Observational Study

Medicine®

OPEN

Genetic polymorphisms in very important pharmacogenomic variants in the Zhuang ethnic group of Southwestern China

A cohort study in the Zhuang population

Jing Li, MD[a], Chenghao Guo, MD[b], Mengdan Yan, MD[a], Fanglin Niu, MD[a], Peng Chen, MD[a], Bin Li, MD[a], Tianbo Jin, MD, PhD[a,c,d,e,*]

Abstract

Pharmacogenomics, the study of the role of genetics in drug response, has recently become a focal point of research. Previous studies showed that genes associated with drug detoxification vary among different populations. However, pharmacogenomic information of the Zhuang ethnic group is scarce. The aim of the present study was to screen members of the Zhuang ethnicity in southwestern China for genotype frequencies of very important pharmacogenomic (VIP) variants and to determine the differences between the Zhuang ethnicity and other human populations.

We genotyped 80 variants of VIP genes in 100 unrelated healthy Zhuang adults from the Yunnan province of China. Next, we analyzed the genotyping data with Structure and *F*-statistics (F_{st}).

We compared our data with those of other populations using the HapMap data set, and observed that the frequency distribution of Zhuang population in Yunnan closely resembles that of JPT. Furthermore, population structure and F_{st} analysis showed that the Zhuang population is closely related to the Shaanxi Han population with respect to genetic background.

Our study supplements existing information on Zhuang population pharmacogenomics and provides an extensive overview for developing personalized medicine.

Abbreviations: ASW = African ancestry in Southwest USA, CEU = Utah, USA residents with Northern and Western European ancestry from the CEPH collection, CHB = Han Chinese in Beijing, CHD = Chinese in metropolitan Denver, F_{st} = *F*-statistics, GIH = Gujarati Indians in Houston, JPT = Japanese in Tokyo, LWK = Luhya in Webuye, Kenya, MEX = Mexican ancestry in Los Angeles, California, USA, MKK = Maasai in Kinyawa, Kenya, TSI = Toscani in Italy, VIP = very important pharmacogenomics, YRI = Yoruba in Ibadan, Nigeria.

Keywords: genetic polymorphisms, pharmacogenomics, VIP variants, Zhuang

1. Introduction

Drug response and reaction vary among individuals. A nationwide study conducted in Spain from 2001 to 2006 showed that 3.5 million people were hospitalized with adverse drug reactions (ADRs), and >5% of these patients eventually died.[1] Pharmacogenomics focuses on the inheritance of individual variations in drug response, and eventually provides guidance to precision medical treatment.[2] Since the term pharmacogenomics appeared in the literature in 1997, the number of articles identifying genetic variations is rapidly increasing.[3] In 2005, a database called Pharmacogenetics and Pharmacogenomics Knowledge Base (PharmGKB) was established for sharing genotype, phenotype, or other data on genetic variation among researchers.[4] Currently, PharmGKB is an easily accessible versatile knowledge database, which contains information on gene variant annotations, drug-centered pathways, and very important pharmacogenes (VIPs).

Evidence shows that genetic variant characteristics vary with populations or ethnicities.[5] For example, CYP2C9, a member of the CYP450 superfamily, is an enzyme related to metabolism of many drugs such as diclofenac and warfarin. Reports show that the allele frequency of the CYP2C9*2 allele is 15% among Caucasians, 1% to 3.6% among African Americans, but 0% among Asians. Thus, one of the major tasks in population pharmacogenetics and pharmacogenomics is to determine the frequencies of polymorphisms in drug detoxification genes among different ethnicities.[6]

China is the most populated country in the world. In addition to the Han people who make up 96% of the country's total population, there are 55 ethnic minority groups in China. Previously, we reported VIP variants in several Chinese ethnic

Editor: Saeed Alzghari.

JL and CG both contributed equally to this study.

This study was supported by grants from the Science and Technology Agency Project of Xizang (Tibet) Autonomous Region (No. 2015ZR-13-11) and graduate student independent innovation project of Northwest University (No. YZZ17162).

The authors have no conflicts of interest to disclose.

[a] Key Laboratory of Resource Biology and Biotechnology in Western China, Northwest University, Ministry of Education, [b] Xi'an 21st Century Precision Medicine Research Institute Co. Ltd, [c] Key Laboratory of Molecular Mechanism and Intervention Research for Plateau Diseases of Tibet Autonomous Region, School of Medicine, Xizang Minzu University, [d] Key Laboratory of High Altitude Environment and Genes Related to Diseases of Tibet Autonomous Region, School of Medicine, [e] Key Laboratory for Basic Life Science Research of Tibet Autonomous Region, School of Medicine, Xizang Minzu University, Xianyang, Shaanxi, China.

[*] Correspondence: Tianbo Jin, 386, #229 North Taibai Road, Xi'an 710069, Shaanxi, China (e-mail: jintianbo@gmail.com).

Medicine (2018) 97:17(e0559)

Received: 25 October 2017 / Accepted: 3 April 2018

http://dx.doi.org/10.1097/MD.0000000000010559

groups, including the Deng, Han, Li, Lohoba, Kyrgyz, Miao, Mongol, Sherpa, Tajik, Uygur, and Tibetan population.[7–17] According to the data of the sixth nationwide population census, the Zhuang, a minority with the largest population in China, has > 16 million people. Most Zhuang people live in Guangxi, Yunnan, Guangdong, Guizhou, and Hunan provinces. However, information on the VIP variants of the Zhuang ethnic group is limited. Therefore, identification of pharmacogenomic variants of Zhuang may extend our understanding of VIP gene variants among different populations.

In this study, we selected and genotyped 80 VIP variants in 100 Zhuang ethnic volunteers from the Yunnan province of China. Next, we compared the frequency differences between the selected Zhuang cohort and 11 major HapMap populations. Finally, F_{st} values were calculated to infer the population structure. Our results will supplement the existing VIP variant data of the Zhuang ethnic group, and may extend our understanding of ethnic diversity and pharmacogenomics.

2. Methods

2.1. Study subjects

We randomly recruited 100 unrelated healthy Zhuang adults (50 males and females each in the age range of 25–40 years) from the Yunnan province of Northwestern China and confirmed their ethnicity from lineage and birth place information. Written informed consent was obtained from all subjects. The study was performed in accordance with the Declaration of Helsinki and approved by the Human Research Committee of the Northwest University for Approval of Research Involving Human Subjects.

2.2. Variant screening and genotyping

We searched the PharmGKB database (https://www.pharmgkb.org/) and selected 80 genetic variants according to available data on frequency, functionality, and linkage based on published research. Genomic DNA was extracted from blood samples using Gold Mag-Mini whole blood genomic DNA purification kit (Gold Mag Ltd., Xi'an, China) according to the manufacturer's protocol. Optical density at 260 nm (OD_{260}) was detected by spectrometry (DU530 UV/VIS spectrophotometer, Beckman Instruments, Fullerton, CA) to estimate DNA concentration. Multiplexed SNP MassEXTEND arrays were designed using the Sequenom MassARRAY Assay Design 3.0 software (San Diego, California).[18] Genotyping of SNPs was conducted by Sequenom MassARRAY RS1000 (San Diego, California) according to manufacturer's instructions. Sequenom Typer 4.0 software was used for data collection and analysis as described previously.[19]

2.3. HapMap genotype data

We downloaded the genotype data of eleven populations from the International HapMap Project website (HapMap_release127) at http://hapmap.ncbi.nlm.nih.gov/biomart/martview/e4f42d4d0acde5ea6c35312381c1e461. The full names of the 11 populations are as follows:African ancestry in Southwest USA (ASW), Utah, USA residents with Northern and Western European ancestry from the CEPH collection (CEU), Han Chinese in Beijing, China (CHB), Chinese in metropolitan Denver, CO (CHD), Gujarati Indians in Houston, TX (GIH), Japanese in Tokyo, Japan (JPT), Luhya in Webuye, Kenya (LWK), Mexican ancestry in Los Angeles, CA (MEX), Maasai in Kinyawa, Kenya (MKK), Toscani in Italy (TSI), and Yoruba in Ibadan, Nigeria (YRI).

