

CARTOGRAPHIC ATLAS OF FREQUENCY VARIATION FOR 45 PHARMACOGENETIC MARKERS IN POPULATIONS OF RUSSIA AND ITS NEIGHBOR STATES

Balanovska EV^{1,2}, Petrushenko VS^{1,3}, Koshel SM^{1,4}, Pocheshkhova EA^{1,5}, Chernevskiy DK¹, Mirzaev KB⁶, Abdullaev SP⁶, Balanovsky OP^{1,2,3}

¹ Research Centre for Medical Genetics, Moscow, Russia

² Biobank of North Eurasia, Moscow, Russia

³ Vavilov Institute of General Genetics, Russian Academy of Sciences, Moscow, Russia

⁴ Lomonosov Moscow State University, Moscow, Russia

⁵ Kuban State Medical University, Krasnodar, Russia

⁶ Russian Medical Academy of Continuous Professional Education, Moscow, Russia

The lack of information about the frequency of pharmacogenetic markers in Russia impedes the adoption of personalized treatment algorithms originally developed for West European populations. The aim of this paper was to study the distribution of some clinically significant pharmacogenetic markers across Russia. A total of 45 pharmacogenetic markers were selected from a few population genetic datasets, including ADME, drug target and hemostasis-controlling genes. The total number of donors genotyped for these markers was 2,197. The frequencies of these markers were determined for 50 different populations, comprised of 137 ethnic and subethnic groups. A comprehensive pharmacogenetic atlas was created, i.e. a systematic collection of gene geographic maps of frequency variation for 45 pharmacogenetic DNA markers in Russia and its neighbor states. The maps revealed 3 patterns of geographic variation. Clinal variation (a gradient change in frequency along the East-West axis) is observed in the pharmacogenetic markers that follow the main pattern of variation for North Eurasia (13% of the maps). Uniform distribution singles out a group of markers that occur at average frequency in most Russian regions (27% of the maps). Focal variation is observed in the markers that are specific to a certain group of populations and are absent in other regions (60% of the maps). The atlas reveals that the average frequency of the marker and its frequency in individual populations do not indicate the type of its distribution in Russia: a gene geographic map is needed to uncover the pattern of its variation.

Keywords: pharmacogenetics, cartographic atlas, populations, ethnic groups, gene pool, gene geography, Russia, North Eurasia

Acknowledgement: the authors thank the donors for their participation, the Institute of General Genetics for the access to the database of genotypes, the Center for Precision Genome Editing and Genetic Technologies for Biomedicine of Pirogov Russian National Research Medical University (Moscow, Russia) for the opportunity to use molecular genetic technologies.

Funding: the study was carried out under the State Assignment of the Russian Ministry of Science and Higher Education for the Research Center for Medical Genetics and Vavilov Institute of General Genetics.

Author contribution: Balanovska EV — data analysis, manuscript draft; Balanovsky OP — study design and supervision; Petrushenko VS — bioinformatic analysis; Koshel SM — map analysis, manuscript editing; Chernevskiy DK, Pocheshkhova EA — tabular data analysis; Mirzaev KB, Abdullaev SP — description of pharmacogenetic markers.

Compliance with ethical standards: the study was approved by the Ethics Committee of the Research Center for Medical Genetics (Protocol № 3/1 dated September 5, 2018) and carried out on the samples obtained during the population genetic study of the gene pools of ethnic groups from Russia and its neighbor states. The donors gave voluntary informed consent to participate.

✉ **Correspondence should be addressed:** Elena B. Balanovska
Moskvorechie, 1, Moscow, 115478; balanovska@mail.ru

Received: 06.11.2020 **Accepted:** 22.11.2020 **Published online:** 18.12.2020

DOI: 10.24075/brsmu.2020.080

КАРТОГРАФИЧЕСКИЙ АТЛАС РАСПРОСТРАНЕНИЯ 45 ФАРМАКОГЕНЕТИЧЕСКИХ МАРКЕРОВ В НАРОДОНАСЕЛЕНИИ РОССИИ И СОПРЕДЕЛЬНЫХ СТРАН

Е. В. Балановская^{1,2}, В. С. Петрушенко^{1,3}, С. М. Кошель^{1,4}, Э. А. Почешхова^{1,5}, Д. К. Черневский¹, К. Б. Мирзаев⁶, Ш. П. Абдуллаев⁶, О. П. Балановский^{1,2,3}

¹ Медико-генетический научный центр имени Н. П. Бочкова, Москва, Россия

² Биобанк Северной Евразии, Москва, Россия

³ Институт общей генетики имени Н. И. Вавилова, Москва, Россия

⁴ Московский государственный университет имени М. В. Ломоносова, Москва, Россия

⁵ Кубанский государственный медицинский университет, Краснодар, Россия

⁶ Российская медицинская академия непрерывного профессионального образования, Москва, Россия

Нехватка информации о распространенности в РФ фармакогенетических маркеров приводит к невозможности внедрения алгоритмов персонализации, разработанных для Западной Европы. Целью работы было систематическое изучение распространенности ряда значимых фармакогенетических маркеров по всей территории России. Из нескольких массивов популяционно-генетических данных отобраны 45 маркеров (ADME-генов; генов, кодирующих фармакодинамические мишени лекарственных средств; генов, кодирующих компоненты системы гемостаза), генотипированных суммарно для 2197 индивидов. Определены частоты этих маркеров в 50 популяциях, включающих информацию о 137 этнических и субэтнических группах. В результате создан фармакогенетический атлас — систематическое собрание географических карт распространенности фармакогенетических ДНК маркеров по всей территории России и сопредельных стран. Атлас выявил три паттерна пространственной изменчивости. Паттерн клинальной изменчивости (градиентного изменения частот по оси «восток–запад») объединяет маркеры, следующие основной закономерности всего генофонда населения Северной Евразии (13% карт атласа). Паттерн равномерного распределения выделяет маркеры, средняя частота которых характерна для большинства регионов России (27% карт атласа). Паттерн «очаговой» изменчивости объединяет фармакогенетические маркеры, характерные только для определенной группы этносов и отсутствующие в других регионах (60% карт атласа). Атлас показывает, что средняя частота маркера и информация о его встречаемости в отдельных популяциях не могут служить указанием на тип его распределения в пространстве РФ — для выявления паттерна изменчивости необходима географическая карта.

Ключевые слова: фармакогенетика, картографический атлас, популяции, этносы, генофонд, геогеография, Россия, Северная Евразия

Благодарности: авторы благодарят доноров образцов, принимавших участие в исследовании, Институт общей генетики им. Н. И. Вавилова РАН за доступ к базе данных генотипов, Центр высокоточного редактирования и генетических технологий для биомедицины РНИМУ им. Н. И. Пирогова (Москва, Россия) за возможность использования молекулярно-генетических технологий.

Финансирование: исследование выполнено при финансовой поддержке Государственного задания Министерства науки и высшего образования РФ для Медико-генетического научного центра имени акад. Н. П. Бочкова и Института общей генетики имени Н. И. Вавилова РАН.

Вклад авторов: Е. В. Балановская — анализ данных, написание текста статьи; О. П. Балановский — дизайн и руководство исследованием; В. С. Петрушенко — биоинформатический анализ; С. М. Кошель — картографический анализ, редактирование рукописи; Д. К. Черневский, Э. А. Почешхова — работа с табличными данными; К. Б. Мирзаев, Ш. П. Абдуллаев — описание фармакогенетических маркеров.

Соблюдение этических стандартов: исследование одобрено этическим комитетом Медико-генетического научного центра (протокол № 3/1 от 5 сентября 2018 г.), все обследуемые подписали добровольное информированное согласие на участие в исследовании.

✉ **Для корреспонденции:** Елена Владимировна Балановская
ул. Москворечье, д. 1, г. Москва, 115478; balanovska@mail.ru

Статья получена: 06.11.2020 **Статья принята к печати:** 22.11.2020 **Опубликована онлайн:** 18.12.2020

DOI: 10.24075/vrgmu.2020.080

The key genes involved in the absorption, distribution, metabolism, and elimination (ADME) of drugs have been described in multiple studies. Although ADME genes are relatively modest in number, each of them is represented by a few functionally relevant polymorphic variants [1, 2]. In addition to ADME genes, research has identified drug target genes and the genes encoding the components of the hemostatic system. Assays and protocols are being developed to guide drug selection and dosing based on the patient's genotype [3–5]. However, genetic differences between ethnicities remain the major hurdle impeding the widespread adoption of such assays: commercial panels designed to screen ADME genes for clinically significant polymorphisms do not account for their variation across races and ethnic groups and thus are poorly adapted for ethnically diverse regions.

Studies conducted in polyethnic regions across the world have demonstrated both quantitative (allele frequencies) and qualitative (population-specific alleles) differences in ADME genes and other pharmacogenetic markers between ethnic groups. Although interethnic differences have been well known since the inception of pharmacogenetics [6, 7], most pharmacogenetic studies are still carried out on Caucasian patients. Consequently, their results cannot be extrapolated to other ethnic groups. For example, it has been shown that standard algorithms for essential drug dosing are often ineffective for African and Hispanic patients, creating life-threatening situations or resulting in death or disability [8, 9].

Pharmacogenetic analysis can be applied not only to an individual but also to an entire population, in which case an average genotype based on the allele frequency in a studied population is analyzed. Indeed, individual genotypes may differ from the average population genotype. Even so, predictions based on the average genotype that reflects the characteristics of the population's gene pool will work better than one-size-fits-all recommendations on drug selection and dosing.

This problem has inspired a wealth of studies researching the possibility of optimizing pharmacogenetic recommendations for a given population based on its genetic characteristics. For example, the average effective warfarin dose was shown to be different between European and Caribbean populations [10, 11]. The situation is further compounded by the fact that many populations are genetically different from the "typical" representatives of their race: many African populations are genetically very different from African Americans; Siberian peoples are different from the Chinese; Eastern Slavs and the populations of the Caucasus differ from white Americans. A striking variation in the frequency of ADME genes is reported between the Han Chinese, who make up 98% of China's population, and other Mongoloid minorities, including Uyghurs, Kyrgyz and Kazakhs [12]. Likewise, black African populations are characterized by great genetic variation in the genes coding for P450 isoenzymes, compared to Caucasian populations [9]. Undescribed or rare alleles of ADME genes can substantially contribute to the efficacy and safety of pharmacotherapy. For example, a recent study has identified a novel allele responsible for up to 31% of interpatient warfarin dose variability in African Americans [8]. In another study of variability in warfarin dose response, rare alleles typically found in African Americans were detected in a Puerto Rican population [13]. Another genetic marker (HLA*15:02) associated with severe allergic response to carbamazepine occurs in 8% of Asians and only 1% of Europeans; the prevalence of carbamazepine allergy differs five- to sixfold between these 2 populations [14]. This indicates that direct extrapolation of data generated by foreign studies to the Russian population is absolutely unacceptable.

Russia is a multiethnic state, hence significant variation in responses to pharmacotherapy and the frequency of ADME, drug target and hemostasis-controlling gene variants in the patients of different ethnic/regional descent. In 2019, a systematic review of publications looking into the geographic distribution of alleles of ADME genes in the Russian population [15] discovered that not all ethnic groups and not all studied polymorphisms were equally represented in the literature. Most studies included in the review focused on the major polymorphisms in the genes coding for cytochrome P450 isoenzymes and some factors of the ADME system [16–25]. Many Russian regions and ethnic groups remain understudied.

