

Технология оценки частот ДНК-маркеров в многонациональных административных единицах по данным о коренном народонаселении в связи с заболеваемостью (на примере сердечно-сосудистых заболеваний)

Балановская Е. В.¹, Потанина А. Ю.¹, Кошель С. М.², Адамов Д. С.¹, Евсюков И. В.¹, Борисова А. Л.³, Покровская М. С.³, Мешков А. Н.³

¹ФГБНУ "Медико-генетический научный центр им. акад. Н.П. Бочкова". Москва; ²ФГБОУ ВО "Московский государственный университет им. М.В. Ломоносова". Москва; ³ФГБУ "Национальный медицинский исследовательский центр терапии и профилактической медицины" Минздрава России. Москва, Россия

Информация о заболеваемости представлена в статистических отчетах суммарно обо всем населении многонациональных субъектов России, но Биобанки народонаселения содержат информацию об отдельных народах.

Цель. Разработать технологию оценки частоты ДНК- (дезоксирибонуклеиновая кислота) маркеров для административных единиц по данным о коренном народонаселении, что позволит оценить связь частоты ДНК-маркеров с заболеваемостью.

Материал и методы. Биобанк Северной Евразии представил частоты 24 ДНК-маркеров, ассоциированных с сердечно-сосудистыми заболеваниями (ССЗ) для 86 коренных народов России (6227 ДНК-образцов). Из переписи 2010г извлечена информация об этническом составе 83 субъектов Российской Федерации (РФ), из данных Минздрава России — заболеваемость ССЗ в 85 субъектах РФ (2020-2021гг).

Результаты. Совокупность данных о генофондах коренного населения и этническом составе субъектов РФ позволила оценить частоты 24 ДНК-маркеров в 85 субъектах РФ. Полученные частоты позволили напрямую оценить их связь с ССЗ в России — хроническими ревматическими болезнями сердца, кардиомиопатией, ишемическими болезнями сердца, гипертензивной болезнью сердца, острым инфарктом. Результаты применения технологии представлены в виде таблиц и геногеографических карт: 1) распространения в субъектах РФ 24 ДНК-маркеров ССЗ; 2) географии заболеваемости в субъектах РФ для пяти ССЗ; 3) корреляции распространения частоты ДНК-маркеров ССЗ и заболеваемости ССЗ среди населения России.

Заключение. Предложенная технология оценки современного генофонда субъектов РФ по данным о генофонде коренных народов

России впервые позволила оценить связь клинически значимых ДНК-маркеров и заболеваемости в РФ.

Ключевые слова: генофонд, ДНК-маркеры, сердечно-сосудистая заболеваемость, геногеография, этносостав, субъекты Российской Федерации, коренное народонаселение.

Отношения и деятельность. Исследование выполнено в рамках Государственного задания Минобрнауки России для ФГБНУ "МГНЦ".

Благодарности. Авторы благодарят АНО "Биобанк Северной Евразии" за предоставление данных о частотах ДНК-маркеров ССЗ в коренном населении России.

Поступила 19/09-2024

Рецензия получена 27/09-2024

Принята к публикации 05/10-2024



Для цитирования: Балановская Е. В., Потанина А. Ю., Кошель С. М., Адамов Д. С., Евсюков И. В., Борисова А. Л., Покровская М. С., Мешков А. Н. Технология оценки частот ДНК-маркеров в многонациональных административных единицах по данным о коренном народонаселении в связи с заболеваемостью (на примере сердечно-сосудистых заболеваний). *Кардиоваскулярная терапия и профилактика*. 2024;23(11):4204. doi: 10.15829/1728-8800-2024-4204. EDN LXRFUJ

*Автор, ответственный за переписку (Corresponding author):

e-mail: potaninaaanna@gmail.com

[Балановская Е. В. — д.б.н., профессор, зав. лабораторией популяционной генетики человека, ORCID: 0000-0002-3882-8300, Потанина А. Ю.* — м.н.с. лаборатории популяционной генетики человека, ORCID: 0009-0000-7911-4673, Кошель С. М. — к.г.н., в.н.с., Географический факультет, ORCID: 0000-0002-4540-2922, Адамов Д. С. — к.ф.-м.н., с.н.с. лаборатории популяционной генетики человека, ORCID: 0009-0009-3527-4195, Евсюков И. В. — н.с. лаборатории популяционной генетики человека, ORCID: 0000-0003-1810-429X, Борисова А. Л. — н.с., руководитель лаборатории "Банк биологического материала" Института персонализированной терапии и профилактики, ORCID: 0000-0003-4020-6647, Покровская М. С. — к.б.н., в.н.с. лаборатории "Банк биологического материала" Института персонализированной терапии и профилактики, ORCID: 0000-0001-6985-7131, Мешков А. Н. — д.м.н., руководитель Института персонализированной терапии и профилактики, ORCID: 0000-0001-5989-6233].

Assessment of the frequency and association with morbidity of DNA markers in multinational administrative divisions based on indigenous population data (based on cardiovascular diseases)

Balanovskaya E. V.¹, Potanina A. Yu.¹, Koshel S. M.², Adamov D. S.¹, Evsyukov I. V.¹, Borisova A. L.³, Pokrovskaya M. S.³, Meshkov A. N.³

¹Bochkov Medical Genetic Research Center. Moscow; ²Lomonosov Moscow State University. Moscow; ³National Medical Research Center for Therapy and Preventive Medicine. Moscow, Russia

Information on morbidity is presented in statistical reports for the entire population of multinational subjects of Russia, but population Biobanks contain information on individual peoples.

Aim. To develop a technology for assessing the frequency of deoxyribonucleic acid (DNA) markers for administrative units based on indigenous population data, which will make it possible to assess the relationship between the frequency of DNA markers and morbidity.

Material and methods. The Biobank of Northern Eurasia presented the frequencies of 24 DNA markers associated with cardiovascular disease (CVD) for 86 indigenous peoples of Russia (6227 DNA samples). Information on the ethnic composition of 83 Russian regions was extracted from the 2010 census, and the CVD incidence in 85 Russian regions (2020–2021) was assessed by data from the Ministry of Health of Russia.

Results. The data on the gene pools of the indigenous population and the ethnic composition of Russian regions made it possible to assess the frequencies of 24 DNA markers in 85 Russian regions. The obtained frequencies allowed us to directly assess their relationship with CVD in Russia — chronic rheumatic heart diseases, cardiomyopathy, coronary artery disease, hypertension, acute myocarditis. Technology application data are presented in the form of tables and genogeographic maps of the distribution of 24 CVD DNA markers in Russian subjects, the geography of morbidity in Russian subjects for five CVDs, and the correlation of the distribution of the frequency of CVD DNA markers and the incidence of CVD among the population of Russia.

Conclusion. The proposed technology for assessing the modern gene pool in Russian subjects based on the data on the gene pool of the indigenous peoples of Russia for the first time made it possible to assess the relationship between clinically significant DNA markers and morbidity in Russia.

Keywords: gene pool, DNA markers, cardiovascular morbidity, genogeography, ethnic composition, subjects of the Russian Federation, indigenous population.

Relationships and Activities. The study was conducted within the State assignment of the Ministry of Education and Science of the Russian Federation for the Medical Genetic Research Center.

Acknowledgements. The authors thank the ANO Biobank of Northern Eurasia for providing data on the frequencies of DNA markers of CVD in the indigenous population of Russia.

Balanovskaya E. V. ORCID: 0000-0002-3882-8300, Potanina A. Yu.* ORCID: 0009-0000-7911-4673, Koshel S. M. ORCID: 0000-0002-4540-2922, Adamov D. S. ORCID: 0009-0009-3527-4195, Evsyukov I. V. ORCID: 0000-0003-1810-429X, Borisova A. L. ORCID: 0000-0003-4020-6647, Pokrovskaya M. S. ORCID: 0000-0001-6985-7131, Meshkov A. N. ORCID: 0000-0001-5989-6233.

*Corresponding author:
potaninaaanna@gmail.com

Received: 19/09-2024

Revision Received: 27/09-2024

Accepted: 05/10-2024

For citation: Balanovskaya E. V., Potanina A. Yu., Koshel S. M., Adamov D. S., Evsyukov I. V., Borisova A. L., Pokrovskaya M. S., Meshkov A. N. Assessment of the frequency and association with morbidity of DNA markers in multinational administrative divisions based on indigenous population data (based on cardiovascular diseases). *Cardiovascular Therapy and Prevention*. 2024;23(11):4204. doi: 10.15829/1728-8800-2024-4204. EDN LXRUFJ

АО — Автономный Округ, ДНК — дезоксирибонуклеиновая кислота, ИБС — ишемическая(ие) болезнь(и) сердца, РФ — Российская Федерация, ССЗ — сердечно-сосудистые заболевания, SNP — Single Nucleotide Polymorphism (однонуклеотидный полиморфизм).

Введение

В настоящее время развитие превентивной и персонализированной медицины невозможно без применения генетических данных. Однако фармакогенетическое тестирование в регионах с недостаточным финансированием здравоохранения недоступно. Решение этой проблемы дает популяционная геногеография, предоставляющая информацию о частоте важнейших ДНК (дезоксирибонуклеиновая кислота)-маркеров фармакогенетики¹ [1–5]. В такой многонациональной стране, как Россия, данные геногеографии могут дать важную информацию для принятия клинических решений. Изучение закономерностей в распределении фармакогенетических маркеров в народонаселении России

выявляет этносы и регионы, где фармакогенетическое тестирование будет клинико-экономически выгодным решением, и предоставляет информацию, на которую может ориентироваться персонализированная медицина.

Исходную информацию популяционная геногеография извлекает из двух типов популяционных биобанков человека. Первый тип — "целевые биобанки" — объединяет медицинские, коммерческие и другие биобанки, формируемые в соответствии с конкретной целью данной организации. Медицинские биобанки хранят образцы, связанные с определенным медицинским диагнозом. Второй тип популяционных биобанков — биобанки народонаселения — хранят биообразцы, не связанные с заболеваниями. Из биобанков народонаселения мировая наука отдает предпочтение биобанкам коренного народонаселения. Их особенностью является тща-

¹ Исследование основано на информации базы данных "Фармакогенетика популяций России и сопредельных стран" (<https://pharma-gene-geography.ru/>).

Ключевые моменты**Что известно о предмете исследования?**

- Биобанки народонаселения содержат уникальную информацию о клинически значимых ДНК- (дезоксирибонуклеиновая кислота) маркерах в коренном народонаселении. Однако показатели заболеваемости представлены не для коренного, а общего населения регионов.

Что добавляют результаты исследования?

- Разработана технология оценки частот ДНК-маркеров в субъектах Российской Федерации по данным о генофондах коренных народов России и этно-демографической информации о субъектах Российской Федерации.
- Разработанный подход позволяет количественно оценить корреляцию клинически значимых ДНК-маркеров с показателями заболеваемости в любых административных единицах, провести расчет для любых регионов и периодов с известным этническим составом и заболеваемостью.

Key messages**What is already known about the subject?**

- Population biobanks contain unique information about clinically significant deoxyribonucleic acid (DNA) markers in the indigenous population. However, the incidence rates are presented not for the indigenous population, but for the general population of the regions.

What might this study add?

- A technology has been developed for assessing the DNA marker frequencies in Russian regions based on data on the gene pools of the indigenous peoples of Russia and ethno-demographic information in regions.
- The developed approach allows for a quantitative assessment of the correlation of clinically significant DNA markers with incidence rates in any administrative units, and for calculations for any regions and periods with a known ethnic composition and incidence.

