Iskakova et al., J Pharmacogenomics Pharmacoproteomics 2014, 5:2
DOI: 10.4172/2153-0645.1000129

Journal of Pharmacogenomics & Pharmacoproteomics

Research Article Open Access

Allele Frequency and Genotype Distribution of 9 SNPs in the Kazakh Population

Aisha N Iskakova[1,2*], Aliya A Romanova[1,2], Elena N Voronina[3], Nurgul S Sikhayeva[1], Liliya A Belozerceva[3], Maksim L Filipenko[3] and Erlan M Ramanculov[1]

[1]National Centre for Biotechnology, Astana, Kazakhstan
[2]General Genetics LLP, Astana, Kazakhstan
[3]Institute of Chemical Biology and Fundamental Medicine, Siberian Branch of the Russian Academy of Sciences (RAS ICBFM), Novosibirsk, Russia

Abstract

Background: Determining the allelic variants of xenobiotic biotransformation genes is important, especially for prescribing personalised drugs. Knowledge of the allele distribution in different populations may be considered when selecting the preferred medication regimen. The frequency of *CYP2C9*, *VKORC1*, *CYP4F2*, *GGCX*, *CYP2D6* and *CYP1A2* genes has been studied in many populations, but the populations in Central Asia have not yet been investigated.

Methods and materials: Using real-time PCR and direct sequencing-based methods, the current study assessed the frequencies of 9 polymorphisms of genes encoding enzymes involved in drug metabolism in 450 healthy individuals from different regions of Kazakhstan and 575 healthy individuals from the West-Siberian region of Russia.

Results: The allele frequencies in the Kazakh population were determined for *CYP2C9*2* (0.02), *CYP2C9*3* (0.03), *VKORC1* c. 173+1369G>C, *VKORC1* c. 173+1000C>T (0.72, *CYP4F2* (0.31), *GGCX* (0.04), *CYP2D6*4* (0.07), *CYP2D6*3* (0.01) and CYP1A2**1F* (0.35). All alleles were in Hardy–Weinberg equilibrium ($p > 0.05$).

The allele frequencies in the Russian population were as follows: *CYP2C9*2*, 0.08; *CYP2C9*3*, 0.08; *VKORC1* (c. 173+1000C>T), 0.40; *VKORC1* (c. 173+1369G>C), 0.41; *CYP4F2* (c. 1297G>A), 0.24; *GGCX* (c. 1913+45G>C), 0.08; *CYP2D6*3*, 0.15; *CYP2D6*4*, 0.22; and *CYP1A2*1F* (c. -9-154C>A), 0.31. All alleles were in Hardy–Weinberg equilibrium ($p>0.05$), except *GGCX* ($p=0.04$).

Conclusion: The Kazakh population allele frequency was between the Caucasian and Asian populations for nearly all of the studied gene allele variants.

Keywords: *CYP2*; *VKORC1*; *GGCX*; Allele frequency; Kazakh; Russian

Introduction

The use of pharmacological agents in the treatment of diseases is becoming increasingly important in modern medicine. However, considerable variability in the efficacy and side effects of medications between different population groups and different individuals is well known from medical practice. The risk of non-standard reactions has increased as result of the production of new medicines and the increasing number of medications taken by patients. Almost 50% of cases of adverse drug reactions or lack of therapy efficacy is caused by the genetic characteristics of a patient [1].

The action of drugs in the human body is due to processes such as absorption, distribution (by organs, cells, and organelles), interaction with receptors, metabolism and excretion.

All stages of the pharmacokinetic process are carried out by means of specific and non-specific enzymes, and xenobiotic biotransformation proteins play a major role in these processes. Mutations in the genes encoding these enzymes may lead to a decrease or increase of their enzymatic activity, which affects the metabolic rate of the corresponding drugs.

Direct (phenotypic) determination of the concentration change rate of a drug or its metabolites in the blood is a time-consuming and complicated procedure, in addition to requiring the patient to take the relevant medication. Determining allelic variants (genotype) of biotransformation genes, however, allows for the identification of patients who are more likely to develop adverse drug reactions.

CYP2C9 gene

The *CYP2C9* gene has high genetic polymorphism. The structural gene polymorphisms *CYP2C9* - R144C (*CYP2C9*2*) and I359L (*CYP2C9*3*) are the most well-studied. The allele *CYP2C9*2* reduces the warfarin dose by 40% for heterozygotes and 68% by homozygotes, and the *CYP2C9*3* allele by 40 and 85%, respectively. This polymorphic variant explained 14.5% of the interindividual variability in the dose of warfarin together with non-genetic predictors, such as age, body weight and height [2].

VKORC1 gene

The *VKORC1* gene on chromosome 16 is one of the major genes associated with effective doses of coumarin anticoagulants. Many mutations are associated with a deficiency in this enzyme. The allelic variant *VKORC1* c.-1639G>A determines up to 30% of the variability in warfarin dosage [3-8].

***Corresponding author:** Aisha N Iskakova, National Centre for Biotechnology, Valikhanov St 13/1, Astana, 010000, Kazakhstan, Tel: 007 701 9521056; E-mail: aishaisk1@gmail.com

Received March 17, 2014; **Accepted** April 22, 2014; **Published** April 30, 2014

Citation: Iskakova AN, Romanova Voronina AA, Voronina EN, Sikhayeva NS, Belozerceva AB, et al. (2014) Allele Frequency and Genotype Distribution of 9 SNPs in the Kazakh Population. J Pharmacogenomics Pharmacoproteomics 5: 129. doi:10.4172/2153-0645.1000129

Copyright: © 2014 Iskakova AN, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

CYP4F2 gene

The *CYP4F2* enzyme of the cytochrome P450 superfamily is responsible for the synthesis of hydroxyeicosatetraenoic acid (20-HETE), which plays an important role in blood vessel function. Therefore, the production efficiency of 20-HETE may cause *CYP4F2* to be associated with the warfarin dose. The *CYP4F2* c. 1297G>A mutation, which leads to the V433M amino acid substitution, affects the activity of the enzyme, reducing 20-HETE synthesis to 60% [9]. In this case, the warfarin dosage should be increased.