In view of the foregoing, we set a goal to explore the frequencies of a few significant ADME gene variants across Russia and the neighbor states and to create a draft version of the pharmacogenetic atlas of Russia. The goal implies a systematic approach that ensures the even, wide coverage of the studied territory and can be implemented owing to the data generated by population studies using genome-wide SNP panels.

METHODS

We analyzed a total of 50 populations inhabiting the main regions of North Eurasia and its neighbor states. Samples that formed the basis for our study were provided by the Biobank of North Eurasia, the largest repository of biological specimens collected from indigenous Russian populations and the populations of Russia's neighbor states [26]. The Biobank provided 2011 samples representing 40 different populations; the average sample size was $n = 50$. The initially small datasets of dozens ethnic groups were pooled to form a larger population sample. By contrast, the initially large sample of Russians was divided into 3 geographically different populations. The resultant datasets were genotyped using an Omni-Exome genome-wide array of 4.5 million SNPs (Illumina; USA).

The list of pharmacogenetic markers was compiled from 2 major sources containing some overlapping data: Very Important Pharmacogene (VIP) summaries [27], which describe pharmacogenetic markers involved in pharmacodynamics, pharmacokinetics and, ultimately, response to pharmaceuticals, and a list of markers from the popular commercial OpenArray™ PGx Express Panel (Thermo Fisher Scientific; USA). The resultant list was extended to include a few promising pharmacogenetic markers characterized by distinct racial/ethnic variation in frequency (*CES1*, *PON1*, *IFNL3*, *ITGB3*) and, thus, comprised a total of 95 pharmacogenetic markers. Of them, 55 were present in the Omni-Exome Illumina panel [28].

This dataset was expanded with the published data on 10 populations from China and the European countries that share a border with Russia [29–44]. The total number of the samples representing those 10 populations was 186; the average sample size was $n = 19$. The samples had been previously genotyped using the genome-wide panels of 600 thousand to 1 million SNPs. In some cases, the data on Russia's ethnic groups generated by the published studies were pooled with the corresponding Biobank data. The analysis of this literature-based dataset revealed the presence of 53 SNPs of 95 pharmacogenetic markers included in our list.

The two datasets (data from the Biobank of North Eurasia and the published genome-wide sequencing data) were compared and found to share 45 pharmacogenetic SNPs polymorphic (frequency above 1%) in the populations of North Eurasia. This paper analyzes their geographic distribution. A total of 2,197 samples representing 50 North Eurasian

Table 1. Characteristics of the studied populations

Arbitrary name of population	Ethnic groups constituting the population	Sample size <i>n</i>	Population coordinates	
			Latitude (degrees)	Longitude (degrees)
Altai	Altaiian	40	52	87
Amur	Nanai, Negidal, Nivkh, Orochi, Udege, Ulchi	52	51	139
Armenian	Armenian	102	40	45
Bashkir	Bashkir	47	54	57
Buryat	Buryat	47	53	110
Upper Volga	Mari, Chuvash	58	56	47
Greece	Greek, Macedonian	29	40	22
North Dagestan	Avar, Dargyn, Kubachi, Kumyk, Lak	52	43	47
South Dagestan	Agul, Lezgin, Rutul, Tabasaran	34	42	48
Ashkenazi Jews	Ashkenazi Jews	18	48	16
Other Jews	Jews of Azerbaijan, Georgia, Sephardi Jews of Uzbekistan	16	38	38
Transcaucasia	Azerbaijani, Georgian	40	42	46
West Siberia	Mansi, Selkup, Khanty	45	63	72
West Caucasus 1	Abazin, Abkhaz, Adyg, Sphapsug	51	44	41
West Caucasus 2	Balkar, Kabardin, Karachai, Cherkess	52	44	43
Italy	Italian	17	42	13
Kamchatka	Itelmen, Koryak, Chukchi	70	61	167
Karakalpakstan	Karakalpak, Turkmen	36	40	60
Karelia	Veps, Vod, Izhora, Karel	61	62	33
Kirghizia	Kyrzyg	38	41	75
China	Chinese (Han)	7	34	120
Komi	Komi, Komi-Permiak	54	60	54
Moldova	Bulgarian, Gagauz, Moldovan	36	45	27
Mongolia	Mongol	103	47	101
Mordovia	Moksha, Shoksha, Erzya	43	54	43
Nogai	Karanogai, Astrakhan Nogai, Kuban Nogai, Stavropol Nogai	34	44	43
Ossetia	Iron, Digor	41	43	44
Baltics	Latvian, Lithuanian, Estonian	19	58	24
North Russia	Russians of Arkhangelsk and Vologda regions	67	61	40
Central Russia	Russians of Kostroma, Novgorod, Pskov, Tver, Yaroslavl regions	97	57	38
South Russia	Russians of Belgorod, Voronezh, Kaluga, Kursk, Nizhniy Novgorod, Orlov, Ryazan, Smolensk, Tambov regions, Kuban Cossack	94	53	35
North Siberia	Nganasan, Nenets, Ket	25	67	83
Northern Europe*	German, Polish, Swedish	19	54,54,59	13,18,18
North China	Yi, Nakhi, Oroqen, Xibo, Tu, Tujia, She	8	48	126
South Slavs	Bosnian, Serb, Slovene, Croat	45	45	14
Central Asia	Kazakh, Uzbek, Uighur	53	46	69
Tadzhikistan 1	Tadzhik	28	37	71
Tadzhikistan 2	Pamir peoples, Pushtun, Yaghnobi	48	38	70
Volga-Ural Tatars	Kazan Tatar, Kryashen Tatar, Mishar Tatar	45	55	52
Siberian Tatars	Siberian Tatar	58	57	67
Tuva	Tofalar, Tuvan	59	52	94
Udmurtia	Besermyan, Udmurt	31	57	53
Ukraine	East and West Ukrainian	77	49	29
Central Europe	Hungarian, Romanian, Slovak	26	47	21
Central Caucasus	Ingush, Chechen	32	43	45
Evenk	Evenk	31	60	111
Even	Even	31	61	145
South-west Europe	Basque, Spanish, French	15	47	3
South Siberia	Khakas, Shor	48	54	89
Yakutia	Dolgan, Yukagir, Yakut	28	66	124
Total 50	Total 137			

Note: * — the population of Northern Europe is represented by 3 dots (Germany, Poland and Sweden).

Table 2. Genetic variation of 45 pharmacogenetic markers across North Eurasia and its neighbor states

Gene	SNP	SNP frequency				F _{ST}	Variation pattern
		MEAN	MIN	MAX	MAX-MIN		
ABCB1	rs1045642-G	0.48	0.36	0.68	0.32	0.0123	Uniform
	rs4148738-G	0.46	0.26	0.69	0.42	0.0122	Uniform
APOE	rs429358-G	0.00	0.00	0.02	0.02	0.0080	Focal
CES1	rs2244613-C	0.37	0.10	0.73	0.63	0.0598	Clinal
COMT	rs4680-A	0.46	0.21	0.66	0.44	0.0248	Clinal
CYP1A2	rs12720461-A	0.01	0.00	0.03	0.03	0.0070	Focal
	rs2069526-C	0.06	0.00	0.21	0.21	0.0129	Uniform
	rs762551-C	0.35	0.21	0.52	0.31	0.0100	Uniform
CYP2B6	rs28399499-G	0.00	0.00	0.02	0.02	0.0069	Focal
CYP2C19	rs28399504-G	0.00	0.00	0.05	0.05	0.0140	Focal
	rs41291556-G	0.00	0.00	0.02	0.02	0.0068	Focal
	rs4244285-A	0.14	0.04	0.25	0.21	0.0098	Uniform
	rs4986893-A	0.02	0.00	0.14	0.14	0.0249	Focal
	rs56337013-A	0.00	0.00	0.02	0.02	0.0060	Focal
	rs6413438-A	0.00	0.00	0.01	0.01	0.0056	Focal
CYP2C9	rs1057910-C	0.08	0.01	0.20	0.19	0.0138	Uniform
	rs1799853-T	0.08	0.00	0.23	0.23	0.0270	Focal
	rs28371685-A	0.00	0.00	0.04	0.04	0.0087	Focal
	rs28371686-G	0.00	0.00	0.07	0.07	0.0341	Focal
	rs56165452-G	0.00	0.00	0.01	0.01	0.0054	Focal
CYP2D6	rs28371725-A	0.02	0.00	0.17	0.17	0.0326	Focal
	rs5030862-A	0.00	0.00	0.03	0.03	0.0098	Focal
	rs59421388-A	0.00	0.00	0.02	0.02	0.0072	Focal
CYP3A4	rs12721629-A	0.00	0.00	0.01	0.01	0.0052	Focal
	rs2242480-A	0.11	0.05	0.23	0.18	0.0088	Uniform
	rs4986910-G	0.00	0.00	0.02	0.02	0.0054	Focal
	rs55785340-G	0.00	0.00	0.01	0.01	0.0036	Focal
	rs62471956-A	0.02	0.00	0.12	0.12	0.0192	Focal
CYP3A5	rs10264272-A	0.00	0.00	0.01	0.01	0.0047	Focal
	rs28365083-A	0.00	0.00	0.02	0.02	0.0050	Focal
	rs41303343-AA	0.00	0.00	0.03	0.03	0.0097	Focal
	rs776746-A	0.10	0.03	0.26	0.23	0.0131	Uniform
CYP4F2	*3 rs2108622-A	0.30	0.18	0.49	0.31	0.0154	Uniform
DPYD	*2A rs3918290-A	0.00	0.00	0.03	0.03	0.0094	Focal
Factor II	rs1799963-A	0.01	0.00	0.10	0.10	0.0176	Focal
Factor V Leiden	rs6025-A	0.02	0.00	0.13	0.13	0.0249	Focal
IFNL3	rs8099917-C	0.14	0.02	0.28	0.25	0.0165	Clinal
ITGB3	rs5918-G	0.11	0.02	0.28	0.26	0.0178	Uniform
MTFHR	rs1801131-C	0.31	0.16	0.49	0.32	0.0100	Clinal
	rs1801133-A	0.27	0.03	0.51	0.48	0.0253	Uniform
PON1	rs662-G	0.37	0.19	0.69	0.51	0.0264	Clinal
SLCO1B1	rs4149056-G	0.15	0.06	0.27	0.21	0.0100	Uniform
TPMT	rs1142345-G	0.01	0.00	0.05	0.05	0.0079	Focal
	rs1800460-A	0.01	0.00	0.04	0.04	0.0063	Focal
VKORC1	rs9923231-T	0.57	0.3	0.95	0.65	0.0699	Clinal

populations were genotyped for those 45 SNPs. Characteristics of the studied populations are provided in Table 1; the SNPs are listed in Table 2.