тельный подбор индивидов, репрезентативно представляющих популяцию определенной этнической или субэтнической группы. Такие биобанки необходимы не только для фундаментальных, но и для медицинских исследований: от формирования контрольной выборки до изучения распространенности клинически значимых ДНК-маркеров. Во всем мире биобанки народонаселения ориентированы на получение биообразцов от всех этнических групп региона, причем преимущественно коренного сельского населения, поскольку стремятся собрать "вневременную" информацию, обеспечивающую устойчивость частот ДНК-маркеров. Для этого отбираются доноры, все предки которых на протяжении минимум трех поколений родились в данной популяции и относились к данному этносу.

Однако для фармакогенетических ДНК-маркеров, полученных благодаря биобанкам коренного населения, невозможно оценить степень корреляции с заболеваемостью в многонациональном регионе: данные о ДНК-маркерах характеризуют отдельные этносы, а данные о заболеваемости даются суммарно по всему многонациональному населению административных субъектов. При этом необходимость в оценке связи ДНК-маркеров и заболеваемости велика, поскольку в генофондах разных по этническому и расовому составу регионах клинически значимые ДНК-маркеры проявляют себя по-разному.

Для решения этой проблемы предлагается технология расчета средней частоты (встречаемости) маркеров в любой многонациональной административной единице (например, субъекте Российской Федерации (РФ)) как взвешенной суммы

с использованием двух массивов данных: частоты ДНК-маркеров в коренном народонаселении и демографической информации об этническом составе этой административной единицы. Частоты ДНК-маркеров в коренных народах берутся каждая со своим весом, равным доле этого народа в этническом составе административной единицы и затем суммируются. Такой подход используется во многих практических областях и дает разумные (а зачастую и единственно правильные) оценки средних показателей. Кроме того, предлагаемая технология позволяет оперативно реагировать на изменения генофонда из-за интенсивных миграционных процессов (частоты можно пересчитать сразу после появления новых данных), намного экономичнее и, скорее всего, точнее, чем непосредственное обследование региона. Полученные средние частоты вместе с данными о заболеваемости в административных единицах уже позволяют оценивать связь между ними статистическими методами. Возможности предлагаемой технологии продемонстрированы на примере расчета частот 24 фармакогенетических маркеров сердечно-сосудистых заболеваний (ССЗ) и выявления их связи с географией заболеваемости населения 85 субъектов России.

Материал и методы

Исследование основано на 6227 ДНК-образцах популяционного Биобанка Северной Евразии (Москва), изученных по 24 биомаркерам ССЗ, данные о частотах которых представлены в базе данных "Фармакогенетика популяций России и сопредельных стран"¹. Коллекции ДНК сформированы по единой технологии в ходе

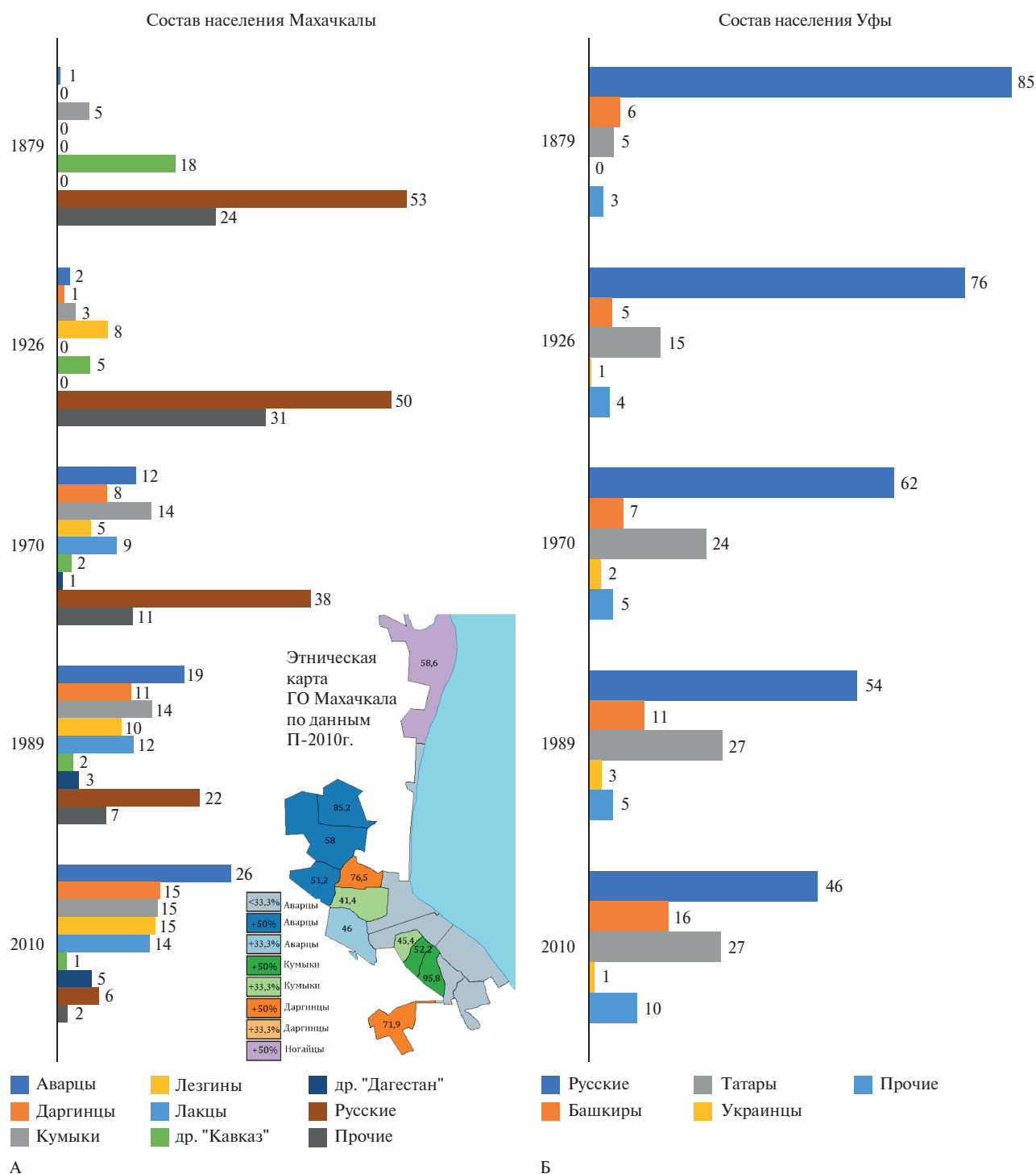


Рис. 1 Динамика изменения этнического состава (в %) населения столиц Дагестана и Башкортостана по данным переписи населения разных лет: А — динамика этнического состава населения Махачкалы; Б — динамика этнического состава населения Уфы. Примечание: цветное изображение доступно в электронной версии журнала.

многолетних экспедиционных обследований коренного населения [6]. В выборки включены образцы неродственных между собой индивидов, предки которых на глубину не менее трех поколений относили себя к данному этносу и данной популяции. Обследование проведено на добровольной основе с письменным информированным согласием, одобренным Этическим комитетом ФГБНУ "Медико-генетический научный центр им. академика Н. П. Бочкова", далее "МГНЦ" (Заключение

от 29.06.2020г). Данные об этническом составе населения субъектов РФ получены из переписей 2010г (для 83 субъектов) и 2020-2021гг (для Республики Крым и г. Севастополь). Данные о заболеваемости ССЗ (хронические ревматические болезни сердца, кардиомиопатия, ишемические болезни сердца (ИБС), гипертензивная болезнь сердца, острый миокардит) в 85 субъектах РФ получены как средние за 2020-2021гг [7]. Геногеографические карты построены с помощью программного пакета GeneGeo

Таблица 1

Этно-демографическая оценка частоты генетического маркера
в населении Башкортостана в 2010г по данным о коренном населении России

Этнос	Русские	Башкиры	Татары	Другие	Сумма
Доля этноса в Башкортостане	0,36	0,30	0,25	0,09	1,00
Этническая частота ДНК-маркера	0,65	0,12	0,21	0,24	
Расчет вклада этноса в частоту данного маркера в Башкортостане	$0,65 \times 0,36$	$0,12 \times 0,30$	$0,21 \times 0,25$	$0,24 \times 0,09$	
Вклад каждого этноса в итоговую частоту данного ДНК-маркера в Башкортостане	0,23	0,04	0,05	0,02	0,34
Искомая частота данного ДНК-маркера в Башкортостане					0,34

[8] методом средневзвешенной интерполяции со степенью весовой функции 3 и радиусом влияния 1800 км.

Результаты

1. Особенности формирования народонаселения субъектов РФ

Возникает закономерный вопрос: почему популяционные биобанки народонаселения собирают информацию о сельском населении, когда доля городского населения России составляет 75% и неуклонно растет (с 1950 до 2000гг доля городского населения выросла с 40 до 73%). Ответ связан с тремя следующими особенностями крупных городов.

Город живет за счет сельского населения. Показано, что крупные города ни сейчас, ни в далеком прошлом не воспроизводили свое население и потому похожи на "черные дыры": они сохраняют свою численность только благодаря постоянному притоку сельского населения [9-13]. Чем крупнее город — тем шире территория, откуда приезжают мигранты, а в мегаполисах ареалы источников миграции уже во многом перекрывают друг друга. В результате генофонды крупных городов становятся похожи: их изучение для целей криминалистики привело к выводу о неотличимости генофондов городских популяций, представляющих центр европейской части России (Москва), ее юг (Белгород, Орел), север (Ярославль), Урал (Оренбург) и Сибирь (Томск) [14]. Города образуют иерархическую систему, на нижнем уровне которой население малых городов интенсивно замещается сельским населением. Перемещение населения обычно идет двумя потоками: часть населения сразу мигрирует на вершину иерархии — в мегаполисы, а часть — в города на один уровень иерархии выше источника миграции. Поэтому чем крупнее город, тем больше он похож на другие города России; чем меньше город — тем больше он отражает особенности местного сельского населения.

Динамичность генофонда городского населения. Зависимость городского населения от притока мигрантов приводит к тому, что состав населения (а значит и его генофонд) может изменяться очень быстро [9-13]. Изменения в политике и экономике могут приводить к быстрому изменению структуры миграций и структуры генофонда: "Различия пара-

метров миграции в двух возрастных группах указывают на возможность динамики этнорегионального состава населения Москвы в последующих поколениях, что непременно вызовет изменение частот многих генетических маркеров" [15].

Насколько критичным для городского генофонда оказывается изменение структуры миграций демонстрирует рисунок 1. Например, данные о генофонде Махачкалы, полученные в 1970г, уже стали непригодны к 1990г. Если в 1970г русские и другие пришлые народы Дагестана составляли в Махачкале 51%, то в 2010г всего 11%. Значительно меняется и соотношение вклада разных народов Дагестана в генофонд Махачкалы (рисунок 1 А). Более того, в разных районах Махачкалы сконцентрированы представители разных народов Дагестана (врезка на рисунок 1 А). Изменение состава городского генофонда прослеживается для всех городов. Например (рисунок 1 Б), в 1970г численность русских в Уфе была больше в 9 раз, чем численность башкир, в 1989г это соотношение упало до 5, а еще через 20 лет — до 3.

Недолговечность биобанков городского населения. Поскольку изменения состава генофонда городского населения столь динамичны и значительны, исследователи приходят к выводу, что при формировании генетических баз данных для населения мегаполиса должны создаваться отдельные "референтные популяции" для каждой этнической группы города [9], причем сохраняется "необходимость постоянного обновления и актуализации таких баз данных и сбора генетико-демографических данных параллельно со сбором биологических образцов с целью изучения генетических маркеров" [15]. Иными словами, биобанки разнородного населения оказываются крайне недолговечными.

