

## Популяционный биобанк как основа для выявления пространственной изменчивости клинически значимых фармакогенетических биомаркеров сердечно-сосудистых заболеваний

Пылёв В. Ю.<sup>1,2</sup>, Агджоян А. Т.<sup>1,2</sup>, Горин И. О.<sup>1</sup>, Петрушенко В. С.<sup>1</sup>, Почешхова Э. А.<sup>2,3</sup>, Мирзаев К. Б.<sup>1,4</sup>, Балановская Е. В.<sup>1,2</sup>

<sup>1</sup>ФГБНУ "Медико-генетический научный центр им. Н. П. Бочкова". Москва; <sup>2</sup>АНО "Биобанк Северной Евразии". Москва;

<sup>3</sup>ФГБОУ ВО "Кубанский государственный медицинский университет" Минздрава России. Краснодар; <sup>4</sup>ФГБОУ ДПО "Российская медицинская академия непрерывного профессионального образования" Минздрава России. Москва, Россия

Внедрение фармакогенетических тестов среди населения России сталкивается с фундаментальным ограничением — выраженными генетическими различиями между популяциями. Снять эти ограничения помогает геногеография фармакогенетических маркеров дезоксирибонуклеиновой кислоты (ДНК-маркеров).

**Цель.** Выявить пространственную изменчивость генофондов коренного населения европейской части России по ДНК-маркерам, значимым для фармакотерапии сердечно-сосудистых заболеваний (ССЗ), используя коллекции популяционного биобанка.

**Материал и методы.** 3170 образцов 61 популяции "Биобанка Северной Евразии", репрезентативно представляющих генофонды коренного населения Восточной Европы, изучены по двум массивам фармакогенетических ДНК-маркеров: 60 наиболее значимым маркерам и 24, связанным с сердечно-сосудистой патологией. Методами многомерной статистики и геногеографии проведено сравнение изменчивости генофонда по двум указанным массивам ДНК-маркеров.

**Результаты.** Создан картографический атлас, включающий карты распространения среди народонаселения Восточной Европы 24 фармакогенетических маркеров ССЗ. Картографические модели атласа, благодаря невербальному представлению большого объема информации, позволяют специалистам разного профиля анализировать закономерности в распространении фармакогенетических маркеров. Общие закономерности дополнены региональными исследованиями на Северном Кавказе, в Приуралье и Русской равнине, выявляющими группы населения со сходным фармакогенетическим статусом. Для каждого региона проведено сравнение изменчивости генофонда по двум массивам фармакогенетических ДНК-маркеров — общему и ССЗ.

**Заключение.** Созданный картографический атлас является основой для развития фармакогенетических исследований, проводимых методами геногеографии по единой панели фарма-маркеров и по репрезентативным выборкам, обеспеченным популяционными биобанками. Надежность результатов обеспечена подробной гене-

алогической и популяционной аннотацией каждого образца биобанка и репрезентативными выборками из популяций.

**Ключевые слова:** популяционный биобанк, ДНК-маркеры, фармакогенетика, сердечно-сосудистые заболевания, популяционная генетика, геногеография.

**Отношения и деятельность.** Исследование выполнено при поддержке гранта РНФ № 21-14-00363 (биоинформатический анализ, картографический анализ) и Государственного задания Министерства науки и высшего образования РФ для Медико-генетического научного центра им. акад. Н. П. Бочкова (статистический анализ, интерпретация результатов, написание статьи).

**Благодарности.** Авторы благодарят АНО "Биобанк Северной Евразии" за предоставление коллекций ДНК. Исследование выполнено при поддержке гранта РНФ № 21-14-00363 (биоинформатический и картографический анализ) и Государственного задания Минобрнауки России для ФГБНУ "МГНЦ" (статистический анализ, интерпретация результатов).

Поступила 22/09-2022

Рецензия получена 03/10-2022

Принята к публикации 23/10-2022



**Для цитирования:** Пылёв В. Ю., Агджоян А. Т., Горин И. О., Петрушенко В. С., Почешхова Э. А., Мирзаев К. Б., Балановская Е. В. 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. EDN ONEXFD

\*Автор, ответственный за переписку (Corresponding author):

e-mail: balanovska@mail.ru

[Пылёв В. Ю. — н.с. лаборатории популяционной генетики человека, генеральный директор, ORCID: 0000-0001-9541-8319, Агджоян А. Т. — к.б.н., в.н.с. лаборатории популяционной генетики человека, н.с., ORCID: 0000-0002-8776-2934, Горин И. О. — н.с. лаборатории популяционной генетики человека, ORCID: 0000-0001-9532-8954, Петрушенко В. С. — н.с. лаборатории популяционной генетики человека, ORCID: 0000-0002-5763-5280, Почешхова Э. А. — учредитель, д.м.н., доцент, профессор кафедры биологии с курсом медицинской генетики, ORCID: 0000-0002-8991-7194, Мирзаев К. Б. — с.н.с. лаборатории популяционной генетики человека, д.м.н., зав. лабораторией геномных предикторов нежелательных лекарственных реакций, проректор по научной работе и инновациям, ORCID: 0000-0002-9307-4994, Балановская Е. В. — д.б.н., профессор, г.н.с., зав. лабораторией популяционной генетики человека, учредитель, ORCID: 0000-0002-3882-8300].

## Population biobank as a basis for determining spatial variation of clinically relevant pharmacogenetic biomarkers of cardiovascular diseases

Pylev V. Yu.<sup>1,2</sup>, Agdzhoyan A. T.<sup>1,2</sup>, Gorin I. O.<sup>1</sup>, Petrushenko V. S.<sup>1</sup>, Pocheshkhova E. A.<sup>2,3</sup>, Mirzaev K. B.<sup>1,4</sup>, Balanovskaya E. V.<sup>1,2</sup>

<sup>1</sup>Bochkov Medical Genetics Research Center. Moscow; <sup>2</sup>Biobank of Northern Eurasia. Moscow; <sup>3</sup>Kuban State Medical University. Krasnodar;

<sup>4</sup>Russian Medical Academy of Continuous Professional Education. Moscow, Russia

The introduction of pharmacogenetic tests among the Russian population faces a fundamental limitation — pronounced genetic differences between populations. The genetic geography of pharmacogenetic markers of deoxyribonucleic acid (DNA) helps to remove these limitations.

**Aim.** To reveal the spatial variation of the gene pools of the indigenous European Russian population in terms of DNA markers that are significant for the pharmacotherapy of cardiovascular diseases (CVDs) using the population biobank collections.

**Material and methods.** A total of 3170 samples from 61 populations of the Biobank of Northern Eurasia, which represents the gene pools of the indigenous Eastern Europe population, were studied using two pharmacogenetic DNA marker arrays as follows: 60 most significant markers and 24 markers associated with CVDs. Using the multivariate statistics and genetic geography, a comparison of gene pool variation was made.

**Results.** A cartographic atlas has been created that includes maps of the distribution among the Eastern Europe population of 24 pharmacogenetic CVD markers. These cartographic models allow various specialists to analyze patterns in the distribution of pharmacogenetic markers. General patterns are supplemented by regional studies in the North Caucasus, the Cisurals and the Russian Plain, which identify population groups with similar pharmacogenetic status. For each region, a comparison of gene pool variation for two arrays of above-mentioned DNA markers was made.

**Conclusion.** The created atlas is the basis for the development of pharmacogenetic studies conducted by genetic geography methods using a single panel of markers and representative samples provided by population biobanks. The reliability of the results is ensured by a detailed genealogical and population annotation of each biobank sample and representative samples from the populations.

**Keywords:** population biobank, DNA markers, pharmacogenetics, cardiovascular diseases, population genetics, genetic geography.

**Relationships and Activities.** The study was supported by the Russian Science Foundation grant № 21-14-00363 (bioinformatic analysis, cartographic analysis) and the State Assignment of the Ministry of Science and Higher Education of the Russian Federation for the Bochkov Medical Genetics Research Center (statistical analysis, interpretation of the results, writing the article).

**Acknowledgments.** The authors are grateful to the Biobank of Northern Eurasia for providing the DNA collections. The study was supported by the Russian Science Foundation grant № 21-14-00363 (bioinformatics and cartographic analysis) and the State Assignment of the Russian Ministry of Education and Science for the Medical Genetics Research Center (statistical analysis, interpretation of the results).

Pylev V. Yu. ORCID: 0000-0001-9541-8319, Agdzhoyan A. T. ORCID: 0000-0002-8776-2934, Gorin I. O. ORCID: 0000-0001-9532-8954, Petrushenko V. S. ORCID: 0000-0002-5763-5280, Pocheshkhova E. A. ORCID: 0000-0002-8991-7194, Mirzaev K. B. ORCID: 0000-0002-9307-4994, Balanovskaya E. V.\* ORCID: 0000-0002-3882-8300.

\*Corresponding author:  
balanovska@mail.ru

**Received:** 22/09-2022

**Revision Received:** 03/10-2022

**Accepted:** 23/10-2022

**For citation:** Pylev V. Yu., Agdzhoyan A. T., Gorin I. O., Petrushenko V. S., Pocheshkhova E. A., Mirzaev K. B., Balanovskaya E. V. 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. doi:10.15829/1728-8800-2022-3430. EDN ONEXFD

ДНК — дезоксирибонуклеиновая кислота, ЦСЗ — сердечно-сосудистые заболевания, SNP — Single Nucleotide Polymorphism.

## Введение

Биобанки, осуществляя сбор и систематизацию биологических материалов, нацелены на анализ данных для решения научных и практических задач. Для такого популярного объекта, как человек, известны два основных типа биобанков.

Первый тип — биобанки образцов пациентов — хранят образцы людей с известным медицинским диагнозом и служат для разработки методов профилактики, диагностики, выявления предрасположенности к заболеваниям и лечения в русле персонализированной медицины.

Биобанки второго типа — популяционные — хранят образцы, представляющие группы населения, неассоциированные с заболеваниями. Обычно популяционные биобанки нацелены на проблемы

фундаментальной науки: междисциплинарного изучения путей формирования генофонда в широком контексте современного и древнего населения. Они имеют непреходящее значение, поскольку хранят информацию о генетическом разнообразии человека: в эпоху глобализации генофонды многих народов исчезают с лица земли или утрачивают генетическую память под давлением миграционных потоков.

К популяционному биобанку, кроме общих требований биобанкирования, предъявляются высокие требования к популяционной репрезентативности коллекций. Ключевое значение имеет первичный этап — корректное формирование выборки из популяций. По мере развития технологий генотипирования этот этап становится все более

### Ключевые моменты

#### Что известно о предмете исследования?

- ДНК-полиморфизмы, значимые при фармакотерапии сердечно-сосудистой патологии, встречаются с разными частотами в разных популяциях России, что затрудняет внедрение фармакогенетических тестов.

#### Что добавляют результаты исследования?

- Созданный геногеографический атлас выявляет закономерности в пространственной изменчивости 24 ДНК-маркеров, информативных при кардиологической патологии (3170 образцов, 61 популяция коренного населения Восточной Европы).
- Пространственная изменчивость популяций Восточной Европы по "кардио"-маркерам иная, чем по совокупности 60 разных фармакогенетических ДНК-маркеров.
- Сравнение карт коренного населения Северного Кавказа, Приуралья и Центральной России показывает отсутствие единого правила при переходе от общей совокупности фармакогенетических маркеров к ДНК-маркерам отдельной группы заболеваний.