2.4. Statistical analysis

We used Excel and SPSS 19.0 statistical packages (SPSS, Chicago, IL) to conduct Hardy–Weinberg equilibrium (HWE) analysis and the χ^2 tests. Validation of the frequency of each variant in the Zhuang people was tested by assessing the departure from HWE using an exact test. In this study, all the P values were calculated 2 sided, and the criterion of statistical significance was $P < .05$, or $P < .000625$ (.05/80) after Bonferroni's multiple adjustment.[20] After χ^2 tests, we selected 2 SNPs which showed more difference between Zhuang and the 11 other populations to perform a global allele frequency analysis. The allele data were downloaded from ALFRED (https://alfred.med.yale.edu/).

2.5. Structure analysis

To analyze the pairwise genetic distance among populations, pairwise F_{st} values were calculated using Arlequin v3.5. Population structure was investigated using the Bayesian clustering algorithm STRUCTURE ver. 2.3.1 (Pritchard Lab, Stanford University, http://pritchardlab.stanford.edu/structure.html).[21] The data of the Deng, Han, Li, Lohoba, Kyrgyz, Miao, Mongol, Sherpa, Tajik, Uygur, and Tibetan populations were obtained from our previous studies. Analyses were performed using the ancestry model with correlated allele frequencies in runs from $K=6$ to $K=10$ (K is the number of genetically distinct clusters). The model choice criterion implemented in structure to detect the true K is an estimate of the posterior probability of the data for a given K, Pr(X|K). This value is obtained by first computing the log likelihood of the data at each step of the MCMC, called "Ln P(D)." To infer the number of clusters, ΔK was calculated using the method of Evanno.[22]. Graphs of STRUCTURE results were generated using Excel.

2.6. Population tree

A population tree was constructed using the F_{st} data and MEGA7 to infer the evolutionary relationship between the 12 populations.[23] The evolutionary history was subsequently inferred using the neighbor-joining method.[24]

3. Results

We successfully genotyped 80 VIP variants selected from PharmGKB VIP in 100 members of the Zhuang population. The basic information of 80 selected variants is listed in Table 1, including those associated with related genes, families, phases, primary locus, alleles, alternative amino acids, and genotype frequencies of 100 Zhuang people. All the variants tested met the HWE.

We first compared the allele frequency differences among the Zhuang ethnic group and the 11 groups selected from the International HapMap project database. In the ASW population, 22 of the selected VIP variants showed differences with Zhuang. The results of the other groups are as follows: CEU, 27; CHB, 11; CHD, 28, GIH, 29; JPT, 8; LWK, 26; MEX, 20; MKK, 24; TSI, 19; YRI, 34 (Table 2). To adjust for multiple comparisons, the level of statistical significance was reduced to 0.000625 (0.05/80), using which, the number of variants obtained with existing

Table 1

Basic information of selected SNP in the Zhuang people.

SNP	Gene	Chr	Allele		Phase	Position	Family	Amino Acid Translation	Zhuang		
			A	B					AA	AB	BB
rs3918290	*DPYD*	chr1	C	T	Phase I[†]	97915614			100	0	0
rs1801131	*MTHFR*	chr1	T	G	Phase I	11854476	Methylenetetrahydrofolate reductase family	Glu429Ala	58	32	10
rs1801133	*MTHFR*	chr1	G	A	Phase I	11856378	Methylenetetrahydrofolate reductase family	Ala222Val	60	32	8
rs689466	*PTGS2*	chr1	T	C	Phase I	186650751			27	50	23
rs1801253	*ADRB1*	chr10	G	C	Phase I	115805056	Adrenergic receptors family	Gly389Arg	6	36	53
rs4986893	*CYP2C19*	chr10	A	G	Phase I	96540410	Cytochrome P450 superfamily	Trp212Null	100	0	0
rs1799853	*CYP2C19*	chr10	C	T	Phase I	96702047	Cytochrome P450 superfamily	Arg144Cys	100	0	0
rs1800497	*ANKK1*	chr11	G	A	Phase I	113270828	Ser/Thr protein kinase family	Glu713Lys	56	36	7
rs6277	*DRD2*	chr11	G	A	Others	113283459	G-protein coupled receptor	Pro319Pro	83	17	0
rs1138272	*GSTP1*	chr11	C	T	Phase II	67353579	Glutathione S-transferase family	Ala114Val	100	0	0
rs1695	*GSTP1*	chr11	A	G	Phase II	67352689	Glutathione S-transferase family	Leu105Val	51	37	12
rs4149056	*SLCO1B1*	chr12	T	C	Others	21331549	Solute carrier family	Val174Ala	83	16	1
rs10735810	*VDR*	chr12	A	G	Others	48272895	Nuclear receptor family		28	51	21
rs11568820	*VDR*	chr12	C	T	Others	48302545	Nuclear receptor family		28	35	17
rs1540339	*VDR*	chr12	C	T	Others	48257326	Nuclear receptor family		12	46	42
rs1544410	*VDR*	chr12	C	T	Others	48239835	Nuclear receptor family		97	3	0
rs2228570	*VDR*	chr12	T	C	Others	48272895	Nuclear receptor family	Met1Thr	21	51	28
rs2239179	*VDR*	chr12	T	C	Others	48257766	Nuclear receptor family		0	0	0
rs2239185	*VDR*	chr12	G	A	Others	48244559	Nuclear receptor family		48	49	3
rs731236	*VDR*	chr12	A	G	Others	48238757	Nuclear receptor family	Ile352Ile	92	8	0
rs7975232	*VDR*	chr12	C	A	Others	48238837	Nuclear receptor family		46	50	4
rs1800566	*NQO1*	chr16	G	A	Phase II	69711242		Pro187Ser	26	57	17
rs1801030	*SULT1A1*	chr16	C	T	Phase II	28617485	sulfotransferase family	Val223Met	0	0	100
rs3760091	*SULT1A1*	chr16	G	C	Phase II	28609479	sulfotransferase family		41	52	7
rs7294	*VKORC1*	chr16	C	T	Phase I	31102321			72	27	1
rs9923231	*VKORC1*	chr16	A	C	Phase I	31096368			100	0	0
rs9934438	*VKORC1*	chr16	G	A	Phase I	31104878			2	26	72
rs151264360	*TYMS*	chr18	–	TTAAAG	Others	673444:673449			100	0	0
rs1801272	*CYP2A6*	chr19	A	T	Phase I	41354533	Cytochrome P450 superfamily	Leu160His	0	0	100
rs28399433	*CYP2A6*	chr19	G	T	Phase I	41356379	Cytochrome P450 superfamily		63	28	8
rs28399444	*CYP2A6*	chr19	G	A	Phase I	41354190	Cytochrome P450 superfamily	Glu197Ser, Glu197Arg	100	0	0
rs28399454	*CYP2A6*	chr19	C	T	Phase I	41351267	Cytochrome P450 superfamily	Val365Met	100	0	0
rs28399499	*CYP2B6*	chr19	T	C	Phase I	41518221	Cytochrome P450 superfamily	Leu328Thr	100	0	0
rs3745274	*CYP2B6*	chr19	G	T	Phase I	41512841	Cytochrome P450 superfamily	Gln172His	62	36	2
rs10929302	*UGT1A1*	chr2	G	A	Phase II	234665782	UDP-glucuronosyltransferase family	86	14	0	
rs4124874	*UGT1A1*	chr2	T	G	Phase II	234665659	UDP-glucuronosyltransferase family	30	49	21	
rs4148323	*UGT1A1*	chr2	G	A	Phase II	234669144	UDP-glucuronosyltransferase family	Gly71Arg	77	23	0
rs5629	*PTGIS*	chr20	G	T	Others	48129706		Arg373Arg	47	47	6
rs1051266	*SLC19A1*	chr21	T	C	Others	46957794	Solute carrier family	His27Arg	24	47	18
rs12659	*SLC19A1*	chr21	C	T	Others	46951556	Solute carrier family	Pro232Pro	21	54	24
rs4680	*COMT*	chr22	G	A	Phase II	19951271		Val158Met	71	26	3
rs16947	*CYP2D6*	chr22	A	G	Phase I	42523943	Cytochrome P450 superfamily	Arg296Cys	83	17	0
rs28371706	*CYP2D6*	chr22	G	A	Phase I	42525772	Cytochrome P450 superfamily	Thr107Ile	100	0	0
rs28371725	*CYP2D6*	chr22	A	G	Phase I	42523805	Cytochrome P450 superfamily		96	4	0
rs5030656	*CYP2D6*	chr22	–	AAG	Phase I	42524176:42524178	Cytochrome P450 superfamily		100	0	0
rs59421388	*CYP2D6*	chr22	C	T	Phase I	42523610	Cytochrome P450 superfamily	Val388Met	100	0	0
rs61736512	*CYP2D6*	chr22	C	T	Phase I	42525134	Cytochrome P450 superfamily	Val136Met	100	0	0
rs3814055	*NR1I2*	chr3	C	T	Others	119500035	Nuclear receptor subfamily		62	31	7
rs1065776	*P2RY1*	chr3	C	T	Others	152553628	G-protein coupled receptor family	Ala19Ala	88	0	0
rs701265	*P2RY1*	chr3	A	G	Others	152554357	G-protein coupled receptor family	Val262Val	76	23	1
rs2046934	*P2RY12*	chr3	G	A	Others	151057642	G-protein coupled receptor family		5	39	56
rs1805124	*SCN5A*	chr3	T	C	Others	38645420	sodium channel gene family	His558Arg	96	4	0
rs6791924	*SCN5A*	chr3	G	A	Others	38674699	sodium channel gene family	Arg34Cys	100	0	0
rs7626962	*SCN5A*	chr3	T	G	Others	38620907	sodium channel gene family	Ser1102Tyr	100	0	0
rs975833	*ADH1A*	chr4	G	C	Phase I	100201739	Alcohol dehydrogenase family		7	32	60
rs1229984	*ADH1B*	chr4	T	C	Phase I	100239319	Alcohol dehydrogenase family	His48Arg	51	37	12
rs2066702	*ADH1B*	chr4	G	A	Phase I	100229017	Alcohol dehydrogenase family	Arg370Cys	100	0	0
rs1042713	*ADRB2*	chr5	G	A	Phase I	148206440	Adrenergic receptors family	Arg16Gly	16	51	33
rs1042714	*ADRB2*	chr5	G	C	Phase I	148206473	Adrenergic receptors family	Gln27Glu	0	15	85
rs1800888	*ADRB2*	chr5	C	T	Phase I	148206885	Adrenergic receptors family	Thr164Ile	100	0	0
rs17238540	*HMGCR*	chr5	G	T	Phase I	74655498			100	0	0
rs17244841	*HMGCR*	chr5	A	T	Phase I	74642855			99	0	0
rs3846662	*HMGCR*	chr5	A	G	Phase I	74651084			19	53	28
rs1142345	*TPMT*	chr6	T	C	Phase II	18130918	Methyltransferase superfamily	Tyr240Cys	86	2	0
rs1800460	*TPMT*	chr6	A	G	Phase II	18139228	Methyltransferase superfamily	Ala154Thr	100	0	0
rs1800462	*TPMT*	chr6	C	G	Phase II	18143955	Methyltransferase superfamily	Ala80Pro	99	0	0