Genome-wide genotyping data were pooled and analyzed; allele frequencies were calculated in PLINK 1.9 [45, 46]. A frequency matrix was generated for 45 pharmacogenetic markers in 50 populations of Russia and its neighbor states

based on the genotypes of 2,197 samples (average sample size $n = 44$). Using the matrix, we constructed the gene geographic maps of allele frequencies in the populations of North Eurasia and bordering states. The maps were created in GeneGeo [47–49] using average weighted interpolation; the radius of influence was set to 2,000 km; the weight function power was set to 3 [48]. On the maps, the regions lying farther than 2,000 km

Table 3. Characteristics and pharmacogenetic significance of 12 ADME markers shown in Fig. 1–3

Gene	SNP	Allele	Chromosome	Position in GRCh37	Position in GRCh38	Function and significance
<i>ABCB1</i>	rs4148738	G	7	87138645	87509329	P-glycoprotein is an ATP-dependent transporter involved in transporting biological substrates across the cell membrane. Its primary function is to regulate the permeation of various compounds, including xenobiotics, across biological barriers. It is also responsible for drug transport across the blood-brain barrier. The rs4148738 polymorphism of the <i>ABCB1</i> gene is associated with plasma concentrations of direct oral anticoagulants, which may affect the efficacy and safety of these drugs
<i>CES1</i>	rs2244613	C	16	55844609	55810697	Carboxylesterases are enzymes that hydrolyze chemical compounds containing complex carboxylic acids, amide and thioester functional groups. They play an important role in hydrolyzing drugs into nonactive metabolites (dabigatran, capecitabine, etc.)
<i>COMT</i>	rs4680	A	22	19951271	19963748	Catechol-O-methyltransferase is an enzyme involved in the regulation of dopamine activity in the prefrontal cortex. It participates in promoting sensitivity to neuroleptics and is associated with adverse effects. The rs4680 polymorphism in the <i>COMT</i> gene determines the efficacy of therapy with nicotine, opioids and some antipsychotic drugs
<i>CYP2C19</i>	rs4986893	A	10	96540410	94780653	Enzymes of the P450 cytochrome family participate in the metabolism of drugs and xenobiotics. Reduced activity of these enzymes may affect the efficacy and safety of proton pump inhibitors, NSAID, clopidogrel, and other drugs
<i>CYP2C9</i>	rs1057910	C	10	96741053	94981296	
<i>Factor II</i>	rs1799963	A	11	46761055	46739505	Prothrombin (blood-clotting factor II) is a vitamin K-dependent glycoprotein synthesized in the liver as an inactive zymogen. It plays a significant role in hemostasis and thrombosis. The rs1799963 allele of the <i>Factor II</i> gene may increase the risk of venous thromboembolism in its carriers exposed to oral contraceptives
<i>Factor V Leiden</i>	rs6025	A	1	169519049	169549811	Proaccelerin (blood-clotting factor V) is a soluble β -globulin. Factor V Leiden increases the risk of primary and recurrent venous thromboembolism three-to-sixfold therefore should be accounted for when prescribing oral contraceptives to the carriers of this mutation
<i>IFNL3</i>	rs8099917	C	19	39743165	39252525	Interferon lambda 3. The rs8099917 polymorphism in the <i>IFNL3</i> gene reduces the efficacy of therapy with interferons and ribavirin in patients with viral liver disease
<i>ITGB3</i>	rs5918	G	17	45360730	47283364	Integrin beta 3 is a component of glycoprotein IIb/IIIa responsible for the aggregation of platelets. Mutations in the <i>ITGB3</i> gene often lead to Glanzmann thrombasthenia, a common inherited blood clotting disorder. Carriers of the rs5918 polymorphism can develop moderate to severe mucosal bleeding. The mutation can determine the efficacy of antiplatelet therapy with clopidogrel and NSAID
<i>MTFHR</i>	rs1801131	C	1	11854476	11794419	MTHFR is an enzyme essential for the metabolism of folates and methionine. Homozygous carriers of its mutant alleles produce only 30% of the normal MTHFR amount. The rs1801131 polymorphism determines the interpersonal variability in the efficacy and safety of methotrexate, capecitabine, fluorouracil and some other drugs
<i>TPMT</i>	rs1800460	A	6	18139228	18138997	TPMT (thiopurine S-methyltransferase) is mostly known for its role in the metabolism of thiopurine derivatives, including azathioprine, 6-mercaptopurine and 6-thioguanine. TPMT catalyzes the S-methylation of thiopurine-based drugs. Mutations in the <i>TPMT</i> gene result in the reduced methylation and poor inactivation of 6-mercaptopurine and, therefore, increase its toxicity
<i>VKORC1</i>	rs9923231	T (A)	16	31107689	31096368	The <i>VKORC1</i> genes encodes the subunit 1 of the vitamin K1 epoxide reductase complex. This enzymatic complex is responsible for the reduction of vitamin K1 2,3-epoxide to its active form, which is crucial for effective blood coagulation. Warfarin dosage should be lowered for the carriers of this allele. In the literature, the allele is described as C>T or G>A (based on the complementary strand)

away from the studied populations are highlighted in white. In population genetics, it is imperative to study an indigenous population in any given territory; it is this population's habitat that is shown on a gene geographic map. The gene pools of immigrant populations can be easily reconstructed from the map using information about their migration sources.

To visualize the cartographic models, we applied a "universal" scale of frequency intervals developed for genetic markers with a broad variation in frequency [50]. Frequencies constituting the main body of the frequency "spectrum" (5–60%) are represented by equally sized intervals that follow each other at an increment of 5%; the space of low frequencies is represented by 3 intervals, while the space of high frequencies, by intervals with 10% increment. This grading approach allows keeping the number of intervals within a reasonable limit (it

would be a challenge to distinguish between over 17 colors on a map). Frequencies below 1% (i.e., below the below the criterion of polymorphism 1%) are shown in the contrasting gray.

RESULTS

The gene geographic maps showing allele frequency variation for 45 pharmacogenetic markers in the populations of North Eurasia and its neighbor states are available on the web site (see the Appendix). The main characteristics of their allele frequency variation are provided in Table 2. Although the geographic distribution of each marker is unique (there are no two identical maps), the analysis of the maps reveals 3 prominent patterns: clinal variation, uniform distribution and focal variation (see Fig. 1–3). Since it would be impossible to describe all the

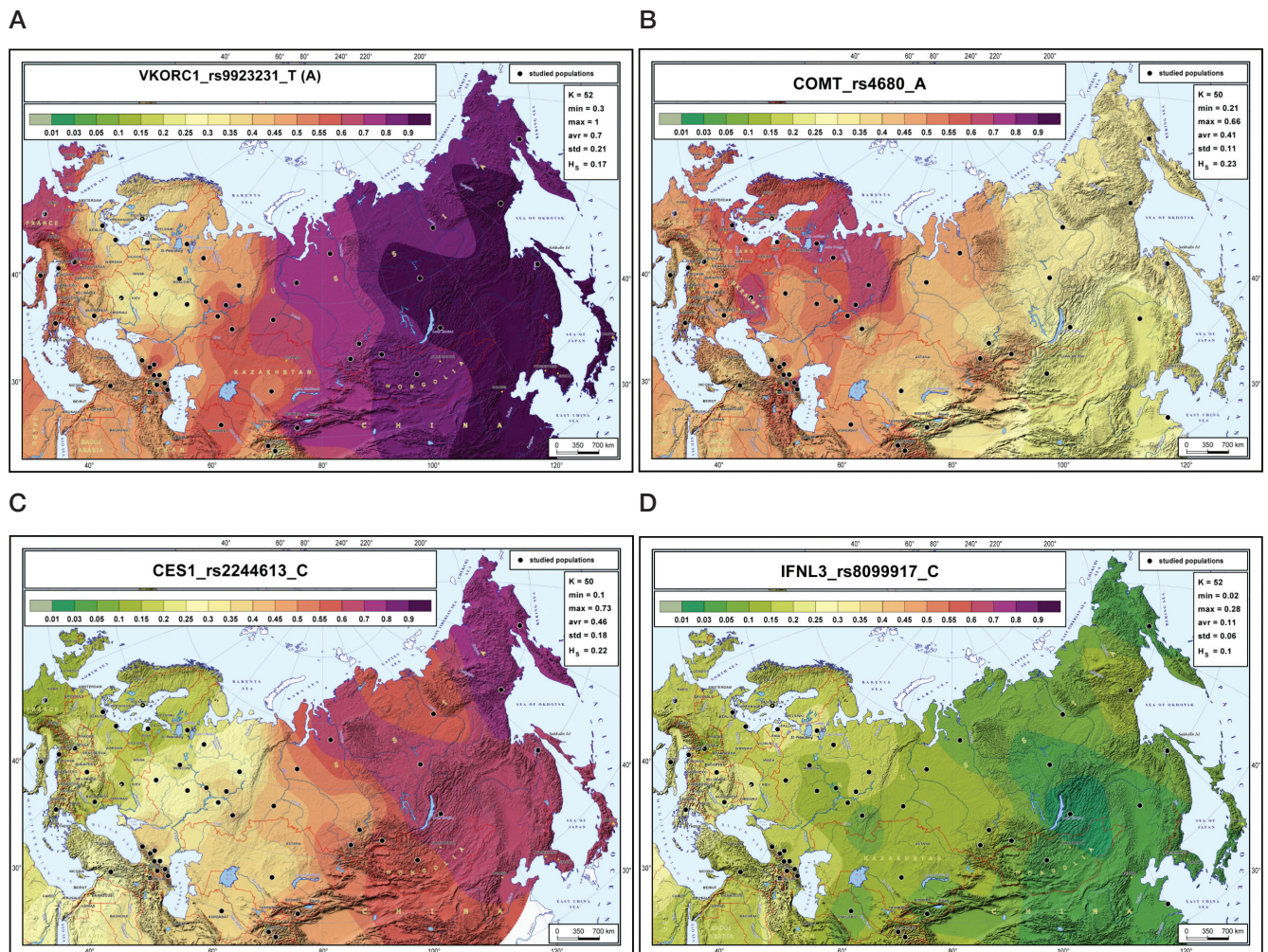


Fig. 1. Maps of frequency distribution for pharmacogenetic markers following the clinal variation pattern: *VKORC1* (rs9923231-T(A)) (A); *COMT* (rs4680-A) (B); *CES1* (rs2244613-C) (C); *IFNL3* (rs809917-C) (D)

maps in this publication, we selected 12 markers to illustrate each pattern (4 maps per pattern). These markers are briefly described in Table 3.

Clinal variation

The clinal (gradient) pattern of allele frequency variation along the East–West axis dominates the gene pool of North Eurasian populations. The pattern was identified during the analysis of classic markers [49, 51] and then corroborated using the data on DNA markers [49]. Moreover, this pattern is typical for the most ancient populations of North Eurasia [49, 52]. This may suggest that the clinal variation pattern emerged in the earliest days of the North Eurasian population and has persisted ever since. Therefore, it is only natural that some of the studied pharmacogenetic markers are characterized by a very distinct clinal variation pattern (see Fig. 1).