### Key messages

#### What is already known about the subject?

- DNA polymorphisms that are significant in the CVD therapy have different prevalence in different populations of Russia, which makes it difficult to implement pharmacogenetic tests.

#### What might this study add?

- The created genetic geographic atlas reveals patterns in the spatial variation of 24 cardiac DNA markers (3170 samples, 61 indigenous populations of Eastern Europe).
- Spatial variation of Eastern Europe populations for cardiac markers is different than for the totality of 60 different pharmacogenetic DNA markers.
- Comparison of maps of the indigenous population of the North Caucasus, the Cisurals and Central Russia does not show a single rule in the transition from the total set of pharmacogenetic markers to DNA markers of a separate disease group.

критичным. Ошибка в формировании выборки может обернуться ложными результатами на всех последующих этапах: её не устранят самые эффективные методы биоинформатики. Поэтому кроме общих правил (письменное информированное согласие, конфиденциальность, этика), при создании популяционных биобанков руководствуются рядом особых правил [1]. Правило "коренное население" следует общепринятому в мировой популяционной генетике требованию изучать именно коренное население региона. Правило "максимум разнообразия" указывает на необходимость формирования выборки из разных подразделений популяции, позволяя охватить ее генетическое разнообразие. Правило "три поколения" означает включение в выборку только тех индивидов, все предки которых относили себя к данной популяции на протяжении, как минимум, 3-х поколений. Правило "не родственники" — исключение родства до 3-го поколения — обеспечивает нескоррелированность генотипов.

Использование популяционных биобанков не исчерпывается проблемами фундаментальной науки. Постоянно расширяется сфера их практического использования для судебно-медицинской экспертизы (формирование референсных баз данных для идентификации дезоксирибонуклеиновой кислоты (ДНК)-идентификации, определение вероятного этногеографического происхождения, поиск

генов, определяющих внешность человека) и здравоохранения.

По мере развития персонализированного подхода в медицине и перехода к геномной медицине, популяционно-генетические исследования играют все более важную роль в сфере здравоохранения: популяционная генетика представляет данные о генетических маркерах, ассоциированных с заболеваниями, но с разной частотой распространенных в популяциях, что важно учитывать при планировании и эффективной организации здравоохранения. Каталог ассоциаций с заболеваниями проекта "1000 геномов" показал, что 35% геномных вариаций не являются общими для населения даже в пределах одного континента. Поэтому необходимо создание региональных биобанков коренного населения, обеспечивающих случай-контроль, непараллельный контроль, когортные исследования, GWAS (genome-wide association studies), EWAS (epigenome-wide association studies) и важных для поиска ассоциаций с заболеваниями, прогноза заболеваемости и планирования их лечения. Кроме решения таких уже ставших традиционными задач, как формирование контрольных выборок, прогноз и мониторинг груза наследственной патологии, анализ географического распространения отдельных генов, ассоциированных с заболеваниями, новым перспективным направлением является фармакогеномика и районирующее генофонда.

Коллекции Биобанка Северной Евразии<sup>1</sup>, созданного под руководством О. П. Балановского и Е. В. Балановской, на протяжении четверти века используются в многочисленных работах о происхождении генофондов как коренного населения региона, так и всего мира (обобщены в монографиях [2-5]). Наряду с решением задач фундаментальной науки данные Биобанка используются в таких областях как криминалистика и медицина: генетический контроль пигментации [6-8], выявление маркеров этногеографического происхождения индивида [9, 10], геногеография фармакогенетических маркеров (фарма-маркеров) [11-14], геногеографии онкогенов [15, 16], трегалазы [17] и маркеров, связанных с тяжелым протеканием COVID-19 (COroNaVIrus Disease) [18, 19].

Фармакогенетика становится важным инструментом персонализированной медицины. Однако для населения России внедрение фармакогенетических тестов сталкивается с фундаментальным ограничением — выраженными генетическими различиями между популяциями. Для снятия этих ограничений необходимо детальное исследование геногеографической изменчивости фармакогенетически значимых ДНК-маркеров, обеспеченное созданием популяционных биобанков.

Важность изучения фарма-маркеров сердечно-сосудистых заболеваний (ССЗ) обусловлена их лидирующим местом среди причин смерти: по данным Всемирной организации здравоохранения<sup>2</sup> ежегодно в мире от ССЗ умирает 17,5 млн человек с наиболее высокими показателями смертности в Восточной Европе и Центральной Азии [20]. В Европе ССЗ составляют 45% смертей: в странах-членах Евросоюза на ССЗ приходится 3,1 млн летальных исходов в год [21, 22], причем ишемическая болезнь сердца и инсульт являются причиной смерти, соответственно, 44 и 21% мужчин, 38 и 26% женщин [22]. Согласно данным Росстата, суммарная смертность от ССЗ в России составляет 856,1 тыс. человек в год, достигая 46,8% от смерти от всех причин [23].

Показано, что геногеографический ландшафт фарма-маркеров отличается [11-14] от паттерна селективно-нейтральных ДНК-маркеров, однако анализ геногеографических закономерностей для конкретных групп заболеваний не проводился.

Цель настоящей работы — выявить пространственную изменчивость генофондов коренного населения европейской части России по панели полиморфизмов, значимых для фармакотерапии ССЗ.

Для этого необходимо создать на основе коллекций Биобанка Северной Евразии картографический атлас пространственной изменчивости 24 SNP

(Single Nucleotide Polymorphism)-маркеров ССЗ; оценить положение генофондов народонаселения европейской части России в многомерном генетическом пространстве как 24 SNP-маркеров ССЗ, так и по спектру 60 основных фарма-маркеров; сравнить паттерны генетического сходства популяций по кардиологическим маркерам (кардио-маркерам) с паттерном по всей панели фарма-маркеров для трех региональных популяций Восточной Европы.

## Материал и методы

Изучены 3170 образцов ДНК индивидов из 61 популяции коренных народов Восточной Европы (Приложение, таблица 1), сформированных в ходе экспедиционных обследований коренного населения по единой технологии [1] и предоставленных Биобанком Северной Евразии (Москва). В выборки включены образцы неродственных между собой индивидов, предки которых на глубину не <3-х поколений относили себя к данному этносу и популяции. Обследование проведено на добровольной основе с письменным информированным согласием, одобренным Этическим комитетом ФГБНУ "МГНЦ" (Заключение от 29.06.2020г).

Генотипирование проведено методом полимеразной цепной реакции в реальном времени с использованием TaqMan-зондов и технологии OpenAmp на амплификаторе QuantStudio 12 Flex по кастомной панели 60 SNP-маркеров, информативных для фармакогенетики (Приложение, таблица 2) и включающих 24 SNP-маркера, значимых при ССЗ (таблица 1). Наибольшее клиническое значение для персонализации терапии ССЗ имеют фармакокинетические и фармакодинамические генетические биомаркеры, кодирующие транспортеры лекарственных средств и ксенобиотиков (ABCB1, ABCG2, SLCO1B1), изоферменты семейства цитохрома P-450 и другие ферменты метаболизма (CYP2C19, CYP2C9, CYP2C8, CYP2D6, CYP3A4, CYP3A5, CYP4F2, CES1, PON1), фармакодинамические мишени и рецепторы (VKORC1, ADRB1, P2RY12), а также другие компоненты метаболических путей<sup>3</sup>.

Картографический анализ проведен с помощью оригинального картографического программного обеспечения GeneGeo [24] аналогично описанному в [11-14]. Карты построены методом средневзвешенной интерполяции со степенной функцией  $k=3$  по данным о популяциях (Приложение, таблица 1) в радиусе R (для Восточной Европы  $R=800$  км).

Карты генетических расстояний для трех регионов (Восточной Европы — от удмуртов и бесермян, Центральной России — от рязанских русских, Северного Кавказа — от осетин) созданы по единому алгоритму. Для популяций каждого региона получена матрица частот по каждому ССЗ-маркеру, по которой рассчитаны генетические расстояния Нея от реперной популяции (удмурты, осетины или рязанские русские) до остальных популяций региона и созданы региональные карты генетических расстояний. Их усреднением получена карта средних генетических расстояний от реперной популяции. Аналогично созданы региональные карты по частотам 60 фарма-маркеров.

<sup>1</sup> <https://биобанк.рф/>.

<sup>2</sup> [https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)/](https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)/).

<sup>3</sup> <https://www.pharmgkb.org/>.



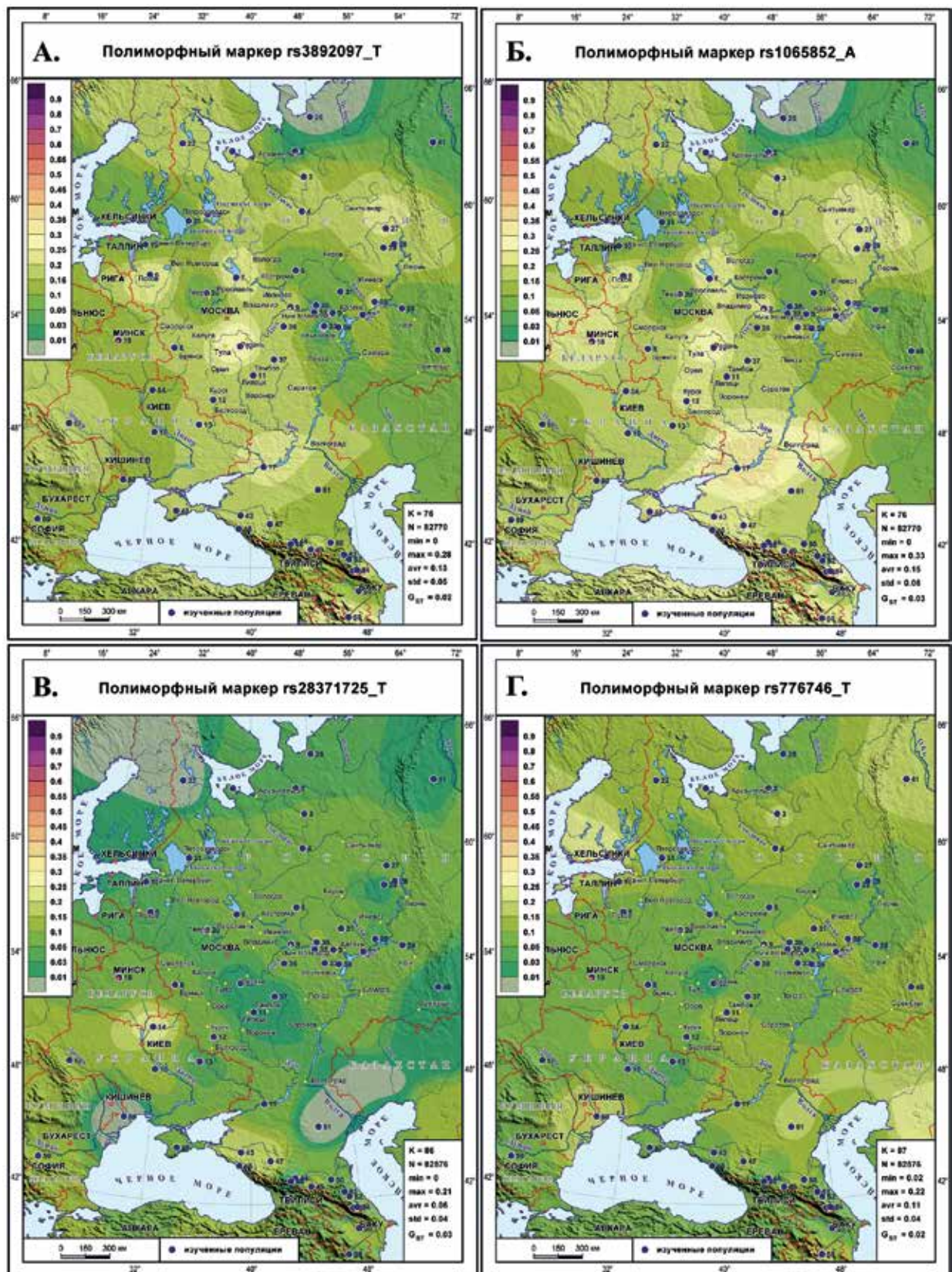


Рис. 1 Генетический ландшафт распространения маркеров CYP2D6 и CYP3A5: А) маркера rs3892097 (Т) гена CYP2D6; Б) маркера rs1065852 (А) гена CYP2D6; В) маркера rs28371725 (Т) гена CYP2D6; Г) маркера rs776746 (Т) гена CYP3A5.

