GENE GEOGRAPHY OF PHARMACOGENETICALLY SIGNIFICANT CYP2C19 CYTOCHROME SUPERFAMILY DNA MARKERS IN THE POPULATIONS OF RUSSIA AND NEIGHBORING COUNTRIES

Balanovska EV[1] ✉, Abdulaev ShP[2], Gorin IO[1], Belov RO[1], Mukatdarova EA[1], Pylev VYu[1,3]

[1] Research Centre for Medical Genetics, Moscow, Russia
[2] Russian Medical Academy of Continuous Professional Education, Moscow, Russia
[3] Biobank of North Eurasia, Moscow, Russia

Genetic testing of each patient aimed at detecting the pharmacogenetic marker carrier state is challenging for healthcare system. However, knowledge about the frequencies of pharmacogenetically important genes enables making decisions about treatment based on the patient's ethnicity. The *CYP2C19* cytochrome gene involved in biotransformation of a broad spectrum of drugs is one of the most important. The study was aimed to determine the frequencies of major *CYP2C19* variants and the patterns of their spatial variability in the population of Russia. The database Pharmacogenetics of the Population of Russia and Neighboring Countries created by the research team was used to determine frequencies of the *CYP2C19 *1, *2, *3, *17* variants and their genotypes: *1 – 53 populations, *n* = 2261 samples; *2 — 79 populations, *n* = 6346; *3 — 92 populations, *n* = 7517; *17 — 35 populations, *n* = 3313. We have created a cartographic atlas that includes the *1, *2, *3, *17 frequency maps, correlation maps, and genotype frequency maps. Specific data on the frequencies of *CYP2C19* variants and their pharmacogenetically significant genotypes in the major ethnic groups of Russia are provided. The cartographic atlas enables prediction of frequencies of significant *CYP2C19* variants and their genotypes in the peoples, information about which is currently missing. The *1 and *2 variants gene geography is characterized by similar pattern: the combination of longitudinal trend of frequency increase from west to southeast and latitudinal variability of frequency increase from north to south in the Asian part of the region. Variant *3 is characterized by the clear longitudinal vector of frequency increase from 0 in the west to the world's maximum in the Amur region. Variant *17 shows a pronounced longitudinal trend with the oppositely directed vector of frequency decrease from west to southeast. The correlation maps indicate regions, where the similarity between core patterns is disrupted.

Keywords: pharmacogenetics, *CYP2C19*, DNA markers, gene pool, gene geography, cartographic atlas, Russia, North Eurasia, populations

Funding: the study was supported by the Russian Science Foundation grant № 21-14-00363 (bioinformatics, statistical and cartographic analysis), State Assignment of the Ministry of Science and Higher Education of the Russian Federation for the Research Centre for Medical Genetics (genealogical analysis, interpretation of the results).

Acknowlegements: the authors would like to thank all sample donors, who took part in the study, and Biobank of North Eurasia for access to DNA collections.

Author contribution: Balanovska EV — data analysis, manuscript writing, research management; Abdulaev ShP — discriptions of pharmacogenetic markers; Gorin IO — bioinformatics analysis; Belov RO — manuscript formatting; Mukatdarova EA — working with genealogical database; Pylev VYu — statistical analysis, cartographic analysis.

Compliance with ethical standards: the study was approved by the Ethics Commitee of the Research Centre for Medical Genetics (protocol № 1 of 29 June 2020); all subjects submitted the informed consent to study participation.

✉ **Correspondence should be addressed:** Elena V. Balanovska
Moskvorechye, 1, 115522, Moscow, Russia; balanovska@mail.ru

Received: 18.09.2023 **Accepted:** 18.10.2023 **Published online:** 31.10.2023

DOI: 10.24075/brsmu.2023.039

ГЕНОГЕОГРАФИЯ ФАРМАКОГЕНЕТИЧЕСКИ ЗНАЧИМЫХ ДНК-МАРКЕРОВ CYP2C19 СУПЕРСЕМЕЙСТВА ЦИТОХРОМОВ В НАРОДОНАСЕЛЕНИИ РОССИИ И СОПРЕДЕЛЬНЫХ СТРАН

Е. В. Балановская[1] ✉, Ш. П. Абдулаев[2], И. О. Горин[1], Р. О. Белов[1], Е. А. Мукатдарова[1], В. Ю. Пылёв[1,3]

[1] Медико-генетический научный центр имени Н. П. Бочкова, Москва, Россия
[2] Российская медицинская академия непрерывного профессионального образования Минздрава России, Москва, Россия
[3] Биобанк Северной Евразии, Москва, Россия

Генетическое тестирование каждого пациента для выявления носительства фармакогенетических маркеров проблемно для системы здравоохранения. Но знание частоты встречаемости фармакогенетически важных генов позволяет принимать решение о терапии в зависимости от этнической принадлежности пациента. Одним из наиболее значимых является ген цитохрома *CYP2C19*, участвующий в биотрансформации широкого спектра лекарственных препаратов. Целью работы было выявить частоты встречаемости основных вариантов *CYP2C19* и паттерны их пространственной изменчивости в народонаселении России. На основе созданной коллективом базы данных «Фармакогенетика популяций России и сопредельных стран» получены частоты вариантов *CYP2C19 *1, *2, *3, *17* и частоты их генотипов: *1 — 53 популяции, *n* = 2261 образец; *2 — 79 популяций, *n* = 6346; *3 — 92 популяции, *n* = 7517; *17 — 35 популяций, *n* = 3313. Создан картографический атлас, включающий карты частоты вариантов *1, *2, *3, *17, их корреляционные карты и карты частоты их генотипов. Представлены конкретные данные о частотах вариантов *CYP2C19* и их фармакогенетически значимых генотипах в основных этнических группах России. Картографический атлас дает прогноз частоты значимых вариантов *CYP2C19* и их генотипов для народов, информация о которых пока отсутствует. Геногеография *1 и *2 характеризуется схожим паттерном: совмещение долготного тренда роста частоты с запада на юго-восток и широтного роста частоты с севера на юг в азиатской части региона. Вариант *3 отличается четкостью долготного вектора роста частоты от 0 на западе до мирового максимума частоты в Приамурье. Вариант *17 имеет выразительный долготный тренд с противоположным вектором падения частоты с запада на юго-восток. Корреляционные карты указывают регионы, в которых нарушено сходство между основными паттернами.

Ключевые слова: фармакогенетика, *CYP2C19*, ДНК-маркеры, генофонд, геногеография, картографический атлас, Россия, Северная Евразия, популяции

Финансирование: РНФ №21-14-00363 (биоинформатический, статистический и картографический анализ) и Государственного задания Министерства науки и высшего образования РФ для Медико-генетического научного центра имени академика Н. П. Бочкова (генеалогический анализ, интерпретация результатов).

Благодарности: авторы благодарят всех доноров образцов, принимавших участие в исследовании, и АНО «Биобанк Северной Евразии» за предоставление коллекций ДНК.

Вклад авторов: Е. В. Балановская — анализ данных, написание текста, руководство исследованием; Ш. П. Абдулаев — описание фармакогенетических маркеров; И. О. Горин — биоинформатический анализ; Р. О. Белов — оформление статьи; Е. А. Мукатдарова — работа с базой генеалогических данных; В. Ю. Пылёв — статистический анализ, картографический анализ.

Соблюдение этических стандартов: исследование одобрено этическим комитетом ФГБНУ «Медико-генетический научный центр имени Н. П. Бочкова» (протокол № 1 от 29 июня 2020 г.); все участники подписали добровольное информированное согласие на участие в исследовании.

✉ **Для корреспонденции:** Елена Владимировна Балановская
ул. Москворечье, д. 1, 115522, г. Москва, Россия balanovska@mail.ru

Статья получена: 18.09.2023 **Статья принята к печати:** 18.10.2023 **Опубликована онлайн:** 31.10.2023

DOI: 10.24075/vrgmu.2023.039

Drug therapy efficacy and safety depend largely on individual differences between patients. This is a pressing issue of modern pharmacotherapy, since the body's genetic status accounts for up to 50% of the pharmacological response individual variability. Selection of the drug and dosage considering the patient's molecular genetic features is the subject of pharmacogenetics [1, 2] aiming to search for the drug dose that would be effective and safe for this particular patient [3].

The CYP450 cytochrome superfamily, in which the *CYP2C19* gene is highly polymorphic, represents one of the pharmacogenes extensively studied from clinical perspective. The *CYP2C19* enzyme is involved in biotransformation of the wide range of drugs, such as clopidogrel, omeprazole, lansoprazole, propranolol, diazepam, imipramine and some other antidepressants [4]. It has been shown that *CYP2C19*2* and *CYP2C19*3* are associated with the enzyme reduced metabolic activity [5], while *CYP2C19*17* (*rs12248560*) is associated with enchanced metabolism of the enzyme substrates [5]. Clopidogrel is a prime example of the drug, for which the guidelines on the treatment regimen and dose adjustment have been developed. In individuals with "normal" *1/*1 genotype, it is used in accordance with the instructions. The *1/*2, *1/*3, *2/*17, *3/*17 genotypes are characterized by smaller decline in platelet aggregation relative to normal, higher residual platelet aggregation, and increased risk of cardiovascular events. Accumulation of "slow" alleles (*2/*2, *2/*3, *3/*3) in the genotype is associated with low clopidogrel efficacy and high residual platelet reactivity. The group of "ultrafast" metabolizers (*1/*17, *17/*17) is characterized by the increased antiplatelet activity and decreased residual platelet aggregation, which can be due to the risk of hemorrhage [5]. The carrier frequency for various *CYP2C19* SNP markers and the associated clopidogrel resistance are characterized by the pronounced ethnoracial heterogeneity [6]: *CYP2C19*2* is found in 15% of Caucasoids, 17% of Negroids and is far more frequent in Mongoloids living in East Asia (31%). The opposite trend has been revealed for *CYP2C19*17*: it is common in Caucasoids (22%) and rare (1.5%) in Mongoloids of East Asia [6]. The *CYP2C19*3* variant is rare: it averagely accounts for 1.4% of global population [6]. In the Russian population the rate of *CYP2C19*2* is about 11%, the rate of *CYP2C19*3* is about 0.34%, and that of *CYP2C19*17* is about 27% [7].

The principles of precision, preventive, and personalized medicine envision using genetic information to make clinical decisions. However, extensive use of pharmacogenetic testing (PGT) in clinical practice has a number of limitations. PGT remains an option that is not available in the regions with inadequate funding of the health sector. An important role is also played by the time of obtaining PGT results, which can be relevant when providing emergency care [8].