2. Технология оценки генофонда разнородного населения по данным о коренном населении

Приведенные данные о городском населении показывают необходимость постоянного обновления, причем для каждого поколения следует собирать данные в каждом городе о каждой этнической группе, ее составляющей. Иными словами, данные собираются с большим трудом, служат всего несколько лет, но все равно вызывают сомнения, поскольку учесть внутреннюю структуру города (разные

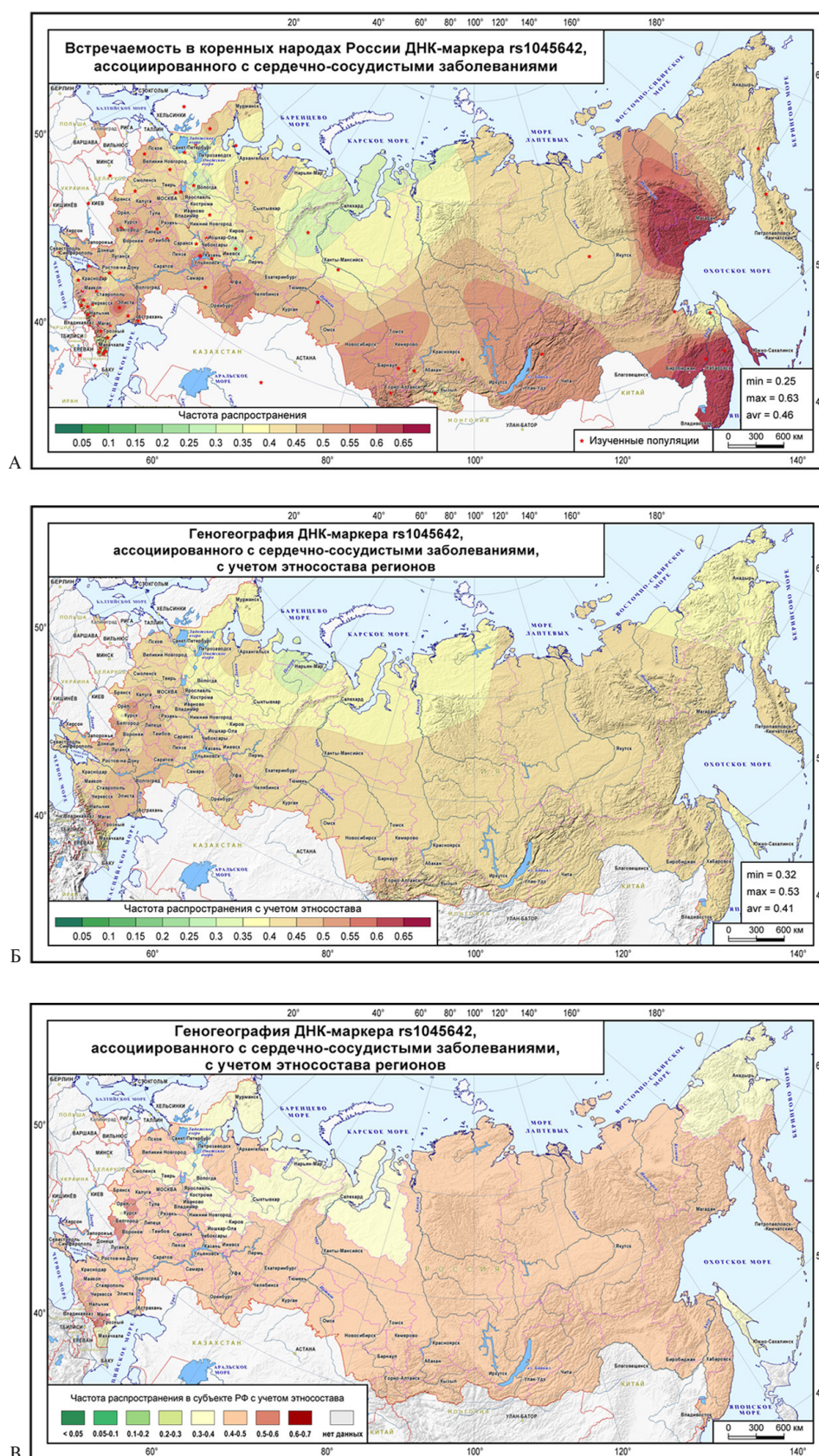


Рис. 2 Распространенность ДНК-маркеров rs1045642 (А-В) и rs12248560 (Г-Д), ассоциированных с сердечно-сосудистыми заболеваниями: А, Г — распространенность ДНК-маркеров в коренных народах России; Б, Д — интерполяционные карты расчетной частоты ДНК-маркеров в полиэтничном населении субъектов РФ; В, Е — картограммы расчетной частоты ДНК-маркеров субъектов РФ в их административных границах.

Примечание: ДНК — дезоксирибонуклеиновая кислота. Цветное изображение доступно в электронной версии журнала.

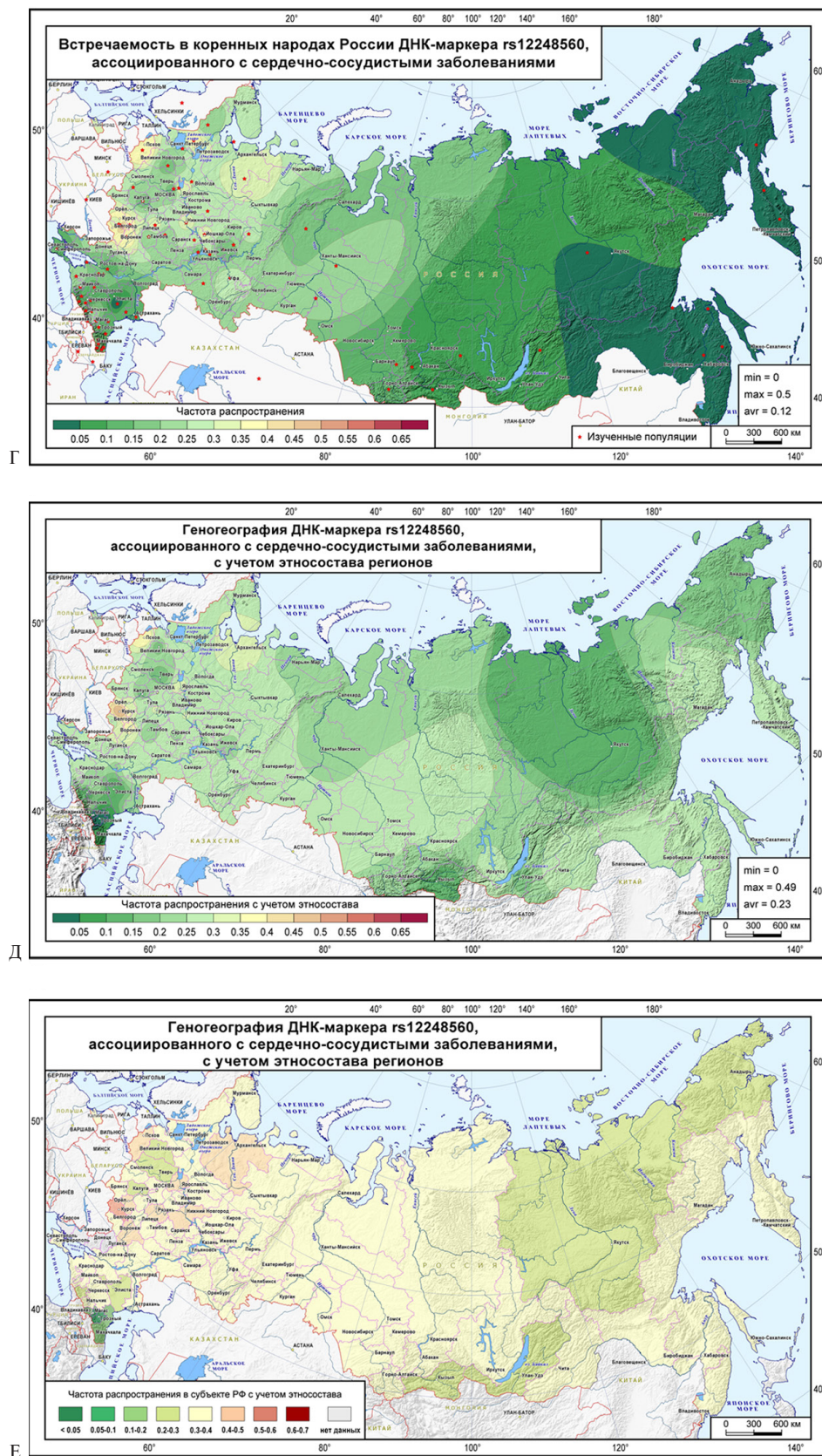


Рис. 2 Продолжение.

этно-демографические группы обычно расселены неравномерно) и собрать репрезентативные выборки для нее чрезвычайно сложно. Еще сложнее собрать

репрезентативную выборку для полноценной характеристики генофонда того или иного административного субъекта — района, области, республики,

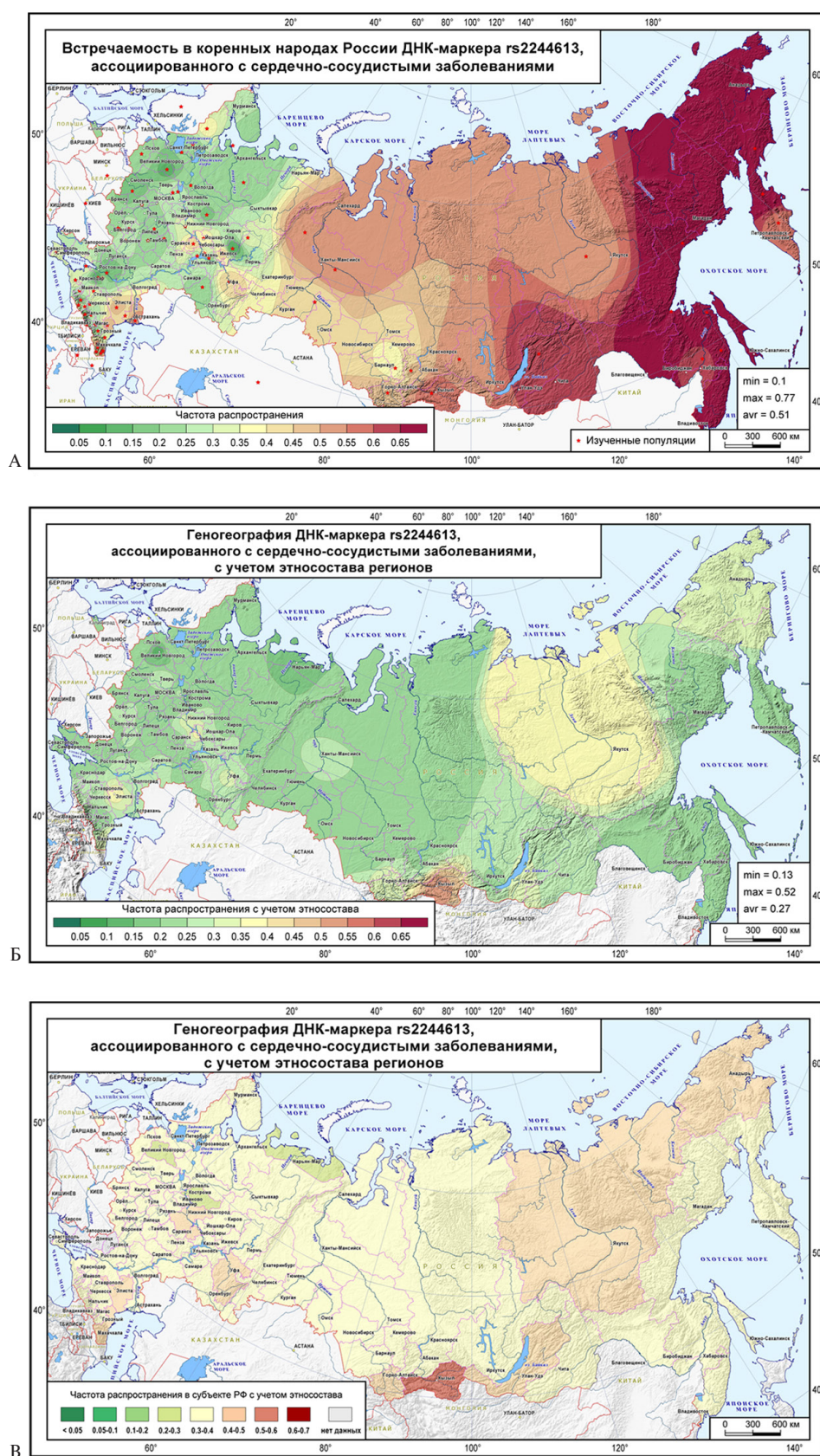


Рис. 3 Распространенность ДНК-маркеров rs2244613 (А-В) и rs3892097 (Г-Д), ассоциированных с сердечно-сосудистыми заболеваниями: А, Г — распространенность ДНК-маркеров в коренных народах России; Б, Д — интерполяционные карты расчетной частоты ДНК-маркеров в полиэтническом населении субъектов РФ; В, Е — картограммы расчетной частоты ДНК-маркеров субъектов РФ в их административных границах.

