

## ГЕНОГЕОГРАФИЯ В РОССИИ И МИРЕ SNP-МАРКЕРОВ ГЕНА *LZTFL1*, АССОЦИИРОВАННЫХ С ТЯЖЕЛЫМ ТЕЧЕНИЕМ COVID-19

Е. В. Балановская<sup>1,2</sup>✉, И. О. Горин<sup>1</sup>, В. С. Петрушенко<sup>1</sup>, Д. К. Черневский<sup>1</sup>, С. М. Кошель<sup>1,3</sup>, И. И. Темирбулатов<sup>1,4</sup>, В. Ю. Пылёв<sup>1,2</sup>, А. Т. Агджоян<sup>1</sup>

<sup>1</sup> Медико-генетический научный центр имени Н. П. Бочкова, Москва, Россия

<sup>2</sup> Автономная некоммерческая организация «Биобанк Северной Евразии», Москва, Россия

<sup>3</sup> Московский государственный университет имени М. В. Ломоносова, Москва, Россия

<sup>4</sup> Российская медицинская академия непрерывного профессионального образования Минздрава России, Москва, Россия

Ранее была обнаружена корреляция между риском смерти от COVID-19 и этногеографическим происхождением пациента. Ген *LZTFL1* отмечается как потенциальный маркер, ассоциированный с двухкратным увеличением риска тяжелого течения COVID-19. Целью исследования было изучить пространственную изменчивость SNP-маркеров гена *LZTFL1* в коренном населении России и мира. Биоинформатическими, статистическими и картографическими методами был проведен анализ пространственной изменчивости полиморфных маркеров гена *LZTFL1* в 28 метапопуляциях (97 этносов) Северной Евразии ( $n = 1980$ ) и 34 метапопуляциях мира ( $n = 3637$ ). В Северной Евразии основные географические векторы изменчивости «север–юг» и «запад–восток» в целом согласуются с антропологическим вектором «европеоидность–монголоидность». Глобальная изменчивость тоже соответствует антропологии: каждый кластер коренного населения включает популяции только «своей» части света — Африки, Азии или Америки. «Индоевропейский» кластер объединяет европеоидные популяции Европы и Азии. Все четыре кластера коренного населения мира отдалены друг от друга, и только огромное генетическое разнообразие народов России и сопредельных стран является мостом, связующим три части света: Европу, Азию и Америку. Создан картографический атлас пространственной изменчивости 11 SNP-маркеров *LZTFL1* в популяциях. Выявлены основные закономерности: а) мировые экстремумы приходятся на коренное население Африки и Америки; 2) Северная Евразия является переходной зоной между мировыми экстремумами, но обладает собственными закономерностями и огромным размахом изменчивости в мировом масштабе; 3) генетический ландшафт России, как правило, органично вписан в Евразийский ландшафт.

**Ключевые слова:** COVID-19, *LZTFL1*, SNP, геогеография, популяции, коренное население, Россия, мир

**Финансирование:** исследование выполнено при поддержке гранта РФФИ №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).

**Благодарности:** авторы благодарят всех участников экспедиционных обследований коренного населения Северной Евразии (доноров образцов) и АНО «Биобанк Северной Евразии» за предоставление коллекций ДНК, М. В. Олькову — за участие в сборе информации о генетических вариантах, связанных с тяжестью протекания COVID-19.