The distribution of *VKORC1* rs9923231-T(A) frequencies on the gene geographic map follows a very distinct clinal pattern (Fig. 1A): the lowest frequencies are observed in the West of European Russia and the Caucasus, increasing gradually toward Eastern Eurasia. Notably, the frequencies lying between the Western and Eastern “extremes” of the frequency spectrum are found in West Siberia (but not in the Ural region), similar to the first PCA component of the North Eurasian gene pool and the map of archeological Paleolithic sites [49, 52]. The lowest frequency is observed in the lands inhabited by Eastern Slavs

and the populations of the Baltic region ($0.3 \leq q < 0.4$; here and below frequency is designated by q). From there, the area of low frequencies stretches eastward to the Volga river and includes the habitat of the indigenous population of Mordovia. The frequency of the polymorphism gradually increases toward the West of Europe and southwards. The domain of high frequencies ($q > 0.80$) begins in the East of North Eurasia with Tuvan and Mongol populations; the polymorphism reaches its frequency peak ($> 90\%$) on the Eastern coast of the continent in the indigenous peoples of the Amur region, Evens, Evenks and continental Buryats.

The map of *COMT* (rs4680-A) frequencies (Fig. 1B) demonstrates the same pattern of clinal variation, but the frequency range is narrower (0.44 vs 0.65 ; see Table 2). Peak values ($0.6 \leq q < 0.7$) are observed in the West; however, in the East they fall by only 20–30%. Consequently, the genetic diversity there is much lower ($F_{ST} = 0.02$; see Table 2) than on the previous map ($F_{ST} = 0.07$; see Table 2). High frequencies ($0.5 \leq q < 0.7$) are observed in the European part of the continent, but their domain expands to the East, covering the Mari, Chuvash, Udmurt, and Tatar populations of the Volga-Ural region. The region of average frequencies spreads further to the East, spanning the Yenisey River basin. But by and large, the area of high frequencies ($0.2 \leq q < 0.35$) corresponds to the area of low frequencies on the previous map.

For *CES1* (rs2244613-C) (Fig. 1C), the pattern of clinal variation is almost a copy of the *VKORC1* (rs9923231-T)

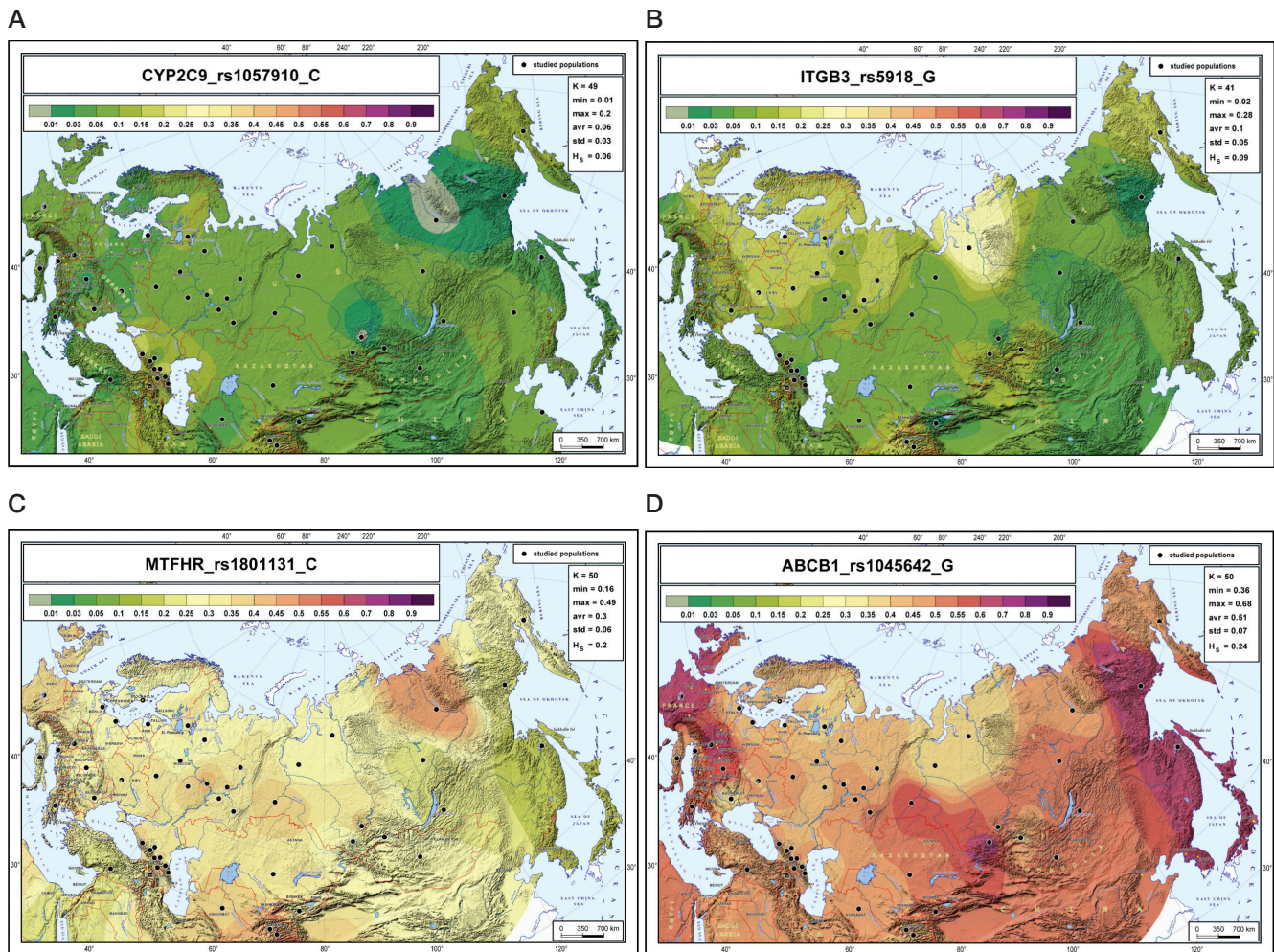


Fig. 2. Maps of frequency distribution for pharmacogenetic markers following the uniform distribution pattern: *CYP2C9* (rs1057910-C) (A); *ITGB3* (rs5918-G) (B); *MTFHR* (rs1801131-C) (C); *ABCB1* (rs1045642-G) (D)

pattern (see Fig. 1A): the domain of maximum frequencies ($0.55 \leq q < 0.8$) is in the East of North Eurasia, but the domain of average frequencies ($0.40 \leq q < 0.50$) is again in West Siberia. Minimum frequency values ($0.10 \leq q < 0.25$) are observed in Western Europe and are lower than in the populations of Eastern Slavs.

For *IFNL3* (rs8099917-C) (Fig. 1D), the pattern of clinal variation is weakened by the narrow frequency range (0.25; see Table 2). However, similar to the previous map in Fig. 1B, the lowest frequencies of this polymorphism are concentrated in the East of North Eurasia, whereas its maximum frequencies ($0.24 \leq q \leq 0.28$) are found in Europe (from Karelia to Italy) and in the Caucasus, although they did not cover the entire European mainland. The narrow range of low frequencies ($0.02 \leq q \leq 0.28$) makes the pattern of clinal variation less distinct due to the growing role of a sampling error.

Uniform distribution

A feature with generally uniform spatial distribution will not necessarily occur at the same frequency at all points in space. Uniform distribution characterizes alleles that occur in almost every population under study but do not exhibit distinct clinal variation on the regional scale, although certain patterns might be identified in some parts of the studied region. Surges or dramatic falls in frequency in some parts of the region suggest a sampling error and the need to revise the number of the populations representing these areas and the sample size.

But if these extreme behaviors are observed in more than one neighboring population, they are not evident of random frequency fluctuations but rather indicate a distinct frequency variation pattern in the studied group of populations.

The distribution of *CYP2C9* (rs1057910-C) across North Eurasia (Fig. 2A) is strikingly uniform. The average frequency ($q = 0.08$; Table 2) of the marker is low, slightly above 5%. But in most ethnic groups inhabiting North Eurasia, the marker occurs at nearly identical frequency which varies within a narrow range ($0.03 \leq q \leq 0.10$). The only identifiable pattern pertains to the increased frequency of this marker observed in almost all populations of the Caucasus and the Transcaucasian region. Single spikes in Karelia, Tadjikistan and Chukotka and a steep decline below 1% (the gray zone of Yakutia) do not form distinct patterns.

Likewise, *ITGB3* (rs5918-G) (Fig. 2B) is characterized by uniform geographic distribution, its frequency varies within a very narrow range ($0.02 \leq q \leq 0.28$), and its average frequency is low ($q = 0.11$). However, deviations from the overall uniform distribution now form a pattern: the area of high frequencies stretches in an almost uninterrupted fashion from the Baltic region to the estuary of the Yenisey River. The second frequency peak occurs in the South of North Eurasia: in Tadjikistan and South Siberia.

The map of *MTFHR* (rs1801131-C) frequency variation (Fig. 2C) shows that the marker is uniformly distributed across the vast territory, although its average frequency ($q = 0.31$; see Table 2) is 3 times higher than the average frequencies from the previous