The population gene geography that reveals the patterns in pharmacogenetic biomarker distribution provides one of the solutions [9–13]. The gene geography data can play an important role in clinical decision-making in such multi-ethnic country as Russia. That is why assessment of polymorphic genes' carrier frequency in Russia is essential for development and implementation of the personalized medicine principles. Exploration and identification of the distribution patterns of important pharmacogenetic markers in the population of Russia make it possible to identify ethnic groups and regions, in which PGT of the large population of patients can be a clinically and economically advantageous solution: in these regions the decision that PGT is essential for therapy personalization can be made based on the patient's ethnicity.

The study was aimed to determine frequencies of major *CYP2C19* cytochrome superfamily DNA markers (*1, *2, *3, *17) in the population of Russia and determine the trends in their gene geographical variability.

METHODS

The study involved analysis of frequencies of the *CYP2C19*1*, *CYP2C19*2*, *CYP2C19*3*, *CYP2C19*17* variants (hereinafter *1, *2, *3, *17) and their genotypes in indigenous peoples of North Eurasia and other regions. These four variants (*1, *2, *3, *17) belong to different *CYP2C19* gene SNP markers and are not alleles of the same SNP; their frequencies were calculated using the PLINK 1.9 tool [14] and Python 3. The *1 variant is the sum of 11 SNPs of "normal" variants. Among them 7 SNPs were found in the databases and had representative frequencies, and the *1 frequency was calculated as the square root of the frequency of the combination of seven homozygotes. The genotype frequencies were calculated based on the *1, *2, *3, *17 variant frequencies in accordance with the Hardy–Weinberg equilibrium.

The analysis involved the use of the database "Pharmacogenetics of Population of Russia and Neighboring Countries" created by the research team and the GG-base (world's populations) [15] compiled in accordance with [16] and assessed using various panels of SNP markers [9–13]. Populations with the sample size of $n < 25$ samples were included in large metapopulations together with other populations based on the commonality of ethnogenesis and region. The data on the peoples of the Caucasus on the scale of North Eurasia were represented by four subregional samples. The total samples for the *CYP2C19* gene SNP variants were as follows: *1 — n = 2261 samples; *2 — n = 6346; *3 — n = 7517; *17 — n = 3313. The results were presented as tables (frequencies of SNP variants and their genotypes in 53 metapopulations of 13 world's regions) and as the gene geographical atlas including maps showing spatial variability of frequencies of SNP markers and their genotypes, as well as correlation maps demonstrating the association of geographical variability of frequencies of all SNP markers. The following SNP variant variability parameters were provided; $\bar{q}$ — average frequency; G_{ST} — interpopulation differences based on certain variant (G_{ST} is the F_{ST} analog for biallelic cases); H_S — heterozygosity level.

The *CYP2C19* gene geographic maps were plotted using the GeneGeo software [17] by the weighted average interpolation with the 2nd degree weighting function, range radius of 1500 km for North Eurasia and 5000 km for the world. In the tables each population was assigned a number that was shown on the maps, which made it possible to clearly identify and distinguish all the studied populations on the maps. As for metapopulations, the trait frequency value was projected onto all coordinates of input local populations. The genotype frequency maps were calculated for each node of the map based on the frequency value in each node of the maps for the *1, *2*, *3, *17 variants in accordance with the Hardy–Weinberg equilibrium. Statistical parameters of the map are shown in specific box of the legend of each map: K — number of input populations for map plotting; min — minimum trait frequency; max — maximum trait frequency; avr — average trait frequency; G_{ST} — interpopulation differences based on this trait; H_S —heterozygosity level.

The correlation maps were created by the 1100 km floating window method using the Kendall rank correlation coefficient. Correlation between two traits was calculated for all nodes falling into the specified window and assigned to the central node. Then this window was moved one node, and calculation was repeated. Thus, correlations were calculated for all nodes

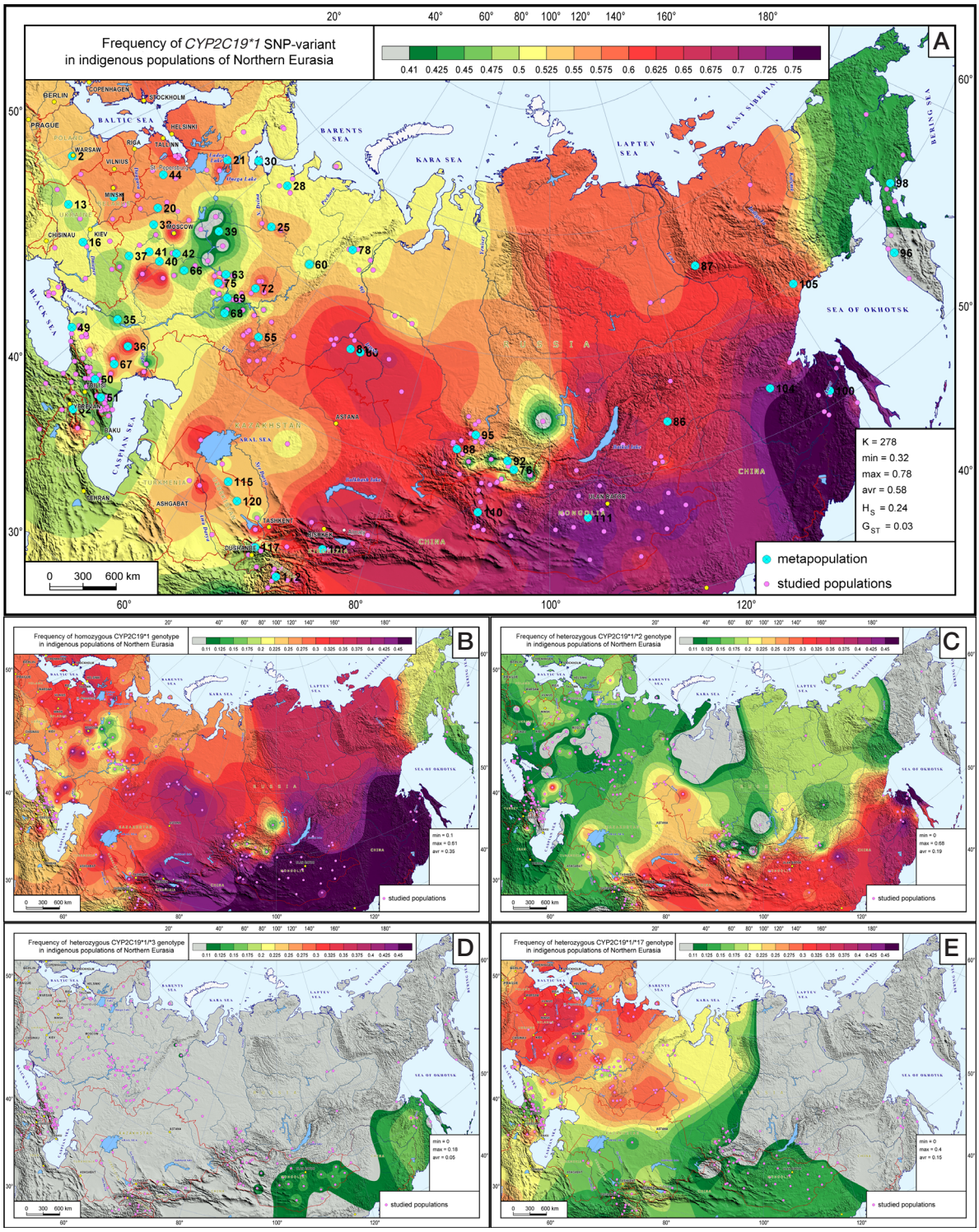

Fig. 1. Spatial *CYP2C19*1* variability in indigenous population of Russia and neighboring countries. **A**. Gene geography of *CYP2C19*1* SNP variant frequency. **B**. Gene geography of *1/*1 genotype frequency. **C**. Gene geography of *1/*2 genotype frequency. **D**. Gene geography of *1/*3 genotype frequency. **E**. Gene geography of *1/*17 genotype frequency

of digital mesh model of the map (81259 nodes) providing the basis for correlation pattern visualization.

RESULTS

Gene geography of *CYP2C19*1* (Fig. 1; Tables 1, 2)

Understanding the *1 "normal" variant gene geography (Fig. 1) is beneficial for identification of patterns in variability of genotypes that are important for pharmacotherapy.

The main longitudinal trend represented by the *1 frequency increase from west to southeast is combined with latitudinal variability in the Asian part of the region. The average *1 frequency ($\bar{q}$ = 0.58) is much higher than that of variants *2, *3, *17 (0.04 < $\bar{q}$ < 0.15). That is why a common frequency interval (0.125) was used in all maps for variant *1 to make all maps comparable (Fig. 1–4), however it was within another frequency range (0.41 < q < 0.75).