Примечание: ДНК — дезоксирибонуклеиновая кислота. Цветное изображение доступно в электронной версии журнала.

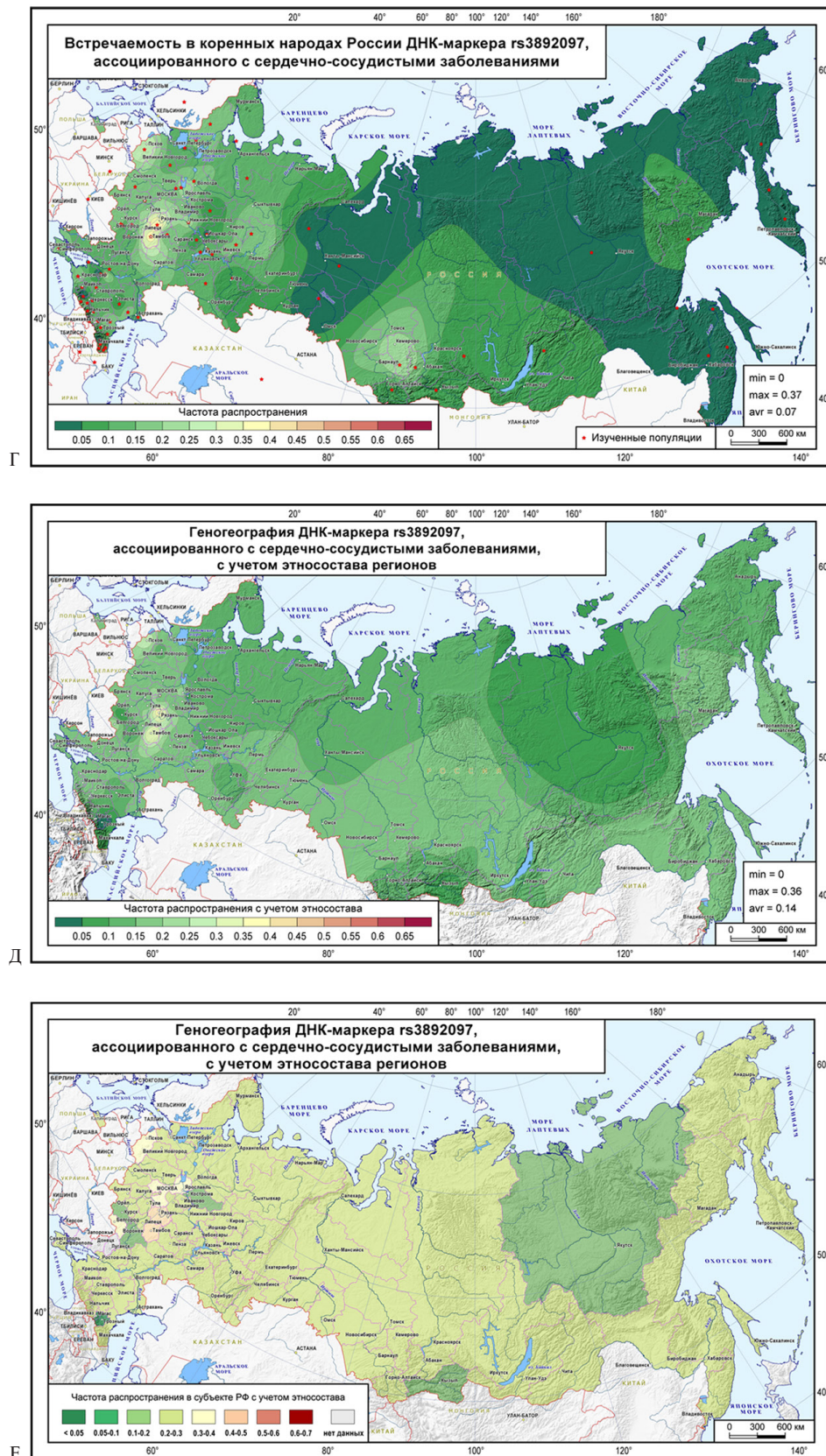


Рис. 3 Продолжение.

поскольку необходимо одновременно собрать представительные выборки биообразцов во всех населенных пунктах каждого субъекта. И все равно результа-

ты, добытые со столь большим трудом, будут крайне недолговечны (в Башкортостане за одно поколение соотношение башкир и татар изменилось с 2,5 до 1,7).

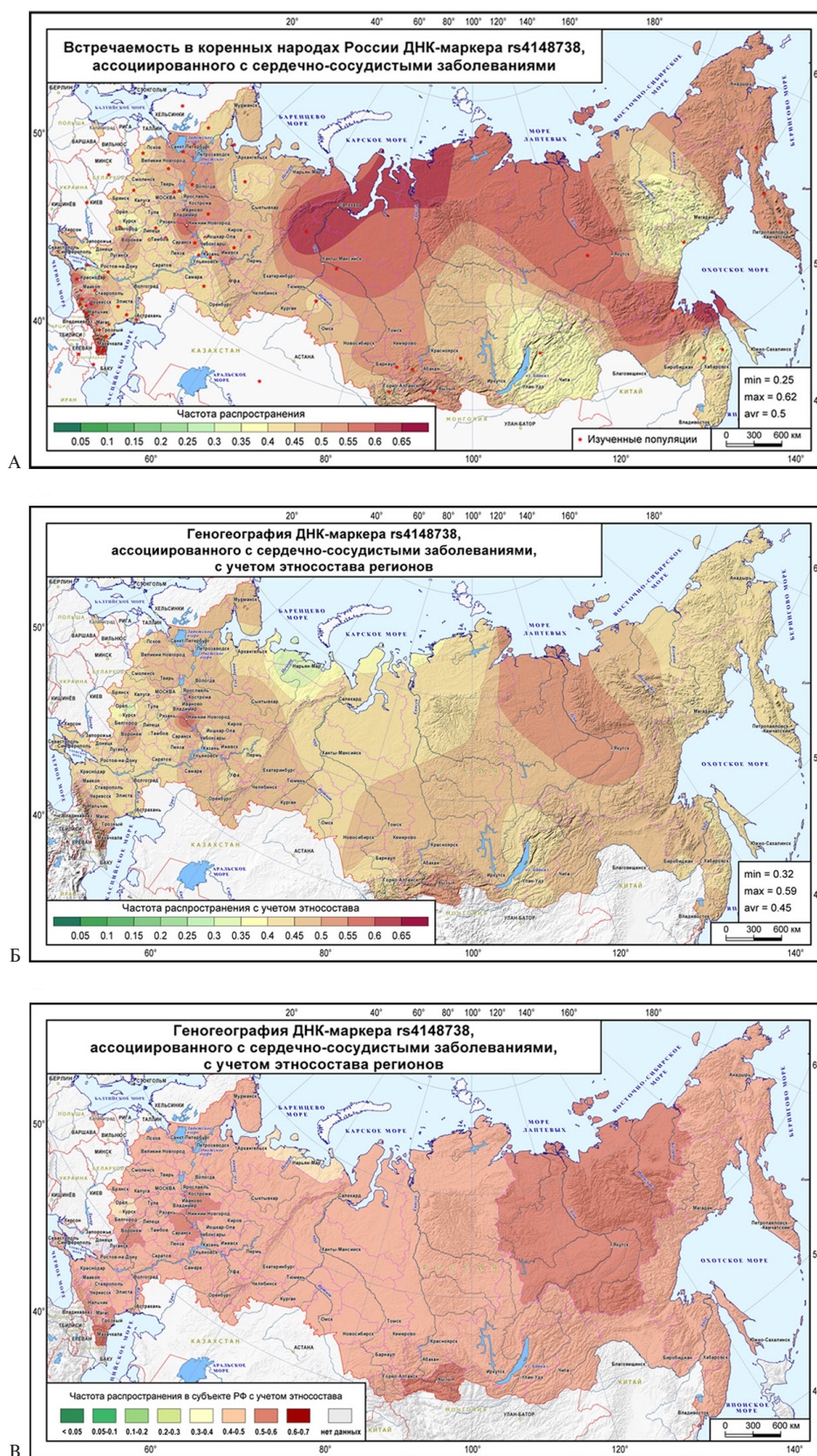


Рис. 4 Распространенность ДНК-маркеров rs4148738 (А-В) и rs9934438 (Г-Д), ассоциированных с сердечно-сосудистыми заболеваниями: А, Г — распространенность ДНК-маркеров в коренных народах России; Б, Д — интерполяционные карты расчетной частоты ДНК-маркеров в полиэтническом населении субъектов РФ; В, Е — картограммы расчетной частоты ДНК-маркеров субъектов РФ в их административных границах.

Примечание: ДНК — дезоксирибонуклеиновая кислота. Цветное изображение доступно в электронной версии журнала.