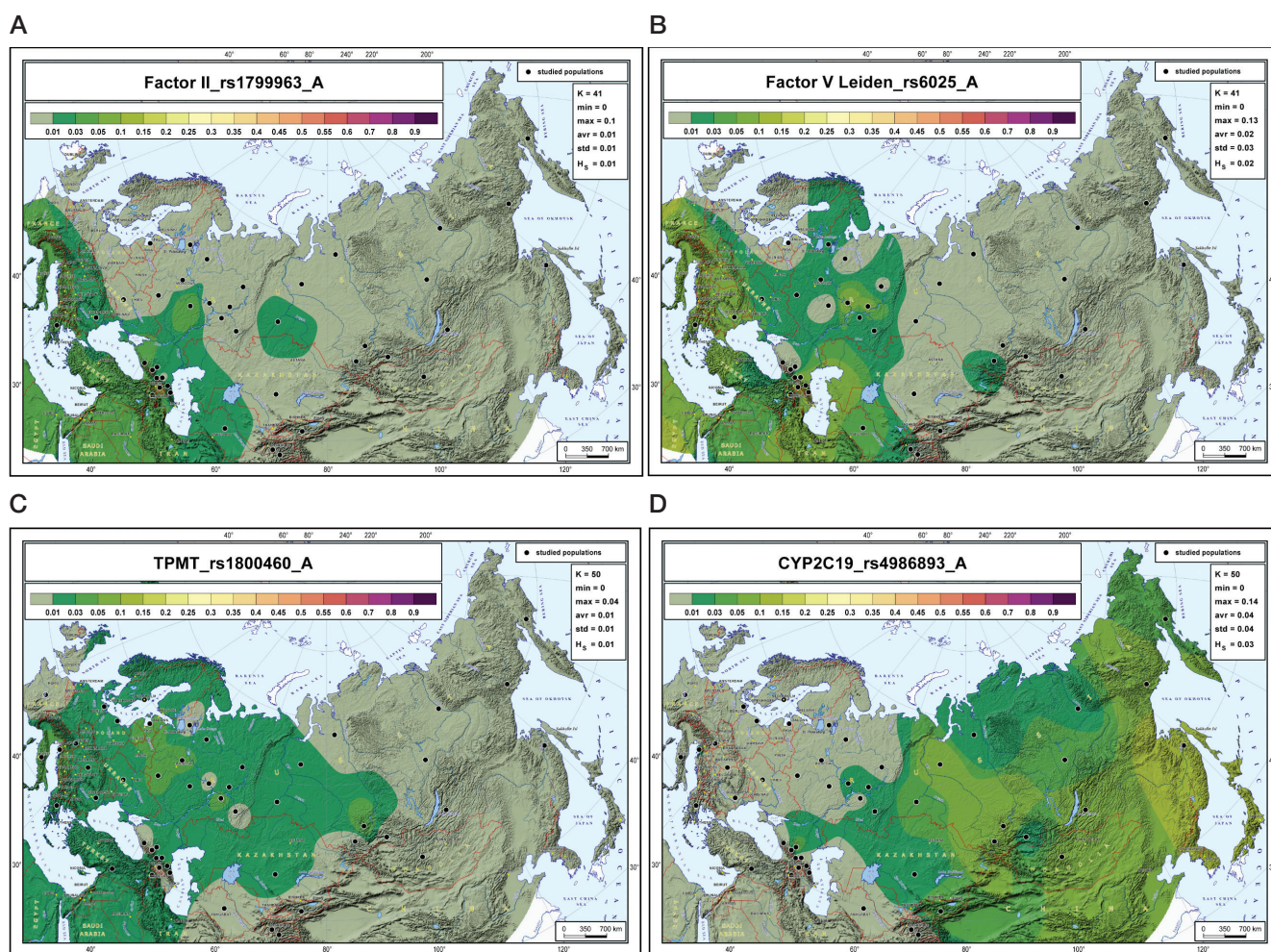


Fig. 3. Maps of frequency distribution for pharmacogenetic markers following the focal variation pattern: *Factor II* (rs1799963-A) (A); *Factor V Leiden* (rs6025-A) (B); *TPMT* (rs1800460-A) (C); *CYP2C19* (rs4986893-A) (D)

maps. The domain of average frequencies ($0.25 \leq q < 0.35$) spreads throughout almost the entire North Eurasia, with two peaks in Yakutia and Tadjikistan.

The average frequency of *ABCB1* (rs1045642-G) (Fig. 2D) is $q = 0.48$ (see Table 2); the scope of frequency variation is the same as on the previous map (0.32; see Table 2), slightly exceeding the scope of frequency variation in Fig. 2A (0.19) and Fig. 2B (0.26). Average frequencies ($0.40 \leq q < 0.55$) cover almost the entire territory of North Eurasia and the neighbor states. Slightly lower frequencies are observed in the North of Middle Siberia in the cluster of Khanty, Mansi, Ket, Nenets and Selkup populations, as well as in two European populations (in Karelia and Moldova). Small areas of increased frequencies are scattered across North Eurasia: frequencies $q > 0.6$ occur in the South-West of Europe, in Ashkenazi Jews, Altaians, Evens and Amur region peoples. These local deviations do not disrupt the generally uniform distribution of the marker.

Similar to the previous set of maps, this set demonstrates that the variation pattern does not depend on the position of the frequency on the frequency spectrum: both high and low frequencies can follow a pattern of uniform distribution across vast territories.

Focal variation

Uniform distribution and clinal variation are typical for all pharmacogenetic markers that occur almost everywhere in North Eurasia and the neighbor states. If a marker occurs in

only one (large as it may be) part of the studied region and is not found in its other parts, the pattern of its distribution can be arbitrarily called "focal variation". The "focal" frequencies are normally low. On the map, populations that either do not carry this marker in their genome or have it at $< 1\%$ frequency are highlighted in gray.

Factor II (rs1799963-A) (Fig. 3A) has a frequency range of $0 \leq q \leq 0.10$ (see Table 2). Its domain stretches uninterruptedly across the Southern regions of North Eurasia, from the Mediterranean to Turkmenia, covering all populations of the Caucasus, excluding the Ingush and Chechen peoples, and reaches its maximum frequency in the Transcaucasia region. Only two populations outside this area are carriers of the marker: Mordovians ($q = 0.05$) and Siberian Tatars ($q = 0.02$).

The frequency of *factor V Leiden* (rs6025-A) (Fig. 3B) varies within almost the same range ($0 \leq q \leq 0.13$; Table 2), but the area is concentrated in its much vaster, covering the territories on the previous map and the substantial share of the population in Eastern Europe and the Volga-Ural region. Peak frequencies are observed in the South of Europe, Caucasus, and the Transural region (Udmurt, Mari and Chuvash peoples). Outside this area, the polymorphism occurs only in the Altai region, where its frequency is below 1%.

Although the frequency range of *TPMT* (rs1800460-A) (Fig. 3C) is extremely narrow ($0 \leq q \leq 0.04$) and its peak frequency is below 5%, its domain stretches across the entire Europe (except for the far South-West, which might be explained by the small sample size representing Basque, Spanish, and

French populations; $n = 29$), Ural, West Siberia, Kazakhstan, to Khakassia. Interestingly, this marker is not found in most populations of the Caucasus, except for the Central Caucasus (the Ingush, Ossetians, Chechens) and North Dagestan.

On its genogeographic map, *CYP2C19* (rs4986893-A) (Fig. 3D) occupies an enormous territory that spans the Asian part of the region, except for Tadjikistan and Turkmenia, protruding into the European mainland, including Ciscaucasia (Nogais), the Caucasus (Ossetians), and the Volga region (the Mari, Udmurt, Tatar, and Chuvash). The frequency peaks in the Far East in the populations of the Amur region ($q = 0.14$), Evens ($q = 0.10$) and in Khakassia ($q = 0.10$). For this marker, the scope of frequency variation is narrow ($0 \leq q \leq 0.14$; see Table 2), and the frequencies themselves are comparable to the markers in Fig. 3 that cover a much smaller area. Maps in Fig. 3 demonstrate that markers with low frequencies and low variation in frequency can cover from small to very extensive areas.

DISCUSSION

There are two types of differences between populations that matter in pharmacogenetics: differences in the presence of pharmacogenetic markers and differences in their frequencies. The first type describes the alleles that are widespread in one population but almost non-existent in the other (some can even be region-specific). The second type pertains to different (sometimes contrasting) allele frequencies of the marker in different populations.

Of 3 patterns identified in our study, clinal variation and uniform distribution describe differences in allele frequencies between populations across the studied territory; importantly, the markers characterized by these two distribution patterns are found in almost every population of the studied territory. The focal variation pattern reflects differences in the presence of pharmacogenetic markers between different populations.

The clinal variation pattern (see Fig. 1) is characterized by the presence of 2 extremes on the frequency spectrum, with an area of intermediate values between them. The extremes are more common for the markers that exhibit great variation in frequency, although this may not always be the case: markers that have a very narrow frequency range often form a geographically distinct gradient pattern. In North Eurasia, the clinal variation pattern was observed for 6 of 45 markers (13% of the maps in the atlas; see Table 2).

With uniform distribution (see Fig. 2), differences in frequencies are either small or their spikes and falls occur sporadically across the entire studied territory. However, there are areas where allele frequencies follow certain patterns, which might become the object of future research. In North Eurasia, 12 of 45 markers followed the pattern of uniform distribution (27% of the maps; see Table 2).

A marker that follows the focal variation pattern is common for one population, but occurs at almost negligible frequency below 1% or is totally absent in another. The area where this marker occurs may vary in size (Fig. 3): from small compact foci to vast territories. Attempts to identify the pattern of marker variation in the "focal" area should account for the sampling error since small sample sizes and low marker frequencies can skew the picture. In our study, focal variation was the most common pattern observed for 27 of 45 markers (60% of the maps; see Table 2).

A cartographic atlas is a systematic collection of intertwined, mutually complementary maps that form a single entity [52]. Our pharmacogenetic atlas includes maps for 45

ADME markers in 21 genes, but the total number of markers with proved pharmacogenetic significance approaches a few hundred [2, 53]. This is one of the reasons why our collection of maps, which meets the formal requirements for a cartographic atlas [52], is referred to in this publication as the first draft version of the atlas. Our atlas covers only the key pharmacogenetic markers included in a genome-wide Illumina panel; other important markers are not included. The second reason is rooted in how comprehensive the study is. Our maps were based on a large number of populations ($K = 50$), but the average sample size was relatively small ($n = 44$); therefore, the maps reveal only general patterns of distribution for every studied marker, and the error in the frequency of a given marker in a given population might be substantial. This indicates the need for further data accumulation, which has already been initiated by a number of researchers [15].

Nevertheless, this first version of the pharmacogenetic atlas provides valuable information on the variation of each studied marker and allows for general conclusions to be drawn. For example, the atlas shows that the average frequency of the marker and its sporadic occurrence in individual populations should not be interpreted as an indication of its geographic distribution pattern. In order to identify the distribution pattern, other tools should be employed, of which a gene geographic map appears to be the most suitable. The analysis of all maps in aggregate can identify areas that display their own patterns of frequency distribution and therefore should be subjected to additional analysis. Among such areas are Western Europe, the Caucasus, the Far East, North Siberia, and some others. Other important patterns, not limited to the 3 patterns described in this paper are expected to be revealed across this vast territory as more populations are studied and more data is accumulated.

Our collection of maps provides a wealth of information that can be exploited in pharmacogenetic research: they show frequencies of each of 45 SNPs in different populations across the entire territory of Russia and the neighbor states. However, caution should be exercised when using the obtained figures, considering the following limitations: the sampling error; the fact that these frequencies represent indigenous populations; and the fact that they represent the entire population but not its single members.

The first limitation is an average sample size. In our study, it was 44 people (88 chromosomes). Consequently, if a marker has a frequency of 25%, the CI will include the margin of error of $\pm 6\%$. This CI is much narrower for the territory of Russia, where the samples are larger, and broader for its neighbor states (the samples of their populations available in the literature are quite small). To account for CI fluctuations, the maps for marker frequencies should be accompanied by reliability maps [55].

The second limitation is related to the fact that each region of the map is represented by the gene pool of its indigenous peoples. Therefore, the frequency of a pharmacogenetic marker in an ethnically Russian population of Siberia should be estimated from the maps of migration sources for this population, e.g. Central Russia, but not the maps of Siberia. Thus, knowing the genealogy of a patient, one can construct their map-based pharmacogenetic portrait in a situation when genotyping cannot be performed.

The third limitation is typical for all pharmacogenetic population studies: the genotype of an individual patient does not necessarily reproduce the average characteristics of the gene pool the patient represents; still, these average characteristics are the best possible approximation of an individual genotype. Therefore, it is possible to apply standard pharmacogenetic protocols to an individual patient, drawing on the average genotype based on the allele frequency

in a given population. This approach has obvious limitations, but there is an internationally held opinion that such protocols are the best solution in cases when an individual genotype of the patient is unknown or genotyping poses a financial burden. In our future studies, we will assess the significance of population frequencies for the economy by investigating the costs of inadequate drug dosing.

CONCLUSION

Based on the data on 50 populations of Russia and its neighbor states, we created a cartographic atlas showing the geographic distribution of 45 pharmacogenetic markers. The gene geographic atlas is composed of 45 maps that demonstrate the frequency of the studied pharmacogenetic markers in a variety of populations across Russia. The maps are straightforward and easy to understand by specialists from different fields and with different qualifications.