The longitudinal trend of frequency decrease from west to southeast is associated with many irregularities. Low

Table 1. Frequencies of CYP2C19*1, *2, *3, 17 SNP variant in indigenous population of Russia and the world

Region	№ on the map	Population	*CYP2C19*1*		*CYP2C19*2*		*CYP2C19*3*		*CYP2C19*17*	
			-		rs4244285		rs4986893		rs12248560	
			n	$\bar{q}$	*n*	$\bar{q}$	*n*	$\bar{q}$	*n*	$\bar{q}$
FOREIGN EUROPE	1	Belarusians	33	0.577	50	0.170	50	0.020	19	0.290
	2	Eastern Europe (Gagauz, Lithuanians, Macedonians)	21	0.535	-	-	-	-	-	-
	3	Northern peoples of Balkans (Hungarians, Romanians, Slovenes)	-	-	-	-	51	0	-	-
	4	Central peoples of Balkans (Bosnians, Kosovars, Serbs, Croats, Montenegrins)	-	-	-	-	80	0.013	-	-
	5	Southern peoples of Balkans (Bulgarians, Greeks, Macedonians)	-	-	70	0.171	129	0	-	-
	6	Peoples of Europe	-	-	503	0.145	503	0	-	-
	7	Peoples of Moldova (Gagauz, Moldovans)	-	-	20	0.175	20	0	-	-
	8	Peoples of Northern Europe (Latvians, Lithuanians, Swedes)	-	-	-	-	34	0	-	-
	9	Peoples of Central Europe (Germans, Poles, Slovaks)	-	-	-	-	45	0	-	-
	10	Peoples of Southern Europe (Spaniards, Italians)	-	-	-	-	36	0	-	-
	11	Peoples of Southern Europe (Macedonians)	-	-	-	-	-	-	49	0.153
	12	Ukrainians (eastern)	-	-	-	-	-	-	64	0.320
	13	Ukrainians (western)	31	0.475	58	0.103	58	0	-	-
	14	Ukrainians (western, central)	-	-	-	-	-	-	37	0.284
	15	Ukrainians (northeastern)	-	-	65	0.154	70	0.043	-	-
	16	Ukrainians (central, eastern)	47	0.505	-	-	-	-	-	-
	17	Ukrainians (central, southeastern)	-	-	70	0.114	72	0	-	-
NORTH OF EUROPEAN PART OF RUSSIA	18	Vepsians	-	-	37	0.189	38	0.013	-	-
	19	Vod, Izhora	-	-	46	0.141	47	0	-	-
	20	Western finnish-speaking peoples (Vepsians, Vod, Izhora, Ingrian Finns)	32	0.612	-	-	-	-	70	0.236
	21	Karelians (North, Tver, South)	26	0.519	-	-	-	-	103	0.238
	22	Karelians (Tver)	-	-	78	0.108	78	0	-	-
	23	Karelians (North), Sami	-	-	20	0.200	21	0	-	-
	24	Karelians (South)	-	-	50	0.100	53	0	-	-
	25	Russians of Arkhangelsk region (Krasnoborsky, Lensky districts)	29	0.557	-	-	-	-	-	-
	26	Russians of Arkhangelsk region (Krasnoborsky district)	-	-	57	0.114	57	0	-	-
	27	Russians of Arkhangelsk region (Leshukonsky, Mezensky districts)	-	-	31	0.177	34	0	-	-
	28	Russians of Arkhangelsk region (Leshukonsky, Pinezhsky, Mezensky districts)	36	0.500	-	-	-	-	-	-
	29	Russians of Arkhangelsk region (Pinezhsky district)	-	-	65	0.115	65	0	-	-
	30	Russians of Arkhangelsk region (Pomors of the Winter, Summer, Onega coasts)	27	0.544	52	0.183	55	0.027	-	-
	31	Northern Russians (Arkhangelsk, Kostroma regions)	-	-	-	-	-	-	46	0.272
	32	Northern Russians (Pinezhsky district of Arkhangelsk region)	-	-	-	-	-	-	45	0.322
	33	Northern Russians (Pomors, Mezensky, Leshukonsky districts of Arkhangelsk region)	-	-	-	-	-	-	40	0.300
	34	Ingrian Finns	-	-	29	0.207	33	0	-	-
CENTER AND SOUTH OF EUROPEAN PART OF RUSSIA	35	Cossacks (Don, Kuban, Nekrasov, Terek)	37	0.435	59	0.085	59	0	-	-
	36	Kalmyks	36	0.667	36	0.264	36	0.028	-	-
	37	Russians of Belgorod, Kursk, Oryol regions	28	0.500	59	0.093	59	0.008	-	-
	38	Russians of Bryansk, Smolensk, Tver regions	42	0.577	62	0.177	62	0	-	-
	39	Russians of Vologda and Kostroma regions	35	0.378	53	0.075	54	0	-	-
	40	Russians of Voronezh and Tambov regions	24	0.677	42	0.143	42	0	-	-
	41	Russians of Kaluga and Ryazan regions	32	0.468	47	0.096	47	0	-	-
	42	Russians of Yaroslavl region	29	0.322	-	-	-	-	-	-
	43	Russians of Nizhny Novgorod region	-	-	28	0.018	28	0	-	-
	44	Russians of Novgorod and Pskov regions	43	0.550	71	0.127	72	0	-	-
	45	Russians of the Yaroslavl region	-	-	68	0.132	68	0.015	-	-

	46	Russians of North-West (Novgorod, Pskov regions)	-	-	-	-	-	-	34	0.309
	47	Central Russians (Nizhny Novgorod, Smolensk, Tver, Yaroslavl regions)	-	-	-	-	-	-	59	0.314
	48	Southern Russians (Belgorod, Voronezh, Kaluga, Kursk, Oryol, Ryazan, Tambov regions)	-	-	-	-	-	-	70	0.293
NORTH CAUCASUS, CRIMEA, TRANSCAUCASIA	49	Western Caucasus and Crimea (Adygea, Kabardino-Balkaria, Karachay-Cherkessia, peoples of Crimea)	107	0.502	390	0.112	390	0.001	264	0.178
	50	Central Caucasus (Ossetia, Ingushetia, Chechnya)	78	0.464	328	0.136	328	0.020	244	0.158
	51	Eastern Caucasus (Dagestan)	129	0.452	655	0.079	1003	0.004	648	0.201
	52	Transcaucasia (Azerbaijan, Armenia, Georgia)	123	0.494	181	0.124	202	0.007	45	0.300
VOLGA-URAL REGION	53	Bashkirs (northern, western)	-	-	63	0.127	63	0.016	-	-
	54	Bashkirs (southeastern)	-	-	37	0.135	37	0.014	-	-
	55	Bashkirs (total)	44	0.584	-	-	-	-	52	0.240
	56	Komi-Permyaks (northern)	-	-	76	0.145	80	0.006	-	-
	57	Komi-Permyaks (southeastern)	-	-	80	0.144	81	0	-	-
	58	Komi-Permyaks (southwestern)	-	-	51	0.137	51	0	-	-
	59	Komi-Permyaks (total)	-	-	-	-	-	-	141	0.252
	60	Komi (Komi-Permyaks, Komi-Zyrians)	49	0.495	-	-	-	-	-	-
	61	Mari (Hill)	-	-	52	0.106	52	0.038	-	-
	62	Mari (Meadow)	-	-	76	0.059	77	0.006	-	-
	63	Mari (total)	31	0.440	-	-	-	-	83	0.313
	64	Peoples of Mordovia (Moksha, Shoksha)	-	-	72	0.125	72	0	-	-
	65	Peoples of Mordovia (Erzya)	-	-	86	0.099	90	0	-	-
	66	Peoples of Mordovia (total)	56	0.463	-	-	-	-	-	-
	67	Nogais (Astrakhan, Kuban, Stavropol, Kara-Nogai people)	34	0.642	34	0.191	35	0.043	-	-
	68	Volga Tatars (Kazan, Astrakhan)	26	0.392	83	0.157	84	0.054	-	-
	69	Volga Tatars (Mishars, Kryashens)	29	0.455	79	0.127	79	0.006	-	-
	70	Volga Tatars (Kazan, Kryashens, Mishars)	-	-	-	-	-	-	95	0.221
	71	Udmurts	-	-	-	-	-	-	47	0.170
	72	Udmurts, Besermyan	51	0.594	112	0.121	113	0.018	-	-
	73	Chuvash (anat jenchi)	-	-	79	0.089	79	0.006	-	-
	74	Chuvash (anatri, virjal)	-	-	55	0.109	55	0	-	-
	75	Chuvash (total)	34	0.594	-	-	-	-	89	0.281
WESTERN AND CENTRAL SIBERIA	76	Buryats (Buryatia, Transbaikal region, Irkutsk region)	32	0.661	41	0.171	41	0.061	-	-
	77	Mansi	-	-	40	0.180	40	0.125	-	-
	78	Peoples of Western Siberia (Mansi, Nenets, Khanty)	59	0.521	-	-	-	-	-	-
	79	Nenets	-	-	21	0.119	21	0	-	-
	80	Siberian Tatars (Baraba, Tobol, Zabolotnie)	48	0.661	49	0.235	49	0.051	-	-
	81	Siberian Tatars (Tyumen-Turin)	35	0.609	45	0.144	45	0.033	-	-
	82	Ugrians (Mansi, Khanty)	-	-	-	-	-	-	40	0.200
	83	Khanty	-	-	56	0.054	56	0.009	-	-
	84	Chukchi people	-	-	35	0	35	0.057	15	0.033
	85	Evenks (Baikal)	-	-	-	-	-	-	29	0.052
	86	Evenks (Baikal), Hamnigan	25	0.566	50	0.060	51	0.088	-	-
	87	Yakuts	39	0.599	41	0.159	41	0.037	-	-
SOUTH SIBERIA	88	Altai people, Shors	56	0.655	-	-	-	-	50	0.120
	89	Nothern Altai people, Shors	-	-	59	0.220	59	0.068	-	-
	90	Southern Altai people	-	-	48	0.156	48	0.042	-	-
	91	Tofalar people	-	-	29	0.052	29	0	-	-
	92	Tuvans, Tofalar people	62	0.381	-	-	-	-	53	0.075
	93	Tuvans (northern)	-	-	43	0.047	43	0.047	-	-
	94	Tuvans (central, southern)	-	-	41	0.049	41	0.024	-	-
	95	Khakas	32	0.559	32	0.141	32	0.063	-	-
FAR EAST	96	Itelmens	29	0.371	29	0.069	29	0	-	-
	97	Koryaks	-	-	60	0.075	60	0.042	-	-
	98	Koryaks, Chukchi people	35	0.447	-	-	-	-	-	-