Примечание: на каждой карте в легенде слева дана шкала цветовых интервалов: красно-фиолетовые тона — высокая частота ( $\geq 50\%$ ), желто-оранжевые — средний диапазон (20-50%), зеленые и салатовые — низкие значения (1-20%), серый цвет  $< 1\%$ . Изученные популяции на картах обозначены темно-синими кружками и сопровождаются номером согласно таблице 1 Приложения. Цветное изображение доступно в электронной версии журнала.



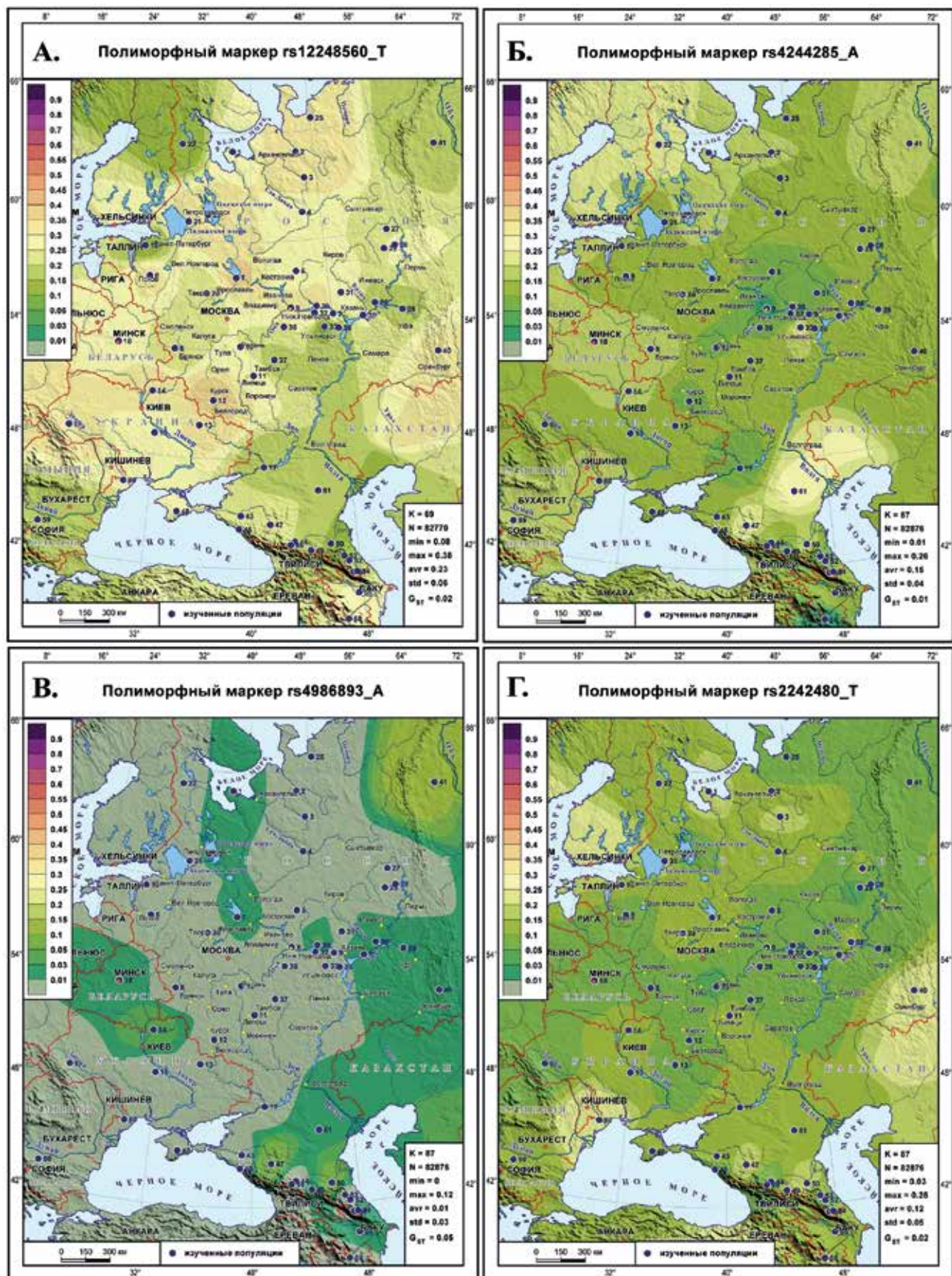


Рис. 2 Генетический ландшафт распространения маркеров CYP2C19 и CYP3A4: А) маркера rs 12248560 (Т) гена CYP2C19; Б) маркера rs4244285 (А) гена CYP2C19; В) маркера rs4986893 (А) гена CYP2C19; Г) маркера rs2242480 (Т) гена CYP3A4.

Примечание: см. рисунок 1. Цветное изображение доступно в электронной версии журнала.

Таблица 1

Изученные 24 SNP-маркера, значимые при персонализации лечения сердечно-сосудистых заболеваний (ССЗ)

Идентификатор	Хромосома	Позиция в сборке GRCh37	Название гена	Препараты для лечения ССЗ
rs3892097	22	42524947	<i>CYP2D6</i>	1.* β-блокаторы
rs1065852	22	42526694	<i>CYP2D6</i>	1. β-блокаторы
rs28371725	22	42523805	<i>CYP2D6</i>	1. β-блокаторы
rs776746	7	99270539	<i>CYP3A5</i>	1. Статины
rs12248560	10	96521657	<i>CYP2C19</i>	1. Антиагреганты, ингибиторы протонной помпы
rs4244285	10	96541616	<i>CYP2C19</i>	1. Антиагреганты, ингибиторы протонной помпы
rs4986893	10	96540410	<i>CYP2C19</i>	1. Антиагреганты, ингибиторы протонной помпы
rs2242480	7	99361466	<i>CYP3A4</i>	1. Статины
rs1057910	10	96741053	<i>CYP2C9</i>	1. Антикоагулянты, НПВС
rs1799853	10	96702047	<i>CYP2C9</i>	1. Антикоагулянты, НПВС
rs2108622	19	15990431	<i>CYP4F2</i>	1. Антикоагулянты непрямые, антиагреганты
rs2244613	16	55844609	<i>CES1</i>	1. Антикоагулянты прямые оральные, антиагреганты
rs7294	16	31102321	<i>VKORC1</i>	1. Антикоагулянты непрямые
rs9923231	16	31107689	<i>VKORC1</i>	1. Антикоагулянты непрямые
rs9934438	16	31104878	<i>VKORC1</i>	1. Антикоагулянты непрямые
rs4149056	12	21331549	<i>SLCO1B1</i>	1. Статины
rs1045642	7	87138645	<i>ABCB1</i>	2. Антиагреганты, гипотензивные препараты
rs4148738	7	87163049	<i>ABCB1</i>	2. Антикоагулянты прямые оральные
rs2046934	3	151057642	<i>P2RY12</i>	2. Антиагреганты
rs662	7	94937446	<i>PON1</i>	2. Антиагреганты
rs2231142	4	89052323	<i>ABCG2</i>	3. Статины
rs2231137	4	89061114	<i>ABCG2</i>	3. Розувастатин
rs1801252	10	115804036	<i>ADRB1</i>	3. β-блокаторы
rs10509681	10	96798749	<i>CYP2C8</i>	3. Антиагреганты

Примечания: \* — цифры 1-3 указывают на степень значимости в ССЗ по 3-балльной шкале; НПВС — нестероидные противовоспалительные средства.

Статистический анализ проведен методом главных компонент по частотам SNP-маркеров (Приложение, таблица 2) в 13 метапопуляциях Восточной Европы (Приложение, таблица 1) с помощью программного обеспечения PLINK 1.9<sup>4</sup> и библиотеки scikit-learn<sup>5</sup> для языка программирования Python 3.

## Результаты

### 1. Картографический атлас изменчивости в народонаселении Восточной Европы частот ДНК-маркеров, связанных с сердечно-сосудистой патологией

Для каждого из 24 ССЗ-маркеров создана карта его распространения в коренном населении Восточной Европы. На рисунках 1-4 приведены карты 16 наиболее значимых кардио-маркеров (таблица 1), остальные 8 приведены в (Приложение, рисунки 1, 2).

Основной особенностью в изменчивости кардио-маркеров оказалось сходство генетических ланд-

шафтов Кавказа и Восточно-Европейской равнины (за исключением паттернов *CYP4F2*, *CES1*, *ADRB1*). Поскольку и селективно-нейтральный генетический ландшафт [3], и ДНК-маркеры, ассоциированные с тяжелым течением COVID-19 [19], фиксируют отличие генофонда Кавказа от остального населения Восточной Европы, важно подчеркнуть выявленное своеобразие в распределении кардио-маркеров.

### Генетический ландшафт распространения маркеров *CYP2D6* и *CYP3A5*

Маркеры гена *CYP2D6* (рисунки 1 А-В) относятся к β-блокаторам и характеризуются сходным паттерном распространения в населении Восточной Европы. Пространственное распределение rs776746\_T гена *CYP3A5* (рисунок 1 Г), функционально связанного с остановкой кровотечений, имеет похожий паттерн. В целом их генетический ландшафт сглажен. При небольшом диапазоне частот географически близкие популяции отличаются друг от друга, маскируя слабо выраженную клинальную изменчивость с вектором "запад-восток".

<sup>4</sup> <https://www.cog-genomics.org/plink/>.

<sup>5</sup> <http://jmlr.csail.mit.edu/papers/v12/pedregosa11a.html/>.