GGCX gene

GGCX is a key factor in the regulation of blood clotting. Several mutations in the *GGCX* gene associated with a decrease in the clotting factor activity have been found. One of these mutations is rs11676382 c. 1913+45G>C, which was detected in intron 14. A study of 186 Caucasians who took thrombosis therapy using warfarin showed that a polymorphic locus accounts for approximately 2% of the variation in drug dosage [10].

CYP2D6 gene

The *CYP2D6* enzyme is encoded by a gene on chromosome 22. Approximately 25% of all drugs are substrates of *CYP2D6*, including beta-blockers, tricyclic antidepressants, neuroleptics, derivatives of morphine and other drugs. At least 40 genetic variants of *CYP2D6* can result in the reduced metabolic activity of this enzyme. However, the major variants of interest are the *1-6, *9, *10 and *17 alleles. Carriers with the *CYP2D6*3* and *CYP2D6*4* alleles have no enzyme activity. For example, a 25% dose reduction is recommended with amitriptyline therapy and for *CYP2D6*4* allele carriers, therapy with other analgesics is recommended rather than using codeine [2].

CYP1A2 gene

The CYP1A subfamily in humans and other mammals consists of two enzymes; according to standard P450 nomenclature, they are denoted as CYP1A1 and CYP1A2. These enzymes have prompted significant interest because they are involved in the metabolic activation of many procarcinogens and are induced by compounds of interest to toxicology, including dioxins. More than 15 alleles of the *CYP1A2* gene have been identified to date (*1B-*16), including several promoter variants [11].

The polymorphic variants g.28338G>A (*CYP1A2*1C*), c.-1635delT (*1D), -739T> G (*1E) and c.-9-154C>A (*1F) located in intron 1 are the most thoroughly studied alleles that lead to a change in the enzymatic activity of CYP1A2. A polymorphic variant *CYP1A2*1F* (c.-9-154C>A) in intron 1 leads to increased enzyme inducibility and increased drug metabolism. For example, increased enzyme inducibility during smoking, leads to low concentrations of clozapine levels in schizophrenic patients undergoing standard treatment. The risk of toxicity increased in allele C carriers during leflunomide and olanzapine therapy [12].

Determining the allelic variants of xenobiotic biotransformation genes is not only important for personalised drug prescription; the knowledge of the allele frequency distribution in different populations may be taken into account when the preferred medication regimen is selected. The frequency of *CYP2C9, VKORC1, CYP4F2, GGCX, CYP2D6,* and *CYP1A2* gene polymorphisms has been studied thoroughly in worldwide populations, except for populations in Central Asia (Tables 1 and 2).

Kazakhs are one of the Turkic peoples of Central Asia and comprise the majority of the population in Kazakhstan. According to the Agency of Statistics of the Republic of Kazakhstan, approximately 11 million Kazakhs live in Kazakhstan, and approximately 3.5 million Kazakhs live in regions neighbouring Kazakhstan and other regions (China, Russia, Uzbekistan, Turkmenistan, Kyrgyzstan, west Mongolia, Turkey) [13]. Kazakhs residing in the territory of Kazakhstan are internally divided into three large groups, the Elder, Middle and Lesser Zhuzes (or Hordes), which have historically demarcated territories. There are several tribes in each Zhuz [14]. Every Kazakh knows to which tribe and Zhuz they belong. Representatives of the same tribe are considered relatives because they

Polymorphism	Allele variant	Amino acide change	Location	Position	Enzyme activity	NCBI dbSNP rs
*CYP2C9*2*	c. 430C>T	R144C	10q23.33	Exon 3	decreased	rs1799853
*CYP2C9*3*	c. 1075A>C	I359L	10q23.33	Exon 7	decreased	rs1057910
VKORC1	c. 173+1000C>T		16p11.2	Intron 1	decreased	rs9934438
VKORC1	c. 173+1369G>C		16p11.2	Intron 2	decreased	rs8050894
CYP4F2	c. 1297G>A	V433M	19p13	Exon 2	decreased	rs2108622
GGCX	c. 1913+45G>C		2p11.2	Intron	decreased	rs11676382
*CYP2D6*3*	c. 622delA	frameshift Arg208Glyfs	22q13.2	Exon 5	none	rs35742686
*CYP2D6*4*	c. 353-1G>A	splicing defect	22q13.2	Intron 3	none	rs3892097
*CYP1A2*1F*	c. -9-154C>A	promoter	15q24.1	Intron	incresed	rs762551

Table 1: Characteristic of studied allele variants of the *CYP2C9, VKORC1, CYP4F2, GGCX, CYP2D6* and *CYP1A2*1F* genes.

Allele	χ^2, p	φ*	Power
*CYP2C9*2* (C430T)	χ^2=28.562; p=0.000	3.09	1
*CYP2C9*3* (A1075C)	χ^2=19.817; p=0.000	2.57	0.98
VKORC 1173	χ^2=142.969; p=0.000	6.90	1
VKORC 1542	χ^2=138.415; p=0.000	6.79	1
CYP4F2	χ^2=6.615; p=0.037	1.48	0.65
GGCX	χ^2=13.016; p=0.002	2.08	0.90
*CYP2D6*3*	χ^2=74.323; p=0.000	4.98	1
*CYP2D6*4*	χ^2=58.649; p=0.000	4.42	1
*CYP1A2*1F*	χ^2=3.537; p=0.171	1.09	0.38

*φ - noncentrality parameter

Table 2: Evaluation of genotype frequency differences between the Kazakh and Russian populations.

are considered to be descended from a common ancestor. Marriages are expected to respect the "seven generations law", that is, marriage between members of the same tribe is only possible when 7 generations from a common ancestor separate the individuals.

The purpose of this study was to determine the frequency of the allelic variants of *CYP2C9 (CYP2C9*2, CYP2C9*3)*, *VKORC1 (*c. 173+1000C>T, c. 173+1369G>C), *CYP4F2* (c. 1297G>A), *GGCX* (c. 1913+45G>C), *CYP2D6 (CYP2D6*3, CYP2D6*4)* and *CYP1A2*1F* (c. -9-154C>A) in Kazakh populations from different regions of Kazakhstan and in Russian populations from the West-Siberian region compared with published frequency data of other populations. The results of allele frequencies obtained by real-time PCR (RT-PCR) and direct sequencing methods from individuals from different regions of Kazakhstan were compared. Because no significant differences were found between these samples, the samples from the three regions were combined. The resulting distributions of allele and genotype frequency in the Kazakh population were compared with the frequencies in a Russian population, which was used as a control group.