	99	Nanai people	-	-	73	0.185	74	0.115	47	0
	100	Peoples of Amur (Nanai people, Nivkh people, Ulchis, Negidals)	49	0.782	-	-	-	-	-	-
	101	Peoples of Amur (Nivkh people, Ulchis, Negidals)	-	-	30	0.433	30	0.083	-	-
	102	Peoples of Kamchatka	-	-	-	-	-	-	36	0.028
	103	Evenks, Evens (Okhotsk)	-	-	-	-	-	-	25	0
	104	Evenks (Okhotsk), Oroch people	25	0.693	34	0.177	35	0.057	-	-
	105	Evens (Okhotsk, Kamchatka)	31	0.596	58	0.112	58	0.078	-	-
CENTRAL AND EAST ASIA	106	Dungan people, Uyghurs	-	-	34	0.250	34	0.059	-	-
	107	Kazakhs	-	-	55	0.164	55	0.045	46	0.141
	108	Kyrgyz	43	0.682	51	0.245	51	0.049	-	-
	109	Peoples of East Asia	-	-	504	0.312	504	0.056	504	0.015
	110	Peoples of Mongolia (other than Khalkha)	70	0.655	94	0.218	94	0.032	-	-
	111	Peoples of Mongolia (Khalkha)	58	0.731	67	0.224	68	0.088	-	-
	112	Peoples of Pamirs (Wakhans, Ryns, Gorans, Shugnans, Rushans, Bartangs Wanchs, Shahdarins, Yazgulyamites)	30	0.548	-	-	-	-	45	0.133
	113	Peoples of Pamirs (northern)	-	-	44	0.159	44	0	-	-
	114	Peoples of Pamirs (southern)	-	-	33	0.106	33	0	-	-
	115	Peoples of Middle Asia (Kazakhs, Karakalpaks, Turkmens)	30	0.516	-	-	-	-	-	-
	116	Peoples of Middle Asia (Karakalpaks, Kyrgyz, Tajiks, Uzbeks)	-	-	-	-	-	-	24	0.146
	117	Peoples of Tajikistan (Tajiks, Yagnobis)	41	0.469	50	0.140	50	0	-	-
	118	Peoples of Central Asia (Dungan people, Mongols, Uyghurs)	-	-	-	-	-	-	55	0.100
	119	Turkmens, Karakalpaks	-	-	30	0.133	30	0.033	-	-
	120	Uzbeks	38	0.628	40	0.200	40	0.038	-	-
SOUTH ASIA	121	Peoples of South Asia (Paniya, Pashtuns, Sakilli, North Kannadi, Hazaras)	-	-	-	-	26	0	-	-
	122	Peoples of South Asia (Ladakh, Tibetans, Farsi)	15	0.365	15	0.067	15	0	-	-
WEST ASIA	123	Arabs	-	-	-	-	50	0		
	124	Ashkenazi Jews	-	-	-	-	29	0	-	-
	125	Jews	-	-	-	-	176	0.003	-	-
	126	Peoples of Anatolia and Levant	-	-	-	-	55	0	-	-
AFRICA	127	Peoples of Africa	-	-	-	-	41	0	-	-
AMERICA	128	Peoples of America	-	-	-	-	39	0	-	-
	129	Peoples of Greenland	-	-	-	-	39	0	-	-

frequencies ($q < 0.44$) are concentrated not in the west of the region, but in the plain stretching from the Russian Vologda and Kostroma regions to peoples of the Volga-Ural region, then to southern Russian populations and after that to peoples of Central and Eastern Caucasus. The second low frequency center is found in the north of the Far East ($0.37 < q < 0.45$). However, when moving southwards from the latter, frequency increases rapidly to maximum values ($0.70 < q < 0.78$) found in Evenks of the Okhotsk Coast and peoples living along Amur. Siberian peoples are characterized by great *1 genetic diversity: high frequencies ($0.61 < q < 0.66$) prevail in Eastern Siberia (Yakut) and in the south of Western Siberia (Siberian Tatars), which decrease ($0.52 < q < 0.56$) in the north of Western Siberia and in Southern Siberia.

The combination of *1 longitudinal and latitudinal variability is even more prominent on the map of "normal" *1/*1 homozygous genotype (Fig. 1B) and the map of "slow" *1/*2 heterozygote (Fig. 1C). An "ultrafast" *1/*17 heterozygote (Fig. 1E) associated with the risk of hemorrhage is characterized by clear, but oppositely directed vector: frequency dropping from west ($q = 0.40$) to east ($q = 0$).

Gene geography of *CYP2C19*2* (Fig. 2; Tables 1, 2)

The gene geographic variability of *2 variant is similar to that of *1. At significantly lower frequencies ($0 < q < 0.43$, $\bar{q} = 0.15$; Fig. 2A) overlapping of two trends is found again. The main trend is longitudinal once again, with the *2 frequency increase from west to southeast, where the primary maximum falls on Central Asia ($0.20 < q < 0.31$) with frequency surge in the Amur region ($q = 0.43$). The latitudinal trend is observed in Siberia: the *2 frequency increases from north to Central Asia. Both trends demonstrate numerous irregularities.

As for European part of the assessed range, additional *2 frequency maximum ($0.19 < q < 0.24$) is found in the northwest in the Veps, Sami, North Karelians, Ingrian Finns. Frequencies above the average ($0.17 < q < 0.18$) are also found in peoples speaking Indo-European languages: in the Russian North and in the west of the European part of the region (Balkans, Belarus, west of Russia, Moldova, Ukraine).

In the Russian populations, the *2 frequency varies within a broad range ($0.02 < q < 0.18$, $\bar{q} = 0.12$). In the Cis-Urals area at $\bar{q} = 0.11$ it varies within a narrower range ($0.06 < q < 0.15$).

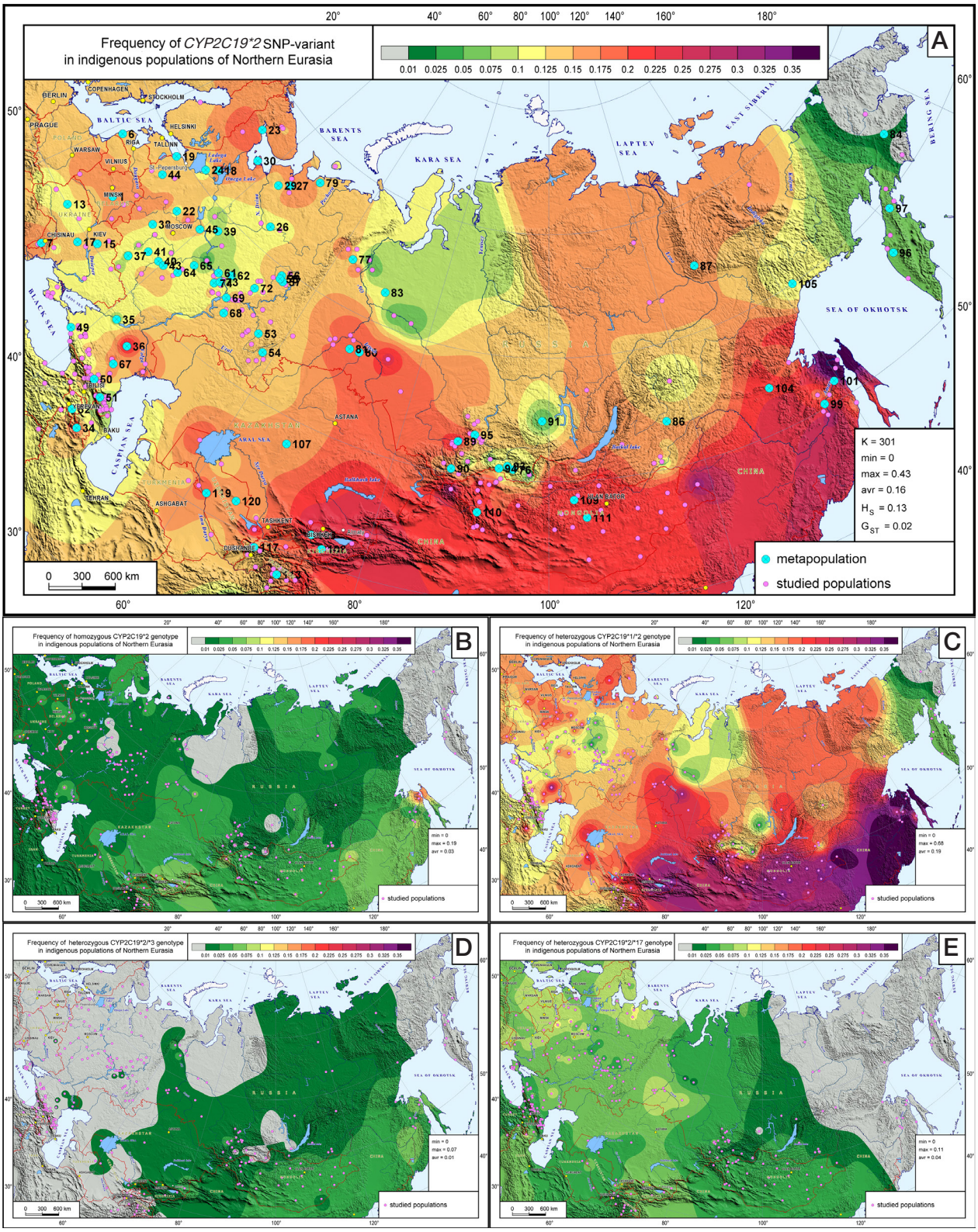

Fig. 2. Spatial *CYP2C19*2* variability in indigenous population of Russia and neighboring countries. **A**. Gene geography of *CYP2C19*2* SNP variant frequency. **B**. Gene geography of **2/*2* genotype frequency. **C**. Gene geography of **2/*1* genotype frequency. **D**. Gene geography of **2/*3* genotype frequency. **E**. Gene geography of **2/*17* genotype frequency.

In the Trans-Urals region, in Ob-Ugrians, the unexpectedly large differences between Khanty (q = 0.05) and Mansi (q = 0.18) have been revealed. In the North Caucasus the **2* frequency varies within a very broad range: between q = 0.08 in Dagestan and q = 0.19 in the Chechens and Ingush. The quite expected frequency increase in the Kalmyks (q = 0.26) results from preserving the memory of their Central Asian ancestral homeland.