**Вклад авторов:** Е. В. Балановская — анализ данных, написание текста, руководство исследованием; И. О. Горин, В. С. Петрушенко — биоинформатический анализ; А. Т. Агджоян, Д. К. Черневский, В. Ю. Пылёв — статистический анализ, И. И. Темирбулатов — описание фармакогенетических подходов; С. М. Кошель — картографический анализ

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

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

**Статья получена:** 13.09.2022 **Статья принята к печати:** 28.09.2022 **Опубликована онлайн:** 23.10.2022

**DOI:** 10.24075/vrgmu.2022.047

## GEOGRAPHIC DISTRIBUTION OF THE *LZTFL1* SNP MARKERS ASSOCIATED WITH SEVERE COVID-19 IN RUSSIA AND WORLDWIDE

Balanovska EV<sup>1,2</sup>✉, Gorin IO<sup>1</sup>, Petrusenko VS<sup>1</sup>, Chernevsky DK<sup>1</sup>, Koshel SM<sup>1,3</sup>, Temirbulatov II<sup>1,4</sup>, Pylev VYu<sup>1,2</sup>, Agdzhoyan AT<sup>1</sup>

<sup>1</sup> Research Centre for Medical Genetics, Moscow, Russia

<sup>2</sup> Autonomous non-profit organization “Biobank of North Eurasia”, Moscow, Russia

<sup>3</sup> Lomonosov Moscow State University, Moscow, Russia

<sup>4</sup> Russian Medical Academy of Continuous Professional Education, Ministry of Healthcare of the Russian Federation, Moscow, Russia

The correlation between the risk of death from COVID-19 and the patient's ethnogeographic origin has been previously detected. *LZTFL1* gene was identified as a potential marker of a two times higher risk of severe COVID-19. The study was aimed to assess spatial variation in the *LZTFL1* SNP markers in indigenous populations of Russia and the world. Spatial variation in the *LZTFL1* polymorphic markers was analyzed in 28 metapopulations (97 ethnic groups) of North Eurasia ( $n = 1980$ ) and 34 world's metapopulations ( $n = 3637$ ) by bioinformatics, statistical and cartographic methods. In North Eurasia, the major geographic variation vectors, North–South and West–East, are generally in line with the Caucasoid–Mongoloid anthropological vector. Global variation also corresponds to anthropological features: each cluster of indigenous populations includes only those from the place where it originates: Africa, Asia, or America. Indo-European cluster integrates Caucasoid populations of Europe and Asia. All four clusters of the world's indigenous population are separated from each other. The huge genetic diversity of Russia peoples and neighboring countries forms a bridge between three continents: Europe, Asia and America. Cartographic atlas for spatial variation in 11 *LZTFL1* markers in the populations has been created. The following major patterns have been revealed: а) the world's extrema fall on the indigenous populations of Africa and America; 2) Eurasia constitutes a transition zone between these two extrema, but has its own patterns and shows enormous scale of variation shows enormous variation on a global scale; 3) the genetic landscape of Russia tends to be seamlessly integrated into the Eurasian landscape.

**Keywords:** COVID-19, *LZTFL1*, SNP, gene geography, populations, indigenous people, Russia, world

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

**Acknowledgements:** the authors express their gratitude to all members of the expedition survey of the North Eurasian indigenous population (sample donors) and the autonomous non-profit organization “Biobank of North Eurasia” for access to DNA collections, and Olkova MV, for her participation in gathering information on the gene variants associated with the COVID-19 severity.

**Author contribution:** Balanovska EV — data analysis, manuscript writing, research management; Gorin IO, Petrusenko VS — bioinformatics analysis; Agdzhoyan AT, Chernevsky DK, Pylev VYu — statistical analysis; Temirbulatov II — explanation of pharmacogenetic approaches; Koshel SM — cartographic analysis.

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

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

**Received:** 13.09.2022 **Accepted:** 28.09.2022 **Published online:** 23.10.2022

**DOI:** 10.24075/brsmu.2022.047

Пандемия COVID-19, с ее высокой смертностью и тяжелыми осложнениями, заставила мировое научное сообщество включиться в поиск ДНК-маркеров, связанных с коронавирусной инфекцией SARS-CoV-2. Поскольку тяжесть заболевания COVID-19 неодинакова у представителей разных популяций мира, международный проект COVID-19 HGI начал сбор информации о встречаемости геномных вариантов ассоциированных с тяжелым течением COVID-19 [1]. Среди них ген *LZTFL1* [2] отмечается как потенциальный маркер, связанный с двукратным увеличением риска тяжелого протекания COVID-19 [3].