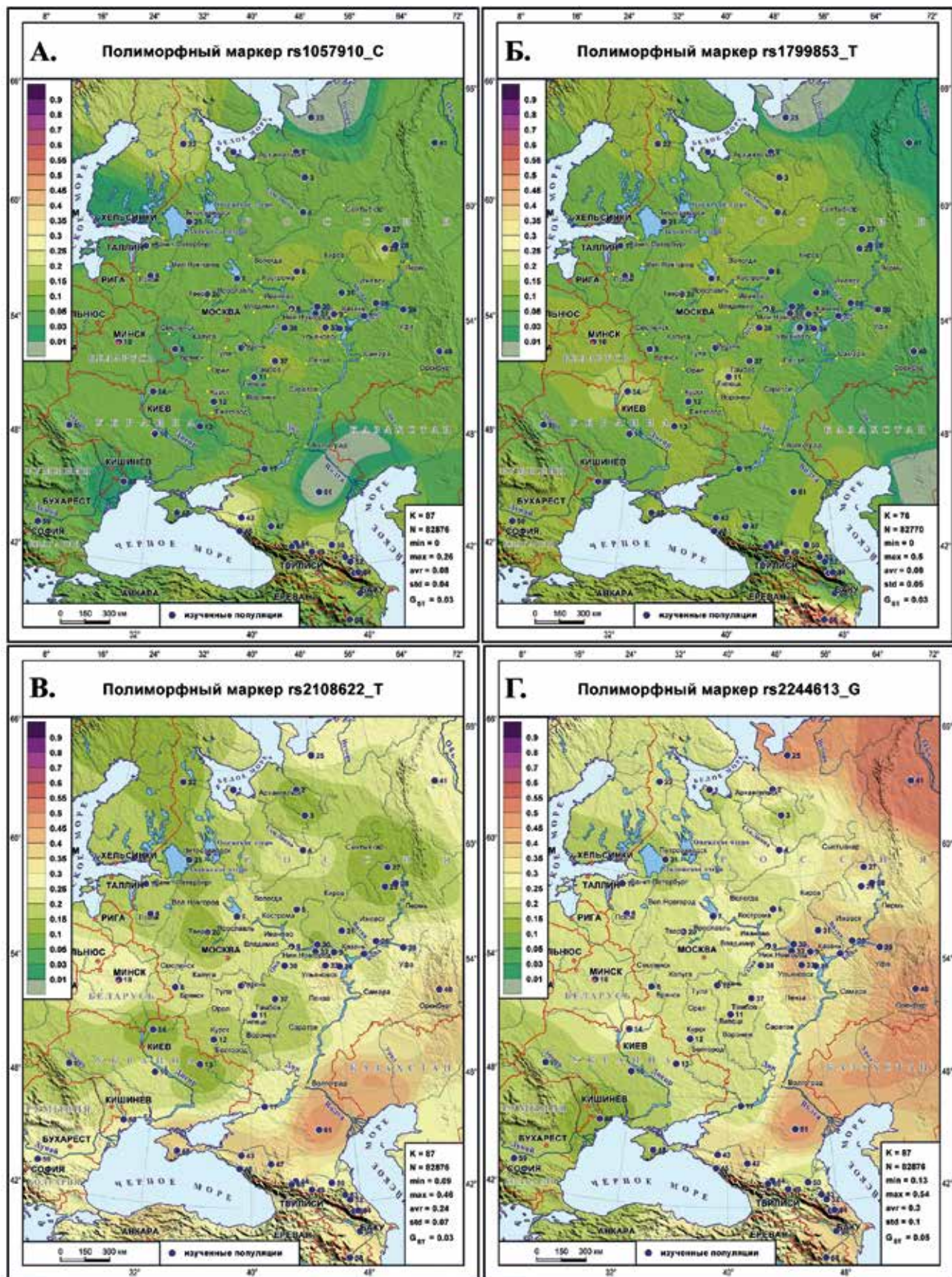


Рис. 3 Генетический ландшафт распространения маркеров CYP2C9, CYP4F2, CES1: А) маркера rs1057910 (C) гена CYP2C9; Б) маркера rs1799853 (T) гена CYP2C9; В) маркера rs2108622 (T) гена CYP4F2; Г) маркера rs2244613 (G) гена CES1.

Примечание: см. рисунок 1. Цветное изображение доступно в электронной версии журнала.



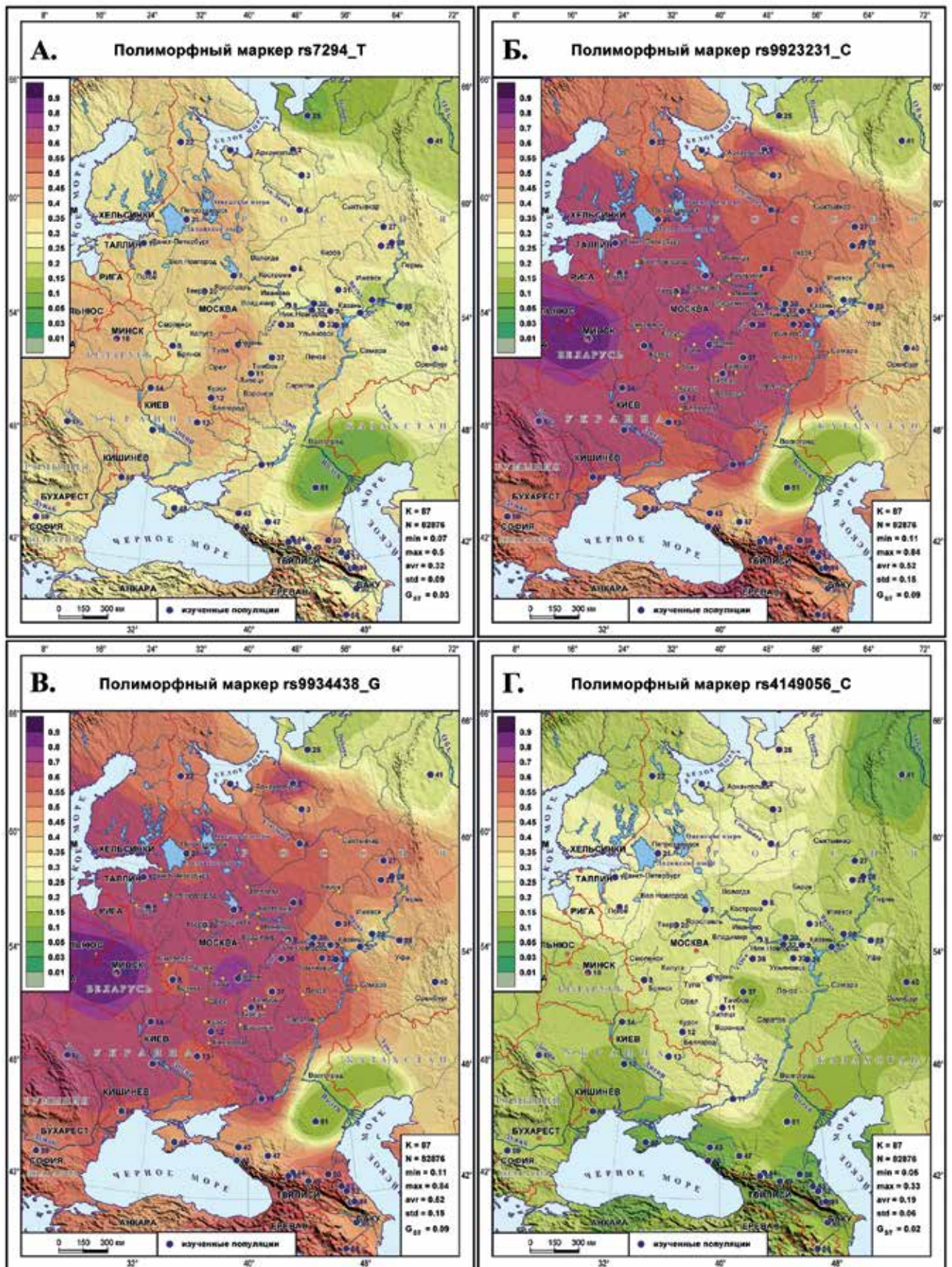


Рис. 4 Генетический ландшафт распространения маркеров VKORC1 и SLCO1B1: А) маркера rs7294 (Т) гена *VKORC1*; Б) маркера rs9923231 (С) гена *VKORC1*; В) маркера rs9934438 (G) гена *VKORC1*; Г) маркера rs4149056 (С) гена *SLCO1B1*.

Примечание: см. рисунок 1. Цветное изображение доступно в электронной версии журнала.



Рис. 5 Положение изученных популяций Восточной Европы в генетическом пространстве 1 и 2 главных компонент: А) по общей панели 60 фармакогенетических маркеров; Б) по панели 24 кардиологических маркеров.

Примечание: на обоих графиках зеленым цветом отмечены метапопуляции славянских народов, красным — финноязычных, фиолетовым — тюркоязычных. Названия метапопуляций расшифрованы в таблице 1 Приложения. Цветное изображение доступно в электронной версии журнала.

**rs3892097 (Т) гена *CYP2D6*** (рисунок 1 А) при слабо выраженной тренде падения частоты с запада на восток обнаруживает максимумы частоты не только в русских популяциях (в Рязанской и Калужской областях 28%, в Ярославской, Новгородской и Псковской — 22%), но и в Приуралье у коми-пермяков (19-21%). На Кавказе также немало максимумов: на Восточном Кавказе — 20%, Западном — 18%, Центральном — 16%. Минимумы (0-5%) сосредоточены не только в Волго-Уральском регионе и южном Урале, но и на северо-востоке у северных русских, ненцев и манси.

**rs1065852 (A) гена *CYP2D6*** (рисунок 1 Б) по своему генетическому ландшафту почти идентичен предыдущему. Можно отметить повышение частоты у белорусов (21%), казаков (30%) и лезгин (33%).

**rs28371725 (Т) гена *CYP2D6*** (рисунок 1 В) обладает иным паттерном, хотя и крайне слабо выраженном из-за малого диапазона частоты и средней частоты почти в 3 раза меньшей, чем на предыдущей карте. Максимумы сосредоточены на западе (украинцы 21%) и на юге (17-18% у народов Западного Кавказа, крымских и астраханских татар), но коридор минимумов тянется от карел до Центрального Кавказа и калмыков, захватывая многие русские популяции, мордву, коми-пермяков. Однако в Волго-Уральском регионе частота выше, чем в окружающем поясе низких частот.

**rs776746 (T) гена CYP3A5** (рисунок 1 Г) обнаруживает нечеткий градиент увеличения частоты в обратном направлении: с запада на восток. Низ-

кие частоты (2-4%) характерны для рязанцев, коми-пермяков, украинцев. Повышенные частоты обнаруживаются и на северо-западе (у финнов — 21%, у северных русских и манси — 15%), и на юге (в Молдавии — 20%, у калмыков — 15%), и в популяциях Кавказа.

### Генетический ландшафт распространения маркеров CYP2C19 и CYP3A4

Рассматриваемые маркеры гена *CYP2C19* относятся к антиагрегантам и ингибиторам протонной помпы, но их генетические ландшафты при отсутствии клинальной изменчивости и сходстве многих очертаний альтернативны — с повышением частоты одного маркера соотносится падение частоты другого.

**rs 12248560 (T) гена *CYP2C19*** (рисунок 2 А) имеет в картографируемом ареале при диапазоне частот 8-38% обширный ареал средних частот ( $\approx 23\%$ ). Общий паттерн без клинальной изменчивости: высокие частоты сосредоточены в центре с неравномерным снижением к окраинам. Однако максимумы встречены и в Украине, и на Русском Севере. Поэтому на севере снижение частоты зафиксировано лишь на западной периферии у северных карел и саамов, а на восточной периферии у манси. В славянских популяциях частота маркера велика — в среднем, 29%. У финно-угров при средней частоте 24% размах различий очень велик — от 11 до 35%. На Кавказе максимумы обнаруживаются и на западе у адыгов, и на востоке у азербайджанцев.