The pharmacogenetic atlas of Russia and its neighbor states reveals 3 major spatial distribution patterns of ADME,

drug target and hemostasis-controlling genes. Some of the studied markers follow the clinal variation pattern (a gradient change in frequency along the East–West axis), which is the main pattern of allele frequency variation in North Eurasian populations. Uniform distribution singles out a group of markers that occur at average frequency in most Russian regions. Focal variation is observed in the markers that are specific to a certain ethnicity/region and are absent on other populations.

The lack of information about the frequency of pharmacogenetic markers in Russia impedes the adoption of personalized treatment algorithms originally developed for West European populations. There are countless examples of insufficient efficacy of standardized dosing protocols in non-European populations. Adaptation of the existing protocols to Russian populations require information about the geographic distribution of clinically significant markers in Russian populations. This study has generated an extensive body of systematized data and thereby may contribute to research in this important clinical field that demands thorough planning and additional large-scale studies.

References

1. Core ADME Gene List [internet]. www.pharmaadme.org. ©2020 — [cited 2020 Nov 4]. Available from: http://pharmaadme.org/joomla/index.php?option=com_content&task=view&id=12&Itemid=27.
2. pharmgkb.org [internet]. ©2001-2020 PharmGKB [cited 2020 Nov 4]. Available from: <https://www.pharmgkb.org>.
3. Scott SA, Sangkuhl K, Stein CM, Hulot JS, Mega JL, Roden DM, et al. Clinical Pharmacogenetics Implementation Consortium guidelines for CYP2C19 genotype and clopidogrel therapy: 2013 update. *Clin Pharmacol Ther.* 2013; 94 (3): 317–23. DOI: 10.1038/clpt.2013.105.
4. Lima JJ, Thomas CD, Barbarino J, Desta Z, Van Driest SL, El Rouby N, et al. Clinical Pharmacogenetics Implementation Consortium (CPIC) Guideline for CYP2C19 and Proton Pump Inhibitor Dosing. *Clin Pharmacol Ther.* 2020. DOI: 10.1002/cpt.2015. Epub ahead of print.
5. Johnson JA, Caudle KE, Gong L, Whirl-Carrillo M, Stein CM, Scott SA, et al. Clinical Pharmacogenetics Implementation Consortium (CPIC) Guideline for Pharmacogenetics-Guided Warfarin Dosing: 2017 Update. *Clin Pharmacol Ther.* 2017; 102 (3): 397–404. DOI: 10.1002/cpt.668.
6. Beutler E. G6PD: population genetics and clinical manifestations. *Blood Rev.* 1996; 10 (1): 45–52. DOI: 10.1016/s0268-960x(96)90019-3.
7. Urban TJ. Race, ethnicity, ancestry, and pharmacogenetics. *Mt Sinai J Med.* 2010; 77 (2): 133–9. DOI:10.1002/msj.20168.
8. Hernandez W, Gamazon AF, Aquino-Michaels K, Smithberger E, O'Brien TJ, Harralson AF, et al. Integrated analysis of genetic variation and gene expression reveals novel variant for increased warfarin dose requirement in African Americans. *J Thromb Haemost.* 2017; 15 (4): 735–43. DOI: 10.1111/jth.13639.
9. Rajman I, Knapp L, Morgan T, Masimirembwa C. African Genetic Diversity: Implications for Cytochrome P450-mediated Drug Metabolism and Drug Development. *EBioMedicine.* 2017; 17: 67–74. DOI: 10.1016/j.ebiom.2017.02.017.
10. Duconge J, Ramos AS, Claudio-Campos K, Rivera-Miranda G, Bermúdez-Bosch L, Renta JY, et al. A Novel Admixture-Based Pharmacogenetic Approach to Refine Warfarin Dosing in Caribbean Hispanics. *PLoS One.* 2016; 11 (1): e0145480. DOI: 10.1371/journal.pone.0145480.
11. Mizzi C, Dalabira E, Kumuthini J, Dzimiri N, Balogh I, Başak N, et al. A European Spectrum of Pharmacogenomic Biomarkers: Implications for Clinical Pharmacogenomics. *PLoS One.* 2016; 11 (9): e0162866. DOI: 10.1371/journal.pone.0162866.
12. Li J, Lou H, Yang X, Lu D, Li S, Jin L, et al. Genetic architectures of ADME genes in five Eurasian admixed populations and implications for drug safety and efficacy. *J Med Genet.* 2014; 51 (9): 614–22. DOI: 10.1136/jmedgenet-2014-102530.
13. Duconge J, Ruaño G. Admixture and ethno-specific alleles: missing links for global pharmacogenomics. *Pharmacogenomics.* 2016; 17 (14): 1479–82. DOI: 10.2217/pgs-2016-0115.
14. Lonjou C, Thomas L, Borot N, Ledger N, de Toma C, LeLouet H, et al. A marker for Stevens-Johnson syndrome ...: ethnicity matters. *Pharmacogenomics J.* 2006; 6 (4): 265–8. DOI: 10.1038/sj.tpj.6500356.
15. Mirzaev KB, Fedorinov DS, Ivashchenko DV, Sychev DA. ADME pharmacogenetics: future outlook for Russia. *Pharmacogenomics.* 2019; 20 (11): 847–65. DOI: 10.2217/pgs-2019-0013.
16. Mustafina OE, Tuktarova IA, Karimov DD, Somova Rs, Nasibullin TR. CYP2D6, CYP3A5, and CYP3A4 gene polymorphisms in Russian, Tatar, and Bashkir populations. *Russ J Genet.* 2015; 51 (1): 98–107. DOI: 10.1134/S1022795415010081.
17. Korytina G, Kochetova O, Akhmadishina L, Viktorova E, Victorova T. Polymorphisms of cytochrome p450 genes in three ethnic groups from Russia. *Balkan Med J.* 2012; 29 (3): 252–60. DOI:10.5152/balkanmedj.2012.039.
18. Kochetova OV, Korytina GF, Akhmadishina LZ, Victorova TV, Iskhakova GM. Analysis of the cytochrome P450 1A1 (CYP1A1) gene polymorphism in the ethnic groups of the republic of Bashkortostan. *Russ J Genet.* 2008; 44 (12): 1454–60. DOI: 10.1134/S1022795408120107.
19. Korytina GF, Celousova OS, Akhmadishina LZ, Kochetova OV, Babenkova LI, Victorova TV. Association of polymorphism of inflammatory mediators (IL1B, TNFA, LTA, IL8, IL6, ILRN, ILR4, TGFB, TLR4, DBF) genes with chronic lung disease in children. *Meditinskaya genetika.* 2008; 7 (2): 17–25. Russian.
20. Korytina GF, Akhmadishina LZ, Victorova TV. Frequencies of CYP1B1 and CYP2F1 polymorphic variants in three ethnic groups of Bashkortostan and in patients with chronic obstructive pulmonary disease. *Mol Biol.* 2010; 44 (1): 28–36. DOI: 10.1134/S002689331001005X.
21. Akhmadishina LZ, Korytina GF, Mingazova SR, Yanbaeva DG, Bakirov AB, Victorova TV. Rol' polimorfizma genov CYP1A1, EPHX1, GSTM1, GSTT1 i GSTP1 v razviti khronicheskikh bronkhitov professional'nogo geneza. *Ekologicheskaya genetika.* 2005; (1): 11–7. Russian.
22. Fedorova YY, Karunas AS, Nurgalieva AKh, Gra OA, Gimalova GF, Ramazanova NN, et al. Role of xenobiotic-metabolizing gene polymorphisms in allergic diseases susceptibility in Tatars.