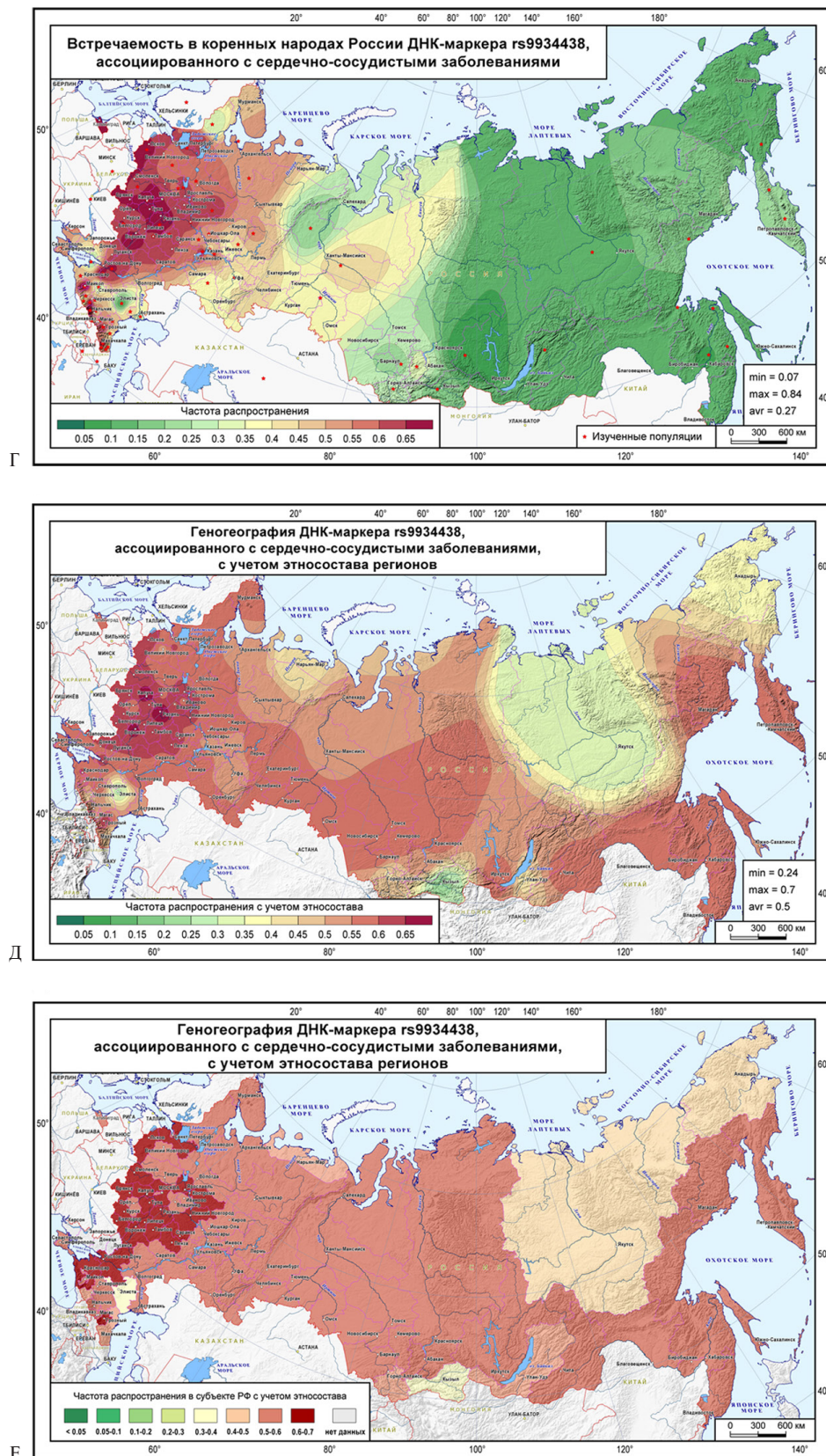


Рис. 4 Продолжение.

Поэтому наиболее достоверной оказывается оценка генофонда любого полиэтнического административного субъекта по совокупности данных

об его этно-демографической структуре (которые постоянно обновляются) и данных Биобанка коренного народонаселения. Зная частоту заданного

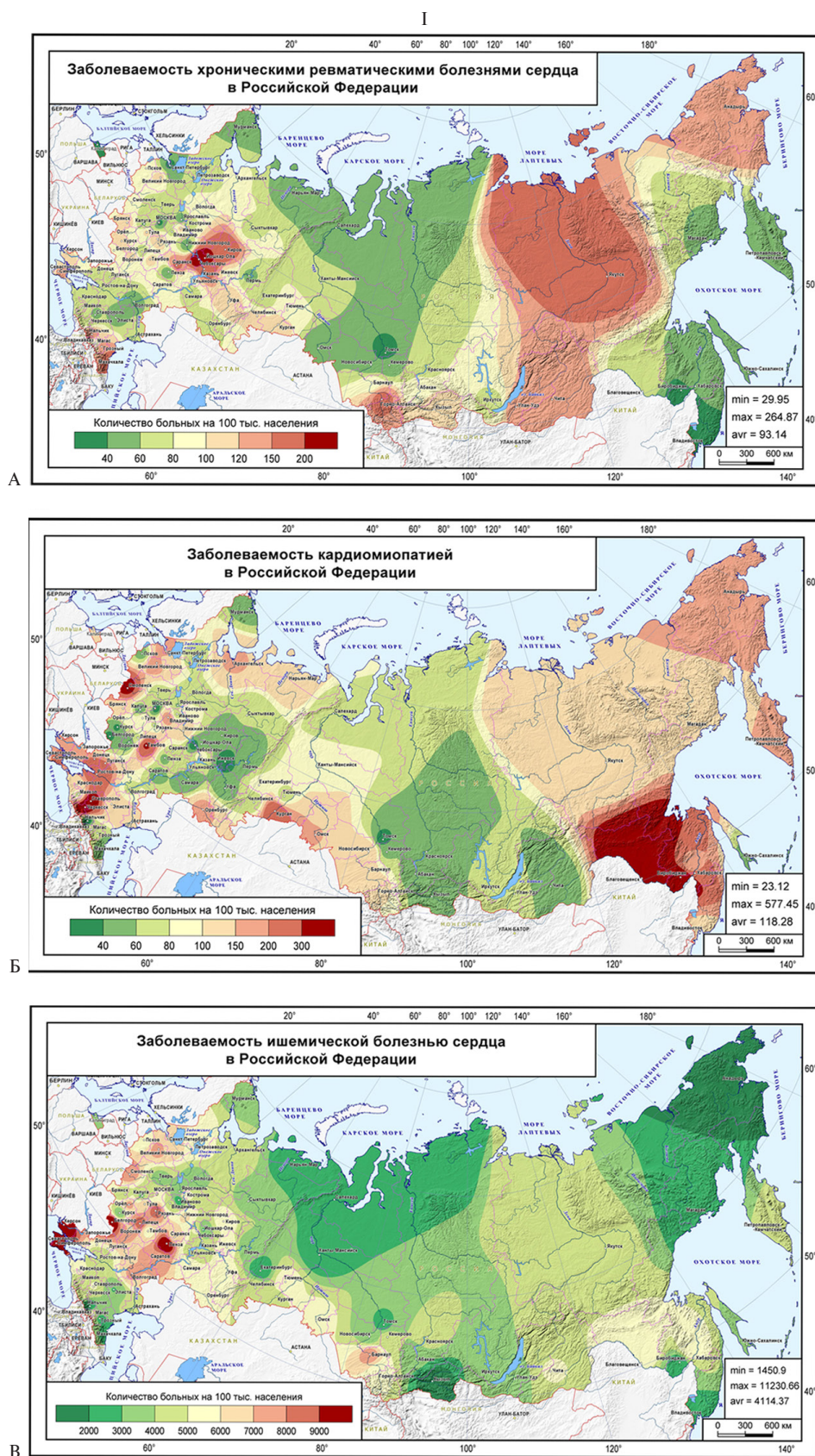
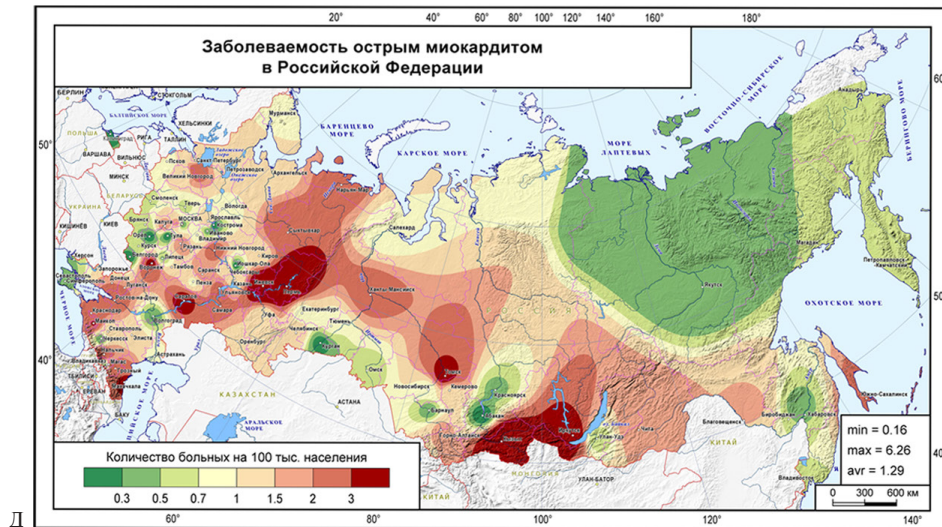


Рис. 5 Заболеваемость ССЗ на территории РФ, представленная в виде интерполяционных карт (I) и картограмм в пределах административных границ субъектов (II): А — хронические ревматические болезни сердца; Б — кардиомиопатия; В — ишемические болезни сердца; Г — гипертензивная болезнь сердца (гипертоническая болезнь с преимущественным поражением сердца); Д — острый миокардит.

I



Г



Д

II



А

Рис. 5 Продолжение.

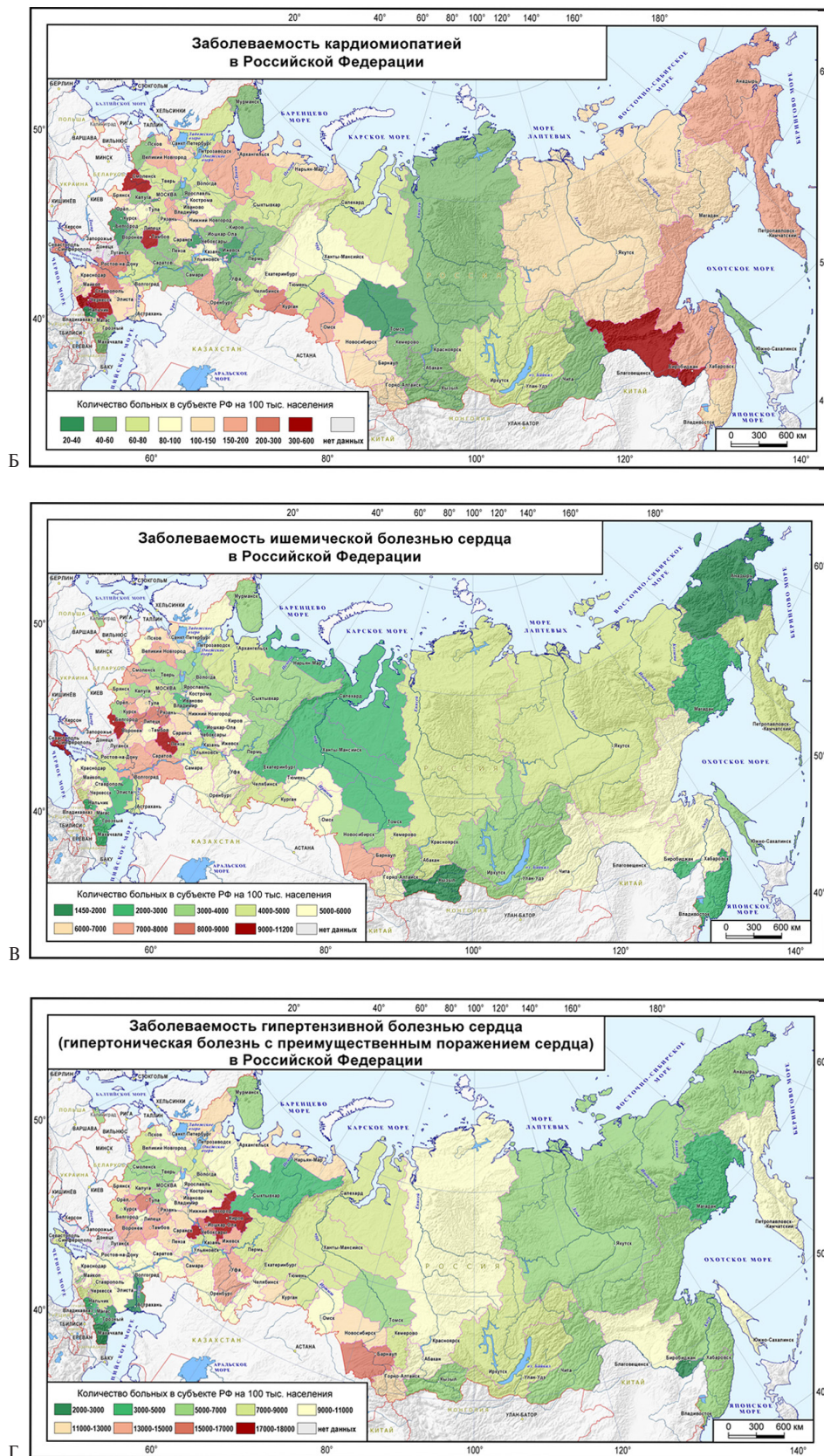


Рис. 5 Продолжение.