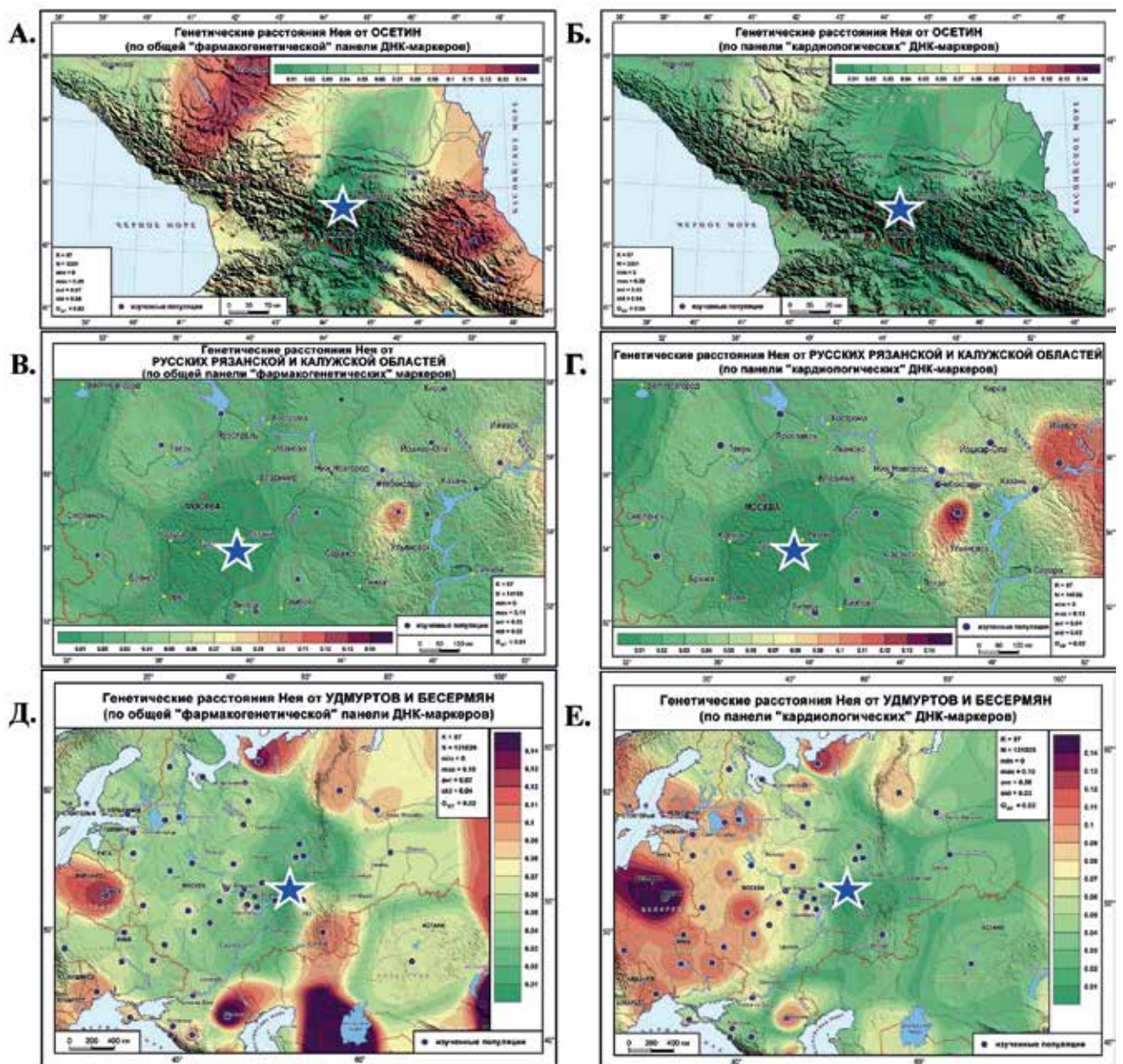


Рис. 6 Карты генетических расстояний по частотам общей фармакогенетической панели 60 ДНК-маркеров (слева) и 24 кардиологических ДНК-маркеров (справа): А, Б — от осетин до популяций Центрального Кавказа; В, Г — от русских Рязанской и Калужской областей до популяций Центральной России; Д, Е — от удмуртов и бесермян до популяций Восточной Европы.

Примечание: звездой обозначены реперные популяции (русские, удмурты, осетины), от которых рассчитаны генетические расстояния Неа до каждой из групп сравнения (обозначены кружками). Минимальные генетические расстояния (сходство с генофондом реперной популяции) обозначены зелеными тонами; максимальные генетические расстояния (несходство с генофондом реперной популяции) обозначены красно-бордовыми тонами; шкала переходов приведена в легенде каждой карты. Цветное изображение доступно в электронной версии журнала.

rs4244285 (А) гена *CYP2C19* (рисунок 2 Б) со средней частотой 15% обладает более сглаженным генетическим рельефом. Но теперь в центре сосредоточены минимумы, а на периферии разбросаны популяции с повышенными частотами — 26% у калмыков, 24% у степных тюрков и северных карел. В славянских популяциях средняя частота всего 12%. У западных финно-угров средняя частота (16%) и размах различий (10-24%) больше, чем у приуральских (12%, 6-15%). Для Кавказа харак-

терны невысокие частоты, лишь у чеченцев и ингушей поднимающиеся до 19%.

rs4986893 (А) гена *CYP2C19* (рисунок 2 В) крайне редок. Но и области нулевой частоты, и наличия маркера являются сплошными, а не случайными пятнами, поэтому ландшафт нельзя объяснить ошибками выборки. Регионы, в которых маркер встречается, вновь находятся на периферии: с максимумом на севере у манси (12%), на юге вновь у тюрков степи (8%), на Кавказе вновь у чеченцев и ингушей (7%).

**rs2242480 (T) гена *CYP3A4*** (рисунок 2 Г) функционально связан с остановкой кровотечений. В его ландшафте вновь отсутствует клинальная изменчивость. Область низких частот тянется от запада до востока. А всплески частоты обнаруживаются на периферии: и на севере у финнов (24%) и в Архангельской области (16%); на юге в Молдавии (25%) и у башкир (21%). У славян и западных финно-угров частота, в среднем (12%), выше, чем у поволжских финно-угров (8%).

#### **Генетический ландшафт распространения маркеров *CYP2C9*, *CYP4F2*, *CES1***

Все картографированные маркеры этих генов относятся к различным антикоагулянтам: *CYP2C9* — к нестероидным противовоспалительным, *CYP4F2* — к непрямым антикоагулянтам, антиагрегантам; *CES1* — к прямым оральным, антиагрегантам. Их генетические ландшафты различаются: у *CYP2C9* клинальная изменчивость практически отсутствует, а у *CYP4F2* и *CES1* выражена весьма убедительно.

**rs1057910 (C) гена *CYP2C9*** (рисунок 3 А) вновь являет сглаженный генетический рельеф с экстремумами на периферии: на севере падение частоты до нуля у ненцев и всплеск у северных карел до 19%, на юге — падение частоты до нуля у калмыков и всплеск частоты у адыгов (26%), и на Кавказе средняя частота (15%) выше среднерегиональной (8%). На остальном пространстве Восточной Европы встречаются лишь локальные невыразительные колебания частоты.

**rs1799853 (T) гена *CYP2C9*** (рисунок 3 Б) хотя имеет в 2 раза больший диапазон частоты (0–50%), но его генетический ландшафт почти столь же невыразителен. Дело в том, что максимальная частота 50% обнаружена лишь у фарсиязычных народов Закавказья. Но на Кавказе средняя частота (9%) такая же, как и в регионе, варьирует в тех же пределах: от 0 до 18% в Дагестане. Хотя просматривается тенденция клинальной изменчивости (снижение частоты с запада на восток), но она скрыта локальными колебаниями частоты.

**rs2108622 (T) гена *CYP4F2*** (рисунок 3 В) демонстрирует четко выраженную клинальную изменчивость увеличения частоты с северо-запада (9%) на восток и юго-восток (46%), нарушенную коридором повышенных частот (20–27%), тянущимся от белорусов через Рязань до юга Архангельской области. Высокие частоты сосредоточены на северо-востоке (поморы, ненцы, манси), левобережье Волги и юге региона, включая Крым и Кавказ.

**rs2244613 (G) гена *CES1*** (рисунок 3 Г) характеризуется выраженным вектором увеличения частоты с северо-запада (18% у карел) на восток (46–54% у ненцев, манси, башкир) и юг (46% у калмыков и в Дагестане). Область высоких частот охватывает популяции правобережья Волги (35–44% у марий-

цев и чувашей), многие народы Северного Кавказа (33–45%) и Закавказья (33–38%).

#### **Генетический ландшафт распространения маркеров *VKORC1* и *SLCO1B1***

Маркеры гена *VKORC1* влияют на эффективность применения антикоагулянтов непрямого действия, например, варфарина. Ген *SLCO1B1* кодирует полипептид-переносчик органических анионов, участвующий в транспорте статинов. В их изменчивости проявляется тенденция падения частоты с северо-запада на юго-восток.

**rs7294 (T) гена *VKORC1*** (рисунок 4 А) характеризуется повышенными частотами на северо-западе (44% у вепсов, 45% у белорусов). Средняя частота маркера одинакова у западных финно-угров и славян (37%). У южных русских ~42%, в Калужской и Рязанской областях поднимается до 50%. Минимумы обнаруживаются на севере у ненцев (9%) и манси (15%), в Приуралье у удмуртов (16%), на юге у калмыков (7%). Но в целом низкие частоты сосредоточены на юго-востоке, формируя вектор падения частоты с северо-запада на юго-восток. На Кавказе обнаружены локальные максимумы (46% у чеченцев и ингушей, 42–44% у армян и грузин).

**rs9923231 (C) гена *VKORC1*** (рисунок 4 Б) полностью вторит предыдущему генетическому ландшафту, но отличается не только в два раза большей средней частотой (62%) но и намного большим разбросом частот (73 вместо 43%). При этом становятся заметнее перепады частоты у русских Архангельской области (48–64%).

**rs9934438 (G) гена *VKORC1*** (рисунок 4 В) обладает генетическим ландшафтом, практически неотличимым от предыдущего.

**rs4149056 (C) гена *SLCO1B1*** (рисунок 4 Г) характеризуется тем же вектором снижения частоты с северо-запада (33% у карел и вепсов) на юг (6–10% в Крыму и на Кавказе) и юго-восток (14–15% у башкир и калмыков). Полоса повышенных частот (28–30%) тянется от нижегородцев к казакам, а с юга ее огибает полоса низких частот (10–14% в Молдавии, у степных тюрков и калмыков). Локальный минимум у манси (5%) соседствует с повышенными частотами (24–27%) на Русском Севере и у ненцев.

## **2. Положение популяций Восточной Европы в генетическом пространстве изменчивости фармакогенетических ДНК-маркеров**

Созданный картографический атлас позволяет выявлять и сравнивать пространственную изменчивость каждого из 24 кардио-маркеров. Но для того, чтобы получить общую картину, необходимы методы многомерной статистики. Из них наиболее подходящим является метод главных компонент, учитывающий скоррелированность маркеров. Для большей репрезентативности анализа были отобраны популяции, для которых проведено генотипирование по единой панели 60 фарма-маркеров, объ-



единены популяции одного этноса и анализ проведен дважды — как для 24 кардио-ДНК-маркеров, так и для общей фарма-панели, включающей кроме кардио-маркеров еще 36 фарма-маркеров (Приложение, таблица 1). В итоге в оба вида анализа включены 13 метапопуляций (рисунок 5) со средней выборкой 158 человек.