Ген *LZTFL1* экспрессируется в клетках легких человека и кодирует белок, участвующий в транспорте других белков к первичным ресничкам клеток мерцательного эпителия [4]. Клиническая значимость гена *LZTFL1* была известна и ранее: он ассоциирован с синдромом Барде–Бидля 17-го типа (OMIM 615994) [5], аутосомно-рецессивным заболеванием из группы цилиопатий [6, 7]. Семь белков BBS и BBIP10 образуют стабильный комплекс, известный как BBSome, обеспечивающий транспортировку белка к цилиарной мембране [8, 9], а снижение функции *LZTFL1* может компенсировать недостаток белков BBS и восстановить подвижность ресничек [10].

В клетках легочного эпителия пациентов с COVID-19 выявлены изменения, связанные с тяжелым течением заболевания COVID-19 и ассоциированные с геном *LZTFL1* [3, 11]: изменения в области хромосомы 3p21.31, включающей ген *LZTFL1*, приводят к двукратному увеличению риска дыхательной недостаточности [12, 13] и более чем двукратному увеличению риска смерти для лиц моложе 60 лет [11]. Исследование полиморфизма одного из вариантов гена *LZTFL1* в населении Великобритании выявило связь риска смерти от COVID-19 с происхождением пациента: для больных из Южной Азии риск в 4 раза больше, чем для больных европейского происхождения. Этими различиями отчасти объясняют более высокий уровень смертности представителей южноазиатских народов в Великобритании [3].

Ассоциация гена *LZTFL1* с тяжелым течением заболевания COVID-19, терапевтический потенциал и этногеографические различия требуют изучения его межпопуляционной изменчивости в населении мира. Научный коллектив располагает информационной базой, включающей массив собственных и литературных данных о геномах народов мира, которая уже обеспечила анализ изменчивости в населении мира (и России более детально) SNP-маркеров (*rs11385942*, *rs657152*), ассоциированных с тяжелым течением COVID-19 [14], и позволяет провести аналогичное исследование для 11 SNP-маркеров *LZTFL1*.

Цель исследования — изучить пространственную изменчивость SNP-маркеров гена *LZTFL1*, ассоциированного с тяжелым течением COVID-19 [15], в народонаселении: 1) провести поиск полиморфных SNP-маркеров гена *LZTFL1*, обеспеченных информацией об их встречаемости в коренном населении мира и Северной Евразии (более детально); 2) сформировать два репрезентативных массива популяционных данных об этих SNP-маркерах; 3) провести их многомерный статистический анализ; 4) создать картографический атлас распространения полиморфных SNP-маркеров *LZTFL1* в популяциях коренного населения Северной Евразии и мира.

## МАТЕРИАЛЫ И МЕТОДЫ

### Два исходных массива ДНК-маркеров

Массив данных о коренном населении Северной Евразии представлен популяциями 97 этносов преимущественно из России, а также из большинства постсоветских стран и Монголии. Образцы ДНК предоставлены АНО «Биобанк Северной Евразии», технология формирования выборок приведена ранее [16]: в выборки включены образцы только неродственных индивидов, все деды и бабушки которых относились к данному этносу. Образцы из географически и исторически близких популяций, но с малыми выборками, объединены в метапопуляции [17]. Результирующий массив данных о коренном народонаселении Северной Евразии включил 28 метапопуляций ( $n = 1980$ ) со средним размером выборки 140 хромосом.

Массив данных о коренном населении других регионов мира ( $n = 1657$ ) включает данные по широкогеномным панелям Illumina, опубликованным в научной литературе и аккумулированным в БД «GG-base» [18]. Географически и исторически близкие группы объединены в метапопуляции с целью формирования репрезентативной выборки. Итоговый набор данных о коренном населении мира ( $n = 3637$ ) включил 34 метапопуляции (64 популяции карт).