The trend becomes latitudinal in the Asian part of the region: the frequency increases from north to south. The range of populations with low frequencies is huge: from Khanty in the west (q = 0.05) to Kamchatka (q = 0.07) and Chukotka (q = 0) in the east. Southwards, it extends to South Siberia (q = 0.05) and the Baikal region (q = 0.06). Increased frequencies are reported in Buryats and Yakuts (q = 0.16). The Central Asian maximum

Table 2. Frequencies of *CYP2C19* genotypes in indigenous population of Russia and the world

Region	Population	*CYP2C19* genotype																			
		*1/*1		*1/*2		*1/*3		*1/*17		*2/*2		*2/*3		*2/*17		*3/*3		*3/*17		*17/*17	
		n	$\bar{q}$	*n*	$\bar{q}$	*n*	$\bar{q}$	*n*	$\bar{q}$	*n*	$\bar{q}$	*n*	$\bar{q}$	*n*	$\bar{q}$	*n*	$\bar{q}$	*n*	$\bar{q}$	*n*	$\bar{q}$
Foreign Europe	Belarusians	33	0.333	33	0.196	33	0.023	19	0.334	50	0.029	50	0.007	19	0.098	50	0	19	0.012	19	0.084
	Eastern Europe	21	0.286	21	0.184	21	0	21	0.164	90	0.030	90	0	49	0.053	149	0	49	0	49	0.023
	Northern peoples of Balkans	-	-	-	-	-	-	-	-	-	-	-	-	-	-	51	0	-	-	-	-
	Central peoples of Balkans	-	-	-	-	-	-	-	-	-	-	-	-	-	-	80	0	-	-	-	-
	Peoples of Europe (total)	-	-	-	-	-	-	-	-	503	0.021	503	0	-	-	503	0	-	-	-	-
	Peoples of Northern Europe	-	-	-	-	-	-	-	-	-	-	-	-	-	-	34	0	-	-	-	-
	Peoples of Central Europe	-	-	-	-	-	-	-	-	-	-	-	-	-	-	45	0	-	-	-	-
	Peoples of Southern Europe	-	-	-	-	-	-	-	-	-	-	-	-	-	-	36	0	-	-	-	-
	Ukrainians	78	0.243	78	0.123	78	0.015	78	0.303	193	0.015	193	0.004	101	0.076	200	0	101	0.009	101	0.094
North of european part of Russia	Western finnish-speaking peoples	32	0.375	32	0.213	32	0.005	32	0.288	112	0.030	112	0.001	70	0.082	118	0	70	0.002	70	0.056
	Karelians, Sami	26	0.269	26	0.122	26	0	26	0.247	148	0.014	148	0	103	0.056	152	0	103	0	103	0.057
	Northern Russians	92	0.282	92	0.150	92	0.008	92	0.316	205	0.020	205	0.002	131	0.084	211	0	131	0.004	131	0.089
Center and south of european part of Russia	Cossacks	37	0.189	37	0.074	37	0	-	-	59	0.007	59	0	-	-	59	0	-	-	-	-
	Kalmyks	36	0.445	36	0.352	36	0.037	-	-	36	0.070	36	0.015	-	-	36	0.001	-	-	-	-
	Northeastern Russians	35	0.143	35	0.057	35	0	-	-	53	0.006	53	0	-	-	54	0	-	-	-	-
	Northwestern Russians	43	0.303	43	0.139	43	0	34	0.340	71	0.016	71	0	34	0.078	72	0	34	0	34	0.095
	Central Russians	71	0.224	71	0.123	71	0.006	59	0.297	158	0.017	158	0.002	59	0.081	158	0	59	0.004	59	0.098
	Southern Russians	84	0.290	84	0.116	84	0.004	70	0.315	148	0.012	148	0.001	70	0.063	148	0	70	0.002	70	0.086
North caucasus. crimea.trans-caucasia	Western Caucasus and Crimea	107	0.252	107	0.112	107	0.001	107	0.179	390	0.012	390	0	264	0.040	390	0	264	0	264	0.032
	Central Caucasus	78	0.215	78	0.126	78	0.018	78	0.146	328	0.018	328	0.005	244	0.043	328	0	244	0.006	244	0.025
	Eastern Caucasus	129	0.204	129	0.072	129	0.004	129	0.181	655	0.006	655	0.001	648	0.032	1003	0	648	0.002	648	0.040
	Transcaucasia	123	0.244	123	0.123	123	0.007	45	0.296	181	0.015	181	0.002	45	0.075	202	0	45	0.004	45	0.090
Volga-ural region	Bashkirs	44	0.341	44	0.152	44	0.018	44	0.281	100	0.017	100	0.004	52	0.063	100	0	52	0.007	52	0.058
	Komi	49	0.245	49	0.141	49	0.002	49	0.249	207	0.020	207	0.001	141	0.072	212	0	141	0.001	141	0.063
	Mari	31	0.194	31	0.069	31	0.017	31	0.276	128	0.006	128	0.003	83	0.049	129	0	83	0.012	83	0.098
	Peoples of Mordovia	56	0.214	56	0.103	56	0	-	-	158	0.012	158	0	-	-	162	0	-	-	-	-
	Nogais	34	0.412	34	0.246	34	0.055	-	-	34	0.037	34	0.016	-	-	35	0.002	-	-	-	-
	Volga Tatars	55	0.181	55	0.121	55	0.026	55	0.188	162	0.020	162	0.009	95	0.063	163	0.001	95	0.014	95	0.049
	Udmurts, Besermyan	51	0.353	51	0.143	51	0.021	47	0.202	112	0.015	112	0.004	47	0.041	113	0	47	0.006	47	0.029
	Chuvash	34	0.353	34	0.115	34	0.004	34	0.334	134	0.009	134	0.001	89	0.055	134	0	89	0.002	89	0.079
Western and central siberia	Buryats	32	0.437	32	0.226	32	0.081	-	-	41	0.029	41	0.021	-	-	41	0.004	-	-	-	-
	Peoples of the north of Western Siberia	59	0.271	59	0.113	59	0.049	40	0.208	117	0.012	117	0.010	40	0.043	117	0.002	40	0.019	40	0.040
	Siberian Tatars	83	0.409	83	0.254	83	0.056	-	-	94	0.039	94	0.017	-	-	94	0.002	-	-	-	-
	Yakuts	39	0.359	39	0.190	39	0.044	-	-	41	0.025	41	0.012	-	-	41	0.001	-	-	-	-
South siberia	Nothern Altai people, Shors	56	0.429	56	0.289	56	0.089	50	0.157	59	0.049	59	0.030	50	0.053	59	0.005	50	0.016	50	0.014
	Southern Altai people	-	-	-	-	-	-	-	-	48	0.024	48	0.013	-	-	48	0.002	-	-	-	-
	Tuvans, Tofalar people	62	0.145	62	0.037	62	0.007	53	0.058	113	0.002	113	0.001	53	0.007	113	0	53	0.001	53	0.006
	Khakas	32	0.312	32	0.157	32	0.070	-	-	32	0.020	32	0.018	-	-	32	0.004	-	-	-	-

Far east	Itelmens	29	0.138	29	0.051	29	0	-	-	29	0.005	29	0	-	-	29	0	-	-	-	-
	Koryaks, Chukchi people	35	0.200	35	0.042	35	0.115	35	0.026	95	0.002	95	0.012	51	0.003	95	0.017	51	0.008	51	0.001
	Peoples of Amur	49	0.612	49	0.402	49	0.165	47	0	103	0.066	103	0.054	47	0	104	0.011	47	0	47	0
	Evenki peoples	81	0.380	81	0.135	81	0.094	54	0.034	142	0.012	142	0.017	54	0.006	144	0.006	54	0.004	54	0.001
Central and east Asia	Peoples of East Asia	-	-	-	-	-	-	-	-	504	0.097	504	0.035	504	0.009	504	0.003	504	0.002	504	0
	Pamir peoples	30	0.300	30	0.149	30	0	30	0.146	77	0.019	77	0	45	0.036	77	0	45	0	45	0.018
	Peoples of Central Asia	152	0.334	152	0.207	152	0.038	70	0.165	226	0.032	226	0.012	70	0.051	226	0.001	70	0.009	70	0.020
	Peoples of Mongolia	128	0.475	128	0.311	128	0.077	55	0.138	195	0.051	195	0.025	55	0.045	196	0.003	55	0.011	55	0.010
South Asia	Peoples of South Asia	15	0.133	15	0.049	15	0	-	-	15	0.004	15	0	-	-	41	0	-	-	-	-
West Asia	Arabs	-	-	-	-	-	-	-	-	-	-	-	-	-	-	50	0	50	0	0	0
	Jews	-	-	-	-	-	-	-	-	-	-	-	-	-	-	205	0	-	-	-	-
	Peoples of Anatolia and Levant	-	-	-	-	-	-	-	-	-	-	-	-	-	-	55	0	-	-	-	-
Africa	Peoples of Africa	-	-	-	-	-	-	-	-	-	-	-	-	-	-	41	0	-	-	-	-
America	Peoples of America	-	-	-	-	-	-	-	-	-	-	-	-	-	-	39	0	-	-	-	-
	Peoples of Greenland	-	-	-	-	-	-	-	-	-	-	-	-	-	-	39	0	-	-	-	-

in the west spans the Dungan people, Kyrgyz and Uighurs (q = 0.25), Mongols, Northern Altai people and Shors (q = 0.22), Siberian Tatars and Uzbeks (q = 0.20). In the Amur region, the *2 frequency in the Nanai people, Oroch people and Evenks is lower (0.18 < q < 0.19), however, an unexpectedly sharp surge is observed in the combined population of the most ancient Far Eastern peoples: Negidals, Nivkh people and Ulchis (q = 0.43).

Gene geography of pharmacogenetically significant genotypes *2/*2 (Fig. 2B) and *2/*3 (Fig. 2D) is monotonous. The homozygote varying over a wide range (0 < q < 0.19) is spread across the region with low frequency ($\bar{q}$ = 0.03) showing a slight increase toward Central Asia and reaching its maximum in the Far East. There is almost no *2/*3 heterozygote in Europe, Urals and Western Siberia, low frequencies (q = 0.07) are reported in Mongolia and the Amur region.

Gene geography of *CYP2C19*3* (Fig. 3; Tables 1, 2)

Spatial variability of the *3 variant differs from previous variants showing a far stronger trend (Fig. 3A). That is why these have the same level of interpopulation differences (G_{ST} = 0.02), despite huge differences in heterozygosity of variants *3 (H_S = 0.04) and *2 (H_S = 0.13), as well as the range of their frequencies (*3 — 0 < q < 0.10; *2 — 0 < q < 0.40).

Furthermore, the maximum *3 frequencies are once again concentrated in the southeast (0.08 < q < 0.12) of Transbaikalia, in Khalkha Mongols and in the Amur region. Some frequency increase is observed in the Evenks of Okhotsk coast and the Chukchi people (q = 0.06) continuing west to the number of peoples of South Siberia (0.06 < q < 0.07) and Central Asia (0.05 < q < 0.06). A sharp frequency surge in the Mansi (q = 0.12) is the exception to this pattern.

In the European part of the region the *3 variant is missing or extremely rare. In Slavic populations, notable frequency have been found only in the Belarusians, Russians of Arkhangelsk region (q = 0.03), Russians of Yaroslavl region, as well as in the Central Caucasus (q = 0.02).

Gene geography of pharmacogenetically significant genotypes is discussed in other sections (Fig. 2D for *2/*3; Fig. 4D for 3/*17).

In general, the *3 variant is characterized by gradual frequency increase from zero values in the west of North Eurasia to low frequency (q = 0.12) in the east and southeast of the region. However, these low frequencies turn out to be maximum frequencies on a global scale (Fig. 5A): high world's frequencies are concentrated in East Asia with their maximum in the Amur region.