В генетическом пространстве главных компонент изменчивости 60 фарма-маркеров выделились три кластера (рисунок 5 А), условно названные "Северный", "Западный" и "Восточный", поскольку согласовываются с географическим пространством (но положение популяций в кластере не всегда соответствует географии). "Северный" кластер включил западные финноязычные популяции (вепсов, воть, ижору, карел, саамов) и русских Архангельской области. "Западный" кластер включил все остальные русские популяции, украинцев и субэтносы мордвы, для которых выявлен значительный вклад в геномы широкого круга русских популяций [13]. "Восточный" кластер включил все тюрко- и финноязычные популяции Приуралья, за исключением самых северных коми-пермяков, приближающихся к "Северному" кластеру. Иными словами, кластеры, выделившиеся при использовании общей фарма-панели, не противоречат географии и приближаются к селективно-нейтральному паттерну [3], но противоречат лингвистике — каждый кластер включает народы, говорящие на языках разных лингвистических семей — индоевропейской (русские, украинцы), уральской (вепсы, карелы, коми-пермяки, марийцы, удмурты) и алтайской (татары, чуваша).

При выделении из широкой фарма-панели только кардио-маркеров (рисунок 5 Б) сохраняются все три кластера, но есть и перемены. "Северный" кластер становится очень "рыхлым" за счет больших отличий карел. "Западный" кластер, напротив, уплотняется. Но оба кластера не меняют состав в отличие от "Восточного", в который вошли коми-пермяки, но ушли "на север" марийцы и чуваша.

## Обсуждение

В генетическом пространстве главных компонент кластеры кардио-панели вступают в противоречие не только с лингвистикой, но и с географией, и с селективно-нейтральным паттерном, хотя сами кластеры и их положение сохраняются такими же, как и при всей фарма-панели. И такое противоречие не удивительно — мы видели его на картах атласа. В чем причина своеобразия кардио-панели? Само сокращение числа маркеров с 60 до 24 вряд ли сказалось, но могло проявить действие отбора из-за функциональной системности кардио-панели. В общую фарма-панель включены маркеры разной направленности, которые, работая как лебедь, рак и щука, не могут отодвинуть "воз" от селективно-

нейтрального паттерна ("Из кожи лезут вон, а возу все нет ходу!"<sup>6</sup>). При более гармоничном действии "специализированных" кардио-маркеров эффект отбора может стать более зримым.

С помощью картографирования можно своими глазами увидеть, насколько отличается изменчивость популяций по маркерам кардио-панели от общей фарма-панели. В работах [12-14] приведены карты генетических расстояний от разных народов (удмуртов, русских и осетин) до окружающих популяций по фарма-панели и проведено сравнение с селективно-нейтральным генетическим ландшафтом. Увеличение фарма-панели (с 45 до 60 маркеров) и средней выборки из популяции (до 145 человек), позволяет теперь сравнить с ней генетический ландшафт по 24 кардио-маркерам (рисунок 6).

На Кавказе по кардио-панели (рисунок 6 Б) все народы генетически сходны — Кавказ окрашен в зеленые тона малых генетических расстояний (небольшие отличия обнаруживаются лишь у степных ногайцев). Но по "общей" фарма-панели (рисунок 6 А) область сходства с генофондом осетин крайне мала — почти все генофонды Кавказа от них резко отличаются. Относительно близкими к осетинам ( $0,06 < d < 0,07$ , желтые тона) оказались только адыги, грузины, ингуши, чеченцы.

Для русских популяций обнаруживаем противоположную картину. Различий по "общей" фарма-панели почти нет (рисунок 6 В) — весь ландшафт окрашен в зеленые тона генетического сходства, охватывающее и русское, и финноязычное, и тюркоязычное население региона (лишь чуваша Татарстана ненамного отличаются от рязанцев). Однако по кардио-панели (рисунок 6 Г) число популяций, генетически несходных с рязанцами, уже заметно больше: и чуваша отличаются значительней, и к ним присоединяются удмурты, бесермяне и горные марийцы ( $d \geq 0,08$ ).

Но еще более яркую метаморфозу демонстрируют карты генетических различий между удмуртами и народами Восточной Европы. Если по "общей" фарма-панели генетически сходные популяции расположены на западе (рисунок 6 Д), то по кардио-панели вектор меняется на противоположный (рисунок 6 Е): западные популяции резко отличны от удмуртов, а область генетического сходства тянется на восток.

Визуализация генетического сходства показывает, что для разных регионов нет единого правила при переходе от совокупности фарма-маркеров к ДНК-маркерам отдельной группы заболеваний. Совокупность результатов позволяет предполагать, что действие отбора различно и по разным группам заболеваний, и по разным факторам воздействия

<sup>6</sup> И. А. Крылов. "Лебедь, Щука и Рак".

на организм в рамках одной группы заболеваний, и в разных регионах. Поэтому основной стратегией должно быть массовое генотипирование по единой обширной панели фарма-маркеров репрезентативных выборок народонаселения России, которое обеспечивается популяционными биобанками.

Создание картографического атласа распространения ДНК-маркеров, связанных с сердечно-сосудистой патологией, позволяет специалисту любого профиля, не прибегая к понятийному аппарату популяционной генетики и фармакогенетики, получить ответы на многие вопросы. Можно извлечь информацию о степени распространения любого из 24 ДНК-маркеров в любой части Восточной Европы, выявить регионы с его наибольшей частотой, определить закономерности в изменчивости и многое другое. Важное преимущество атласа — создание всех карт по единой технологии, единым параметрам и в одной шкале — позволяет корректно сравнивать любые карты друг с другом, получая некое многомерное пространство изменчивости изучаемых ДНК-маркеров.

Однако, для полноценного использования атласа важно обсудить две проблемы.

Во-первых, проблема соотношения коренного и "тотального" населения. Популяционная генетика во всем мире работает с коренным населением, и именно его частоты представлены на картах. Медицина обычно работает с "тотальным" населением, не оценивая его этническую разнородность. Как разрешить это противоречие? С помощью данных демографии, которая указывает, какую долю составляют разные этносы в "тотальной" популяции: зная частоту ДНК-маркера в каждом из этносов и взяв как вес долю этого этноса в тотальной популяции, с легкостью оцениваем искомую частоту в "тотальном" населении.

Вторая проблема связана с надежностью исходных данных, используемых для построения карт. В отличие от статистического анализа, для картографирования важно иметь как можно больше дробных популяций, покрывающих пространство карты. Но при таком дроблении снижается размер выборки и растет случайная ошибка. Отчасти она компенсируется за счет интерполяционной процедуры построения карты — соседние популяции как бы "корректируют" эту случайную ошибку. Но основным выходом служит увеличение объема выбор-

ки, что обеспечивается развитием популяционных биобанков.

## Заключение

Изучение частоты носительства полиморфных генов является важным этапом внедрения фармакогенетического тестирования в клиническую практику. Решение о необходимости фармакогенетического тестирования для персонализации терапии может приниматься в зависимости от этнической принадлежности пациента, что исключает необходимость избыточного необоснованного назначения тестирования пациентов. Проведение широких популяционных фармакогенетических исследований необходимо для выявления частоты носительства клинически значимых полиморфных генов и определения этнических групп, в которых проведение фармакогенетического тестирования может применяться для снижения нежелательных реакций и/или увеличения эффективности фармакотерапии и будет экономически обоснованно.

Созданный картографический атлас изменчивости кардиологических маркеров в населении Восточной Европы является стартовой площадкой для развития геногеографических исследований разных групп населения в разных регионах, которые должны базироваться на изучении популяций по единой панели фармакогенетических маркеров и по репрезентативным выборкам народонаселения, которое обеспечивается популяционными биобанками.

**Благодарности.** Авторы благодарят АНО "Биобанк Северной Евразии" за предоставление коллекций ДНК. Исследование выполнено при поддержке гранта РНФ № 21-14-00363 (биоинформатический и картографический анализ) и Государственного задания Минобрнауки России для ФГБНУ "МГНЦ" (статистический анализ, интерпретация результатов).

**Отношения и деятельность.** Исследование выполнено при поддержке гранта РНФ № 21-14-00363 (биоинформатический анализ, картографический анализ) и Государственного задания Министерства науки и высшего образования РФ для Медико-генетического научного центра им. академика Н. П. Бочкова (статистический анализ, интерпретация результатов, написание статьи).