(*continued*)

Table 1

(continued).

SNP	Gene	Chr	Allele		Phase	Position	Family	Amino Acid Translation	Zhuang		
			A	B					AA	AB	BB
rs1045642	*ABCB1*	chr7	A	G	Others	87138645	ATP-binding cassette (ABC)	Ile1145Ile	17	50	33
rs1128503	*ABCB1*	chr7	A	G	Others	87179601	(ABC)transporters superfamily	Gly412Gly	41	49	10
rs2032582	*ABCB1*	chr7	A	C	Others	87160618	ATP-binding cassette, sub-family B	Ser893Ala	20	43	20
rs2066853	*AHR*	chr7	G	A	Others	17379110		Arg554Lys	35	46	19
rs12721634	*CYP3A4*	chr7	C	T	Phase I	99381661	Cytochrome P450 superfamily	Leu15Pro	100	0	0
rs2740574	*CYP3A4*	chr7	A	G	Phase I	99382096	Cytochrome P450 superfamily		97	3	0
rs4986909	*CYP3A4*	chr7	G	A	Phase I	99359670	Cytochrome P450 superfamily	Pro415Leu	100	0	0
rs4986910	*CYP3A4*	chr7	A	G	Phase I	99358524	Cytochrome P450 superfamily	Met444Thr	100	0	0
rs4986913	*CYP3A4*	chr7	G	A	Phase I	99358459	Cytochrome P450 superfamily	Pro466Ser	100	0	0
rs10264272	*CYP3A5*	chr7	C	T	Phase I	99262835	Cytochrome P450 superfamily	Lys208Lys	100	0	0
rs12720441	*KCNH2*	chr7	G	A	Others	150647304	Eag family	Arg784Gly	100	0	0
rs36210421	*KCNH2*	chr7	G	T	Others	150644428	Eag family	Arg1047Leu	100	0	0
rs3807375	*KCNH2*	chr7	C	T	Others	150667210	Eag family		5	30	65
rs6151031	*ALDH1A1*	chr9	–	SEQ*	Others	75568383:75568383	Aldehyde dehydrogenase family		97	2	0

* The SEQ in table represents CTGGTGAGGAGAGAACC.

† Phase I and Phase II represent that the gene is involved in drug phase I metabolisms and drug phase II metabolisms, respectively.

Table 2

Genotype frequency differences between Zhuang and 11 populations after multiple adjustment.

