analyses of cisplatin-induced nephrotoxicity indicate a renoprotective effect of ERCC1 polymorphisms. *Pharmacogenomics*. 2011;12:1417-1427.
41. Fagerholm R, Hofstetter B, Tommiska J, et al. NAD(P)H:quinone oxidoreductase 1 NQO1*2 genotype (P187S) is a strong prognostic and predictive factor in breast cancer. *Nat Genet*. 2008;40:844-853.
42. Oliveira AL, Rodrigues FF, Santos RE, et al. GSTT1, GSTM1, and GSTP1 polymorphisms and chemotherapy response in locally advanced breast cancer. *Genet Mol Res*. 2010;9:1045-1053.
43. Han JY, Shin ES, Lee YS, et al. A genome-wide association study for irinotecan-related severe toxicities in patients with advanced non-small-cell lung cancer. *Pharmacogenomics J*. 2013;13:417-422.
44. Zgheib NK, Akra-Ismail M, Aridi C, et al. Genetic polymorphisms in candidate genes predict increased toxicity with methotrexate therapy in Lebanese children with acute lymphoblastic leukemia. *Pharmacogenet Genomics*. 2014;24:387-396.
45. Nakajima Y, Saito Y, Shiseki K, et al. Haplotype structures of EPHX1 and their effects on the metabolism of carbamazepine-10,11-epoxide in Japanese epileptic patients. *Eur J Clin Pharmacol*. 2005;61:25-34.
46. Xing Q, Qian X, Li H, et al. The relationship between the therapeutic response to risperidone and the dopamine D2 receptor polymorphism in Chinese schizophrenia patients. *Int J Neuropsychopharmacol*. 2007;10:631-637.
47. Pu M, Zhang Z, Xu Z, et al. Influence of genetic polymorphisms in the glutamatergic and GABAergic systems and their interactions with environmental stressors on antidepressant response. *Pharmacogenomics*. 2013;14:277-288.
48. Puranik YG, Birnbaum AK, Marino SE, et al. Association of carbamazepine major metabolism and transport pathway gene polymorphisms and pharmacokinetics in patients with epilepsy. *Pharmacogenomics*. 2013;14:35-45.
49. Winterer G, Mittelstrass K, Giegling I, et al. Risk gene variants for nicotine dependence in the CHRNA5-CHRNA3-CHRNB4 cluster are associated with cognitive performance. *Am J Med Genet B*. 2010;153B:1448-1458.
50. Crist RC, Clarke TK, Ang A, et al. An intronic variant in OPRD1 predicts treatment outcome for opioid dependence in African-Americans. *Neuropsychopharmacology*. 2013;38:2003-2010.
51. Hernandez-Avila CA, Covault J, Wand G, Zhang H, Gelernter J, Kranzler HR. Population-specific effects of the Asn40Asp polymorphism at the mu-opioid receptor gene (OPRM1) on HPA-axis activation. *Pharmacogenet Genomics*. 2007;17:1031-1038.
52. Rakvåg TT, Ross JR, Sato H, Skorpen F, Kaasa S, Klepstad P. Genetic variation in the catechol-O-methyltransferase (COMT) gene and morphine requirements in cancer patients with pain. *Mol Pain*. 2008;4:64.
53. Nishizawa D, Fukuda K, Kasai S, et al. Genome-wide association study identifies a potent locus associated with human opioid sensitivity. *Mol Psychiatry*. 2014;19:55-62.
54. Schroner Z, Javorsky M, Tkacova R, et al. Effect of sulphonylurea treatment on glycaemic control is related to TCF7L2 genotype in patients with type 2 diabetes. *Diabetes Obes Metab*. 2011;13:89-91.
55. Duan QL, Lasky-Su J, Himes BE, et al. A genome-wide association study of bronchodilator response in asthmatics. *Pharmacogenomics J*. 2014;14:41-47.
56. Iannaccone CK, Lee YC, Cui J, et al. Using genetic and clinical data to understand response to disease-modifying anti-rheumatic drug therapy: data from the Brigham and Women's Hospital rheumatoid arthritis sequential study. *Rheumatology*. 2011;50:40-46.
57. Susser S, Herrmann E, Lange C, et al. Predictive value of interferon-lambda gene polymorphisms for treatment response in chronic hepatitis C. *PLoS One*. 2014;9:e112592.
58. Riva E, Scagnolari C, Monteleone K, et al. Interleukin-28B (IL-28B) single-nucleotide polymorphisms and interferon plus ribavirin treatment outcome in Italian chronically HCV-infected patients. *J Viral Hepat*. 2012;19:650-653.
59. Niioka T, Satoh S, Kagaya H, et al. Comparison of pharmacokinetics and pharmacogenetics of once- and twice-daily tacrolimus in the early stage after renal transplantation. *Transplantation*. 2012;94:1013-1019.
60. Khan AR, Raza A, Firasat S, Abid A. CYP3A5 gene polymorphisms and their impact on dosage and trough concentration of tacrolimus among kidney transplant patients: a systematic review and meta-analysis. *Pharmacogenomics J*. 2020;20:553-562.
61. Rodrigues-Soares F, Suarez-Kurtz G. Pharmacogenomics research and clinical implementation in Brazil. *Basic Clin Pharmacol Toxicol*. 2019;124:538-549.