- Meditsinskaya genetika. 2010; 9 (6): 28–35. Russian.
23. Polonikov A, Kharchenko A, Bykanova M, Sirotina S, Ponomarenko I, Bocharova A, et al. Polymorphisms of CYP2C8, CYP2C9 and CYP2C19 and risk of coronary heart disease in Russian population. *Gene*. 2017; 627: 451–9. DOI: 10.1016/j.gene.2017.07.004.
 24. Gaikovitch EA, Cascorbi I, Mrozikiewicz PM, Brockmüller J, Frötschl R, Köpke K, et al. Polymorphisms of drug-metabolizing enzymes CYP2C9, CYP2C19, CYP2D6, CYP1A1, NAT2 and of P-glycoprotein in a Russian population. *Eur J Clin Pharmacol*. 2003; 59 (4): 303–12. DOI: 10.1007/s00228-003-0606-2.
 25. Gra O, Mityaeva O, Berdichevets I, Kozhekbaeva Z, Fesenko D, Kurbatova O, et al. Microarray-based detection of CYP1A1, CYP2C9, CYP2C19, CYP2D6, GSTT1, GSTM1, MTHFR, MTRR, NQO1, NAT2, HLA-DQA1, and ABO allele frequencies in native Russians. *Genet Test Mol Biomarkers*. 2010; 14 (3): 329–42. DOI: 10.1089/gtmb.2009.0158.
 26. Balanovska EV, Zhabagin MK, Agdzhoyan AT, Chukhryaeva MI, Markina NV, Balaganskaya OA, et al. Population biobanks: Organizational models and prospects of application in gene geography and personalized medicine. *Russ J Genet*. 2016; 52 (12): 1227–43. DOI: 10.1134/S1022795416120024.
 27. Whirl-Carrillo M, McDonagh EM, Hebert JM, Gong L, Sangkuhl K, Thorn CF, et al. Pharmacogenomics Knowledge for Personalized Medicine. *Clinical Pharmacology & Therapeutics*. 2012; 92 (4): 414–17. pharmgkb.org [internet]. ©2001–2020 «Very Important Pharmacogene» (VIP) [cited 2020 Nov 4]. Available from: <https://www.pharmgkb.org/vips/>.
 28. Balanovsky OP, Gorin IO, Zapisetskaya YuS, Golubeva AA, Kostyukova ES, Balanovska EV. Interactions between gene pools of Russian and Finnish-speaking populations from Tver region: analysis of 4 million SNP markers *Bulletin of RSMU*. 2020; 6 (in print).
 29. Behar DM, Yunusbayev B, Metspalu M, Metspalu E, Rosset S, Parik J, et al. The genome-wide structure of the Jewish people. *Nature*. 2010; 466 (7303): 238–42. DOI: 10.1038/nature09103.
 30. Behar DM, Metspalu M, Baran Y, Kopelman NM, Yunusbayev B, Gladstein A, et al. No evidence from genome-wide data of a Khazar origin for the Ashkenazi Jews. *Hum Biol*. 2013; 85 (6): 859–900. DOI: 10.3378/027.085.0604.
 31. Chaubey G, Metspalu M, Choi Y, Mägi R, Romero IG, Soares P, et al. Population genetic structure in Indian Austroasiatic speakers: the role of landscape barriers and sex-specific admixture. *Mol Biol Evol*. 2011; 28 (2): 1013–24. DOI: 10.1093/molbev/msq288.
 32. Di Cristofaro J, Pennarun E, Mazières S, Myres NM, Lin AA, Temori SA, et al. Afghan Hindu Kush: where Eurasian sub-continent gene flows converge. *PLoS One*. 2013; 8 (10): e76748. DOI: 10.1371/journal.pone.0076748.
 33. Fedorova SA, Reidla M, Metspalu E, Metspalu M, Rootsi S, Tambets K, et al. Autosomal and uniparental portraits of the native populations of Sakha (Yakutia): implications for the peopling of Northeast Eurasia. *BMC Evol Biol*. 2013; 13: 127. DOI: 10.1186/1471-2148-13-127.
 34. Flegontov P, Changmai P, Zidkova A, Logacheva MD, Altınışık NE, Flegontova O, et al. Genomic study of the Ket: a Paleo-Eskimo-related ethnic group with significant ancient North Eurasian ancestry. *Sci Rep*. 2016; 6: 20768. DOI: 10.1038/srep20768.
 35. Haber M, Mezzavilla M, Xue Y, Comas D, Gasparini P, Zalloua P, et al. Genetic evidence for an origin of the Armenians from Bronze Age mixing of multiple populations. *Eur J Hum Genet*. 2016; 24 (6): 931–6. DOI: 10.1038/ejhg.2015.206.
 36. Kovacevic L, Tambets K, Ilumäe AM, Kushniarevich A, Yunusbayev B, Solnik A, et al. Standing at the gateway to Europe—the genetic structure of Western balkan populations based on autosomal and haploid markers. *PLoS One*. 2014; 9 (8): e105090. DOI: 10.1371/journal.pone.0105090.
 37. Kushniarevich A, Utevska O, Chuhryaeva M, Agdzhoyan A, Dibirova K, Uktveryte I, et al. Genetic heritage of the Balto-Slavic speaking populations: a synthesis of autosomal, mitochondrial and Y-chromosomal data. *PLoS One*. 2015; 10 (9): e0135820. DOI: 10.1371/journal.pone.0135820.
 38. Li JZ, Absher DM, Tang H, Southwick AM, Casto AM, Ramachandran S, et al. Worldwide human relationships inferred from genome-wide patterns of variation. *Science*. 2008; 319 (5866): 1100–4. DOI: 10.1126/science.1153717.
 39. Raghavan M, Skoglund P, Graf KE, Metspalu M, Albrechtsen A, Moltke I, et al. Upper Palaeolithic Siberian genome reveals dual ancestry of Native Americans. *Nature*. 2014; 505 (7481): 87–91. DOI: 10.1038/nature12736.
 40. Raghavan M, DeGiorgio M, Albrechtsen A, Moltke I, Skoglund P, Korneliusen TS, et al. The genetic prehistory of the New World Arctic. *Science*. 2014; 345 (6200): 1255832. DOI: 10.1126/science.1255832.
 41. Raghavan M, Steinrücken M, Harris K, Schiffels S, Rasmussen S, DeGiorgio M, et al. POPULATION GENETICS. Genomic evidence for the Pleistocene and recent population history of Native Americans. *Science*. 2015; 349 (6250): aab3884. DOI: 10.1126/science.aab3884
 42. Rasmussen M, Li Y, Lindgreen S, Pedersen JS, Albrechtsen A, Moltke I, et al. Ancient human genome sequence of an extinct Palaeo-Eskimo. *Nature*. 2010; 463 (7282): 757–62. DOI: 10.1038/nature08835.
 43. Yunusbayev B, Metspalu M, Järve M, Kutuev I, Rootsi S, Metspalu E, et al. The Caucasus as an asymmetric semipermeable barrier to ancient human migrations. *Mol Biol Evol*. 2012; 29 (1): 359–65. DOI: 10.1093/molbev/msr221.
 44. Yunusbayev B, Metspalu M, Metspalu E, Valeev A, Litvinov S, Valiev R, et al. The genetic legacy of the expansion of Turkic-speaking nomads across Eurasia. *PLoS Genet*. 2015; 11 (4): e1005068. DOI: 10.1371/journal.pgen.1005068.
 45. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience*. 2015; 4 (1): 7. DOI: 10.1186/s13742-015-0047-8.
 46. Purcell S, Chang C. PLINK: Whole genome association analysis toolset. Version 1.9 [software]. [cited 2020 Nov 4]. Available from: www.cog-genomics.org/plink/1.9/.
 47. Balanovsky O, Dibirova K, Dybo A, Mudrak O, Frolova S, Pocheshkhova E, et al. Parallel Evolution of Genes and Languages in the Caucasus Region. *Mol Biol Evol*. 2011; 28 (10): 2905–20. DOI: 10.1093/molbev/msr126.
 48. Koshel SM. Geoinformatsionnye tekhnologii v genogeografii. V sbornike: Lur'e I. K., Kravtsova V. I., redaktory. *Sovremennaya geograficheskaya kartografiya*. M.: Data+, 2012; p. 158–66. Russian.
 49. Balanovska EV, Balanovsky OP. *Russkiy genofond na Russkoy ravnine*. M.: Luch, 2007; 416 p. Russian.
 50. Balanovsky O, Zhabagin M, Agdzhoyan A, Chukhryaeva M, Zaporozhchenko V, Utevska O, et al. Deep phylogenetic analysis of haplogroup G1 provides estimates of SNP and STR mutation rates on the human Y-chromosome and reveals migrations of Iranic speakers. *PLoS One*. 2015; 10 (4): e0122968. DOI: 10.1371/journal.pone.0122968.
 51. Rychkov YuG, Balanovska EV. Genofond i genogeografiya naseleniya SSSR. *Genetika*. 1992; (28): 52–75. Russian.
 52. Balanovska EV, Grekhova LV, Rychkov YuG. *Komp'yuternaya genogeografiya i arkheologiya: metody kartograficheskogo modelirovaniya rasprostraneniya material'noy kul'tury*. V sbornike: *Gorizonty antropologii*. M.: Russkiy mir, 1997; p. 54–62. Russian.
 53. Berlyant AM. *Kartografiya: Uchebnik dlya vuzov*. M.: Aspekt Press, 2002; 336 p. Russian.
 54. Genes-Drugs [internet]. CPIC. 2020 [cited 2020 Nov 4]. Available from: <https://cpicpgx.org/genes-drugs>.
 55. Nurbayev SD, Balanovska EV. *Genogeografiya i genofond. Otsenivanie nadezhnosti karty*. V sbornike: *Novye metody — novye podkhody v sovremennoy antropologii*. M.: Staryy sad, 1997; p. 116–32. Russian.

Литература

1. Core ADME Gene List [internet]. www.pharmaadme.org. ©2020 — [cited 2020 Nov 4]. Available from: <http://pharmaadme.org/>
2. joomla/index.php?option=com_content&task=view&id=12&Itemid=27. pharmgkb.org [internet]. ©2001–2020 PharmGKB [cited 2020