Gene geography of *CYP2C19*17* (Fig. 4; Tables 1, 2).

The trend of *17 variant variability is much stronger, and the vector is oppositely directed: the natural frequency drop from maximum values (q = 0.32) in the west of North Eurasia to zero frequency in the east and southeast of the region. This clear variability, even with low average frequency ($\bar{q}$ = 0.13), results in the interpopulation difference value (G_{ST} = 0,08) 4 times higher than that of other variants.

The maximum frequency range (0.27 < q < 0.32) included 12 populations out of 35 assessed using this SNP marker: all Slavic populations (Belarusians, Russians, Ukranians), Mari and Chuvash of the Ural region, peoples of Transcaucasia.

The next interval (0.20 < q < 0.25) brought together the Finnish-speaking peoples (all western Finnish-speaking populations and Komi-Permyaks), Ob-Ugrians, Urals Turkic (Bashkir and Volga Tatars), and peoples of Dagestan.

Only two European populations (Udmurts and peoples of South Europe) have been found among other populations showing above the average frequencies (0.13 < q < 0.18), and the trend is shifted southeastward: here peoples of Kazakhstan, Pamirs and Central Asia coexist with the populations of Western and Central Caucasus.

Only Asian peoples are found among populations showing below the average frequencies (0 < q < 0.12): South Siberia (Altai people, Tofalars, Tuvans, Shors), Baikal region (Evenks), Far East (Nanai people, peoples of Kamchatka, Evenks, Evens, Chukchi people), East and Central Asia.

Both genotypes of "ultrafast" metabolizers (*17/*17, Fig. 4B; *1/*17, Fig. 4C) are characterized by the same variability trend: obvious frequency drop from west to east. However, their variability ranges are different: the range of variant *17/*17 is small (0 < q < 0.10) and that of variant *1/*17 is 4 times larger (0 < q < 0.40).

The *3/*17 heterozygote (Fig. 4E) encoding intermediate metabolizers has been found showing extremely low frequencies

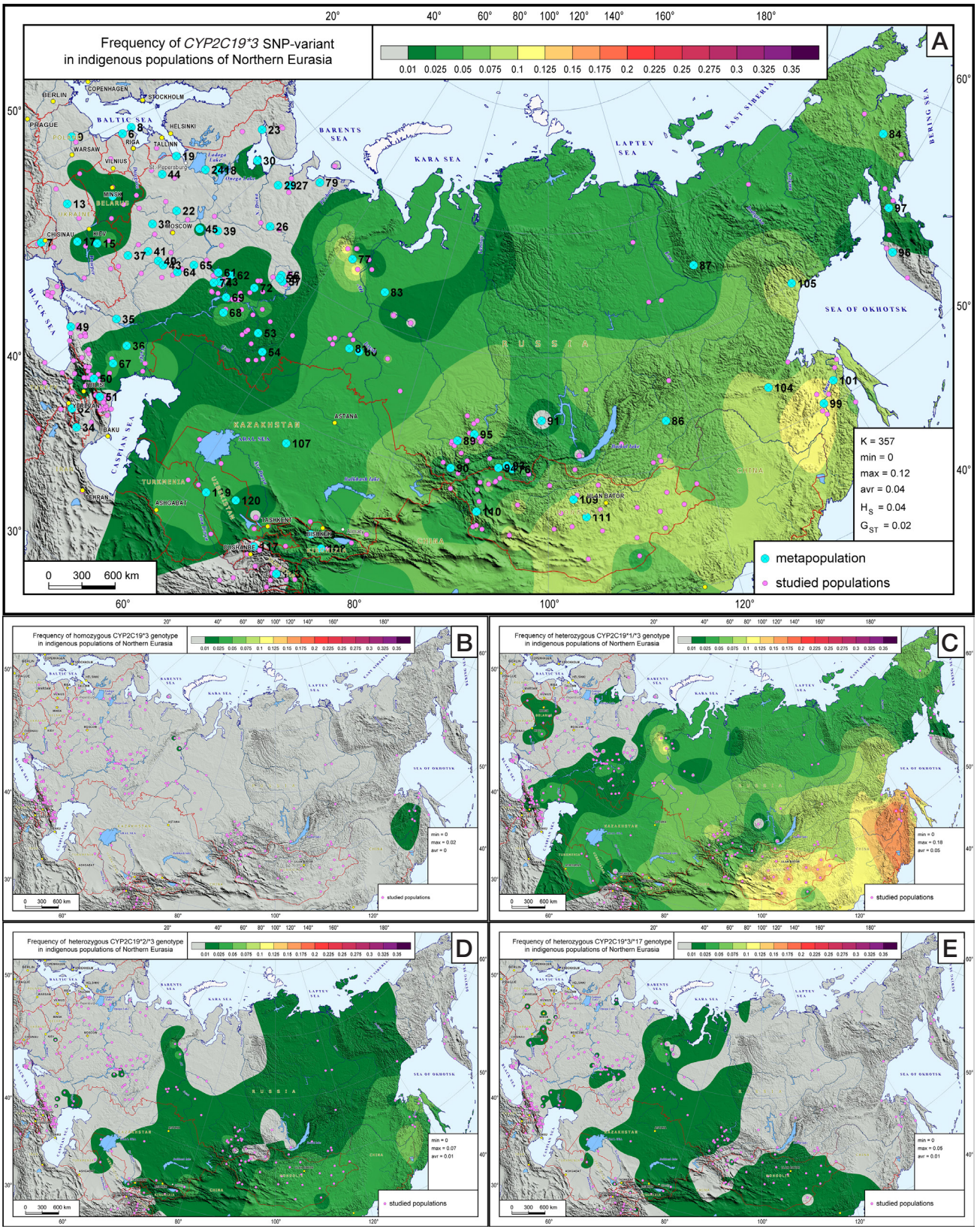

Fig. 3. Spatial *CYP2C19*3* variability in indigenous population of Russia and neighboring countries. **A**. Gene geography of *CYP2C19*3* SNP variant frequency. **B**. Gene geography of *3/*3 genotype frequency. **C**. Gene geography of *3/*1 genotype frequency. **D**. Gene geography of *3/*2 genotype frequency. **E**. Gene geography of *3/*17 genotype frequency

(0 < q < 0.05), it is almost absent in the west and east of North Eurasia. It has unique geography: the range where *3/*17 is present forms a continuous strip stretching from Ob-Ugrians in the north to peoples of Central Asia, then stretching to the east, to Mongolia and East Asia.

DISCUSSION

The population of North Eurasia is the area of most ancient (since Paleolithic times) interactions between major racial branches of the humankind: western (Caucasoids) and eastern

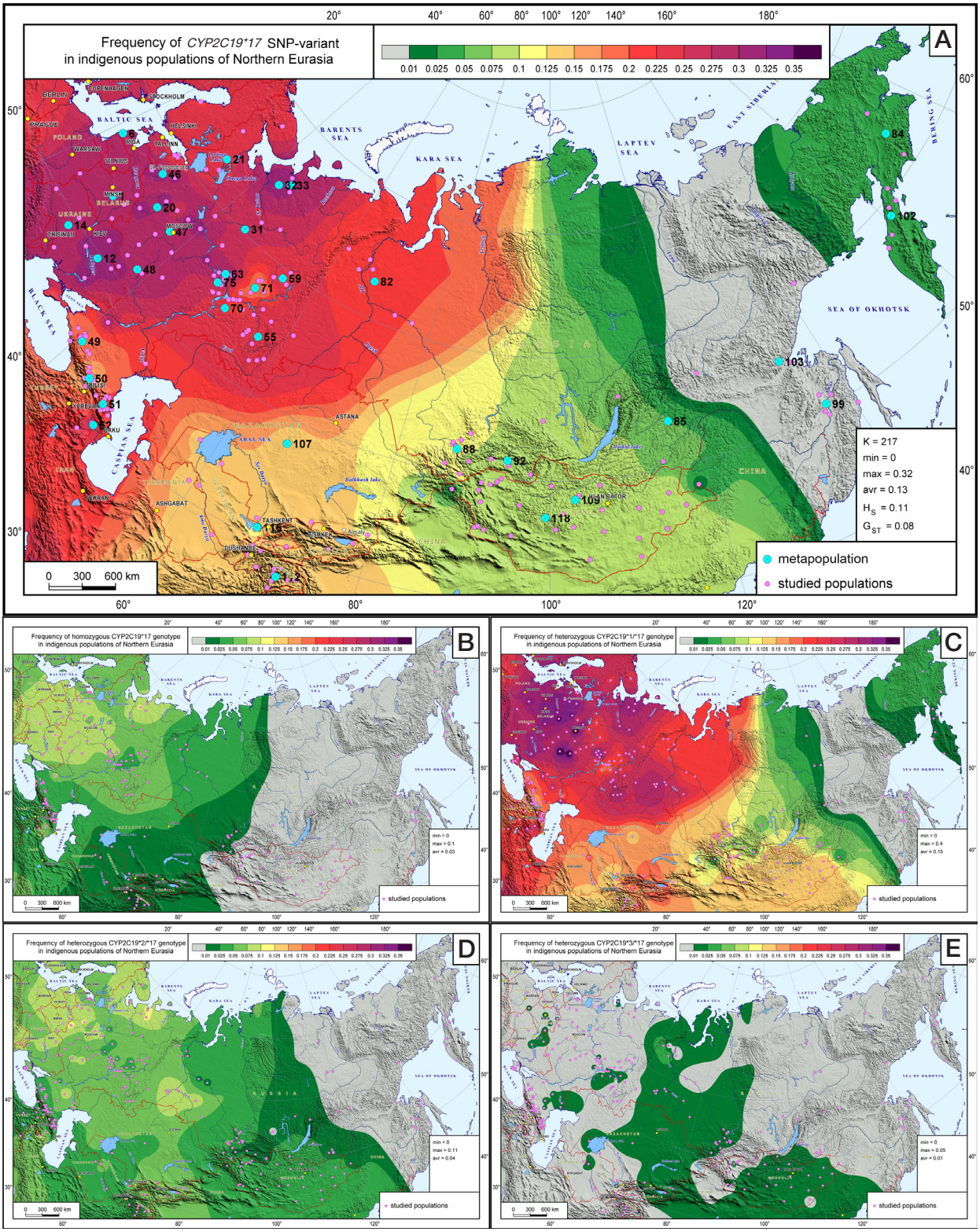

Fig. 4. Spatial *CYP2C19*17* variability in indigenous population of Russia and neighboring countries. **A**. Gene geography of *CYP2C19*17* SNP variant frequency. **B**. Gene geography of *17/*17 genotype frequency. **C**. Gene geography of *17/*1 genotype frequency. **D**. Gene geography of *17/*2 genotype frequency. **E**. Gene geography of *17/*3 genotype frequency

(Mongoloids). These interactions are clearly reflected by the *CYP2C19* SNP marker maps (Fig. 1–4). However, it has been no less convincingly shown how tough and imprecise is confining the real picture showing variability of these variants to the straightforward scheme of two racial poles.