Materials and Methods

Characteristics of the study populations

Kazakh and Russian populations were included (Table 3). The first group included people of Kazakh nationality. Each volunteer filled out a questionnaire in which they indicated their nationality and the nationality of their parents and grandparents, in addition to standard personal data. The concept of "Zhety ata" exists in the traditions of the Kazakh people. It means that each Kazakh should know 7 generations of their ancestors. Even today, this tradition is strongly maintained. Therefore, by conducting this survey, we were able to trace the ancestors of the participants to nearly 7 generations. As mentioned above, data on the participants' ancestors to the 2nd generation were collected through the questionnaire. Additionally, a verbal survey of the volunteers was performed to establish the ethnicity their ancestors from the third to the seventh generation. When any ancestors were found to be of other (not Kazakh) ethnicity, blood samples of these volunteers were excluded from the study. In addition to ethnicity, tribal affiliations of volunteers were specified. Kazakhs residing in the territory of Kazakhstan have an internal division into three large groups, the Elder, Middle and Lesser Zhuzes, which have historically demarcated territories. There are several tribes in each Zhuz.

	Kazakh	Russian
Number of samples	450	575
Sex (% male/female)	31/69	65/35
Mean age, years	40.6±13.2	33.5±11
Mean weight, kg	68±14	
Mean height, cm	167±8.9	
Number of samples by regions: North Kazakhstan North-East Kazakhstan South Kazakhstan West-Siberian	 161 176 113	 575
Number of samples by Zhuzes: Elder Zhuze Middle Zhuze Lesser Zhuze	 91 266 93	

Table 3: Characteristics of the studied groups.

The study group included 450 healthy individuals of Kazakh nationality from different regions of Kazakhstan, and a written informed consent was obtained from all of the participants. The age of the participants ranged from 17 to 70 years (mean age, 40.6 ± 13.2 years) (Table 3). Venous blood samples (5 ml) in tubes with EDTA were collected at clinical bases in the cities of Astana and Kokshetau (North Kazakhstan), Pavlodar (Northeast Kazakhstan) and Taraz (South Kazakhstan). DNA from blood was extracted by the salting-out method [15].

The second group was used as the control and included Russian residents of the West-Siberian region (the Novosibirsk and Kemerovo regions and the Altai Krai) who donated blood in the regional blood transfusion station and passed the standard medical examination but were not permanent donors. The age of the participants ranged from 18 to 55 years (mean age, 33.5 ± 11 years) (Table 3). DNA was extracted from venous blood using the phenol-chloroform method [16].

Ethical approval was received from the Ethics Committee of the National Centre for Biotechnology of the Republic of Kazakhstan, Astana, Kazakhstan (No. 10,14.02.2010).

Genotyping

Genotyping was performed using two methods: RT-PCR and

Alleles	Primers sequence (5'- 3')	Probe sequence
*CYP2C9*2*	ctgcggaattttgggatg	5'- r6g - cattgaggaccgtgttcaag -bhq-3'
	taaggtcagtgatatggagtaggg	5'- fam - cattgaggactgtgttcaagag -bhq-3'
*CYP2C9*3*	caaatgccctacacagatgc	5'- r6g - ccagagataccttgaccttctc -bhq-3'
	gatactattaatttggggacttcg	5'- fam - ccagagatacattgaccttctc -bhq-3'
VKORC 1173	acctgggctatcctctgttc	5'- fam - cctagtccaagagtcgatgatc -bhq-3'
	gggaggataggtcagtg	5'- r6g - ctagtccaagggtcgatgatc -bhq-3'
VKORC1 1542	tcagccccactccatacaatc	5'- r6g - ctcatcacggagcgtcctg -bhq-3'
	ccagttagttacctccccacatc	5'- fam- ctcatcaccgagcgtcctg -bhq-3'
CYP4F2	tgcctcatcagtgttttcgg	5'- r6g - caacccagctgtgtggcc -bhq-3'
	ttgagggaggtgatgttggatac	5'- fam - caacccagctatgtggccg -bhq-3'
GGCX	tatgtctttcgccgcaggt	5'- r6g - tgttggcaagcttggtaact-bhq-3'
	gaagaatggcaggaaaagatac	5'- fam - tgttggcaaccttggtaact-bhq-3'
*CYP2D6*3*	tggcaatgtcctacgcttc	5'- fam - tgagcacaggatgacct -bhq-3'
	ctctcaccttctccatctctgc	5'- r6g - tgagcacggatgacct -bhq-3'
*CYP2D6*4*	ggcaagaagtcgctggaccag	5'- r6g- cccccaggacgcccct-bhq-3'
	ttgctcacggctttgtccagg	5'- fam- cccccaagacgcccctt-bhq-3'
*CYP1A2*1F*	attctgtgatgctcaaagggtg	5'- fam - ctgtgggcacaggacgca -bhq-3'
	aaggagggactaggctgagg	5'- r6g - ctgtgggcccaggacgc -bhq-3'

Table 4: Primers and probes using for genotyping by real-time PCR.

direct sequencing. The characteristics of the studied allele variants of the *CYP2C9*, *VKORC1*, *CYP4F2*, *GGCX*, *CYP2D6* and *CYP1A2*1F* genes are shown in Table 1. A small amount of DNA was genotyped by direct sequencing and RT-PCR. Thus, the obtained DNA sample genotypes were used as a control for the further genotyping by RT-PCR. DNA samples were genotyped by direct sequencing if genotype determination by RT-PCR was difficult.