### Отбор полиморфных маркеров гена *LZTFL1*

При биоинформатическом анализе обоих массивов данных выявлено 10 SNP-полиморфизмов гена *LZTFL1*, репрезентативно представленных в коренном населении Северной Евразии и остальных регионах мира. Из них два SNP-маркера указаны как достоверно ассоциированные с COVID-19: *rs1058961* (3'-нетранслируемая область) и *rs12493471* (2-й интрон) [1]. Остальные восемь SNP-маркеров исследовали впервые: *rs11130077* (3-й интрон), *rs17078408*, *rs1860264*, *rs2191031*, *rs2236938*, *rs6441929*, *rs7614952* и *rs9842595* (2-й интрон). О маркере *rs17713054*, ассоциированном с тяжелым течением заболевания COVID-19 [3] (энхансер *LZTFL1*), отсутствует информация в БД «GG-base», поэтому его анализ проведен только для Северной Евразии.

Оценка неравновесия по сцеплению (LD R<sup>2</sup>) исследуемых 11 SNP-маркеров гена *LZTFL1* дана по массиву данных Северной Евразии (приложение, табл. 1). Для маркера *rs17713054* [3] выявлено выраженное сцепление с *rs12493471*, ассоциированным с тяжелым течением COVID-19 [1], а также с *rs2191031* и *rs11130077*.

### Статистический и картографический анализ

Многомерный статистический анализ проведен по частотам тех 10 SNP-маркеров *LZTFL1*, для которых есть информация о популяциях и Северной Евразии, и мира (приложение, табл. 2, 3): для Северной Евразии — методом многомерного шкалирования (MDS) по частотам как всех 10 SNP, так и только 5 SNP, наименее сцепленных между собой (приложение, табл. 1); для мира — по частотам всех 10 маркеров методами анализа главных компонент (PCA) и MDS. MDS рассчитаны по матрицам попарных генетических расстояний Нея (d).

По частотам SNP-маркеров гена *LZTFL1* построены карты их распространения в популяциях коренного населения Северной Евразии (11 SNP-маркеров) и мира (10 SNP-маркеров) с помощью оригинального ПО

GeneGeo. На картах и в таблицах (приложение, табл. 2, 3; рис. В) указан номер каждой популяции, по которому ее легко идентифицировать. Карты созданы на основе сеточной цифровой модели, представляющей собой матрицу интерполированных значений частоты маркера в узлах регулярной сетки, вычисленных процедурой средневзвешенной интерполяции с весами, убывающими пропорционально кубу расстояния, по значениям во всех опорных точках, попавших в круг радиуса  $R$  (для Северной Евразии  $R = 3000$  км, для мира  $R = 4200$  км). Единые цветовая и числовая шкалы всех карт создают единство атласа.

## РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

### Гетерогенность населения Северной Евразии по SNP-маркерам гена *LZTFL1*

Данные о частоте встречаемости отдельных SNP-маркеров в популяциях Северной Евразии представлены в табл. 2 в приложении. Но до анализа генетического ландшафта каждого из SNP-маркеров гена *LZTFL1* необходимо распознать закономерности в изменчивости всей их совокупности (рис. 1).

Наиболее важной чертой взаиморасположения популяций России и сопредельных стран на графике многомерного шкалирования (рис. 1) оказалось их следование географической изменчивости по обеим осям — «север–юг» и «запад–восток».

«Северный» кластер, объединивший самые северные популяции Западной Сибири (ненцев, манси, хантов) и Дальнего Востока (ительменов, коряков, чукчей), противостоит всем остальным кластерам, расположившимся строго по оси «запад–восток». Однако в пределах каждого из «южных» кластеров популяции не следуют их географическому положению.

«Западный» кластер включил в себя и всех восточных славян (белорусов, русских, украинцев), и финноязычные

народы России (бесермяне, карелы, коми-пермяки, мордва, удмурты).