Рис. 5 Продолжение.

генетического маркера в каждом этносе коренного населения и долю этого этноса в составе данного субъекта в данный момент, можно рассчитать среднюю частоту маркера в данном субъекте в данное время. Простой пример такого расчета приведен для населения Башкортостана (таблица 1).

Получить информацию о частотах генетического маркера в населении каждого субъекта РФ в один и тот же момент времени невозможно. И в любом случае, такая информация устареет через 10 лет. Однако, используя данные Биобанка коренного народонаселения, можно получить одновременную характеристику встречаемости ДНК-маркера для всех субъектов РФ (и табличную, и в виде карт). Поскольку данные об этническом составе административных единиц имеются для различных периодов времени, можно реконструировать и динамику генофондов во времени.

Этно-демографическая оценка современного состояния генофонда города или иной административной единицы необходима как для криминалистики, так и для медицинских исследований, поскольку данные о заболеваемости обычно представляются для субъектов РФ. Мы получаем возможность проследить во времени динамику связи клинически значимых ДНК-маркеров и заболеваемости, поскольку данные об этническом составе и заболеваемости представляются регулярно.

3. Оценка частоты ДНК-маркеров, ассоциированных с ССЗ, для населения субъектов РФ по данным о коренном населении

С помощью описанной выше технологии проведен расчет частот 24 ДНК-маркеров, ассоциированных с заболеваниями ССЗ, для каждого из 85 субъектов (аналогично таблице 1). Для этого использованы два массива данных.

Первый массив — демографические данные о составе этносов в каждом из 85 субъектов РФ (для

Донецкой народной республики, Луганской народной республики, Запорожской и Херсонской областей сопоставимые данные отсутствуют) согласно наиболее корректной переписи 2010г. Второй массив — данные о частотах в коренном населении России 24 ДНК-маркеров, ассоциированных с ССЗ [4, 16]. За четверть века в интенсивных экспедиционных обследованиях была добыта информация практически обо всех коренных народах России. Для каждого народа известен его основной этнический ареал. Так, для русского этноса — это Русская равнина и Русский Север [17], а Кавказ, Приуралье, Урал и вся азиатская часть России принадлежат этническим ареалам других народов России. В пределах каждого этнического ареала собирались биообразцы от представителей только коренного этноса, причем так, чтобы максимально охватить все генетическое разнообразие этноса. Итогом стала коллекция, содержащая 30 тыс. биообразцов коренного народонаселения "Биобанка Северной Евразии", часть которой, изученная по ДНК-маркерам ССЗ, и легла в основу оценки их геногеографии в субъектах РФ.

Генофонды многих народов России гетерогенны. Например, генофонд центральных и южных русских популяций значительно отличается от Русского Севера, чем от популяций Белоруссии и Польши. Поэтому для расчета выбиралась не средняя этническая частота ДНК-маркера, а его частота в тех популяциях данного этноса, которые наиболее географически и исторически близки к изучаемому субъекту РФ и с наибольшей вероятностью участвовали в формировании его генофонда. Например, для учета в других субъектах РФ русские популяции объединены в метапопуляции южных, центральных и северных русских, а популяции татар — в метапопуляции поволжских, сибирских и астраханских татар. Для каждой метапопуляции были рассчитаны частоты ДНК-маркеров,

Таблица 2

Значимые для персонализации лечения CC3 SNP-маркеры
и степень корреляции их частот с заболеваемостью пяти CC3

Идентификатор	Название гена	Средняя частота по табличным данным	Средняя частота по интерполяционной карте	Хронические ревматические болезни сердца	Кардио-миопатия	Ишемическая болезнь сердца	Гипертензивная болезнь сердца (гипертоническая болезнь с преимущественным поражением сердца)	Острый миокардит	Препараты для лечения CC3
rs1045642	ABCB1	0,43	0,41	0,086	-0,001	0,315	0,150	-0,018	Антиагреганты, гипотензивные препараты
rs10509681	CYP2C8	0,07	0,06	-0,170	0,024	0,215	0,046	-0,101	Антиагреганты
rs1057910	CYP2C9	0,07	0,06	0,109	-0,056	-0,046	-0,267	0,000	Антикоагулянты, нестероидные противовоспалительные средства
rs1065852	CYP2D6	0,17	0,15	-0,211	0,180	0,365	0,236	-0,096	β-блокаторы
rs12248560	CYP2C19	0,25	0,23	-0,160	-0,055	0,223	0,397	-0,142	Антиагреганты, ингибиторы протонной помпы
rs1799853	CYP2C9	0,09	0,08	-0,227	0,099	0,067	0,075	-0,102	Антикоагулянты, нестероидные противовоспалительные средства
rs1801252	ADRB1	0,15	0,15	-0,152	0,049	0,072	0,203	-0,115	β-блокаторы
rs2046934	P2RY12	0,13	0,12	-0,101	0,030	0,164	0,179	0,156	Антиагреганты
rs2108622	CYP4F2	0,22	0,23	0,248	-0,005	-0,121	-0,281	0,164	Антикоагулянты непрямого действия
rs2231137	ABCG2	0,10	0,13	0,262	-0,142	-0,315	-0,055	0,118	Розувастатин
rs2231142	ABCG2	0,09	0,10	-0,010	-0,153	-0,060	0,095	-0,095	Статины
rs2242480	CYP3A4	0,09	0,09	0,050	0,393	0,112	-0,221	0,165	Статины
rs2244613	CES1	0,25	0,27	0,393	-0,141	-0,232	-0,059	0,235	Антикоагулянты прямые пероральные, антиагреганты
rs28371725	CYP2D6	0,06	0,06	0,112	-0,003	-0,080	-0,126	-0,083	β-блокаторы
rs3892097	CYP2D6	0,16	0,14	-0,221	0,165	-0,256	0,027	0,171	β-блокаторы
rs4148738	ABCB1	0,45	0,45	0,165	-0,092	-0,046	-0,098	0,218	Антикоагулянты прямые пероральные
rs4149056	SLCO1B1	0,23	0,22	-0,256	0,213	0,246	0,331	-0,168	Статины
rs4244285	CYP2C19	0,12	0,12	0,027	0,049	-0,115	-0,098	-0,044	Антиагреганты, ингибиторы протонной помпы
rs4986893	CYP2C19	0,01	0,01	0,171	-0,127	-0,141	-0,212	0,110	Антиагреганты, ингибиторы протонной помпы
rs662	PON1	0,27	0,28	0,303	-0,145	-0,168	-0,266	0,227	Антиагреганты
rs7294	VKORC1	0,34	0,30	-0,074	-0,018	0,336	0,113	-0,096	Антикоагулянты непрямого действия
rs776746	CYP3A5	0,09	0,09	-0,004	-0,120	0,033	0,090	0,052	Статины
rs9923231	VKORC1	0,56	0,50	-0,211	0,028	0,409	0,243	-0,180	Антикоагулянты непрямого действия
rs9934438	VKORC1	0,56	0,50	-0,216	0,022	0,401	0,230	-0,174	Антикоагулянты непрямого действия
Число достоверных корреляций			8		1	11	6	3	

Примечание: оранжевым фоном отмечены значимые корреляции Пирсона при $\alpha=0,01$ (порог значимости корреляции =0,278), желтым — значимые при $\alpha=0,05$ (порог значимости корреляции =0,213), серым — ДНК-маркеры, геногеография которых представлена на рисунках 2-4, CC3 — single nucleotide polymorphism (однонуклеотидный полиморфизм). Цветное изображение доступно в электронной версии журнала.

которые и использовались за пределами русского или татарского этнических ареалов. Для ногайцев использовались частоты популяций (астраханские, кубанские, ставропольские ногайцы, караногайцы), географически наиболее близких к конкретному субъекту РФ. Из расчета исключались редкие этносы ("другие") и те немногие, для которых не было информации в Биобанке, но в расчеты включены в среднем >95% этносов субъекта, что обеспечило приемлемую точность результатов.

В результате получен массив искомым частот в 85 субъектах РФ 24 ДНК-маркеров ССЗ в виде таблицы данных 24×85, на основе которой созданы геногеографические карты встречаемости ДНК-маркеров ССЗ в субъектах РФ. На рисунках 2-4 приведены по три карты для шести ДНК-маркеров. Сначала (на рисунках А, Г) даны исходные карты частоты ДНК-маркера — в коренном народонаселении России; далее приведены по две карты (на рисунках Б, В, Д, Е) расчетной частоты этого ДНК-маркера в полиэтничном населении всех субъектов РФ с указанием административных границ 89 субъектов РФ. На рисунках Б и Д эта информация представлена в виде интерполяционных карт (исходные данные помещены в административные центры субъектов). На рисунках В и Е эта же информация дана в виде картограмм, где частота ДНК-маркера в каждом субъекте занимает площадь точно в административных границах данного субъекта РФ (указаны на картах рисунков Б, В, Д, Е светлокрасными линиями). Поскольку эти границы не служат "заборами" для распространения генов, интерполяционные карты (рисунки Б и Д) оказываются более корректными и близкими к реальности: они позволяют проследивать реальные тренды в изменчивости ДНК-маркеров. Для корректности сравнения все интерполяционные карты (рисунки А, Б, Г, Д) и коренного, и полиэтничного населения приведены в единой шкале. Все картограммы (рисунки В, Е) также выполнены в собственной единой шкале, по цветовой гамме, максимально приближенной к шкале интерполяционных карт.

Задачей настоящей работы является изложение самой технологии оценки частоты ДНК-маркера в любой административной единице по совокупности данных о его этносоставе и о генофондах коренного населения. Поэтому ограничимся лишь описанием трех особенностей карт на рисунках 2-4.

Во-первых, выразительны отличия между картами для коренного и общего (разнородного) населения. Эти различия связаны с долей пришлого населения (таблица Приложения). В азиатской части России резко преобладает пришлое население (Алтайский край — 99,9%; Ханты-Мансийский Автономный Округ (АО) — 97,9%; Ямало-Ненецкий АО — 97,8%), и именно оно формирует полиэтничный генофонд субъектов РФ. Напротив, на Кавка-

зе доля пришлого населения минимальна (в Чечне — 4,7%, в Ингушетии — 5,9), и поэтому там нет расхождений по частоте ДНК-маркера между коренным населением и субъектом. Похожая ситуация наблюдается в областях русского ареала с преобладанием коренного русского населения (3-8% пришлого населения). Во-вторых, мы видим не менее яркие отличия по размеру территории, занимаемой субъектами. Например, площадь Якутии больше суммарной площади 50 субъектов РФ, входящих в состав Центрального, Северо-Западного и Приволжского федеральных округов. Совокупность этих двух параметров — вариации доли пришлого населения и различий по площади субъектов РФ — приводит к третьей особенности геногеографии ДНК-маркеров, а именно: к различной гетерогенности распределения (дисперсии) ДНК-маркера в европейской и в азиатской частях РФ (рисунки 2-4).