Genotyping by RT-PCR

RT-PCR genotyping was performed using TaqMan probes (Table 4). The composition of the PCR mixture was as follows: a buffer (65 mM Tris-HCl (pH 8.9); 0.05% Tween 20; 16 mM $(NH_4)_2SO_4$; 3.5 mM $MgCl_2$), 0.1-0.2 μM Taqman probe with FAM, 0.1-0.2 μM Taqman probe with R6G, 0.2 mM dNTPs, 0.3 μM primers and 0.5 units of thermally stable Taq polymerase. The reaction volume was 25.5 μl, and each reaction mixture was covered by an equal volume of mineral oil. The PCR conditions were as follows: denaturation at 96°C for 2 min and then 45 two-step cycles, which included 96°C for 30 s, 58-60°C for 40 s and 25°C for 10 s. The amplification was performed using the iCycler iQ5 PCR machine (Bio-Rad, United States).

Alleles	Primers sequence (5'- 3')
*CYP2C9*2*	tagtttcgtttctcttcctgtta aaatgtttccaagaatgtcagta
*CYP2C9*3*	ctaaagtccaggaagagattga atgatactatgaatttggggact
CYP4F2	agcggataacgtgttttcggaacccatcac agcggataacgccttggaatggacaaaaac
GGCX	gcttcttgttgcgaaagctctat caaacacttgggaacagttagct
*CYP2D6*3*	cccgttctgtctggtgta tgtcccagcaaagttcat
*CYP2D6*4*	cttctccgtgtccaccttt ccttcttacagtggggtct

Table 5: The primers used to determine the nucleotide sequence by direct sequencing method.

Genotyping by the direct sequencing method

Fragments of the gene region containing the analysed SNPs were obtained by PCR using oligonucleotide pairs of primers (Table 5). The amplification was performed in a multichannel thermocycler "PTC-0240 DNA EngineTetrad2 Cycler" (BioRad, Foster city, California, USA). The composition of the PCR mixture was as follows: a buffer (65 mM Tris-HCl (pH 8.9); 0.05% Tween 20, 16 mM $(NH_4)_2SO_4$; 2.5 mM $MgCl_2$), 0.2 mM dNTPs, 0.4 μM primers and 0.5 units of thermostable Taq-polymerase. The reaction volume was 25 ml. The PCR conditions were as follows: initial denaturation for 5 min at 95°C and then 35 cycles of three steps, 30 seconds of denaturation at 95°C, 30 seconds of primer annealing at 50-60°C and 30 seconds elongation at 72°C, and then 7 min at 72°C. The identification of SNPs was carried out based on DNA sequencing using the Big Dye Terminanor v3.1 Cycle Sequencing Kit (3730XL Genetic Analyzer, Applied Biosystems ABI) in accordance with the manufacturer's protocol. Sequence Scanner v1.0 software was used for analysing data.

Statistical analyses

The statistical analysis was performed using SPSS software (v16.0). The correspondence of the distribution of the genotype frequencies to the Hardy–Weinberg equilibrium was assessed using the χ^2 criterion (α=0.05, df=2). Significant differences in genotype frequencies between Kazakh and Russian populations were assessed according to the χ^2 criterion (α=0.05, df=2). The analysis of linkage disequilibrium of genes *VKORC1, CYP2C9,* and *CYP2D6* was performed using the available program at http://www.oege.org/software/cubex.

Results

Allele and genotype frequency data were obtained for 9 SNPs in

Polymorphisms	Number of samples	Hardy – Weinberg equilibrium	Allele	n[a]	Frequency	Genotype	n[b]	Frequency
CYP2C9*2	437	p=0.66	C T	856 18	0.98 0.02	CC CT TT	419 18 0	0.96 0.04 0.00
CYP2C9*3	444	p=0.54	A C	863 25	0.97 0.03	AA AC CC	419 25 0	0.94 0.06 0.00
VKORC1 1173	286	p=0.76	C T	162 410	0.28 0.72	CC CT TT	24 114 148	0.08 0.40 0.52
VKORC1 1542	259	p=0.65	G C	142 376	0.28 0.72	GG GC CC	18 106 135	0.07 0.41 0.52
CYP4F2	284	p=0.26	G A	396 172	0.70 0.31	GG GA AA	134 128 22	0.47 0.45 0.08
GGCX	266	p=0.57	C G	514 18	0.97 0.04	CC CG GG	248 18 0	0.93 0.07 0.00
CYP2D6*3	287	p=0.86	A del	568 6	0.99 0.01	AA A/del del/del	281 6 0	0.98 0.02 0.00
CYP2D6*4	343	p=0.15	G A	637 49	0.93 0.07	GG GA AA	294 49 0	0.86 0.14 0.00
CYP1A2*1F	257	p=0.63	A C	332 182	0.65 0.35	AA AC CC	109 114 34	0.42 0.44 0.13

[a]number of chromosomes; [b]number of alleles

Table 6: Allele frequency and genotype distribution in the Kazakh population

the Kazakh and Russian populations. The allele frequencies (MAF) in the Kazakh population were as follows: *CYP2C9*2*, 0.02; *CYP2C9*3*, 0.03; *VKORC1* (c. 173+1000C>T), 0.72; *VKORC1* (c. 173+1369G>C), 0.72; *CYP4F2* (c. 1297G>A), 0.31; *GGCX* (c. 1913+45G>C), 0.04; *CYP2D6*3*, 0.01; *CYP2D6*4*, 0.07; and *CYP1A2*1F* (c. -9-154C>A), 0.35. The allele and genotype frequencies of the Kazakh population are summarised in Table 6. All alleles were in Hardy-Weinberg equilibrium (p>0.05). DNA samples from individuals of Kazakh nationality were collected from different regions of Kazakhstan. Therefore, a comparative analysis of allele frequencies by regions using the statistical criterion χ^2 was assessed. Significant differences in the Kazakh population according to region (North, Northeast and South Kazakhstan) are presented in Table 7. A significant difference in the frequency of genotypes was not found (p>0.05). Thus, the DNA samples from different regions of Kazakhstan were combined into one group.