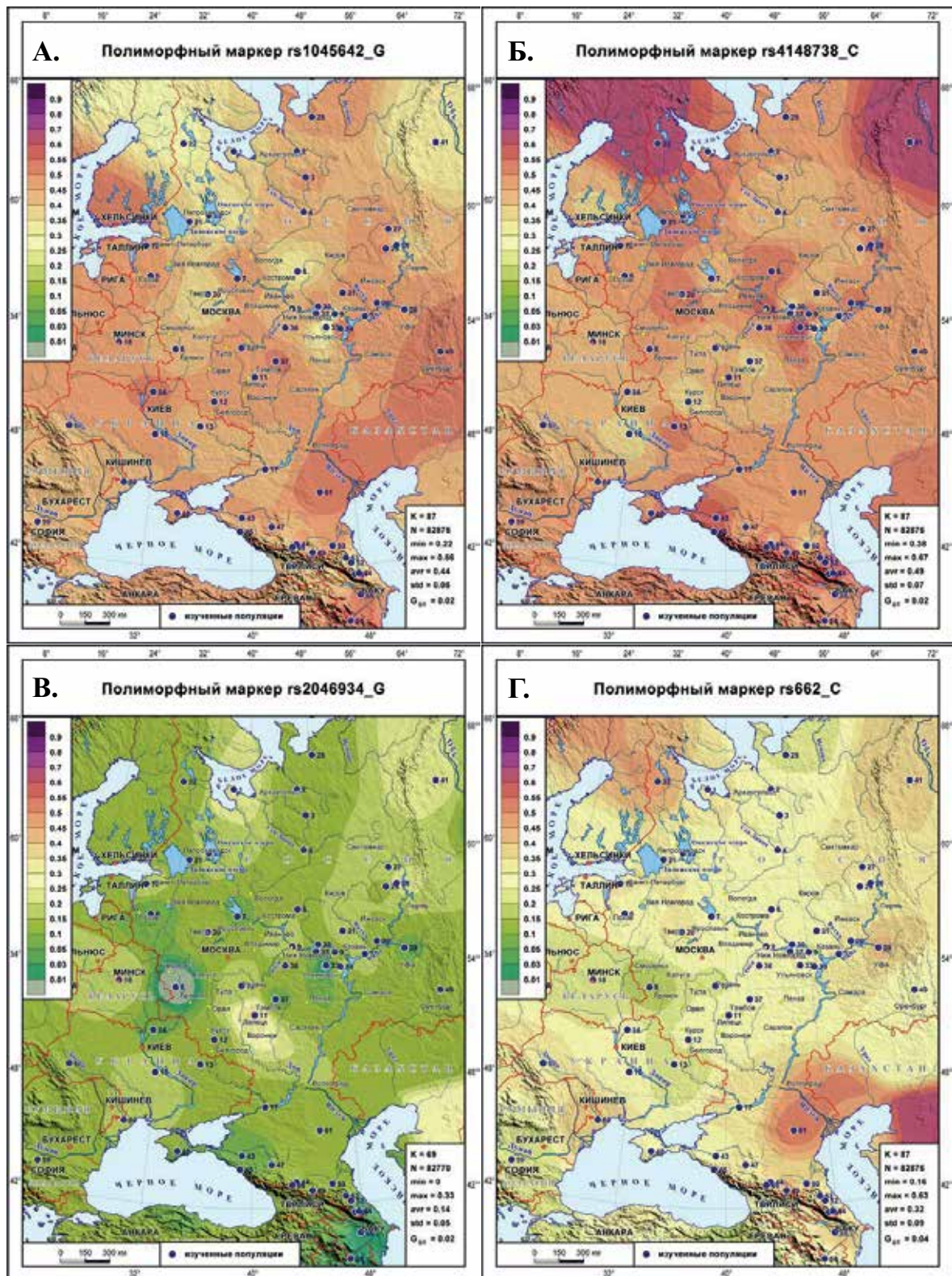
## Литература/References

1. Balanovska EV, Zhabagin MK, Agdzhoyan AT, et al. Population biobanks: organizational models and prospects of application in gene geography and personalized medicine. *Russ J 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.
2. Balanovska EV, Balanovsky OP. Russian gene pool on the Russian Plain. Moscow: Luch, 2007. p. 412. (In Russ.) Балановская Е.В., Балановский О.П. Русский генофонд на Русской равнине. М.: Луч, 2007. 412 с. ISBN: 978-5-87140-267-2.
3. Balanovsky OP. The Gene Pool of Europe. Moscow: Tovari-shhestvo nauchnykh izdaniy KMK, 2015. p. 354 (In Russ.) Балановский О.П. Генофонд Европы. М.: Товарищество научных изданий КМК, 2015. 354 с. ISBN: 978-5-9907157-0-7.
4. Balanovsky O, Dibirova K, Dybo A, et al. Parallel evolution of genes and languages in the Caucasus region. *Mol Biol Evol*. 2011;28(10):2905-20. doi:10.1093/molbev/msr126.
5. The Human Population Genetics. In: Medical genetics: National guidelines. Moscow: GEHOTAR-Media, 2022:216-54. (In Russ.) Популяционная генетика человека. В кн: Медицинская генетика: национальное руководство. М.: ГЭОТАР-Медиа, 2022:216-54. doi:10.33029/9704-6307-9-GEN-2022-1-896. ISBN: 978-5-9704-6307-9.
6. Balanovska E, Lukianova E, Kagazheva J, et al. Optimizing the genetic prediction of the eye and hair color for North Eurasian populations. *BMC Genomics*. 2020;21(Suppl 7):527. doi:10.1186/s12864-020-06923-1.
7. Balanovsky OP, Petruschenko VS, Gorin IO, et al. The accuracy of predicting eye and hair pigmentation based on genetic markers in Russian populations. *Bulletin of RSMU*. 2019;(5):25-41. (In Russ.) Балановский О.П., Петрушенко В.С., Горин И.О. и др. Точность предикции пигментации волос и глаз по генетическим маркерам для популяций России. *Вестник РГМУ*. 2019;5:25-41. doi:10.24075/vrgmu.2019.069.
8. Balanovska EV, Gorin IO, Koshel' SM, et al. Gene geographic atlas of DNA markers controlling the color of human eyes and hair. *Russ J Genetics*. 2021;57(12):1356-75. (In Russ.) Балановская Е.В., Горин И.О., Кошель С.М. и др. Геногеографический атлас ДНК-маркеров, контролирующих цвет глаз и волос человека. *Генетика*. 2021;57(12):1356-75. doi:10.31857/S0016675821120031.
9. Gorin IO, Petruschenko VS, Zapisetskaya YS, et al. Application of the population biobank for analysis of the distribution of the clinically significant DNA markers in the Russian populations: 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 распространенности клинически значимых ДНК-маркеров в населении России: биоинформатические аспекты. *Кардиоваскулярная терапия и профилактика*. 2020;19(6):2732. doi:10.15829/1728-8800-2020-2732.
10. Gorin IO, Balanovsky OP, Kozlov OV, et al. Determining the area of ancestral origin for individuals from North Eurasia based on 5,000 SNP markers. *Front. Genet*. 2022;13:902309. doi:10.3389/fgene.2022.902309.
11. Balanovska EV, Petruschenko 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 РГМУ*. 2020;(6):39-52. doi:10.24075/vrgmu.2020.080.
12. Balanovska EV, Napol'skikh VV, Churakov VS, et al. Gene pools of Udmurts and Besermians in the context of Finno-Ugric and other neighboring ethnic groups: the genome-wide and pharmacogenetic data. *Yearbook of Finno-Ugric Studies*. 2022;16(20):328-46. (In Russ.) Балановская Е.В., Напольских В.В., Чураков В.С. и др. Генофонды удмуртов и бесермян в контексте финно-угорских и других окружающих народов: полногеномные и фармакогенетические данные. *Ежегодник финно-угорских исследований*. 2022;16(2):328-46. doi:10.35634/2224-9443-2022-16-2-328-346.
13. Balanovska EV, Gorin IO, Ponomarev GY, 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. (In Russ.) Балановская Е.В., Горин И.О., Пономарев Г.Ю. и др. Следы взаимодействия финноязычного, славянского и тюркоязычного населения в современном генофонде и их отражение в фармакогенетике. *Вестник РГМУ*. 2022;2:20-9. doi:10.24075/vrgmu.2022.019.
14. Balanovska EV, Agdzhoyan AT, Gorin IO, et al. In search of the Alan heritage: the genetic history of the North Caucasus according to genome-wide data on the autosomal gene pool. *Moscow University Anthropology Bulletin. Series 23*. 2022;3(2022):48-62. (In Russ.) Балановская Е.В., Агджоян А.Т., Горин И.О. и др. В поисках аланского следа: генетическая история Северного Кавказа по полногеномным данным об аутосомном генофонде. *Вестник Московского университета. Серия XXIII. Антропология*. 2022;3(2022):48-62. doi:10.32521/2074-8132.2022.3.048-062.
15. Ol'kova MV, Petruschenko VS, Ponomarev GY. Analysis of 13 TP53 and WRAP53 polymorphism frequencies in Russian populations. *Bulletin of RSMU*. 2021;(1):33-42. (In Russ.) Олькова М.В., Петрушенко В.С., Пономарев Г.Ю. Анализ частот 13 полиморфизмов в генах TP53 и WRAP53 в российских популяциях. *Вестник РГМУ*. 2021;(1):33-42. doi:10.24075/vrgmu.2021.001.
16. Ol'kova MV, Balanovsky OP. 10 CDKN2A and RB1 genes' markers frequency research in Russian populations. *Bulletin of Experimental Biology and Medicine*. 2021;172(9):347-55. (In Russ.) Олькова М.В., Балановский О.П. Исследование частот 10 полиморфных маркеров генов CDKN2A и RB1 в российских популяциях. *Бюллетень экспериментальной биологии и медицины*. 2021;172(9):347-55. doi:10.47056/0365-9615-2021-172-9-347-355.
17. Kozlov AI, Balanovsky OP, Vershubskaya GG, et al. Genetically determined trehalase deficiency in various population groups of Russia and neighboring countries. *Vopr Pitan*. 2021;90(5):96-103. (In Russ.) Козлов А.И., Балановский О.П., Вершубская Г.Г. и др. Генетически детерминированная недостаточность трегалазы в различных группах населения России и сопредельных стран. *Вопросы питания*. 2021;90(5):96-103. doi:10.33029/0042-8833-2021-90-5-96-103.
18. Balanovsky O, Petruschenko V, Mirzaev K, et al. The variation of genome sites associated with severe COVID-19 across populations the worldwide and national patterns. *Pharmgenom Pers Med*. 2021;14:1391-402. doi:10.2147/PGPM.S320609.

19. Balanovska EV, Gorin IO, Petrushenko VS, et al. Gene Geography in Russia and the world of SNP markers of the LZTFL1 gene associated with severe COVID-19. Bulletin of RSMU. 2022;(5):31-40. (In Russ.) Балановская Е. В., Горин И. О., Петрушенко В. С. и др. Геногеография в России и мире SNP-маркеров гена LZTFL1, ассоциированных с тяжелым течением COVID-19. Вестник РГМУ. 2022;(5):31-40. doi:10.24075/vrgmu.2022.047.
20. Vos T, Lim S S, Abbafati C, et al. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet. 2020;396(10258):1204-22. doi:10.1016/S0140-6736(20)30925-9.
21. Piepoli MF, Abreu A, Albus C, et al. Update on cardiovascular prevention in clinical practice: A position paper of the European Association of Preventive Cardiology of the European Society of Cardiology. Eur J Prev Cardiol. 2020;27(2):181-205. doi:10.1177/2047487319893035.
22. Timmis A, Townsend N, Gale CP, et al. European Society of Cardiology: Cardiovascular Disease Statistics 2019. Eur Heart J. 2020;41(1):12-85. doi:10.1093/eurheartj/ehz859.
23. Health care in Russia. 2019: Statistical compendium. Moscow. 2019. p. 170. (In Russ.) Здравоохранение в России. 2019: Статистический сборник. Москва. 2019. 170 с. ISBN: 978-5-89476-470-2.
24. Koshel SM. Geoinformation technologies in Gene Geography. In: Sovremennaja geograficheskaja kartografija. Moscow: Data+. 2012;158-66. (In Russ.) Кошель С. М. Геоинформационные технологии в геногеографии. В кн: Современная географическая картография. М.: Дата+. 2012;158-66. ISBN: 978-5-94559-012-0. EDN: TBEVEF.