- Nov 4]. Available from: <https://www.pharmgkb.org>.
3. Scott SA, Sangkuhl K, Stein CM, Hulot JS, Mega JL, Roden DM, et al. Clinical Pharmacogenetics Implementation Consortium guidelines for CYP2C19 genotype and clopidogrel therapy: 2013 update. *Clin Pharmacol Ther.* 2013; 94 (3): 317–23. DOI: 10.1038/clpt.2013.105.
 4. Lima JJ, Thomas CD, Barbarino J, Desta Z, Van Driest SL, El Rouby N, et al. Clinical Pharmacogenetics Implementation Consortium (CPIC) Guideline for CYP2C19 and Proton Pump Inhibitor Dosing. *Clin Pharmacol Ther.* 2020. DOI: 10.1002/cpt.2015. Epub ahead of print.
 5. Johnson JA, Caudle KE, Gong L, Whirl-Carrillo M, Stein CM, Scott SA, et al. Clinical Pharmacogenetics Implementation Consortium (CPIC) Guideline for Pharmacogenetics-Guided Warfarin Dosing: 2017 Update. *Clin Pharmacol Ther.* 2017; 102 (3): 397–404. DOI: 10.1002/cpt.668.
 6. Beutler E. G6PD: population genetics and clinical manifestations. *Blood Rev.* 1996; 10 (1): 45–52. DOI: 10.1016/s0268-960x(96)90019-3.
 7. Urban TJ. Race, ethnicity, ancestry, and pharmacogenetics. *Mt Sinai J Med.* 2010; 77 (2): 133–9. DOI:10.1002/msj.20168.
 8. Hernandez W, Gamazon ER, Aquino-Michaels K, Smithberger E, O'Brien TJ, Harralson AF, et al. Integrated analysis of genetic variation and gene expression reveals novel variant for increased warfarin dose requirement in African Americans. *J Thromb Haemost.* 2017; 15 (4): 735–43. DOI: 10.1111/jth.13639.
 9. Rajman I, Knapp L, Morgan T, Masimirembwa C. African Genetic Diversity: Implications for Cytochrome P450-mediated Drug Metabolism and Drug Development. *EBioMedicine.* 2017; 17: 67–74. DOI: 10.1016/j.ebiom.2017.02.017.
 10. Duconge J, Ramos AS, Claudio-Campos K, Rivera-Miranda G, Bermúdez-Bosch L, Renta JY, et al. A Novel Admixture-Based Pharmacogenetic Approach to Refine Warfarin Dosing in Caribbean Hispanics. *PLoS One.* 2016; 11 (1): e0145480. DOI: 10.1371/journal.pone.0145480.
 11. Mizzi C, Dalabira E, Kumuthini J, Dzimiri N, Başak N, et al. A European Spectrum of Pharmacogenomic Biomarkers: Implications for Clinical Pharmacogenomics. *PLoS One.* 2016; 11 (9): e0162866. DOI: 10.1371/journal.pone.0162866.
 12. Li J, Lou H, Yang X, Lu D, Li S, Jin L, et al. Genetic architectures of ADME genes in five Eurasian admixed populations and implications for drug safety and efficacy. *J Med Genet.* 2014; 51 (9): 614–22. DOI: 10.1136/jmedgenet-2014-102530.
 13. Duconge J, Ruaño G. Admixture and ethno-specific alleles: missing links for global pharmacogenomics. *Pharmacogenomics.* 2016; 17 (14): 1479–82. DOI: 10.2217/pgs-2016-0115.
 14. Lonjou C, Thomas L, Borot N, Ledger N, de Toma C, LeLouet H, et al. A marker for Stevens-Johnson syndrome ...: ethnicity matters. *Pharmacogenomics J.* 2006; 6 (4): 265–8. DOI: 10.1038/sj.tpj.6500356.
 15. Mirzaev KB, Fedorinov DS, Ivashchenko DV, Sychev DA. ADME pharmacogenetics: future outlook for Russia. *Pharmacogenomics.* 2019; 20 (11): 847–65. DOI: 10.2217/pgs-2019-0013.
 16. Мустафина О. Е., Туктарова И. А., Каримов Д. Д., Сомова Р. Ш., Насибуллин Т. Р. Полиморфизм генов CYP2D6, CYP3A5, и CYP3A4 в популяциях русских, татар и башкир. *Генетика.* 2015; 51 (1): 109–19. DOI: 10.7868/S0016675815010087.
 17. Korytina G, Kochetova O, Akhmadishina L, Viktorova E, Victorova T. Polymorphisms of cytochrome p450 genes in three ethnic groups from Russia. *Balkan Med J.* 2012; 29 (3): 252–60. DOI:10.5152/balkanmedj.2012.039.
 18. Кочетова О. В., Корытина Г. Ф., Ахмадишина Л. З., Исхакова Г. М., Викторова Т. В. Анализ полиморфизма гена цитохрома P450 1A1 (CYP1A1) в этнических группах республики Башкортостан. *Генетика.* 2008; 44 (12): 1677–83.
 19. Корытина Г. Ф., Целоусова О. С., Ахмадишина Л. З., Кочетова О. В., Бабенкова Л. И., Викторова Т. В. Анализ ассоциации полиморфных маркеров генов медиаторов воспаления (IL1B, tNfA, LTA, IL8, IL6, IL1RN, IL10, TGFb, TLR4, DBP) с развитием хронических заболеваний респираторной системы у детей. *Медицинская генетика.* 2008; 7 (2): 17–25.
 20. Корытина Г. Ф., Ахмадишина Л. З., Викторова Т. В. Частоты полиморфных вариантов генов CYP1B1 и CYP2F1 в трех этнических группах жителей Республики Башкортостан и у больных хронической обструктивной болезнью легких. *Молекулярная биология.* 2010; 44 (1): 33–41.
 21. Ахмадишина Л. З., Корытина Г. Ф., Мингазова С. Р., Янбаева Д. Г., Бакиров А. Б., Викторова Т. В. Роль полиморфизма генов CYP1A1, EPHX1, GSTM1, GSTT1 и GSTP1 в развитии хронических бронхитов профессионального генеза. *Экологическая генетика.* 2005; (1): 11–7.
 22. Федорова Ю. Ю., Карунас А. С., Нургулиева А. Х., Гра О. А., Гималова Г. Ф., Рамазанова Н. Н. и др. Роль полиморфных вариантов генов системы биотрансформации в развитии аллергических заболеваний у татар. *Медицинская генетика.* 2010; 9 (6): 28–35.
 23. Polonikov A, Kharchenko A, Bykanova M, Sirotina S, Ponomarenko I, Bocharova A, et al. Polymorphisms of CYP2C8, CYP2C9 and CYP2C19 and risk of coronary heart disease in Russian population. *Gene.* 2017; 627: 451–9. DOI: 10.1016/j.gene.2017.07.004.
 24. Gaikovitch EA, Cascorbi I, Mrozikiewicz PM, Brockmöller J, Frötschl R, Köpke K, et al. Polymorphisms of drug-metabolizing enzymes CYP2C9, CYP2C19, CYP2D6, CYP1A1, NAT2 and of P-glycoprotein in a Russian population. *Eur J Clin Pharmacol.* 2003; 59 (4): 303–12. DOI: 10.1007/s00228-003-0606-2.
 25. Gra O, Mityaeva O, Berdichevets I, Kozhekbaeva Z, Fesenko D, Kurbatova O, et al. Microarray-based detection of CYP1A1, CYP2C9, CYP2C19, CYP2D6, GSTT1, GSTM1, MTHFR, MTRR, NQO1, NAT2, HLA-DQA1, and ABO allele frequencies in native Russians. *Genet Test Mol Biomarkers.* 2010; 14 (3): 329–42. DOI: 10.1089/gtmb.2009.0158.
 26. Балановская Е. В., Жабалин М. К., Агджоян А. Т., Чухряева М. И., Маркина Н. В., Балаганская О. А. и др. Популяционные биобанки: принципы организации и перспективы применения в геногеографии и персонализированной медицине. *Генетика.* 2016; 52 (12): 1371–87. DOI: 10.7868/S001667581612002X.
 27. Whirl-Carrillo M, McDonagh EM, Hebert JM, Gong L, Sangkuhl K, Thorn CF, et al. Pharmacogenomics Knowledge for Personalized Medicine. *Clinical Pharmacology & Therapeutics.* 2012; 92 (4): 414–17. [pharmgkb.org](http://www.pharmgkb.org) [internet]. ©2001–2020 «Very Important Pharmacogene» (VIP) [cited 2020 Nov 4]. Available from: <https://www.pharmgkb.org/vips/>.
 28. Балановский О. П., Горин И. О., Записецкая Ю. С., Голубева А. А., Кострюкова Е. С., Балановская Е. В. Взаимодействие генофондов русского и финноязычного населения Тверской области: анализ 4 млн SNP-маркеров. *Вестник РГМУ.* 2020; 6 (в печати).
 29. Behar DM, Yunusbayev B, Metspalu M, Metspalu E, Rosset S, Parik J, et al. The genome-wide structure of the Jewish people. *Nature.* 2010; 466 (7303): 238–42. DOI: 10.1038/nature09103.
 30. Behar DM, Metspalu M, Baran Y, Kopelman NM, Yunusbayev B, Gladstein A, et al. No evidence from genome-wide data of a Khazar origin for the Ashkenazi Jews. *Hum Biol.* 2013; 85 (6): 859–900. DOI: 10.3378/027.085.0604.
 31. Chaubey G, Metspalu M, Choi Y, Mägi R, Romero IG, Soares P, et al. Population genetic structure in Indian Austroasiatic speakers: the role of landscape barriers and sex-specific admixture. *Mol Biol Evol.* 2011; 28 (2): 1013–24. DOI: 10.1093/molbev/msq288.
 32. Di Cristofaro J, Pennarun E, Mazières S, Myres NM, Lin AA, Temori SA, et al. Afghan Hindu Kush: where Eurasian sub-continent gene flows converge. *PLoS One.* 2013; 8 (10): e76748. DOI: 10.1371/journal.pone.0076748.
 33. Fedorova SA, Reidla M, Metspalu E, Metspalu M, Rootsi S, Tambets K, et al. Autosomal and uniparental portraits of the native populations of Sakha (Yakutia): implications for the peopling of Northeast Eurasia. *BMC Evol Biol.* 2013; 13: 127. DOI: 10.1186/1471-2148-13-127.
 34. Flegontov P, Changmai P, Zidkova A, Logacheva MD, Altınışık NE, Flegontova O, et al. Genomic study of the Ket: a Paleo-Eskimo-related ethnic group with significant ancient North Eurasian ancestry. *Sci Rep.* 2016; 6: 20768. DOI: 10.1038/srep20768.
 35. Haber M, Mezzavilla M, Xue Y, Comas D, Gasparini P, Zalloua P, et al. Genetic evidence for an origin of the Armenians from Bronze Age mixing of multiple populations. *Eur J Hum Genet.* 2016; 24 (6): 931–6. DOI: 10.1038/ejhg.2015.206.

36. Kovacevic L, Tambets K, Ilumäe AM, Kushniarevich A, Yunusbayev B, Solnik A, et al. Standing at the gateway to Europe--the genetic structure of Western balkan populations based on autosomal and haploid markers. *PLoS One*. 2014; 9 (8): e105090. DOI: 10.1371/journal.pone.0105090.
37. Kushniarevich A, Utevska O, Chuhryaeva M, Agdzhoyan A, Dibirova K, Uktveryte I, et al. Genetic heritage of the Balto-Slavic speaking populations: a synthesis of autosomal, mitochondrial and Y-chromosomal data. *PLoS One*. 2015; 10 (9): e0135820. DOI: 10.1371/journal.pone.0135820.
38. Li JZ, Absher DM, Tang H, Southwick AM, Casto AM, Ramachandran S, et al. Worldwide human relationships inferred from genome-wide patterns of variation. *Science*. 2008; 319 (5866): 1100–4. DOI: 10.1126/science.1153717.
39. Raghavan M, Skoglund P, Graf KE, Metspalu M, Albrechtsen A, Moltke I, et al. Upper Palaeolithic Siberian genome reveals dual ancestry of Native Americans. *Nature*. 2014; 505 (7481): 87–91. DOI: 10.1038/nature12736.
40. Raghavan M, DeGiorgio M, Albrechtsen A, Moltke I, Skoglund P, Korneliusson TS, et al. The genetic prehistory of the New World Arctic. *Science*. 2014; 345 (6200): 1255832. DOI: 10.1126/science.1255832.
41. Raghavan M, Steinrücken M, Harris K, Schiffels S, Rasmussen S, DeGiorgio M, et al. POPULATION GENETICS. Genomic evidence for the Pleistocene and recent population history of Native Americans. *Science*. 2015; 349 (6250): aab3884. DOI: 10.1126/science.aab3884
42. Rasmussen M, Li Y, Lindgreen S, Pedersen JS, Albrechtsen A, Moltke I, et al. Ancient human genome sequence of an extinct Palaeo-Eskimo. *Nature*. 2010; 463 (7282): 757–62. DOI: 10.1038/nature08835.
43. Yunusbayev B, Metspalu M, Järve M, Kutuev I, Rootsi S, Metspalu E, et al. The Caucasus as an asymmetric semipermeable barrier to ancient human migrations. *Mol Biol Evol*. 2012; 29 (1): 359–65. DOI: 10.1093/molbev/msr221.
44. Yunusbayev B, Metspalu M, Metspalu E, Valeev A, Litvinov S, Valiev R, et al. The genetic legacy of the expansion of Turkic-speaking nomads across Eurasia. *PLoS Genet*. 2015; 11 (4): e1005068. DOI: 10.1371/journal.pgen.1005068.
45. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience*. 2015; 4 (1): 7. DOI: 10.1186/s13742-015-0047-8.
46. Purcell S, Chang C. PLINK: Whole genome association analysis toolset. Version 1.9 [software]. [cited 2020 Nov 4]. Available from: www.cog-genomics.org/plink/1.9/.
47. Balanovsky O, Dibirova K, Dybo A, Mudrak O, Frolova S, Pocheshkhova E, et al. Parallel Evolution of Genes and Languages in the Caucasus Region. *Mol Biol Evol*. 2011; 28 (10): 2905–20. DOI: 10.1093/molbev/msr126.
48. Кошель С. М. Геоинформационные технологии в геногеографии. В сборнике: Лурье И. К., Кравцова В. И., редакторы. Современная географическая картография. М.: Дата+, 2012; с. 158–66.
49. Балановская Е. В., Балановский О. П. Русский генофонд на Русской равнине. М.: Луч, 2007; 416 с.
50. Balanovsky O, Zhabagin M, Agdzhoyan A, Chukhryaeva M, Zaporozhchenko V, Utevska O, et al. Deep phylogenetic analysis of haplogroup G1 provides estimates of SNP and STR mutation rates on the human Y-chromosome and reveals migrations of Iranic speakers. *PLoS One*. 2015; 10 (4): e0122968. DOI: 10.1371/journal.pone.0122968.
51. Рычков Ю. Г., Балановская Е. В. Генофонд и геногеография населения СССР. *Генетика*. 1992; (28): 52–75.
52. Балановская Е. В., Грехова Л. В., Рычков Ю. Г. Компьютерная геногеография и археология: методы картографического моделирования распространения материальной культуры. В сборнике: Горизонты антропологии. М.: Русский мир, 1997; с. 54–62.
53. Берлянт А. М. Картография: Учебник для вузов. М.: Аспект Пресс, 2002; 336 с.
54. Genes-Drugs [internet]. CPIC. 2020 [cited 2020 Nov 4]. Available from: <https://cpicpgx.org/genes-drugs>.
55. Нурбаев С. Д., Балановская Е. В. Геногеография и генофонд. Оценивание надежности карты. В сборнике: Новые методы — новые подходы в современной антропологии. М.: Старый сад, 1997; с.116–32.