DNA samples from a Russian population were used as the control. The allele frequencies (MAF) in the Russian population were as follows: *CYP2C9*2*, 0.08; *CYP2C9*3*, 0.08; *VKORC1* (c. 173+1000C>T), 0.40; *VKORC1* (c. 173+1369G>C), 0.41; *CYP4F2* (c. 1297G>A), 0.24; *GGCX* (c. 1913+45G>C), 0.08; *CYP2D6*3*, 0.15; *CYP2D6*4*, 0.22; and *CYP1A2*1F* (c. -9-154C>A), 0.31. The allele and genotype frequencies of the Russian population are summarised in Table 8. All alleles except *GGCX* (p=0.04) were in Hardy-Weinberg equilibrium (p>0.05) in the Russian population.

Allele	Allele frequency			Significant differences
	North Kazakhstan *n=72*	North-east Kazakhstan *n=146*	South Kazakhstan *n=70*	
*CYP2C9*2* (C430T)	C=0.94 T=0.06	C=0.98 T=0.02	C=0.97 T=0.03	χ^2=3.982; p=0.137
*CYP2C9*3* (A1075C)	A=0.96 C=0.04	A=0.95 C=0.05	A=0.96 C=0.04	χ^2=0.427; p=0.808
VKORC 1173	C=0.31 T=0.69	C=0.27 T=0.73	C=0.28 T=0.72	χ^2=4.007; p=0.405
VKORC 1542	G=0.29 C=0.71	G=0.27 C=0.73	G=0.26 C=0.74	χ^2=0.701; p=0.951
CYP4F2	G=0.64 A=0.36	G=0.72 A=0.28	G=0.73 A=0.27	χ^2=6.175; p=0.186
GGCX	C=0.95 G=0.05	C=0.97 G=0.03	C=0.98 G=0.02	χ^2=0.968; p=0.616
*CYP2D6*3*	A=0.99 del=0.01	A=0.99 del=0.01	A=1.00 del=0.00	χ^2=1.940; p=0.379
*CYP2D6*4*	G=0.94 A=0.06	G=0.90 A=0.10	G=0.91 A=0.09	χ^2=2.417; p=0.299
*CYP1A2*1F*	A=0.63 C=0.38	A=0.63 C=0.37	A=0.69 C=0.31	χ^2=5.710; p=0.222

Table 7: Evaluation of allele and genotype frequency differences according to region (North, Northeast and South Kazakhstan).

Polymorphisms	Number of samples	Hardy – Weinberg equilibrium	Allele	n[a]	Frequency	Genotype	n[b]	Frequency
CYP2C9*2	284	p=0.11	C T	521 47	0.92 0.08	CC CT TT	241 39 4	0.85 0.14 0.01
CYP2C9*3	283	p=0.86	A C	521 45	0.92 0.08	AA AC CC	240 41 2	0.85 0.14 0.01
VKORC1 1173	575	p=0.73	C T	685 465	0.60 0.40	CC CT TT	202 281 92	0.35 0.49 0.16
VKORC1 1542	574	p=0.53	G C	677 471	0.59 0.41	GG GC CC	196 285 93	0.34 0.50 0.16
CYP4F2	284	p=0.73	G A	433 135	0.76 0.24	GG GA AA	164 105 15	0.58 0.37 0.05
GGCX	565	p=0.04	C G	1041 89	0.92 0.08	CC CG GG	476 89 0	0.84 0.16 0.00
CYP2D6*3	350	p=0.59	A del	596 104	0.85 0.15	AA A/del del/del	255 86 9	0.73 0.25 0.03
CYP2D6*4	352	p=0.84	G A	551 153	0.78 0.22	GG GA AA	215 121 16	0.61 0.34 0.05
CYP1A2*1F	341	p=0.25	A C	468 214	0.69 0.31	AA AC CC	156 156 29	0.46 0.46 0.09

[a]number of chromosomes; [b]number of alleles

Table 8: Allele frequency and genotype distribution in the Russian population.

LD	*VKORC1 1542*	*CYP2C9*2*	*CYP2D6*3*
VKORC1 1173	D'=0.96 r^2=0.9094		
*CYP2C9*3*		D'=-1.0 r^2=0.0006	
*CYP2D6*4*			D'=-1.0 r^2=0.001

Table 9: Evaluation of linkage disequilibrium of the *VKORC1, CYP2C9* and *CYP2D6* genes in the Kazakh population.

Significant differences in genotype frequencies between Kazakh and Russian populations were assessed according to the χ^2 criterion. The overall difference (between Kazakh and Russian populations) in allele frequencies were highly significant ($p<0.05$), except for *CYP1A2*1F* ($p=0.17$) (Table 2).

Furthermore, the linkage disequilibrium of the *VKORC1*, *CYP2C9* and *CYP2D6* genes in the Kazakh population was analysed. We identified the linkage disequilibrium of the *VKORC1* c. 173+1369G>C and *VKORC1* c. 173+1000C>T (D'=0.96) allele polymorphisms (Table 9).

Discussion

Kazakhs residing in Kazakhstan are internally divided into three large groups, the Elder, Middle and Lesser Zhuzes, which have historically demarcated territories. There are several tribes in each Zhuz.

The metabolism of xenobiotics may vary depending on the individual's ethnicity and race. Therefore, we compared allele frequencies from the different regions of Kazakhstan. Blood samples were collected from the northern and southern part of Kazakhstan because the climatic conditions differ from north to south. Northern and Northeastern Kazakhstan are traditionally the territory of the Middle Zhuze, and Southern Kazakhstan is traditionally the territory of the Elder Zhuze. However, significant differences between genotype frequencies were not detected ($p>0.05$) (Table 7).

Tribes within Zhuzes are considered related to each other, and marriage between Zhuzes is encouraged. Therefore, the distribution of allele frequencies between all three Zhuzes was nearly equal, which allowed the DNA samples from different regions of Kazakhstan to be combined into one group, representative of the Kazakh population.

The previous literature has examined Kazakh populations living in areas other than those included here. For example, Wang et al. studied Kazakhs living in China [17]; ethnicity was only tracked until the third generation. Thus, we could not use these data to assess differences. Magalon et al. studied only 26 DNA samples and did not indicate whether they monitored the ethnicity of the participants' ancestors, which does not allow an adequate comparative analysis [18]. Tarlykov et al. investigated the hypervariable segment I of mitochondrial DNA and Y-STR loci in DNA samples of individuals living in isolation in the same area and who were members of the same tribe. However, there was no comparative analysis with other groups living in other regions of Kazakhstan or who were representatives of other tribes and Zhuzes [19].