Information about the fact that *CYP2C19*2* is two times less often found in Caucasoids (15%) than in Mongoloids of East Asia (31%) [5] is obviously insufficient. True situation is much more complex: the actual *2 variability (Fig. 2) represents the imposition of the latitudinal vector and additional maximum in Europe on the common longitudinal vector in Siberia. That is why the use of specific data on the peoples of Russia and neighboring countries in pharmacogenetics is so popular. Such data on the frequencies of four *CYP2C19* variants in multiple

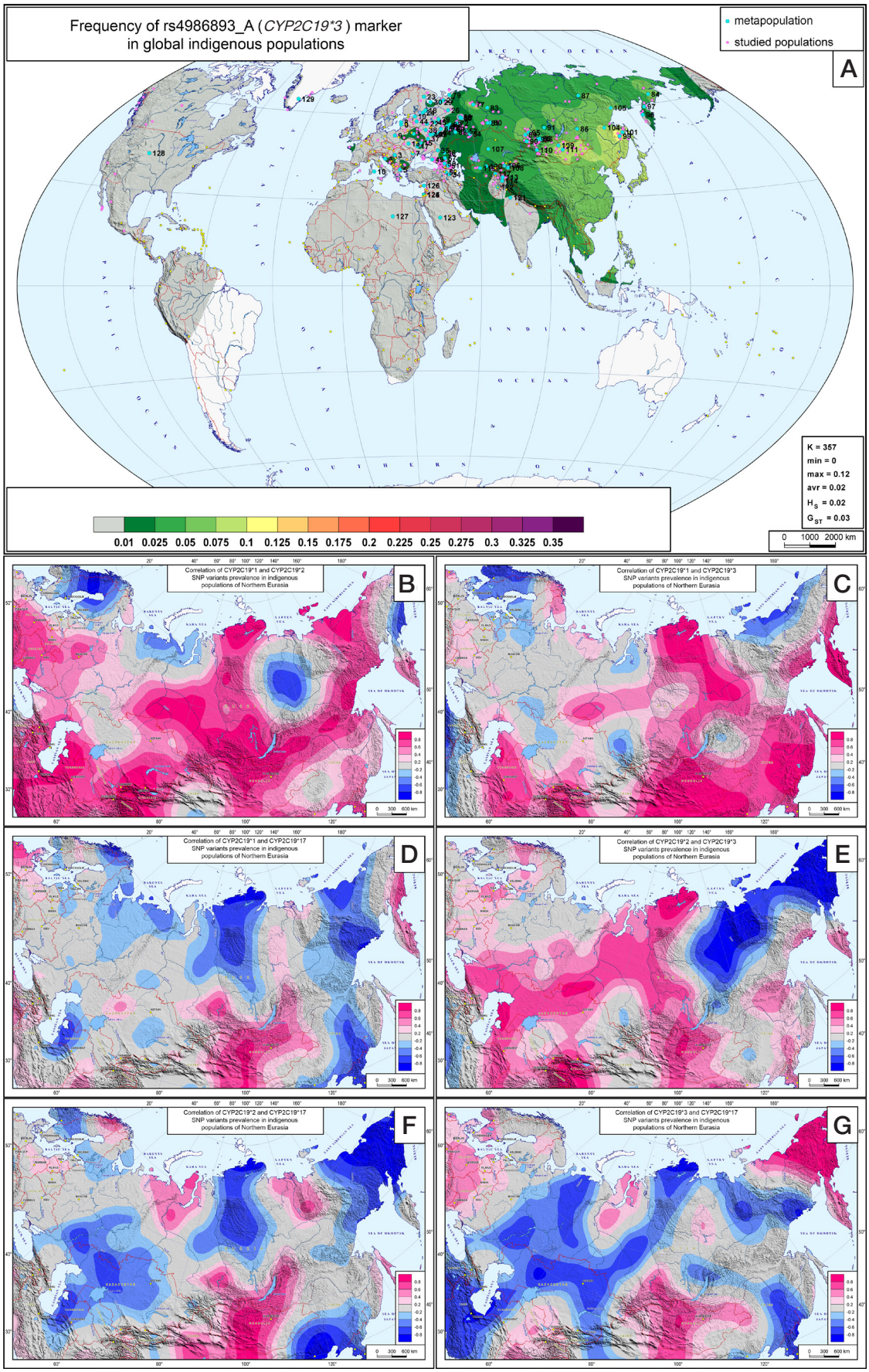

Fig. 5. Maps of correlations between patterns of *1, *2, *3, *17 variant frequency variability and gene geography of *CYP2C19*3* in indigenous population of the world. **A**. Spatial *CYP2C19*3* variability in indigenous population of the world. B. Correlation map for patterns of *1 and *2 variant frequencies. **C**. Correlation map for patterns of *1 and *3 variant frequencies. **D**. Correlation map for patterns of *1 and *17 variant frequencies. **E**. Correlation map for patterns of *2 and *3 variant frequencies. **F**. Correlation map for patterns of *2 and *17 variant frequencies. **G**. Correlation map for patterns of *3 and *17 variant frequencies.

ethnic groups or their sets are provided in Table 1: information about the *2 variant spans 79 populations with the total sample of 6346 individuals. With the *2 frequency variability span between q = 0 and q = 0.43, both its minimum (Chukchi people, q = 0) and maximum (Negidals, Nivkh people, Ulchis, q = 0.43) are in the same region, in the Far East of Russia. This precedent demonstrates that it is impossible to think in terms of generalized ethnoracial categories and there is a need for a differentiated approach. The *CYP2C19*2* frequency in the Russian populations is about 11% [6], but it varies over a very broad range: between q = 0.02 in the Nizhny Novgorod region and q = 0.18 in the Arkhangelsk, Bryansk and Smolensk regions.

The sharper ethnoracial differences are reported for *17 variant in literature: between q = 0.22 in Caucasoids and q = 0.02 in Mongoloids of East Asia [5]. As for European populations, the papers report the increase in *17 frequency in Central and Eastern Europe ($0.25 < q < 0.33$) with the decrease in the north ($0.19 < q < 0.22$), south (up to q = 0.18) and west (q = 0.17) of Europe [18, 19]. Our data on the frequency of *17 variant provided in Table 1 for 35 populations with the total sample of 3313 individuals demonstrate a similar Eurasian*17 frequency range ($0 < q < 0.32$). Furthermore, in Caucasoids, frequency varies within the range between q = 0.15 in peoples of South Europe and q = 0.32 in the north in Russians of Arkhangelsk region. A wide variety of populations conditionally classified as Mongoloids has an equally large span: from 0 in the Far East to q = 0.31 in Mari of the Cis-Urals region. That is why it is necessary to assess the real picture of geographical variability instead of using the straightforward Caucasoid-Mongoloid scheme.

The data reported (Tables 1, 2) provide important information about many ethnic groups of Russia and neighboring countries, populations of which have migrated en masse to Russia. However, these data cover only a part of the genetic diversity of peoples of our countries. That is why the knowledge of the gene geographical variability (Fig. 1–5) predicting frequencies of the *CYP2C19* clinically significant variants for peoples, information about which is currently missing in the databases and published papers, is so important.

Given the common patterns, first of all, it is necessary to emphasize that the geographical trend clarity is not dependent on the abundance of this or that *CYP2C19* variant or its variability span (Fig. 1–4).

Variants *2 and *17 are characterized by similar variability span ($0 < q < 0.4$) and similar average heterozygosity ($0.11 < H_S < 0.12$). However, while *17 demostrates a strong trend of frequency decrease from west to east (Fig. 4), the *2 variability is much more complex (Fig. 2). This apparent difference between *2 and *17 is also indicated by the interpopulation variability (G_{ST}) value: *2 shows interpopulation differences ($G_{ST} = 0.02$) that are 4 times lower than that of *17 ($G_{ST} = 0.08$). However, a less frequent ($0 < q < 0.1$) marker *3 shows a very strong trend (Fig. 3), which results in the same interpopulation difference span ($G_{ST} = 0.02$) as in *2.

The correlation maps (Fig. 5) demonstrate similarity zones in the patterns of the *1, *2, *3, *17 frequency variability (positive correlations are highlighted in red) and in the area of the oppositely directed vectors of their variability (negative correlations are highlighted in blue). The set of six correlation maps (Fig. 5B–G) demonstrates that, given overall similarity of the *1, *2, *3, *17 frequency variability patterns, there are always regions, in which the general pattern is replaced by the correlation with opposite sign. The combination of longitudinal and latitudinal trends (red color) makes similarity of the *1 and *2 gene geography obvious (Fig. 5B), however, a number of exceptions are found in the north showing negative correlation between maps of variants *1 and *2 (blue color). The pronounced similarity of the patterns of maps for *2 and *3 (Fig. 5E) is disrupted in the northeast of the region (negative correlation highlighted in blue). Despite the fact that the *3 and *17 frequency change vectors are generally oppositely directed (negative correlation), but not always alternative to each other, in some regions (foreign Europe, Ob-Ugrians, Middle Asia, South Siberia, north of the Far East) there is a positive correlation between these two maps (Fig. 5G). In general, the correlation maps convincingly show that even when the common trends in pharmacogenetic biomarker variability have been found, it is necessary to continue studying each of the peoples living in multi-ethnic Russia. The upcoming publication will show how high such variability can be on the example of peoples of the Caucasus region. That is why it is necessary to get closer to the real picture of the pharmacogenetic biomarker gene geographical variability and create cartographic atlases for various regions of Russia.

Study limitations

Certain study limitations are related to limited sample sizes of assessed biomaterial samples from some populations (sample sizes are provided in Tables 1, 2).