SNP ID	Gene	P < 0.000625										
		ASW	CEU	CHB	CHD	GIH	JPT	LWK	MEX	MKK	TSI	YRI
rs1045642	*ABCB1*	0.0007575	0.0041215	0.3341309	0.5728694	0.6843787	0.5250199	–	0.6761743	**1.65E−09**	0.6042449	**1.97E−11**
rs1128503	*ABCB1*	**5.74E−13**	**6.33E−05**	0.3649974	–	**3.02E−08**	0.3049784	**4.93E−22**	0.0037474	**4.85E−24**	**4.64E−05**	**1.21E−23**
rs2032582	*ABCB1*	**2.18E−10**	0.737558	0.2341574	–	–	0.3587152	**1.59E−21**	0.4932877	**6.57E−19**	0.4888005	–
rs975833	*ADH1A*	–	**9.89E−13**	0.3383337	–	–	0.5148173	–	–	–	–	**1.83E−12**
rs1229984	*ADH1B*	–	**3.32E−26**	0.3484655	–	–	0.3576529	–	–	–	–	**5.65E−26**
rs2066702	*ADH1B*	**1.06E−10**	–	–	**1.06E−19**	**2.44E−05**	–	**4.08E−07**	–	–	–	**5.65E−15**
rs1801253	*ADRB1*	–	0.3246496	0.915698	–	–	0.1095975	–	–	–	–	0.0148695
rs1042713	*ADRB2*	0.9036078	**2.12E−05**	0.5576991	–	–	0.018263	0.3309662	0.0766787	0.6237045	**1.10E−05**	0.0996356
rs1042714	*ADRB2*	–	**4.06E−11**	–	**2.25E−11**	**2.33E−20**	0.2766107	–	–	–	–	0.0177478
rs1800888	*ADRB2*	–	–	–	–	–	–	–	–	–	–	–
rs2066853	*AHR*	0.4069935	**6.82E−12**	0.9798667	–	–	0.6705716	0.334035	**3.40E−06**	0.5712862	**4.72E−11**	0.8615206
rs6151031	*ALDH1A1*	–	–	–	–	–	–	–	–	–	–	–
rs1800497	*ANKK1*	0.027553	0.3408199	0.0043298	**1.70E−05**	0.0150345	0.0074833	0.0539363	0.0118319	0.0438344	0.5893711	0.0026336
rs4680	*COMT*	0.0941232	**7.30E−10**	0.0101734	–	–	0.0123149	0.0023659	**0.0001355**	0.0151898	**8.25E−09**	0.0008761
rs1801272	*CYP2A6*	–	**1.80E−35**	–	**6.73E−41**	**7.70E−40**	**5.38E−32**	–	–	–	–	–
rs28399433	*CYP2A6*	–	–	–	–	–	–	–	–	–	–	–
rs28399444	*CYP2A6*	–	–	–	–	–	–	–	–	–	–	–
rs28399454	*CYP2A6*	–	–	–	–	–	–	–	–	–	–	–
rs28399499	*CYP2B6*	**0.0003479**	–	–	–	–	–	–	–	0.1678005	–	**2.04E−06**
rs3745274	*CYP2B6*	0.2028649	0.1456051	0.4024842	–	–	0.2771727	0.0208504	0.1642739	**0.0001615**	0.0938538	**7.96E−06**
rs1799853	*CYP2C19*	–	–	–	–	–	–	–	–	–	–	–
rs4986893	*CYP2C19*	–	–	–	–	–	–	–	–	–	–	–
rs16947	*CYP2D6*	–	–	–	0.0023995	**2.91E−08**	–	–	–	–	–	–
rs28371706	*CYP2D6*	–	–	–	–	–	–	–	–	–	–	–
rs28371725	*CYP2D6*	–	–	–	–	–	–	–	–	–	–	–
rs5030656	*CYP2D6*	–	–	–	**2.37E−13**	**7.20E−13**	–	–	–	–	–	–
rs59421388	*CYP2D6*	–	–	–	–	–	–	–	–	–	–	–
rs61736512	*CYP2D6*	–	–	–	–	–	–	–	–	–	–	–
rs12721634	*CYP3A4*	–	–	–	**9.80E−37**	**3.44E−16**	–	–	–	–	–	–
rs2740574	*CYP3A4*	–	–	–	–	–	–	–	–	–	–	–
rs4986909	*CYP3A4*	–	–	–	–	–	–	–	–	–	–	–
rs4986910	*CYP3A4*	–	–	–	–	–	–	–	–	–	–	–
rs4986913	*CYP3A4*	–	–	–	–	–	–	–	–	–	–	–
rs10264272	*CYP3A5*	–	–	–	**7.70E−31**	**2.01E−21**	–	**1.45E−12**	-	**8.82E−08**	-	**8.95E−09**
rs3918290	*DPYD*	–	–	–	**3.13E−18**	**1.27E−33**	–	–	–	–	–	–
rs6277	*DRD2*	–	**4.21E−16**	–	–	–	–	–	–	–	–	–
rs1138272	*GSTP1*	–	–	–	–	–	–	–	0.0006473	–	–	–
rs1695	*GSTP1*	0.0257726	0.014474	0.0281926	–	–	**1.65E−05**	**5.18E−05**	0.000767	0.2991854	0.4146075	0.0956162
rs17238540	*HMGCR*	–	–	–	–	–	–	–	–	–	–	–
rs17244841	*HMGCR*	–	–	–	–	–	–	–	–	–	–	–
rs3846662	*HMGCR*	**6.17E−08**	0.0324162	0.7978304	–	**1.57E−25**	0.9176502	**8.02E−19**	0.0124242	**2.29E−11**	0.1206333	**1.89E−20**
rs12720441	*KCNH2*	–	–	–	–	–	–	–	–	–	–	–
rs36210421	*KCNH2*	–	–	–	**1.95E−07**	**2.09E−20**	–	–	–	–	–	–
rs3807375	*KCNH2*	0.0627252	**6.64E−16**	0.2065308	**1.63E−18**	**1.60E−11**	0.2922671	0.3893726	0.0007278	0.0894274	**1.40E−15**	0.8461238
rs1801131	*MTHFR*	0.1799368	0.0943318	0.2221122	0.1376385	0.0726819	0.0998872	0.1563887	0.4424836	0.7514338	0.510967	0.0009447
rs1801133	*MTHFR*	0.0067256	0.1474965	**4.06E−05**	–	–	0.0207001	0.0007812	0.0049762	**3.86E−05**	**0.0002038**	**0.0002906**
rs1800566	*NQO1*	**5.35E−05**	**9.59E−09**	0.619526	–	–	0.060109	**2.53E−08**	0.0716608	**6.76E−11**	**5.31E−06**	**2.44E−08**
rs3814055	*NR1I2*	0.2112222	0.0338047	0.4145432	**2.32E−08**	**1.64E−13**	0.5564439	0.3437412	0.0732008	0.0964769	0.0021053	0.3219651
rs1065776	*P2RY1*	–	–	–	**5.31E−23**	**3.82E−31**	–	–	–	–	–	–

(*continued*)

Table 2

(continued).

SNP ID	Gene	P < 0.000625										
		ASW	CEU	CHB	CHD	GIH	JPT	LWK	MEX	MKK	TSI	YRI
rs701265	*P2RY1*	**2.67E−16**	0.2072429	**0.0002343**	0.2020855	**1.56E−24**	0.0027077	**1.92E−27**	0.2720617	**2.01E−31**	0.3087642	**4.56E−30**
rs2046934	*P2RY12*	–	0.8701002	0.6884447	–	-	0.5923626	–	–	–	–	0.5434161
rs5629	*PTGIS*	0.0260789	0.0278114	0.2018212	–	-	0.2516244	**6.49E−07**	0.3698856	**3.98E−06**	0.2994548	**0.0001086**
rs689466	*PTGS2*	**1.54E−09**	**2.91E−10**	0.8617374	**4.27E−05**	**2.12E−06**	0.1645723	**2.15E−19**	0.0007389	**1.67E−26**	**1.22E−08**	**6.91E−16**
rs1805124	*SCN5A*	**2.29E−09**	**1.79E−07**	–	**5.05E−25**	**1.59E−15**	–	**2.29E−11**	**0.0001002**	**1.67E−16**	**4.82E−09**	**3.96E−14**
rs6791924	*SCN5A*	–	–	–	**2.61E−22**	**1.63E−07**	–	–	–	–	–	–
rs7626962	*SCN5A*	–	–	–	**4.88E−15**	**7.05E−29**	–	–	–	–	–	0.0012409
rs1051266	*SLC19A1*	0.5813304	0.1185185	0.4865896	–	–	0.8046724	**0.000257**	0.0035212	**9.72E−07**	0.2060935	0.002182
rs12659	*SLC19A1*	–	–	–	–	–	–	–	–	–	–	–
rs4149056	*SLCO1B1*	0.3113879	0.0996441	0.1659976	**3.26E−07**	**1.81E−10**	0.7390691	0.0087002	0.7772447	0.6642217	0.0022041	**0.0005021**
rs1801030	*SULT1A1*	–	–	–	**5.98E−36**	**3.57E−30**	–	–	–	–	–	–
rs3760091	*SULT1A1*	–	–	–	**9.44E−13**	0.4179507	–	–	–	–	–	–
rs1142345	*TPMT*	–	–	–	**4.96E−15**	**5.79E−18**	–	**0.0004355**	0.0243167	–	–	0.1973033
rs1800460	*TPMT*	–	–	–	–	–	–	–	–	–	–	–
rs1800462	*TPMT*	–	–	–	**1.86E−22**	**9.16E−13**	–	–	–	–	–	–
rs151264360	*TYMS*	–	–	–	–	–	–	–	–	–	–	–
rs10929302	*UGT1A1*	–	**6.29E−06**	–	**1.71E−08**	**7.70E−13**	–	–	–	–	–	**1.66E−08**
rs4124874	*UGT1A1*	**1.30E−06**	0.9831273	0.0015632	**1.19E−07**	**4.79E−19**	0.0569758	**1.52E−15**	0.4675602	**3.70E−15**	0.7885351	**5.80E−20**
rs4148323	*UGT1A1*	–	–	0.0117568	0.2798952	–	0.5489341	–	–	–	–	–
rs10735810	*VDR*	**6.16E−08**	0.030309	0.0373621	–	–	**9.36E−05**	**2.35E−12**	0.9064841	**2.39E−12**	0.0091999	**2.05E−11**
rs11568820	*VDR*	**6.08E−05**	**1.81E−05**	0.9231801	–	–	0.8421516	**5.17E−14**	0.0008767	**4.95E−11**	0.0011594	**2.10E−25**
rs1540339	*VDR*	**1.05E−09**	**1.55E−07**	0.4169708	**1.71E−26**	**3.68E−05**	0.1649253	**2.63E−19**	**0.000146**	**2.83E−20**	**2.62E−07**	**1.65E−16**
rs1544410	*VDR*	–	**2.91E−20**	–	–	–	–	**7.69E−11**	**1.62E−09**	**4.94E−18**	**6.86E−19**	**7.10E−13**
rs2228570	*VDR*	–	–	–	**1.33E−08**	0.7086021	–	–	–	–	–	–
rs2239179	*VDR*	–	–	–	–	–	–	–	–	–	–	–
rs2239185	*VDR*	–	–	0.0267499	–	–	0.0516003	–	–	–	–	**9.89E−07**
rs731236	*VDR*	**1.07E−06**	**4.19E−17**	–	–	–	–	**3.14E−08**	**2.56E−07**	**5.85E−22**	**1.01E−15**	**6.12E−11**
rs7975232	*VDR*	**7.57E−09**	**1.16E−08**	0.0687814	–	–	0.1440584	**7.63E−16**	0.008202	**3.75E−15**	**5.62E−09**	**2.76E−11**
rs7294	*VKORC1*	**3.05E−10**	**2.50E−06**	0.0126767	0.0026499	0.0013092	0.3353498	**1.78E−09**	0.0006995	**5.52E−13**	**3.87E−05**	**1.28E−13**
rs9923231	*VKORC1*	–	–	–	**1.11E−40**	**1.33E−10**	–	–	–	–	–	–
rs9934438	*VKORC1*	**1.12E−24**	**3.90E−16**	0.0111364	–	–	0.2333683	**1.49E−33**	**4.69E−10**	**5.05E−35**	**4.91E−12**	**1.31E−41**
Total number of significant variants	17	21	2	26	27	3	22	7	22	15	27	