Обсуждение

Результаты расчетов средневзвешенной частоты ДНК-маркеров, ассоциированных с тем или иным заболеванием, полученные для субъектов РФ, позволяют оценить величину связи с заболеваемостью, поскольку она фиксируется для всех 85 субъектов РФ по единой схеме учета. Геногеография показателей заболеваемости для пяти ССЗ приведена на рисунке 5: в левом столбце (I) даны интерполяционные карты (более реалистичные); в правом столбце (II) — картограммы. Данные по заболеваемости получены как средние за 2020 и 2021гг [7], поскольку с 2022г могли возникнуть нестандартные отклонения от обычных показателей. Заболеваемость ССЗ чрезвычайно варьирует (рисунок 5). В статистической легенде интерполяционных карт указаны средние значения количества больных/100 тыс. населения: от 7884 (гипертоническая болезнь с преимущественным поражением сердца) до 1,3 для острого миокардита.

Поскольку и относительное количество больных, и частота ДНК-маркеров теперь, благодаря технологии, представлены для каждого из 85 субъектов РФ, появилась возможность впервые оценить для населения РФ связь между каждым из пяти заболеваний ССЗ и каждым из 24 ДНК-маркеров, ассоциированных с этими заболеваниями (таблица 2). Из полученных 120 коэффициентов корреляции достоверными оказались 32: наибольшее число ДНК-маркеров достоверно коррелируют с ИБС, наименьшее — с кардиомиопатией.

Впервые полученные данные о реальной связи между частотами ДНК-маркеров в населении РФ и заболеваемостью в субъектах РФ уникальны, поскольку имеющаяся в литературе информация об ассоциации ДНК-маркеров с ССЗ представляет преимущественно население Западной Европы, ге-

нофонд которой существенно отличается от генофонда населения РФ.

В то же время следует помнить о значительной гетерогенности всех показателей в пределах обширной территории РФ: и самих ДНК-маркеров (ри-

сунки 2-4, таблица 2), и показателей заболеваемости ССЗ (рисунок 5). Поэтому полученные коэффициенты связи между частотами ДНК-маркеров и заболеваемостью являются "средней температурой по больнице": многие субъекты РФ обладают

Таблица 3

Коэффициенты ранговой корреляции Спирмена
между тремя ДНК-маркерами и тремя ССЗ для 4-х федеральных округов РФ

Федеральный округ РФ	Число субъектов РФ	Ранговый коэффициент корреляции Спирмена		
		между rs12248560 и гипертензивной болезнью сердца ¹	между rs2244613 и хроническими ревматическими болезнями сердца ²	между rs9934438 и ишемической болезнью сердца ³
Северо-Западный*	10	-0,309	-0,018	0,515
Дальневосточный**	10	0,061	0,828	0,242
Сибирский	10	0,285	0,622	0,471
Северо-Кавказский***	10	0,638	0,313	0,227

Примечание: А) оранжевым фоном отмечены значимые корреляции при $\alpha=0,01$, желтым — при $\alpha=0,05$; Б) * — без Калининградской области; ** — без Бурятии; *** — добавлены Краснодарский край, Адыгея, Крым; В) ¹ — карта согласованности векторов изменчивости для rs12248560 и гипертензивной болезни сердца; ² — для rs2244613 и хронических ревматических болезней сердца; ³ — для rs9934438 и ишемической болезни сердца. ДНК — дезоксирибонуклеиновая кислота, ССЗ — сердечно-сосудистые заболевания. Цветное изображение доступно в электронной версии журнала.

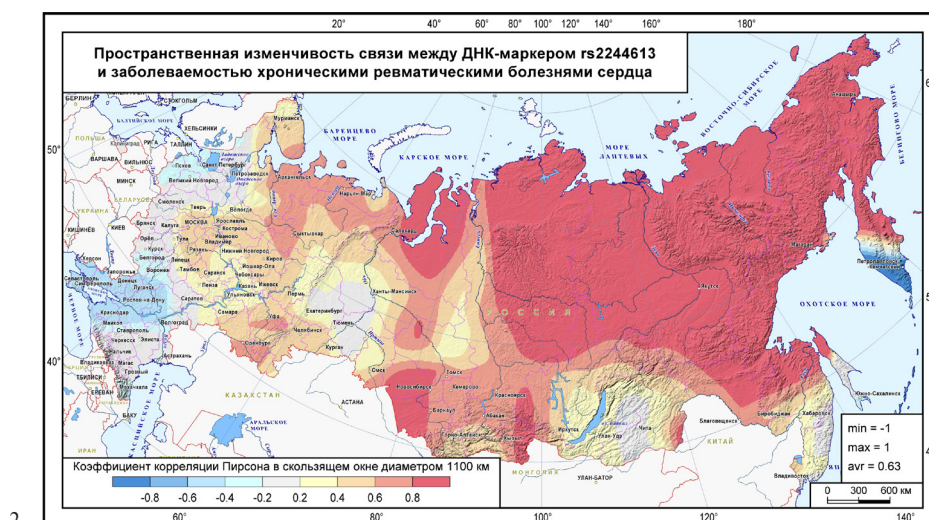
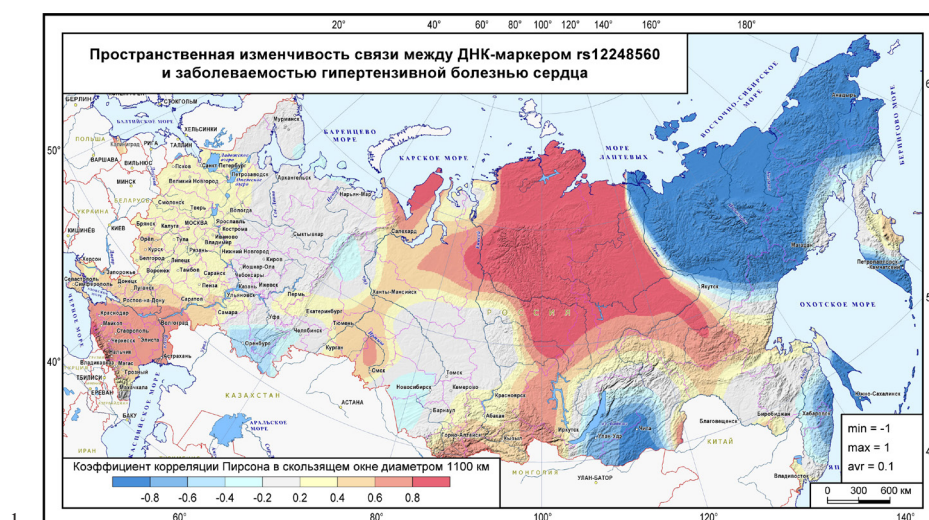
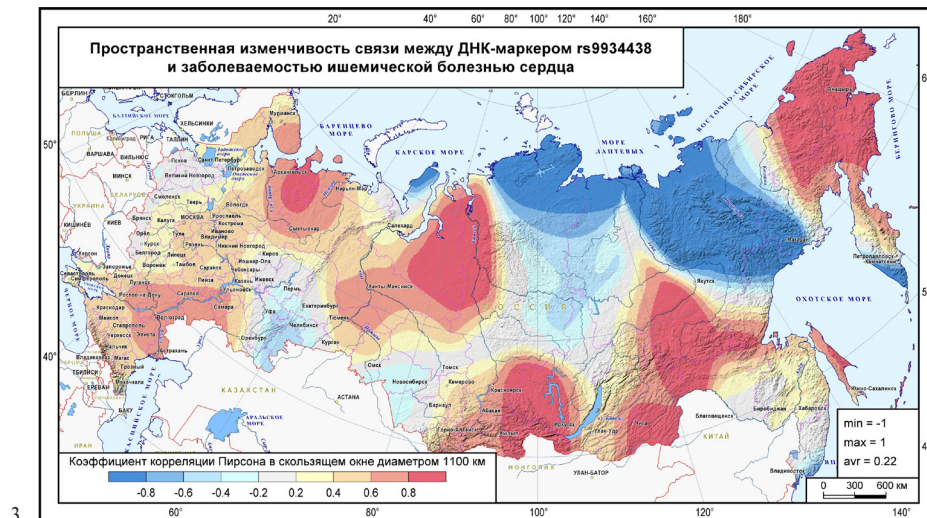


Таблица 3. Продолжение



различными генофондами, и для разных субъектов показатели связи между ДНК-маркерами и заболеваемостью могут существенно различаться. К сожалению, из-за отсутствия данных для более дробного административного деления рассчитать корреляции для отдельного субъекта нельзя, но их можно рассчитать для федеральных округов. В таблице 3 для 4-х федеральных округов приведены результаты расчета корреляции между тремя ДНК-маркерами и тремя заболеваниями: для корректности сравнения в каждом из федеральных округов учтены ровно по 10 субъектов РФ. При том, что для всех трех пар признаков "средняя температура по больнице" оценивалась как +0,4 (таблица 2) в федеральных округах показатель связи крайне изменчив. Размах изменения связи между rs12248560 и гипертензивной болезнью сердца варьирует максимально (от -0,3 до +0,6). Между rs2244613 и хроническими ревматическими болезнями сердца колебания показателя связи также очень велики (от 0 до 0,8). А вот между rs9934438 и ИБС связь наиболее устойчива — она варьирует от 0,2 до 0,5. Это означает, что реальная ассоциация ДНК-маркеров с заболеваниями в РФ отличается не только от населения Западной Европы, но различна в разных регионах РФ, что требует дифференцированного подхода для разных субъектов РФ.

Конечно, любая оценка требует проверки и уточнения. К сожалению, прямую проверку можно провести только при субтотальном обследовании самых малых административных единиц. Чем меньше административная единица — тем более надежные данные можно собрать во всех ее субпопуляциях: и о частотах ДНК-маркеров, и об этническом составе, и о заболеваемости; и тем точнее будет оценка связи ДНК-маркеров и заболеваемости. Для малых административных единиц также можно

получить надежную информацию о реальных источниках миграции (например, для русского этноса — из точно известных областей), что позволит точнее прогнозировать частоты ДНК-маркеров.

В крупных административных единицах такое субтотальное обследование, да еще и одновременно во всех субпопуляциях, провести нельзя. Однако и для них прогноз можно уточнять. Для этого следует, во-первых, провести расширенное генотипирование по исследуемым ДНК-маркерам, увеличив выборки для каждой популяции, биообразцы которой уже имеются в Биобанке коренного населения. Во-вторых, необходимо максимально пополнять коллекции Биобанка за счет субпопуляций этносов с широким этническим ареалом. Иными словами, именно пополнение информации, получаемой из Биобанка коренного населения, позволит сделать все более точным прогноз о частоте в населении ДНК-маркеров, ассоциированных с различными заболеваниями.

Заключение

Предложена технология оценки современного состояния генофонда населения для любых административных единиц России (с разнородным и быстро меняющимся этническим составом) с использованием данных о генофонде коренных народов России. Она позволяет оценить связь между частотой ДНК-маркера, ассоциированного с заболеванием, и заболеваемостью им в России. Для этого используются три источника данных: данные о клинически значимых ДНК-маркерах практически для всех народов России; данные об этническом составе для субъектов РФ; данные о заболеваемости в субъектах РФ. Совокупность первых двух источников данных — о генофондах коренного населения и этническом составе субъектов РФ — позволила про-

гнозировать частоту клинически значимых ДНК-маркеров в современном населении субъектов РФ. Полученные частоты клинически значимых ДНК-маркеров позволили напрямую анализировать их связь с заболеваемостью в субъектах РФ.

Использование метода проиллюстрировано на примере ДНК-маркеров сердечно-сосудистой патологии и географии заболеваемости населения во всех субъектах РФ пятью ССЗ: это хронические ревматические болезни сердца, кардиомиопатия, ИБС, гипертензивная болезнь сердца, острый миокардит.

Результаты применения предлагаемой технологии представлены в виде трех серий таблиц и геногеографических карт: 1) частоты в населении России 24 ДНК-маркеров сердечно-сосудистой патологии; 2) частоты сердечно-сосудистой заболеваемости во всех субъектах РФ; 3) корреляции между частотами ДНК-маркеров и сердечно-сосудистой заболеваемостью в России.