## Приложение

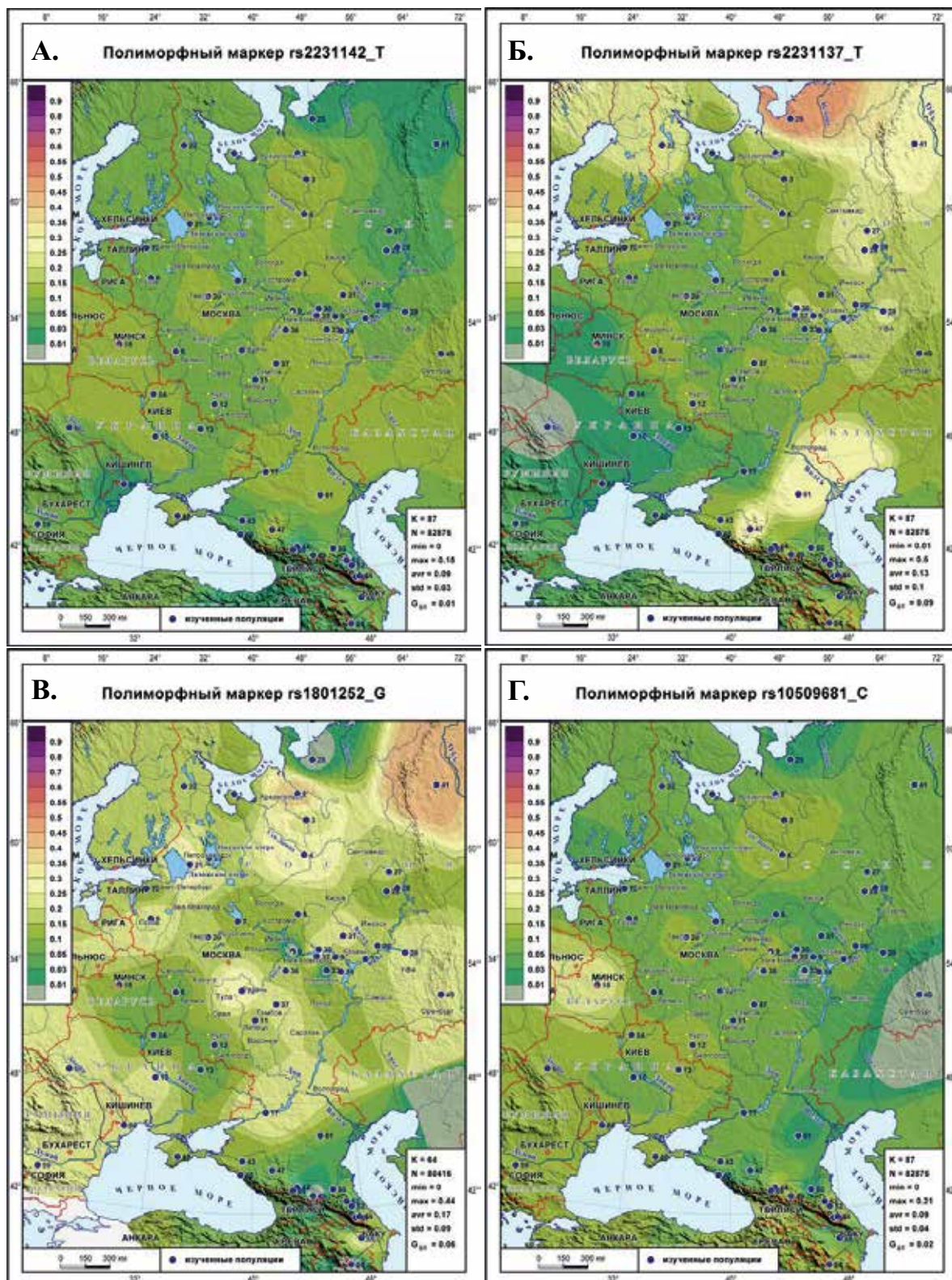


Приложение. Рис. 1 Генетический ландшафт распространения маркеров ABCB1, P2RY12 и PON1: А) маркера rs1045642 (G) гена *ABCB1*; Б) маркера rs4148738 (C) гена *ABCB1*; В) маркера rs2046934 (G) гена *P2RY12*; Г) маркера rs662 (C) гена *PON1*.

Примечание: на каждой карте в легенде слева дана шкала цветовых интервалов: красно-фиолетовые тона — высокая частота ( $\geq 50\%$ ), желто-оранжевые — средний диапазон (20-50%), зеленые и салатовые — низкие значения (1-20%), серый цвет  $< 1\%$ . Изученные популяции на картах обозначены темно-синими кружками и сопровождаются номером согласно таблице 1 Приложения. Цветное изображение доступно в электронной версии журнала.



## Приложение



Приложение. Рис. 2 Генетический ландшафт распространения маркеров ABCG2, ADRB1 и CYP2C8: А) маркера rs2231142 (Т) гена *ABCG2*; Б) маркера rs2231137 (Т) гена *ABCG2*; В) маркера rs1801252 (G) гена *ADRB1*; Г) маркера rs10509681 (C) гена *CYP2C8*. Примечание: см. рисунок 1 Приложения. Цветное изображение доступно в электронной версии журнала.



Приложение. Таблица 1

## Изученные популяции

Номер популяции*	Популяция	Объем выборки из популяции (п, человек)	Название метапопуляции на графике главных компонент**	Объем выборки из метапопуляции (п, человек)	Лингвистическая принадлежность метапопуляций
1	Русские северные поморы (Приморский и Онежский р-ны Архангельской обл.)	55	Рус-Арх	213	Славяне
2	Русские северные Лешуконского р-на Архангельской обл.	35	Рус-Арх		Славяне
3	Русские северные Пинежского р-на Архангельской обл.	66	Рус-Арх		Славяне
4	Русские северные Красноборского р-на Архангельской обл.	57	Рус-Арх		Славяне
5	Русские северные Вологодской и Костромской обл.	55	Рус-север	205	Славяне
6	Русские центральные Новгородской и Псковской обл.	77	Рус-север		Славяне
7	Русские центральные Ярославской обл.	73	Рус-север		Славяне
8	Русские центральные Тверской, Брянской и Смоленской обл.	63	Рус-центр	142	Славяне
9	Русские центральные Нижегородской обл.	30	Рус-центр		Славяне
10	Русские южные Калужской и Рязанской обл.	49	Рус-центр		Славяне
11	Русские южные Тамбовской и Воронежской обл.	42	Рус-юг	166	Славяне
12	Русские южные Орловской, Курской и Белгородской обл.	60	Рус-юг		Славяне
17	Казачи (донские, кубанские, некрасовские, терские)	64	Рус-юг		Славяне
13	Украинцы юго-восточные	36	Украинцы	206	Славяне
14	Украинцы северо-восточные	72	Украинцы		Славяне
15	Украинцы центральные	38	Украинцы		Славяне
16	Украинцы западные	60	Украинцы		Славяне
18	Белорусы	50	—	—	—
20	Карелы тверские	80	Карелы	157	Финно-угры
21	Карелы Карелии южные	55	Карелы		Финно-угры
22	Карелы Карелии северные, саамы	22	Карелы		Финно-угры
23	Вепсы	39	Вепсы	87	Финно-угры
19	Ижора ивудь	48	Вепсы		Финно-угры
24	Финны	37	—	—	—
25	Ненцы	21	—	—	—
26	Удмурты и бесермяне	90	Удмурты	90	Финно-угры
27	Коми-пермяки северные	86	Коми-перм	218	Финно-угры
28	Коми-пермяки юго-восточные	81	Коми-перм		Финно-угры
29	Коми-пермяки юго-западные	51	Коми-перм		Финно-угры
30	Мари горные	52	Мари	129	Финно-угры
31	Мари луговые	77	Мари		Финно-угры
32	Чуваши верховые	37	Чуваши	134	Тюрки
33	Чуваши низовые	18	Чуваши		Тюрки
34	Чуваши средненизовые	79	Чуваши		Тюрки
35	Татары казанские и кряшены Татарстана, татары Башкирии	85	Татары	156	Тюрки
36	Татары мишари	71	Татары		Тюрки
37	Мордовия (мокша, шокша)	65	Мордва	155	Финно-угры
38	Мордовия (эрзя)	90	Мордва		Финно-угры
39	Башкиры западные и северные	63	—	—	—
40	Башкиры юго-восточные	37	—	—	—
41	Манси	40	—	—	—
42	Ханты	56	—	—	—
43	Адыгейцы	45	—	—	—

Приложение. Таблица 1. Продолжение

Номер популяции*	Популяция	Объем выборки из популяции (п, человек)	Название метапопуляции на графике главных компонент**	Объем выборки из метапопуляции (п, человек)	Лингвистическая принадлежность метапопуляций
44	Кабардинцы, черкесы, абазины	48	—	—	—
45	Шапсуги, убыхи	34	—	—	—
46	Тюрки Кавказа (карачаевцы, балкарцы, кумыки)	40	—	—	—
47	Ногайцы (кубанские, караногайцы, ставропольские, астраханские)	56	—	—	—
48	Крым (крымские татары, караимы, греки)	39	—	—	—
49	Осетины	39	—	—	—
50	Нахские народы (чеченцы, ингуши)	45	—	—	—
51	Рутульцы, цахуры	40	—	—	—
52	Даргинцы, кубачинцы, кайтагцы, лакцы	36	—	—	—
53	Аварцы, дидойцы, гинухцы	44	—	—	—
54	Агульцы, табасаранцы, лезгины	35	—	—	—
55	Азербайджанцы	74	—	—	—
56	Армяне	37	—	—	—
57	Грузия	32	—	—	—
58	Фарсиязычные народы, курды, езиды, талыши	38	—	—	—
59	Славяне Балкан	70	—	—	—
60	Молдавия (молдаване, гагаузы)	20	—	—	—
61	Калмыки	36	—	—	—

Примечания: \* — номер популяции на картах рисунков 1-5 статьи, рисунков 1 и 2 Приложения; \*\* — метапопуляции на графиках главных компонент на рисунке 5 статьи.

Приложение. Таблица 2

Состав общей фармакогенетической панели ("фарма"-панели)

Идентификатор	Хромосома	Позиция в сборке GRCh37	Название гена/фрагмента генома	Приоритет при CC3
rs3892097	22	42524947	CYP2D6	1
rs1065852	22	42526694	CYP2D6	1
rs28371725	22	42523805	CYP2D6	1
rs776746	7	99270539	CYP3A5	1
rs12248560	10	96521657	CYP2C19	1
rs4244285	10	96541616	CYP2C19	1
rs4986893	10	96540410	CYP2C19	1
rs2242480	7	99361466	CYP3A4	1
rs1057910	10	96741053	CYP2C9	1
rs1799853	10	96702047	CYP2C9	1
rs2108622	19	15990431	CYP4F2*3	1
rs2244613	16	55844609	CES1	1
rs7294	16	31102321	VKORC1	1
rs9923231	16	31107689	VKORC1	1
rs9934438	16	31104878	VKORC1	1
rs4149056	12	21331549	SLCO1B1	1
rs1045642	7	87138645	ABCB1	2
rs4148738	7	87163049	ABCB1	2
rs2046934	3	151057642	P2RY12	2
rs662	7	94937446	PON1	2
rs2231142	4	89052323	ABCG2	3
rs2231137	4	89061114	ABCG2	3
rs1801252	10	115804036	ADRB1	3
rs10509681	10	96798749	CYP2C8	3



Приложение. Таблица 2. Продолжение

Идентификатор	Хромосома	Позиция в сборке GRCh37	Название гена/фрагмента генома	Приоритет при CC3
rs657152	9	136139265	<i>ABO</i> (Intron Variant)	
rs11385942	3	45876460	<i>LZTFL1</i> (Intron Variant)	
rs2032582	7	87160618	<i>ABCB1</i>	
rs1128503	7	87179601	<i>ABCB1</i>	
rs4680	22	19951271	<i>COMT</i>	
rs28399433	19	41356379	<i>CYP2A6</i>	
rs5031016	19	41349774	<i>CYP2A6</i>	
rs8192726	19	41354496	<i>CYP2A6</i>	
rs3745274	19	41512841	<i>CYP2B6</i>	
rs2279343	19	41515263	<i>CYP2B6</i>	
rs3211371	19	41522715	<i>CYP2B6</i>	
rs11572080	10	96827030	<i>CYP2C8</i>	
rs1801160	1	97770920	<i>DPYD</i>	
rs1801159	1	97981395	<i>DPYD</i>	
rs1801265	1	98348885	<i>DPYD</i>	
rs6277	11	113283459	<i>DRD2</i>	
rs1801028	11	113283484	<i>DRD2</i>	
rs1800497	11	113270828	<i>DRD2</i>	
rs1695	11	67352689	<i>GSTP1</i>	
rs1138272	11	67353579	<i>GSTP1</i>	
rs1801131	1	11854476	<i>MTFHR</i>	
rs1801133	1	11856378	<i>MTFHR</i>	
rs4271002	8	18248268	<i>NAT2</i>	
rs1041983	8	18257795	<i>NAT2</i>	
rs1801280	8	18257854	<i>NAT2</i>	
rs1799929	8	18257994	<i>NAT2</i>	
rs1799930	8	18258103	<i>NAT2</i>	
rs1208	8	18258316	<i>NAT2</i>	
rs1799931	8	18258370	<i>NAT2</i>	
rs1495741	8	18272881	<i>NAT2</i>	
rs4646244	8	18247718	<i>NAT2</i>	
rs116855232	13	48619855	<i>NUDT15</i>	
rs1051266	21	46957794	<i>SLC19A1</i>	
rs4148323	2	234669144	<i>UGT1A1</i>	
rs3064744	2	234668881	<i>UGT1A1</i>	
rs2236757	21	34624917	<i>IFNAR2</i> (Intron Variant)	