The significant variants (after multiple adjustment) are in bold.

ASW = African ancestry in Southwest USA; CEU = Utah, USA residents with Northern and Western European ancestry from the CEPH collection; CHB = Han Chinese in Beijing; CHD = Chinese in metropolitan Denver; GIH = Gujarati Indians in Houston; JPT = Japanese in Tokyo; LWK = Luhya in Webuye, Kenya; MEX = Mexican ancestry in Los Angeles, California, USA; MKK = Maasai in Kinyawa, Kenya; TSI = Toscani in Italy; YRI = Yoruba in Ibadan, Nigeria.

significant differences were as follows: ASW, 17; CEU, 21; CHB, 2; CHD, 26; GIH, 27; JPT, 3; LWK, 22; MEX, 7; MKK, 22; TSI, 15; YRI, 27. Obviously, the frequency distribution of the Zhuang population in Yunnan was similar to that of CHB, followed by JPT and MEX. In these different loci, rs7294 and rs689466 (located in *VKORC1* and *PTGS2*, respectively) were significantly different in the Zhuang population compared to in other populations.

Among the 80 variants listed in Table 1, 67 variants could be classified as specific families. Based on the classification, the number of differing loci changed slightly as follows (after adjustment): ASW, 12; CEU, 15; CHB, 2; CHD, 23; GIH, 23; JPT, 3; LWK, 16; MEX, 4; MKK, 16; TSI, 9; YRI, 21 (Table 3).

To further verify the ubiquitous differences between different groups, we downloaded the data pertaining to rs7294 and rs689466 from ALFRED (https://alfred.med.yale.edu/) and performed a global frequency analysis using the new data.

Figure 1A shows the global allele frequencies of rs7294. We observed that the A allele frequencies in the East Asian populations were lower than those in most of the other populations of the world. The frequencies of the Zhuang people (0.145) were relatively closer to those of the East Asian populations, especially the southern Chinese populations such as Yi (0.15) and Lahu (0.15). In addition, the Zhuang people showed frequency similar to those of other East Asian populations for rs689466 (Fig. 1B).

Pairwise FST values were used to estimate and assess the magnitude of differentiation among geographic populations (0 indicating no divergence, 1 indicating complete separation). As shown in Table 4, the F_{st} values of the Zhuang and CHD population were the smallest ($F_{ST}=0.00884$), followed by those of CHB ($F_{ST}=0.01701$), and JPT ($F_{ST}=0.02057$), indicating that the allele frequencies of the Zhuang and these 3 populations are similar. In addition, the highest divergence was observed for the MKK ($F_{ST}=0.21627$) population.

Combining these results with our previous data on 11 Chinese ethnicities and 11 populations from the International HapMap project, a genetic structure was derived using STRUCTURE 2.3.1 (Fig. 2A). Individuals were divided into K clusters to display the genetic components. At $K=6$, population structure was almost in accordance with the major geographic regions, and populations could be divided into 6 subgroups (subgroup 1: Zhuang, Shaanxi Han, CHB, CHD, and JPT; subgroup 2: ASW, LWK, MKK, and YRI; subgroup 3: Kyrgyz, Tajik, Mongol, and Uygur; subgroup 4: CEU, GIH, MEX, and TSI; subgroup 5: Deng, Lohoba, and Sherpa; subgroup 6: Miao, Li, and Tibetan) according to the clusters in each population. In addition, we inferred that Zhuang is most closely related to Shaanxi Han, followed by 3 other East Asian populations (CHD, CHB, and JPT) (Fig. 2A).

Since individuals from the same population show similar ancestry proportions, statistical evaluation of the genetic relationships among populations is important. Therefore, we constructed a phylogenetic tree using the neighbor-joining method (Fig. 2B). Results indicate that ASW are located near the root of the tree, and JPT shares the closest evolutionary relationship with Zhuang.

Table 3

Significant variants in Zhuang compared to the other 11 populations after classification.