A Russian population was used as a control group. A comparative analysis of the allele frequency of the Russian population between the studied samples and published data was carried out. The findings correspond to published data (Table 10).

Furthermore, a comparative analysis of the allele frequency of the Kazakh population between studied samples and published data from the other populations was conducted (Table 10).

CYP2C9 gene

Wide inter-ethnic variability was observed for the *CYP2C9* gene polymorphic variants. On average, the frequency of the *CYP2C9*2* allele was 11-15% in the Caucasian population and 2-3% in the African and Asian populations, and for the *CYP2C9*3* allele, the averages were 5-7% and 2-4%, respectively [3,4,20]. Thus, the *2 and *3 alleles of the *CYP2C9* gene are much less common among individuals in Asia and Africa than among Caucasians.

The allele frequency of *CYP2C9*2* (0.02) and *CYP2C9*3* (0.03) in the Kazakh population is near the frequency of these polymorphic loci in Asian populations ($p=0.27$ and $p=0.30$) and significantly different from that of Caucasian populations ($p=0.00$).

In the Russian population, the frequency of *CYP2C9*3* (0.08) is close to that of the Caucasian population ($p=0.45$) and significantly different from the frequency in the Asian and African-American populations ($p=0.00$).

Thus, in future pharmacogenetic studies in Kazakh populations, close attention should be paid to the polymorphic variants of the *CYP2C9* gene, which affect the dose of warfarin and are characteristic of Asian populations.

VKORC1 gene

The distribution of *VKORC1* allele polymorphisms in populations is interesting. The C. 173+1000C>T allele frequency is as high as 92% in Asian populations [4], whereas in Caucasian populations, it is approximately 42% [5]. The allele frequency in the Kazakh population ($f=0.72$, $p=0.00$) studied here was between that of Asian and Caucasian populations.

The allele frequency in the Russian population was 0.41 ($p=0.09$), which agrees with published data and corresponds to the Caucasoid populations.

According to Rieder et al., the *VKORC1* c.-1639G>A mutation is linked to mutations in *VKORC1* c. 173+1369G>C and *VKORC1* c.

	2c9*2	2c9*3	VKORC rs9934438	VKORC rs8050894	CYP4F2	GGCX	2d6*3	2d6*4	1a2*1F
Kazakh	**0.02**	**0.03**	**0.72**	**0.72**	**0.31**	**0.04**	**0.01**	**0.07**	**0.35**
Russian	**0.08**	**0.08**	**0.40**	**0.41**	**0.24**	**0.08**	**0.15**	**0.22**	**0.31**
African-American (1000 Genomes)	0.02	0.01	0.07	0.21	0.09	0.01	0.0	0.06	0.46
African-American*	0.01-0.027	0.005-0.02	0.02-0.13	0.19-0.28	0.05-0.1	0.0-0.02	0.0-0.01	0.01-0.12	0.35-0.54
Caucasians (1000 Genomes)	0.12	0.06	0.4	0.41	0.27	0.09	0.02	0.19	0.31
Caucasians*	0.11-0.2	0.06-0.16	0.39-0.48	0.37-0.41	0.23-0.32	0.05-0.13	0.0-0.04	0.07-0.21	0.22-0.52
Asians (1000 Genomes)	0.0	0.02	0.92	0.92	0.21	0.0	0.0	0.0	0.37
Asians*	0.0-0.05	0.02-0.1	0.9-0.95	0.89-0.94	0.19-0.34	0.0	0.0	0.0-0.15	0.33-0.61

*Different sources with a population sample of more than 100 people (including HapMap data) [3-8,10-12,20-48].

Table 10: A comparative table of frequencies of genes in different populations.

173+1000C>T [6]. We also identified a linkage disequilibrium between the *VKORC1* c. 173+1369G>C and *VKORC1* c. 173+1000C>T (D'=0.96) polymorphisms (Table 9).

CYP4F2 gene

The *CYP4F2* c. 1297G>A allele frequency in both Asian and Caucasian populations is approximately 25% on average [7,8] and is 8% in African-Americans [7]. The allele frequency was 31% in the Kazakh population, which corresponds with the frequencies in Asian and Caucasian populations (p=0.25). The allele frequency in the Russian population was 0.24, which is near to that of the Japanese population (p=0.38).

GGCX gene

A mutant version of the *GGCX* gene (rs11676382) occurs at a frequency of 10% in Caucasian populations and nearly never occurs in the Asian and African-American populations [21]. The allele frequency in Kazakh population was 4%, which is between the Caucasoid and Asian populations.

The allele frequency in the Russian population was 8%, which corresponds with the Caucasian population (p=0.77).

CYP2D6 gene

Ethnic differences in the frequencies of polymorphisms of this gene are very high. The *CYP2D6*3* and *CYP2D6*4* alleles are almost never found in Asian and African populations. *CYP2D6*3* occurs at a frequency of 0.01-0.02 and *CYP2D6*4* at a frequency of 0.07-0.20 in the Caucasian population [22-24].

The frequency of *CYP2D6*4* (0.07) in the Kazakh population is significantly different from that of the Caucasian population (p=0.00) and corresponds to the frequencies in Asian populations. At the same time, the frequency of *CYP2D6*3* is closer to that of Caucasian populations (p=0.93).

In the studied Russian population, the frequency of *CYP2D6*3* (0.15) is higher than in all analysed populations and does not correspond well with the literature data (p=0.00). The frequencies of *CYP2D6*4* (0.22) found in our study are similar to those of similar populations reported in the literature (p=0.85).

CYP1A2 gene

The *CYP1A2*1F* allelic polymorphism frequency varies widely between populations, ranging from 22% in the Caucasian population to 61% in the Japanese population. The average frequency of the *CYP1A2*1F* allele is 31% in the Caucasian population, 37% in the Asian population, and reaches 46% in the African-American population [20,25]. In the Kazakh population, the frequency was 35%, which is between the Caucasian and Asian populations.

The allele frequency in the Russian population was 31%, which is typical for Caucasian populations (p=0.14). The allele frequency of the studied Kazakh and Russian population was nearly the same (p=0.17).