Предложенная технология оценки частоты ДНК-маркера в полиэтнических субъектах РФ по

данным о коренном народонаселении позволяет впервые оценить их связь с заболеваемостью в современном населении России. Получение таких данных иным способом не представляется возможным ввиду динамичного изменения этносовета населения под действием миграционных потоков, а также невозможности проведения синхронного субтотального обследования во всех субъектах РФ. В то же время, данная технология позволяет оценить связь ДНК-маркеров с заболеваемостью в любые периоды времени, для которых имеются данные об этническом составе любого региона и заболеваемости в нем.

Благодарности. Авторы благодарят АНО "Биобанк Северной Евразии" за предоставление данных о частотах ДНК-маркеров ССЗ в коренном населении России.

Отношения и деятельность. Исследование выполнено в рамках Государственного задания Минобрнауки России для ФГБНУ "МГНЦ".

Литература/References

1. Mirzaev KB, Fedorinov DS, Ivashchenko DV, et al. Multi-Ethnic Analysis of Cardiac Pharmacogenetic Markers of Cytochrome p450 and Membrane Transporters Genes in the Russian Population. *Rational Pharmacotherapy in Cardiology*. 2019;15(3):393-406. (In Russ.) Мирзаев К. Б., Федоринов Д. С., Иващенко Д. В. и др. Мультиэтнический анализ кардиологических фармакогенетических маркеров генов цитохрома Р450 и мембранных транспортеров в российской популяции. *Рациональная фармакотерапия в кардиологии*. 2019;15(3):393-406. doi:10.20996/1819-6446-2019-15-3-393-406.
2. Gorin IO, Petrushenko VS, Zapisetskaya YuS, et al. Population-based biobank for analyzing the frequencies of clinically relevant DNA markers in the Russian population: bioinformatic aspects. *Cardiovascular Therapy and Prevention*. 2020;19(6):2732. (In Russ.) Горин И. О., Петрушенко В. С., Записецкая Ю. С. и др. Применение популяционного биобанка для анализа частот клинически значимых ДНК-маркеров у населения России: биоинформатические аспекты. *Кардиоваскулярная терапия и профилактика*. 2020;19(6):2732. doi:10.15829/1728-8800-2020-2732.
3. Balanovska EV, Petrushenko VS, Koshel SM, et al. Cartographic atlas of frequency variation for 45 pharmacogenetic markers in populations of Russia and its neighbor states. *Bulletin of RSMU*. 2020;(6):39-52. (In Russ.) Балановская Е. В., Петрушенко В. С., Кошель С. М. и др. Картографический атлас распространения 45 фармакогенетических маркеров в народонаселении России и сопредельных стран. *Вестник Российского государственного медицинского университета*. 2020;(6):39-52. doi:10.24075/vrgmu.2020.080.
4. Pylev VYu, Agdzhoian AT, Gorin IO, et al. Population biobank as a basis for determining spatial variation of clinically relevant pharmacogenetic biomarkers of cardiovascular diseases. *Cardiovascular Therapy and Prevention*. 2022;21(11):3430. (In Russ.) Пылёв В. Ю., Агджоян А. Т., Горин И. О. и др. Популяционный биобанк как основа для выявления пространственной изменчивости клинически значимых фармакогенетических биомаркеров сердечно-сосудистых заболеваний. *Кардиоваскулярная терапия и профилактика*. 2022;21(11):3430. doi:10.15829/1728-8800-2022-3430.
5. Balanovsky O, Petrushenko V, Mirzaev K, et al. Variation of genomic sites associated with severe Covid-19 across populations: global and national patterns. *Pharmacogenomics Pers Med*. 2021;4:1391-402. doi:10.2147/PGPM.S320609.
6. Balanovska EV, Zhabagin MK, Agdzhoian AT, et al. Population biobanks: organizational models and prospects of application in gene geography and personalized medicine. *Russian Journal of Genetics*. 2016;52(12):1371-87. (In Russ.) Балановская Е. В., Жабагин М. К., Агджоян А. Т. и др. Популяционные биобанки: принципы организации и перспективы применения в геногеографии и персонализированной медицине. *Генетика*. 2016;52(12):1371-87. doi:10.7868/S001667581612002X.
7. Kotova EG, Kobayakova OS, Aleksandrova GA, et al. Morbidity of the entire population of Russia in 2021: statistical materials. Moscow: RIH of the Ministry of Health of Russia, 2022. p. 145. (In Russ.) Котова Е. Г., Кобякова О. С., Александрова Г. А. и др. Заболеваемость всего населения России в 2021 году: статистические материалы. М.: ЦНИИОИЗ Минздрава России, 2022. p.145. ISBN: 978-5-94116-072-3.
8. Koshel SM. Geoinformation technologies in genogeography. In: *Modern geographical cartography*. 2012;158-66. (In Russ.) Кошель С. М. Геоинформационные технологии в геногеографии. В кн.: *Современная географическая картография*. (под ред. Лурье И. К., Кравцова В. И.). М.: Дата+. 2012;158-66.
9. Kurbatova OL, Pobedonostseva EY, Veremeychik VM, et al. Genetic demography of populations of three megalopolises in relation to the problem of creating genetic databases. *Russian Journal of Genetics*. 2013;49(4):513-22. (In Russ.) Курбатова О. Л., Победоносцева Е. Ю., Веремейчик В. М. и др. Особенности генетико-демографических процессов в населении трех мегаполисов в связи с проблемой создания генетических баз данных. *Генетика*. 2013;49(4):513-22. doi:10.7868/S0016675813040085.
10. Kurbatova OL, Yankovsky NK. Migration as the main factor of population dynamics of urban population in Russia. *Russian*

- Journal of Genetics. 2016;52(7):831-51. (In Russ.) Курбатова О.Л., Янковский Н.К. Миграция — основной фактор популяционной динамики городского населения России. Генетика. 2016;52(7):831-51. doi:10.7868/S0016675816070067.
11. Kurbatova OL, Udina IG, Gracheva AS, et al. Genetic and demographic parameters of the population of Novosibirsk. Russian Journal of Genetics. 2018;54(13):74-84. (In Russ.) Курбатова О.Л., Удина И.Г., Грачева А.С. и др. Генетико-демографические параметры населения г. Новосибирска. Генетика. 2018;54(13):74-84. doi:10.1134/S0016675818130106.
 12. Kurbatova OL, Udina IG, Gracheva AS, et al. Genetic and demographic parameters of the population of St.Petersburg. Migration processes. Russian Journal of Genetics. 2019;55(9):1071-82. (In Russ.) Курбатова О.Л., Удина И.Г., Грачева А.С. и др. Генетико-демографические параметры населения г. Санкт-Петербурга. Миграционные процессы. Генетика. 2019;55(9):1071-82. doi:10.1134/S001667581909008X.
 13. Kurbatova OL, Gracheva AS, Pobedonostseva EY, et al. Genetic and demographic parameters of the population of Moscow. Migration processes. Russian Journal of Genetics. 2021;56(12):1438-49. (In Russ.) Курбатова О.Л., Грачева А.С., Победоносцева Е.Ю. и др. Генетико-демографические параметры населения г. Москвы. Миграционные процессы. Генетика. 2021;56(12):1438-49. doi:10.31857/S0016675821120080.
 14. Stepanov VA, Balanovsky OP, Melnikov AV, et al. Characteristics of Russian Federation populations by a panel of fifteen loci used for DNA identification and forensic examination. Acta Naturae. 2011;3(2):59-71. doi:10.32607/20758251-2011-3-2-56-67.
 15. Udina IG, Gracheva AS, Kurbatova OL. Frequencies of Y-chromosome haplogroups and migration processes in three generations of Moscow residents. Russian Journal of Genetics. 2022;58(11):1325-33. (In Russ.) Удина И.Г., Грачева А.С., Курбатова О.Л. Частоты гаплогрупп Y-хромосомы и процессы миграции в трех поколениях жителей Москвы. Генетика. 2022;58(11):1325-33. doi:10.31857/S001667582110121.
 16. Balanovska EV, Gorin IO, Ponomarev GYu, et al. Genogeographic technologies of a population biobank as a tool for assessing selection effects (using the example of pharmacogenetic biomarkers of cardiovascular diseases). Cardiovascular Therapy and Prevention. 2023;22(11):3773. (In Russ.) Балановская Е.В., Горин И.О., Пономарёв Г.Ю. и др. Геногеографические технологии популяционного биобанка как инструмент оценки эффектов отбора (на примере фармакогенетических биомаркеров сердечно-сосудистых заболеваний). Кардиоваскулярная терапия и профилактика. 2023;22(11):3773. doi:10.15829/1728-8800-2023-3773.
 17. Balanovska EV, Balanovsky OP. Russian gene pool in the Russian Plain. Moscow: Luch, 2007. p. 412. (In Russ.) Балановская Е.В., Балановский О.П. Русский генофонд на Русской равнине. М.: Луч, 2007. с. 412. ISBN: 978-5-87140-267-2.

Приложение

Таблица

Доля пришлого населения в субъектах РФ

Субъект РФ	Доля пришлого населения, %	Субъект РФ	Доля пришлого населения, %	Субъект РФ	Доля пришлого населения, %
Алтайский край	99,87	Ленинградская область	7,16	Республика Северная Осетия	34,93
Амурская область	99,82	Липецкая область	3,74	Республика Татарстан	46,76
Архангельская область без Ненецкого АО	99,83	Магаданская область	97,82	Республика Тыва	17,96
Астраханская область	99,17	Московская область	7,08	Республика Хакасия	87,85
Белгородская область	5,63	Мурманская область	99,80	Ростовская область	9,66
Брянская область	3,30	Ненецкий АО	81,38	Рязанская область	4,89
Владимирская область	4,42	Нижегородская область	4,85	Самарская область	14,45
Волгоградская область	9,99	Новгородская область	4,94	Саратовская область	12,45
Вологодская область	2,70	Новосибирская область	6,90	Сахалинская область	99,38
Воронежская область	4,50	Омская область	14,17	Свердловская область	9,35
Город Москва	8,35	Оренбургская область	24,12	Смоленская область	5,34
Город Санкт-Петербург	7,52	Орловская область	3,96	Ставропольский край	19,11
Город Севастополь	99,35	Пензенская область	13,20	Тамбовская область	3,01
Еврейская автономная область	7,26	Пермский край	12,89	Тверская область	6,55
Забайкальский край	93,07	Приморский край	99,92	Томская область	7,95
Ивановская область	4,43	Псковская область	4,99	Тульская область	4,69
Иркутская область	96,63	Республика Адыгея	74,84	Тюменская область без Ханты-Мансийского АО — Югра и Ненецкого АО	99,91
Кабардино-Балкарская Республика	30,16	Республика Алтай	62,21	Удмуртская Республика	72,00
Калининградская область	99,18	Республика Башкортостан	70,51	Ульяновская область	26,42
Калужская область	6,92	Республика Бурятия	69,96	Хабаровский край	98,25
Камчатский край	95,11	Республика Дагестан	13,42	Ханты-Мансийский АО	97,90
Карачаево-Черкесская Республика	47,13	Республика Ингушетия	5,88	Челябинская область	16,20
Кемеровская область	6,30	Республика Калмыкия	42,63	Чеченская Республика	4,73
Кировская область	8,12	Республика Карелия	92,62	Чувашская Республика	32,30
Костромская область	3,40	Республика Коми	76,31	Чукотский АО	64,85
Краснодарский край	99,93	Республика Крым	85,87	Ямало-Ненецкий АО	91,84
Красноярский край	99,41	Республика Марий Эл	56,08	Ярославская область	4,03
Курганская область	7,53	Республика Мордовия	59,94		
Курская область	3,52	Республика Саха (Якутия)	50,09		

Примечание: АО — автономный округ.