SNP ID	Gene	P < 0.000625										
		ASW	CEU	CHB	CHD	GIH	JPT	LWK	MEX	MKK	TSI	YRI
rs1045642	*ABCB1*	0.0007575	0.0041215	0.3341309	0.5728694	0.6843787	0.5250199	–*	0.6761743	**1.65E−09**†	0.6042449	**1.97E−11**
rs1128503	*ABCB1*	**5.74E−13**	**6.33E−05**	0.3649974	–	**3.02E−08**	0.3049784	**4.93E−22**	0.0037474	**4.86E−24**	**4.64E−05**	**1.21E−23**
rs2032582	*ABCB1*	**2.18E−10**	0.737558	0.2341574	–	–	0.3587152	**1.59E−21**	0.4932877	**6.57E−19**	0.4888005	–
rs975833	*ADH1A*	–	**9.89E−13**	0.3383337	–	–	0.5148173	–	–	–	–	**1.83E−12**
rs1229984	*ADH1B*	–	**3.32E−26**	0.3484655	–	–	0.3576529	–	–	–	–	**5.65E−26**
rs2066702	*ADH1B*	**1.07E−10**	–	–	**1.06E−19**	**2.44E−05**	–	**4.08E−07**	–	–	–	**5.65E−15**
rs1042713	*ADRB2*	0.9036078	**2.12E−05**	0.5576991	–	–	0.018263	0.3309662	0.0766787	0.6237045	**1.10E−05**	0.0996356
rs1042714	*ADRB2*	–	**4.06E−11**	–	**2.25E−11**	**2.34E−20**	0.2766107	–	–	–	–	0.0177478
rs1800497	*ANKK1*	0.027553	0.3408199	0.0043298	**1.70E−05**	0.0150345	0.0074833	0.0539363	0.0118319	0.0438344	0.5893711	0.0026336
rs1801272	*CYP2A6*	–	**1.81E−35**	–	**6.73E−41**	**7.70E−40**	**5.38E−32**	–	–	–	–	–
rs28399499	*CYP2B6*	**0.0003479**	–	–	–	–	–	–	–	0.1678005	–	**2.04E−06**
rs3745274	*CYP2B6*	0.2028649	0.1456051	0.4024842	–	–	0.2771727	0.0208504	0.1642739	**0.0001615**	0.0938538	**7.96E−06**
rs16947	*CYP2D6*	–	–	–	0.0023995	**2.91E−08**	–	–	–	–	–	–
rs5030656	*CYP2D6*	–	–	–	**2.37E−13**	**7.20E−13**	–	–	–	–	–	–
rs12721634	*CYP3A4*	–	–	–	**9.80E−37**	**3.44E−16**	–	–	–	–	–	–
rs10264272	*CYP3A5*	–	–	–	**7.70E−31**	**2.01E−21**	–	**1.45E−12**	–	**8.82E−08**	–	**8.95E−09**
rs6277	*DRD2*	–	**4.21E−16**	–	–	–	–	–	–	–	–	-
rs1695	*GSTP1*	0.0257726	0.014474	0.0281926	–	–	**1.65E−05**	**5.18E−05**	0.000767	0.2991854	0.4146075	0.0956162
rs36210421	*KCNH2*	–	–	–	**1.95E−07**	**2.09E−20**	–	–	–	–	–	–
rs3807375	*KCNH2*	0.0627252	**6.64E−16**	0.2065308	**1.63E−18**	**1.60E−11**	0.2922671	0.3893726	0.0007278	0.0894274	**1.40E−15**	0.8461238
rs1801133	*MTHFR*	0.0067256	0.1474965	**4.06E−05**	–	–	0.0207001	0.0007812	0.0049762	**3.86E−05**	**0.0002038**	**0.0002906**
rs3814055	*NR1I2*	0.2112222	0.0338047	0.4145432	**2.32E−08**	**1.64E−13**	0.5564439	0.3437412	0.0732008	0.0964769	0.0021053	0.3219651
rs1065776	*P2RY1*	–	–	–	**5.31E−23**	**3.82E−31**	–	–	–	–	–	–
rs701265	*P2RY1*	**2.67E−16**	0.2072429	**0.0002343**	0.2020855	**1.56E−24**	0.0027077	**1.92E−27**	0.2720617	**2.01E−31**	0.3087642	**4.56E−30**
rs1805124	*SCN5A*	**2.29E−09**	**1.79E−07**	–	**5.05E−25**	**1.59E−15**	–	**2.29E−11**	**0.0001002**	**1.67E−16**	**4.82E−09**	**3.96E−14**
rs6791924	*SCN5A*	–	–	–	**2.61E−22**	**1.63E−07**	–	–	–	–	–	–
rs7626962	*SCN5A*	–	–	–	**4.88E−15**	**7.06E−29**	–	–	–	–	–	0.0012409
rs1051266	*SLC19A1*	0.5813304	0.1185185	0.4865896	–	–	0.8046724	**0.000257**	0.0035212	**9.72E−07**	0.2060935	0.002182
rs4149056	*SLCO1B1*	0.3113879	0.0996441	0.1659976	**3.26E−07**	**1.81E−10**	0.7390691	0.0087002	0.7772447	0.6642217	0.0022041	**0.0005021**
rs1801030	*SULT1A1*	–	–	–	**5.98E−36**	**3.57E−30**	–	–	–	–	–	–
rs3760091	*SULT1A1*	–	–	–	**9.44E−13**	0.4179507	–	–	–	–	–	–
rs1142345	*TPMT*	–	–	–	**4.96E−15**	**5.79E−18**	–	**0.0004355**	0.0243167	–	–	0.1973033
rs1800462	*TPMT*	–	–	–	**1.86E−22**	**9.16E−13**	–	–	–	–	–	–
rs10929302	*UGT1A1*	–	**6.29E−06**	–	**1.71E−08**	**7.70E−13**	–	–	–	–	–	**1.67E−08**
rs4124874	*UGT1A1*	**1.30E−06**	0.9831273	0.0015632	**1.19E−07**	**4.79E−19**	0.0569758	**1.52E−15**	0.4675602	**3.70E−15**	0.7885351	**5.80E−20**
rs10735810	*VDR*	**6.16E−08**	0.030309	0.0373621	–	–	**9.36E−05**	**2.35E−12**	0.9064841	**2.39E−12**	0.0091999	**2.05E−11**
rs11568820	*VDR*	**6.08E−05**	**1.81E−05**	0.9231801	–	–	0.8421516	**5.17E−14**	0.0008767	**4.95E−11**	0.0011594	**2.10E−25**
rs1540339	*VDR*	**1.05E−09**	**1.55E−07**	0.4169708	**1.71E−26**	**3.68E−05**	0.1649253	**2.63E−19**	**0.000146**	**2.83E−20**	**2.63E−07**	**1.65E−16**
rs1544410	*VDR*	–	**2.91E−20**	–	–	–	–	**7.69E−11**	**1.62E−09**	**4.94E−18**	**6.86E−19**	**7.10E−13**
rs2228570	*VDR*	–	–	–	**1.33E−08**	0.7086021	–	–	–	–	–	–
rs2239185	*VDR*	–	–	0.0267499	–	–	0.0516003	–	–	–	–	**9.89E−07**
rs731236	*VDR*	**1.07E−06**	**4.19E−17**	–	–	–	–	**3.14E−08**	**2.56E−07**	**5.85E−22**	**1.01E−15**	**6.12E−11**
rs7975232	*VDR*	**7.57E−09**	**1.16E−08**	0.0687814	–	–	0.1440584	**7.63E−16**	0.008202	**3.75E−15**	**5.62E−09**	**2.76E−11**
Total number	12	15	2	23	23	3	16	4	16	9	21	

* Results of calculation is meaningless.

† The significant variants (after multiple adjustment) are in a bold.

ASW = African ancestry in Southwest USA, CEU = Utah, USA residents with Northern and Western European ancestry from the CEPH collection, CHB = Han Chinese in Beijing, CHD = Chinese in metropolitan Denver, GIH = Gujarati Indians in Houston, JPT = Japanese in Tokyo, LWK = Luhya in Webuye, Kenya, MEX = Mexican ancestry in Los Angeles, California, USA, MKK = Maasai in Kinyawa, Kenya, TSI = Toscani in Italy, YRI = Yoruba in Ibadan, Nigeria.

4. Discussion

As early as 2002, a World Health Organization (WHO) report showed that more than half of the medicines were prescribed, dispensed, or sold inappropriately, while 50% of the patients did not consume them correctly. For safety issues, approximately 150 prescription drugs have been removed from the market since 1960, and certain new expensive drugs do not appear in the market.[25] This is largely because of the individual differences in toxicity and response caused by pharmacogenomic variants. For providing more information on VIP variants of different ethnicities, we selected and genotyped 80 VIP variants in 100 Zhuang people from the Yunnan province of China, and conducted a series of statistical analyses. According to F_{st} values and STRUCTURE analysis, we speculated that Zhuang is most closely related to the Shaanxi Han population.

Comparison of genotype frequency distribution obtained from χ^2 tests showed that rs7294, a SNP located in *VKORC1* on chromosome 16, differs significantly among populations. *VKORC1*, a gene encoding vitamin K epoxide reductase complex 1, is the target of warfarin,[26] a widely used anticoagulant prescribed for chronic atrial fibrillation, mechanical valves, pulmonary embolism, and dilated cardiomyopathy.[27] To achieve the same anticoagulant effect (normally defined by the international normalized ratio, INR), individuals harboring polymorphisms in dosage-related genes require lower or higher warfarin doses than people carrying the wild type genes.[28] For example, an investigation involving 279 patients of European ancestry on warfarin medication indicated that individuals with the TT genotype of rs7294 required 53% higher dose than individuals with the CC genotype.[29] Similarly, reports show that in Chinese populations, patients with allele T of rs7294 require higher plasma concentration to achieve similar INR.[30] Our previous study showed that the allele frequency of rs7294-T in Uygur from northwestern China and the Han ethnic group of Shaanxi province is lower than that in most of the other populations in the world. The same trend was observed in Zhuang people, suggesting that most Zhuang individuals may require lower dosage of warfarin.[10,14]

In addition to rs7294, the genotype frequency of rs689466 in the Zhuang population also shows notable differences with other