Conclusions

The Kazakh population exhibits allele frequencies at an intermediate level between Caucasian and Asian populations in nearly all of the studied gene allele variants. This distribution of frequencies can be explained by recent studies that have suggested that the Kazakh population was formed by the mixing of Asian and Caucasoid populations [14]. Currently, individuals exhibit distinctive Asian and/or Caucasoid traits, depending on the region of Kazakhstan. Based on our data, we can conclude that further pharmacogenetic studies are required.

Availability of Supporting Data

The datasets used in this article are included here.

Competing Interests

The authors declare that they have no competing interests.

Authors' Contributions

I.A.N. extracted DNA, performed *SNP* genotyping, summarised genotyping data, analysed data, carried out statistical analysis and drafted the manuscript. R.A.A. extracted DNA and performed SNP genotyping. V.E.N. helped edit the manuscript and developed the SNP typing methods (RT-PCR). S.N.S. carried out the statistical analysis. B.L.A. performed *SNP* genotyping for Russian populations. R.E.M. and F.M.L. helped edit the manuscript and made suggestions on study design. All authors read and approved the final manuscript.

Acknowledgement

This work was supported by grant No. 1.04.01 from the Ministry of Education and Science of the Republic of Kazakhstan.

References

1. Sychev DA (2005) Clinical pharmacogenetic as a way to personalized medicine: Are expectations justified? Clinical Pharmacology and Therapy 14: 77-83.
2. The Pharmacogenomics Knowledge Base.
3. Limdi NA, McGwin G, Goldstein JA, Beasley TM, Arnett DK, et al. (2008) Influence of CYP2C9 and VKORC1 1173C/T genotype on the risk of hemorrhagic complications in African-American and European-American patients on warfarin. Clin Pharmacol Ther 83: 312-321.
4. Wang TL, Li HL, Tjong WY, Chen QS, Wu GS, et al. (2008) Genetic factors contribute to patient-specific warfarin dose for Han Chinese. Clin Chim Acta 396: 76-79.
5. Takahashi H, Wilkinson GR, Nutescu EA, Morita T, Ritchie MD, et al. (2006) Different contributions of polymorphisms in VKORC1 and CYP2C9 to intra- and inter-population differences in maintenance dose of warfarin in Japanese, Caucasians and African-Americans. Pharmacogenet Genomics 16: 101-110.
6. Rieder MJ, Reiner AP, Gage BF, Nickerson DA, Eby CS, et al. (2005) Effect of VKORC1 haplotypes on transcriptional regulation and warfarin dose. N Engl J Med 352: 2285-2293.
7. Singh O, Sandanaraj E, Subramanian K, Lee LH, Chowbay B (2011) Influence of CYP4F2 rs2108622 (V433M) on warfarin dose requirement in Asian patients. Drug Metab Pharmacokinet 26: 130-136.
8. Liang R, Li L, Li C, Gao Y, Liu W, et al. (2012) Impact of CYP2C9*3, VKORC1-1639, CYP4F2rs2108622 genetic polymorphism and clinical factors on warfarin maintenance dose in Han-Chinese patients. J Thromb Thrombolysis 34: 120-125.
9. Stec DE, Roman RJ, Flasch A, Rieder MJ (2007) Functional polymorphism in human CYP4F2 decreases 20-HETE production. Physiol Genomics 30: 74-81.
10. Rieder MJ, Reiner AP, Rettie AE (2007) Gamma-glutamyl carboxylase (GGCX) tagSNPs have limited utility for predicting warfarin maintenance dose. J Thromb Haemost 5: 2227-2234.
11. Jaquenoud Sirot E, Knezevic B, Morena GP, Harenberg S, Oneda B, et al. (2009) ABCB1 and cytochrome P450 polymorphisms: clinical pharmacogenetics of clozapine. J Clin Psychopharmacol 29: 319-326.
12. Zhou SF, Yang LP, Zhou ZW, Liu YH, Chan E (2009) Insights into the substrate specificity, inhibitors, regulation, and polymorphisms and the clinical impact of human cytochrome P450 1A2. AAPS J 11: 481-494.
13. Agency of statistics of the Republic of Kazakhstan.
14. Iskakova MK (2012) Kazakh's tamga. Saryarka Press, Astana.

15. Salting-out extraction method.

16. Phenol-chloroform extraction method.

17. Wang SM, Zhu AP, Li D, Wang Z, Zhang P, et al. (2009) Frequencies of genotypes and alleles of the functional SNPs in CYP2C19 and CYP2E1 in mainland Chinese Kazakh, Uygur and Han populations. J Hum Genet 54: 372-375.

18. Magalon H, Patin E, Austerlitz F, Hegay T, Aldashev A, et al. (2008) Population genetic diversity of the NAT2 gene supports a role of acetylation in human adaptation to farming in Central Asia. Eur J Hum Genet 16: 243-251.

19. Tarlykov PV, Zholdybayeva EV, Akilzhanova AR, Nurkina ZM, Sabitov ZM, et al. (2013) Mitochondrial and Y-chromosomal profile of the Kazakh population from East Kazakhstan. Croat Med J 54: 17-24.

20. Marie-Genica Consortium on Genetic Susceptibility for Menopausal Hormone Therapy Related Breast Cancer Risk1 (2010) Genetic polymorphisms in phase I and phase II enzymes and breast cancer risk associated with menopausal hormone therapy in postmenopausal women. Breast Cancer Res Treat 119: 463-474.

21. King CR, Deych E, Milligan P, Eby C, Lenzini P, et al. (2010) Gamma-glutamyl carboxylase and its influence on warfarin dose. Thromb Haemost 104: 750-754.

22. Sachse C, Brockmöller J, Bauer S, Roots I (1997) Cytochrome P450 2D6 variants in a Caucasian population: allele frequencies and phenotypic consequences. Am J Hum Genet 60: 284-295.

23. Brown MA, Edwards S, Hoyle E, Campbell S, Laval S, et al. (2000) Polymorphisms of the CYP2D6 gene increase susceptibility to ankylosing spondylitis. Hum Mol Genet 9: 1563-1566.