CONCLUSIONS

The paper provides data in the *CYP2C19*1, *2, *3, *17* SNP marker frequencies and pharmacogenetically significant genotypes in the major ethnic groups of Russia and neighboring countries. The gene geographical variability of *CYP2C19*1* (based on the data on 2261 individuals of 53 populations) combines a longitudinal trend of frequency increase from west to southeast of North Eurasia and latitudinal variability of frequency increase from north to south in the Asian part of the region. The *CYP2C19*2* spatial variability (6346 individuals, 79 populations) is characterized by variability similar to that of *1, however, both trends, longitudinal and latitudinal, are interrupted by local extrema. Gene geography of *CYP2C19*3* (7517 individuals, 92 populations) shows a stronger longitudinal trend of natural frequency increase from 0 in the west to 12% in the east and southeast of North Eurasia. This is a world's maximum: the high frequency area is located in East Asia with the peak frequency in the Amur region. The *CYP2C19*17* gene geographical variability (3313 individuals, 35 populations) is different from that of previous variants, it shows a strong oppositely directed longitudinal trend of frequency decrease from west to southeast. The correlation maps of the *CYP2C19*1, *2, *3, *17* variant frequencies demonstrate regions, in which there is no similarity between the main frequency variability patterns of these *CYP2C19* gene variants. The fact is important for practical use in pharmacogenomics. So long as the currently available data do not cover all peoples of Russia, the gene geographical variability maps are first to predict the *CYP2C19*1, *2, *3, *17* variant frequency and pharmacogenetically significant genotypes for the populations, information about which is missing.

References

1. Mini E, Nobili S. Pharmacogenetics: implementing personalized medicine. Clinical cases in mineral and bone metabolism. 2009; 6 (1):17.
2. Samani NJ, Tomaszewski M, Schunkert H. The personal genome — the future of personalised medicine? The Lancet. 2010; 375 (9725): 1497–8.
3. Collins FS, Varmus H. A new initiative on precision medicine. New England journal of medicine. 2015; 372 (9): 793–5.
4. Sinitsina II, Boyarko AV, Temirbulatov II, Sychev DA, Akmalova KA, Sozaeva ZA, et al. CYP2C9 gene polymorphisms influence on antihypertensive effectiveness and hypouricemic effect of losartan among patients with arterial hypertension: an observational study. Drug Metabolism and Personalized Therapy. 2022; 38 (2): 163–8.
5. Lee CR, Luzum JA, Sangkuhl K, Gammal RS, Sabatine MS, Stein CM, et al. Clinical pharmacogenetics implementation consortium guideline for CYP2C19 genotype and clopidogrel therapy: 2022 update. Clinical Pharmacology & Therapeutics. 2022; 112 (5): 959–67.
6. Niu X, Mao L, Huang Y, Baral S, Li JY, Gao Y, et al. CYP2C19 polymorphism and clinical outcomes among patients of different races treated with clopidogrel: a systematic review and meta-analysis. Journal of Huazhong University of Science and Technology [Medical Sciences]. 2015; 35 (2): 147–56.
7. 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. European journal of clinical pharmacology. 2003; 59: 303–12.
8. 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. Clinical Pharmacology & Therapeutics. 2017; 102 (3): 397–404.
9. Pyljov VJu, Agdzhojan AT, Gorin IO, Petrushenko VS, Pocheshhova JeA, Balanovskaja EV i dr. Populjacionnyj biobank kak osnova dlja vyjavlenija prostranstvennoj izmenchivosti klinicheski znachimyh farmakogeneticheskih biomarkerov serdechno-sosudistyh zabolevanij. Kardiovaskuljarnaja terapija i profilaktika. 2022; 21 (11): 114–34. Russian.
10. Balanovska EV, Petrushenko VS, Koshel SM, Pocheshkhova EA, Chernevskiy DK, Mirzaev KB, Abdullaev ShP, Balanovsky OP, et al. Cartographic atlas of frequency variation for 45 pharmacogenetic markers in populations of Russia and its neighbor states. Bulletin of RSMU. 2020 (6): 38–51.
11. Balanovska EV, Gorin IO, Ponomarev GYu, Pylev VYu, Petrushenko VS, Markina NV, Mamaeva AD, Larin AK, Agdzhoyan AT et al. Footprints of interaction among Finnic-speaking, Slavic, and Turkic-speaking populations in modern gene pool and their reflection in pharmacogenetics. Bulletin of RSMU. 2022; 2: 20–9.
12. Balanovska EV, Napolskih VV, Churakov VS, Pislegin NV, Zapiseckaja YuS, Balanovsky OP, i dr. Genofondy udmurtov i besermjan v kontekste finno-ugorskih i drugih okruzhajushhih narodov: polnogenomnye i farmakogeneticheskie dannye. Ezhegodnik finno-ugorskih issledovanij. 2022; 2: 328–46. Russian.
13. Balanovsky O, Petrushenko V, Gorin I, Agdzhoyan A, Balanovska E, Sychev D, et al. Variation of genomic sites associated with severe Covid-19 across populations: global and national patterns. Pharmacogenomics and Personalized Medicine. 2021: 1391–402.
14. 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): s13742–015.
15. GG-base [cited 2022 Sep 10]. Available from: https://gg-base.org/.
16. Balanovska EV, Zhabagin MK, Agdzhoyan AT, Chuhryaeva MI, Markina NV, Asylguzhin RR, i dr. Populjacionnye biobanki: principy organizacii i perspektivy primenenija v genogeografii i personalizirovannoj medicine. Genetika. 2016; 52 (12): 1371–87. Russian.
17. Koshel SM. Geoinformacionnye tehnologii v genogeografii. Sovremennaja geograficheskaja kartografija. 2012; c. 158–8. Russian.
18. Petrović J, Pešić V, Lauschke VM. Frequencies of clinically important CYP2C19 and CYP2D6 alleles are graded across Europe. European journal of Human Genetics. 2020; 28 (1): 88–94.
19. Pedersen RS, Brasch-Andersen C, Sim SC, et al. Linkage disequilibrium between the CYP2C19*17 allele and wildtype CYP2C8 and CYP2C9 alleles: identification of CYP2C haplotypes in healthy Nordic populations. European journal of clinical pharmacology. 2010; 66: 1199–1205.20.

Литература

1. Mini E, Nobili S. Pharmacogenetics: implementing personalized medicine. Clinical cases in mineral and bone metabolism. 2009; 6 (1):17.
2. Samani NJ, Tomaszewski M, Schunkert H. The personal genome — the future of personalised medicine? The Lancet. 2010; 375 (9725): 1497–8.
3. Collins FS, Varmus H. A new initiative on precision medicine. New England journal of medicine. 2015; 372 (9): 793–5.
4. Sinitsina II, Boyarko AV, Temirbulatov II, Sychev DA, Akmalova KA, Sozaeva ZA, et al. CYP2C9 gene polymorphisms influence on antihypertensive effectiveness and hypouricemic effect of losartan among patients with arterial hypertension: an observational study. Drug Metabolism and Personalized Therapy. 2022; 38 (2): 163–8.
5. Lee CR, Luzum JA, Sangkuhl K, Gammal RS, Sabatine MS, Stein CM, et al. Clinical pharmacogenetics implementation consortium guideline for CYP2C19 genotype and clopidogrel therapy: 2022 update. Clinical Pharmacology & Therapeutics. 2022; 112 (5): 959–67.
6. Niu X, Mao L, Huang Y, Baral S, Li JY, Gao Y, et al. CYP2C19 polymorphism and clinical outcomes among patients of different races treated with clopidogrel: a systematic review and meta-analysis. Journal of Huazhong University of Science and Technology [Medical Sciences]. 2015; 35 (2): 147–56.
7. 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. European journal of clinical pharmacology. 2003; 59: 303–12.
8. 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. Clinical Pharmacology & Therapeutics. 2017; 102 (3): 397–404.
9. Пылёв В. Ю., Агджоян А. Т., Горин И. О., Петрушенко В. С., Почешхова Э. А., Балановская Е. В., и др. Популяционный биобанк как основа для выявления пространственной изменчивости клинически значимых фармакогенетических биомаркеров сердечно-сосудистых заболеваний. Кардиоваскулярная терапия и профилактика. 2022; 21 (11): 114–34.
10. Балановская Е. В., Петрушенко В. С., Кошель С. М., Почешхова Э. А., Черневский Д. К., Балановский О. П., и др. Картографический атлас распространения 45 фармакогенетических маркеров в народонаселении России и сопредельных стран. Вестник Российского государственного медицинского университета. 2020 (6): 39–52.
11. Балановская Е. В., Горин И. О., Пономарев Г. Ю., Пылёв В. Ю., Петрушенко В. С., Агджоян А. Т., и др. Следы взаимодействия финноязычного, славянского и тюркоязычного населения в современном генофонде и их отражение в фармакогенетике. Вестник Российского государственного медицинского университета. 2022; 2: 20–9.

12. Балановская Е. В., Напольских В. В., Чураков В. С., Пислегин Н. В., Записецкая Ю. С., Балановский О. П., и др. Генофонды удмуртов и бесермян в контексте финно-угорских и других окружающих народов: полногеномные и фармакогенетические данные. Ежегодник финно-угорских исследований. 2022; 2: 328–46.
13. Balanovsky O, Petrushenko V, Gorin I, Agdzhoyan A, Balanovska E, Sychev D, et al. Variation of genomic sites associated with severe Covid-19 across populations: global and national patterns. Pharmacogenomics and Personalized Medicine. 2021: 1391–402.
14. 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): s13742–015.
15. GG-base [cited 2022 Sep 10]. Available from: https://gg-base.org/.
16. Балановская Е. В., Жабагин М. К., Агджоян А. Т., Чухряева М. И., Маркина Н. В., Асылгужин Р. Р., и др. Популяционные биобанки: принципы организации и перспективы применения в геногеографии и персонализированной медицине. Генетика. 2016; 52 (12): 1371–87.
17. Кошель С. М. Геоинформационные технологии в геногеографии. Современная географическая картография. 2012; с. 158–8.
18. Petrović J, Pešić V, Lauschke VM. Frequencies of clinically important CYP2C19 and CYP2D6 alleles are graded across Europe. European journal of Human Genetics. 2020; 28 (1): 88–94.
19. Pedersen RS, Brasch-Andersen C, Sim SC, et al. Linkage disequilibrium between the CYP2C19*17 allele and wildtype CYP2C8 and CYP2C9 alleles: identification of CYP2C haplotypes in healthy Nordic populations. European journal of clinical pharmacology. 2010; 66: 1199–1205.20.