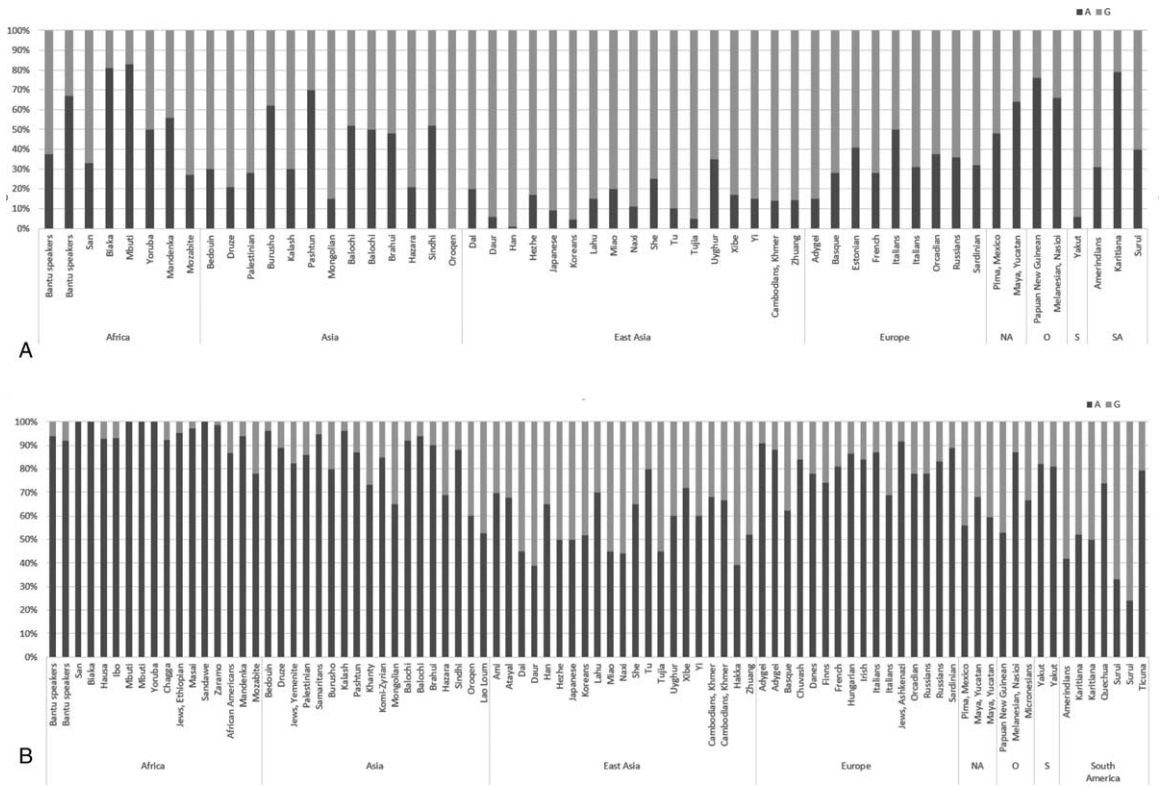

Figure 1. (A) Global allele frequency of rs7294. NA is short for North America; O is for Oceanic; S is for Siberia; SA is for South America. (B) Global allele frequency of rs689466.

populations. Rs689466 or COX-2 -1195G>A, is a SNP located in *COX-2/PTGS2* (chromosome 1q25.2-3). COX-2 (cyclooxygenase-2), also known as prostaglandin-endoperoxide synthase 2, is composed of 10 exons and is a key enzyme that converts arachidonic acid to prostaglandins, which is involved in several important biological processes such as inflammation, immune function, cell proliferation, and angiogenesis.[31,32] Evidence show that COX-2 is over-expressed in tumor tissue, whereas it is rarely detected in normal tissue.[33] According to a meta-analysis involving 50,672 subjects in 2015, the A allele of rs689466 was associated with higher risk of cancer; in addition, a subgroup analysis by ethnicity showed it to be associated with high cancer risk in Asians.[32] In contrast, the G allele is associated with high blood pressure in Asian individuals such as Koreans and Japanese.[34,35] Consistent with the results obtained with the East Asian population, our data showed that the rs689466 A allele frequency of Zhuang (0.52) is relatively lower than those of other populations, especially European and African populations, indicating that Zhuang may have lower risk of cancer. Nevertheless, owing to their similarity with the Tibetan people, the G allele frequency of Zhuang is relatively higher, highlighting that Zhuang people should be cautious about the risk of

Table 4

Pairwise F_{st} distances among the 12 populations.

	Zhuang	CHB	CHD	JPT	CEU	GIH	MEX	TSI	ASW	LWK	MKK	YRI
Zhuang	0											
CHB	0.01701	0										
CHD	0.00884	−0.00161	0									
JPT	0.02057	0.00586	0.00761	0								
CEU	0.12232	0.13026	0.12708	0.11499	0							
GIH	0.14187	0.15697	0.15321	0.14338	0.03311	0						
MEX	0.07546	0.08424	0.07821	0.08033	0.02248	0.05258	0					
TSI	0.11758	0.11524	0.11626	0.10172	0.00012	0.04047	0.02447	0				
ASW	0.17776	0.1955	0.19394	0.17125	0.12124	0.08173	0.11144	0.12461	0			
LWK	0.24885	0.26654	0.26764	0.23703	0.18539	0.14618	0.18563	0.19061	0.01719	0		
MKK	0.21627	0.23189	0.23406	0.19985	0.13638	0.10553	0.15181	0.14253	0.01888	0.01336	0	
YRI	0.25329	0.26827	0.27045	0.23703	0.19138	0.14351	0.19235	0.1978	0.01513	0.00383	0.01359	0

F_{st} values were displayed by color schemes with dark gray for large F_{st} (0.2–0.3), neutral gray (0.1–0.2) for intermediate F_{st}, light gray (0–0.1) for small F_{st}. ASW = African ancestry in Southwest USA, CEU = Utah, USA residents with Northern and Western European ancestry from the CEPH collection, CHB = Han Chinese in Beijing, CHD = Chinese in metropolitan Denver, GIH = Gujarati Indians in Houston, JPT = Japanese in Tokyo, LWK = Luhya in Webuye, Kenya, MEX = Mexican ancestry in Los Angeles, California, USA, MKK = Maasai in Kinyawa, Kenya, TSI = Toscani in Italy, YRI = Yoruba in Ibadan, Nigeria.

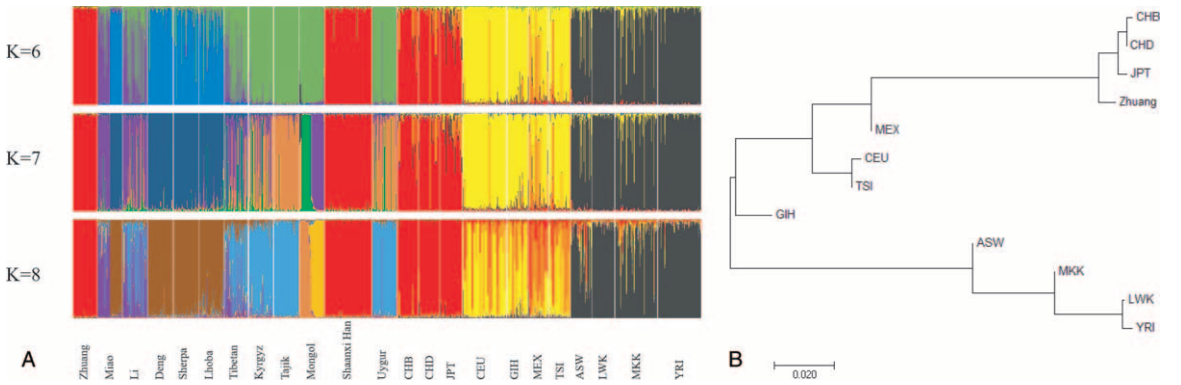

Figure 2. (A) Structure analysis of 23 populations. Each individual is represented by a vertical line which was partitioned into colored segments. *K* is the number of estimated clusters. ASW = African ancestry in Southwest USA; CEU = Utah, USA residents with Northern and Western European ancestry from the CEPH collection; CHB = Han Chinese in Beijing; CHD = Chinese in metropolitan Denver; GIH = Gujarati Indians in Houston; JPT = Japanese in Tokyo; LWK = Luhya in Webuye, Kenya; MEX = Mexican ancestry in Los Angeles, California, USA; MKK = Maasai in Kinyawa, Kenya; TSI = Toscani in Italy; YRI = Yoruba in Ibadan, Nigeria. Among them, CHB, CHD And JPT come from Asia; CEU, GIH, MEX and ASW come from America; TSI come from Europe; LWK, MKK and YRI come from Africa. (B) Evolutionary relationships of populations. The evolutionary history was inferred using the neighbor-joining method. The optimal tree with the sum of branch length = 0.26960109 is shown. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. Evolutionary analyses were conducted in MEGA7.

developing hypertension, especially that caused by side effects of non-steroidal anti-inflammatory drugs (NSAIDs).[11]

Populations are usually defined based on culture or geography but not genetic relationships. Therefore, analysis of population structure may assist in investigating human evolutionary history.[36,37] Comparison with 11 Chinese ethnicities showed that the genetic background of Zhuang closely resembles that of the Shaanxi Han population. In addition, we observed that some populations such as Miao have partial membership in multiple clusters, which may be caused by continuous gradations in allele frequencies across regions or admixture of neighboring groups.[36]

The Clinical Pharmacogenetics Implementation Consortium (CPIC) is a shared project of PharmGKB and the Pharmacogenomics Research Network (PGRN),[38] which was established in 2009 for reducing the barrier in translating genetic laboratory test results into actionable prescribing decisions for affected drugs. Currently, CPIC guidelines are being developed and updated using established methods, including a rigorous review and grading of the relevant scientific literature.[39] Our study may provide information for developing CPIC guidelines.