24. Kubota T, Yamaura Y, Ohkawa N, Hara H, Chiba K (2000) Frequencies of CYP2D6 mutant alleles in a normal Japanese population and metabolic activity of dextromethorphan O-demethylation in different CYP2D6 genotypes. Br J Clin Pharmacol 50: 31-34.

25. Hamdy SI, Hiratsuka M, Narahara K, Endo N, El-Enany M, et al. (2003) Genotyping of four genetic polymorphisms in the CYP1A2 gene in the Egyptian population. Br J Clin Pharmacol 55: 321-324.

26. Ingelman-Sundberg M, Sim SC, Gomez A, Rodriguez-Antona C (2007) Influence of cytochrome P450 polymorphisms on drug therapies: pharmacogenetic, pharmacoepigenetic and clinical aspects. Pharmacol Ther 116: 496-526.

27. Nebert DW, McKinnon RA (1994) Cytochrome P450: evolution and functional diversity. Prog Liver Dis 12: 63-97.

28. Meyer UA (1994) Pharmacogenetics: the slow, the rapid, and the ultrarapid. Proc Natl Acad Sci U S A 91: 1983-1984.

29. Limdi NA, Beasley TM, Crowley MR, Goldstein JA, Rieder MJ, et al. (2008) VKORC1 polymorphisms, haplotypes and haplotype groups on warfarin dose among African-Americans and European-Americans. Pharmacogenomics 9: 1445-1458.

30. Watson MA, Stewart RK, Smith GBJ, Massey TE, Bell DA (1998) Human glutathione S-transferase P1 polymorphisms: relationship to lung tissue enzyme activity and population frequency distribution. Carcinogenesis 19: 275-280.

31. Rathore SS, Agarwal SK, Pande S, Mittal T, Mittal B (2010) Frequencies of VKORC1 -1639 G>A, CYP2C9*2 and CYP2C9*3 genetic variants in the Northern Indian population. Biosci Trends 4: 333-337.

32. Munshi A, Sharma V, Kaul S, Al-Hazzani A, Alshatwi AA, et al. (2012) Association of 1347 G/A cytochrome P450 4F2 (CYP4F2) gene variant with hypertension and stroke. Mol Biol Rep 39: 1677-1682.

33. Bravo-Villalta HV, Yamamoto K, Nakamura K, Bayá A, Okada Y, et al. (2005) Genetic polymorphism of CYP2C9 and CYP2C19 in a Bolivian population: an investigative and comparative study. Eur J Clin Pharmacol 61: 179-184.

34. Mushiroda T, Ohnishi Y, Saito S, Takahashi A, Kikuchi Y, et al. (2006) Association of VKORC1 and CYP2C9 polymorphisms with warfarin dose requirements in Japanese patients. J Hum Genet 51: 249-253.

35. Rusdiana T, Araki T, Nakamura T, Subarnas A, Yamamoto K (2013) Responsiveness to low-dose warfarin associated with genetic variants of VKORC, CYP2C9, CYP2C19, and CYP4F2 in an Indonesian population. Eur J Clin Pharmacol 69: 395-405.

36. Sayed SK, Imam HMK (2012) Study of the Association of CYP2D6*4 Polymorphism with the Susceptibility of HCV- Related Liver Cirrhosis and Liver Cancer. Life Science Journal 9:1571-1577.

37. Qumsieh RY, Ali BR, Abdulrazzaq YM, Osman O, Akawi NA, et al. (2011) Identification of new alleles and the determination of alleles and genotypes frequencies at the CYP2D6 gene in Emiratis. PLoS One 6: e28943.

38. Menoyo A, del Rio E, Baiget M (2006) Characterization of variant alleles of cytochrome CYP2D6 in a Spanish population. Cell Biochem Funct 24: 381-385.

39. Gaikovitch EA, Cascorbi I, Mrozikiewicz PM, Brockmöller J, Frötschl R, et al. (2003) Polymorphisms of drug-metabolizing enzymes CYP2C9, CYP2C19, CYP2D6, CYP1A, NAT2 and of P-glycoprotein in a Russian population. Eur J Clin Pharmacol 59: 303-312.

40. Ahmadishina LZ (2007) Monooxygenase enzyme gene polymorphism and the antioxidant defense system in patients with chronic diseases of the respiratory system in the Republic of Bashkortostan, Abstract.

41. Santos PC, Dinardo CL, Schettert IT, Soares RA, Kawabata-Yoshihara L, et al. (2013) CYP2C9 and VKORC1 polymorphisms influence warfarin dose variability in patients on long-term anticoagulation. Eur J Clin Pharmacol 69: 789-797.

42. Branco CC, Pereirinha T, Cabral R, Pacheco PR, Mota-Vieira L (2009) Thrombotic genetic risk factors and warfarin pharmacogenetic variants in São Miguel's healthy population (Azores). Thromb J 7: 9.

43. Roco A, Quiñones L, Agúndez JAG, García-Martín E, Squicciarini V, et al. (2012) Frequencies of 23 functionally significant variant alleles related with metabolism of antineoplastic drugs in the Chilean population: comparison with Caucasian and Asian populations. Front Genet 3: 229.

44. Zainuddin Z, Teh LK, Suhaimi AW, Ismail R (2006) Malaysian Indians are genetically similar to Caucasians: CYP2C9 polymorphism. J Clin Pharm Ther 31: 187-191.

45. Cornelis MC, El-Sohemy A, Kabagambe EK, Campos H (2006) Coffee, CYP1A2 genotype, and risk of myocardial infarction. JAMA 295: 1135-1141.

46. Suarez-Kurtz G (2005) Pharmacogenomics in admixed populations. Trends Pharmacol Sci 26: 196-201.

47. Sistonen J, Sajantila A, Lao O, Corander J, Barbujani G, et al. (2007) CYP2D6 worldwide genetic variation shows high frequency of altered activity variants and no continental structure. Pharmacogenet Genomics 17: 93-101.

48. Solus JF, Arietta BJ, Harris JR, Sexton DP, Steward JQ, et al. (2004) Genetic variation in eleven phase I drug metabolism genes in an ethnically diverse population. Pharmacogenomics 5: 895-931.