5. Conclusions

We identified the features of 80 VIP variants of Zhuang people from southwestern China, and observed that the genetic background of the Zhuang population of Yunnan is closest to that of the Shaanxi Han population. This study supplements the existing knowledge regarding different pharmacogenomic variants and may provide guidance for developing personalized medicine in future. However, our study has certain limitations. Our sample size was relatively small, and further studies with larger groups are necessary to verify the accuracy of our study.

Acknowledgment

We are grateful to the patients and control subjects for their participation in this study. We also thank the clinicians and hospital staff who obtained the blood samples and collected data for this study.

6. Author contributions

Formal analysis: Jing Li, Tianbo Jin.
Investigation: Jing Li, Mengdan Yan, Fanglin Niu.
Data curation: Chenghao Guo.
Writing – original draft: Chenghao Guo.
Writing – review & editing: Chenghao Guo.
Project administration: Mengdan Yan.
Conceptualization: Peng Chen, Bin Li, Tianbo Jin.
Supervision: Tianbo Jin.

References

[1] Carrasco-Garrido P, de Andres LA, Barrera VH, et al. Trends of adverse drug reactions related-hospitalizations in Spain (2001–2006). BMC Health Serv Res 2010;10:287.
[2] Relling MV, Evans WE. Pharmacogenomics in the clinic. Nature 2015;526:343–50.
[3] Shan B, Cai JH, Yang SY, et al. Association of DENND1A gene polymorphisms with polycystic ovary syndrome: a meta-analysis. J Clin Res Pediatr Endocrinol 2015;8:135–43.
[4] Thorn CF, Klein TE, Altman RB. PharmGKB: the pharmacogenetics and pharmacogenomics knowledge base. Methods Mol Biol 2005;311:179–91.
[5] Halushka MK, Fan JB, Bentley K, et al. Patterns of single-nucleotide polymorphisms in candidate genes for blood-pressure homeostasis. Nat Genet 1999;22:239–47.
[6] Lamba JK, Lin YS, Thummel K, et al. Common allelic variants of cytochrome P4503A4 and their prevalence in different populations. Pharmacogenetics 2002;12:121–32.
[7] Ding Y, He P, He N, et al. Genetic polymorphisms of pharmacogenomic VIP variants in Li nationality of southern China. Environ Toxicol Pharmacol 2016;42:237–42.
[8] Jin T, Aikemu A, Zhang M, et al. Genetic polymorphisms analysis of pharmacogenomic VIP variants in Miao Ethnic Group of southwest China. Med Sci Monit 2015;21:3769–76.
[9] Jin T, Shi X, Wang L, et al. Genetic polymorphisms of pharmacogenomic VIP variants in the Mongol of Northwestern China. BMC Genet 2016;17:70.
[10] Jin T, Zhao R, Shi X, et al. Genetic polymorphisms study of pharmacogenomic VIP variants in Han ethnic of China's Shaanxi province. Environ Toxicol Pharmacol 2016;46:27–35.
[11] Jin TB, Xun XJ, Shi XG, et al. Genetic polymorphisms in very important pharmacogenomic (VIP) variants in the Tibetan population. Genet Mol Res: GMR 2015;14:12497–504.
[12] He Y, Yang H, Geng T, et al. Genetic polymorphisms of pharmacogenomic VIP variants in the Ihoba population of southwest China. Int J Clin Exp Pathol 2015;8:13293.

[13] Shi X, Wang L, Du S, et al. Genetic polymorphism of pharmacogenomic VIP variants in the Deng people from the Himalayas in Southeast Tibet. Biomarkers 2015;20:275–86.
[14] Wang L, Aikemu A, Yibulayin A, et al. Genetic polymorphisms of pharmacogenomic VIP variants in the Uygur population from northwestern China. BMC Genet 2015;16:66.
[15] Wang L, Ren Y, Shi X, et al. The population genetics of pharmacogenomics VIP variants in the Sherpa population. Drug Metab Pharmacokinet 2016;31:82–9.
[16] Yunus Z, Liu L, Wang H, et al. Genetic polymorphisms of pharmacogenomic VIP variants in the Kyrgyz population from northwest China. Gene 2013;529:88–93.
[17] Zhang J, Jin T, Yunus Z, et al. Genetic polymorphisms of VIP variants in the Tajik ethnic group of northwest China. BMC Genet 2014;15:102.
[18] Gabriel S, Ziaugra L, Tabbaa D. SNP genotyping using the Sequenom MassARRAY iPLEX platform. Curr Protoc Hum Genet 2009;2.12.11–12.12.16.
[19] Thomas RK, Baker AC, Debiasi RM, et al. High-throughput oncogene mutation profiling in human cancer. Nat Genet 2007;39:347–51.
[20] Song M, Lin F, Ward S, et al. Composite variables: when and how. Nurs Res 2013;62:
[21] Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics 2000;155:945–59.
[22] Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 2005;14:2611.
[23] Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 2016;33:1870.
[24] Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 1987;4:406.
[25] Zhang W, Roederer MW, Chen WQ, et al. Pharmacogenetics of drugs withdrawn from the market. Pharmacogenomics 2012;13:223–31.
[26] Rieder MJ, Reiner AP, Gage BF, et al. Effect of VKORC1 haplotypes on transcriptional regulation and warfarin dose. N Engl J Med 2005;352:2285–93.
[27] Piatkov I, Rochester C, Jones T, et al. Warfarin toxicity and individual variability-clinical case. Toxins 2010;2:2584–92.
[28] Mazzaccara C, Conti V, Liguori R, et al. Warfarin anticoagulant therapy: a southern Italy pharmacogenetics-based dosing model. PLoS One 2013;8:e71505.
[29] Botton MR, Bandinelli E, Rohde LE, et al. Influence of genetic, biological and pharmacological factors on warfarin dose in a Southern Brazilian population of European ancestry. Br J Clin Pharmacol 2011;72:442–50.
[30] Li S, Zou Y, Wang X, et al. Warfarin dosage response related pharmacogenetics in Chinese population. PLoS One 2015;10:e0116463.
[31] Zhao D, Xu D, Zhang X, et al. Interaction of cyclooxygenase-2 variants and smoking in pancreatic cancer: a possible role of nucleophosmin. Gastroenterology 2009;136:1659–68.
[32] Wang Y, Jiang H, Liu T, et al. Cyclooxygenase-2-1195G>A (rs689466) polymorphism and cancer susceptibility: an updated meta-analysis involving 50,672 subjects. Int J Clin Exp Med 2015;8:12448–62.
[33] Eberhart CE, Coffey RJ, Radhika A, et al. Up-regulation of cyclooxygenase 2 gene expression in human colorectal adenomas and adenocarcinomas. Gastroenterology 1994;107:1183–8.
[34] Iwai N, Tago N, Yasui N, et al. Genetic analysis of 22 candidate genes for hypertension in the Japanese population. J Hypertens 2004;22:1119–26.
[35] Jin HS, Hong KW, Lim JE, et al. Association between prostaglandin-endoperoxide synthase 2 (PTGS2) polymorphisms and blood pressure in Korean population. Genomics Inform 2008;6:110–6.
[36] Rosenberg NA, Pritchard JK, Weber JL, et al. Genetic structure of human populations. Science 2002;298:2381–5.
[37] Foster MW, Sharp RR. Race, ethnicity, and genomics: social classifications as proxies of biological heterogeneity. Genome Res 2002;12:844.
[38] Relling MV, Klein TE. CPIC: clinical pharmacogenetics implementation consortium of the pharmacogenomics research network. Clin Pharmacol Ther 2011;89:464.
[39] Caudle KE, Klein TE, Hoffman JM, et al. Incorporation of pharmacogenomics into routine clinical practice: the Clinical Pharmacogenetics Implementation Consortium (CPIC) guideline development process. Curr Drug Metab 2014;15:209–17.