«Урало-кавказский» кластер включил все народы Кавказа, Закавказья, Таджикистана и тюрков Урала (башкиры, поволжские татары, чувашаи), а также финноязычных марийцев, вошедших в одну метапопуляцию с чувашами.

«Мезокластер» визуализирует переход генофондов от «Западного» кластера к «Восточному», объединяя те народы Южной Сибири и Евразийской степи, генофонды которых включают как древний европеоидный пласт, так и мощные поздние наслоения монголоидного населения (алтайцы, казахи, каракалпаки, киргизы, ногайцы, сибирские татары, туркмены, узбеки, уйгуры, хакасы).

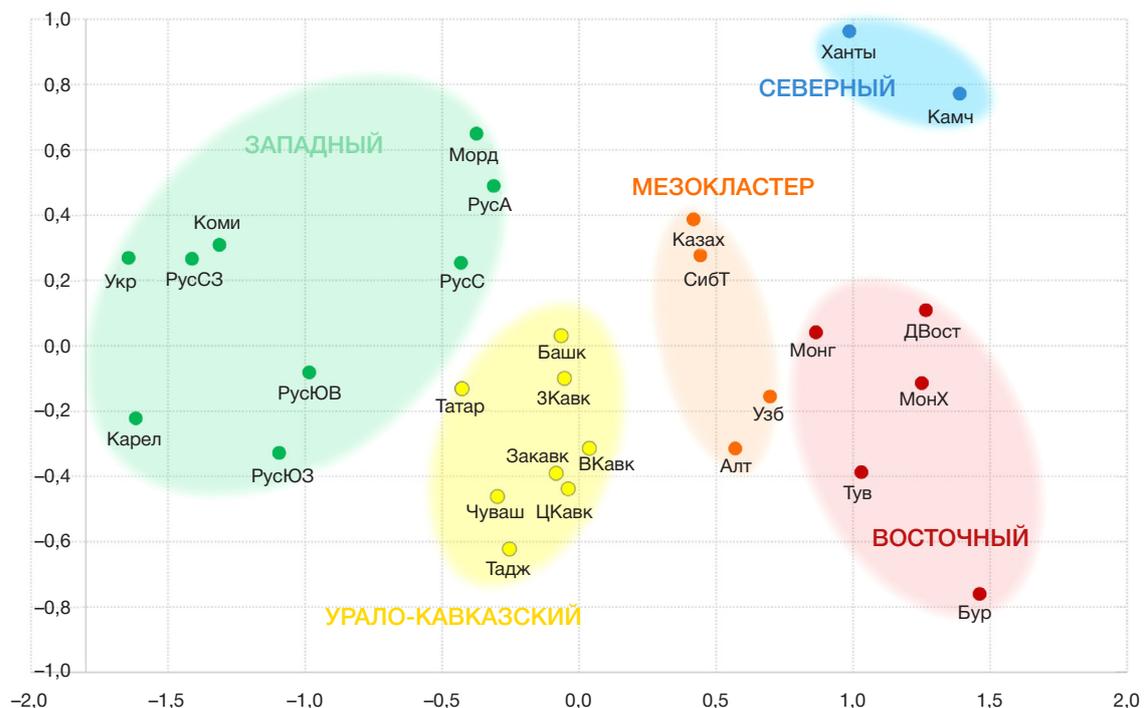
«Восточный» кластер объединил все монголоязычные народы Северной Евразии (буряты, калмыки, народы Монголии), тувинцев (до середины XX века входивших в состав Монголии) и малочисленные народы Дальнего Востока (нанайцы, нивхи, ульчи, эвены).

Аналогичный анализ, проведенный по частотам пяти наименее сцепленных ( $LD R^2 < 0,2$ ) SNP-маркеров, выявляет такую же структуру (приложение, рис. А) за исключением перехода поволжских татар в «Западный» кластер.

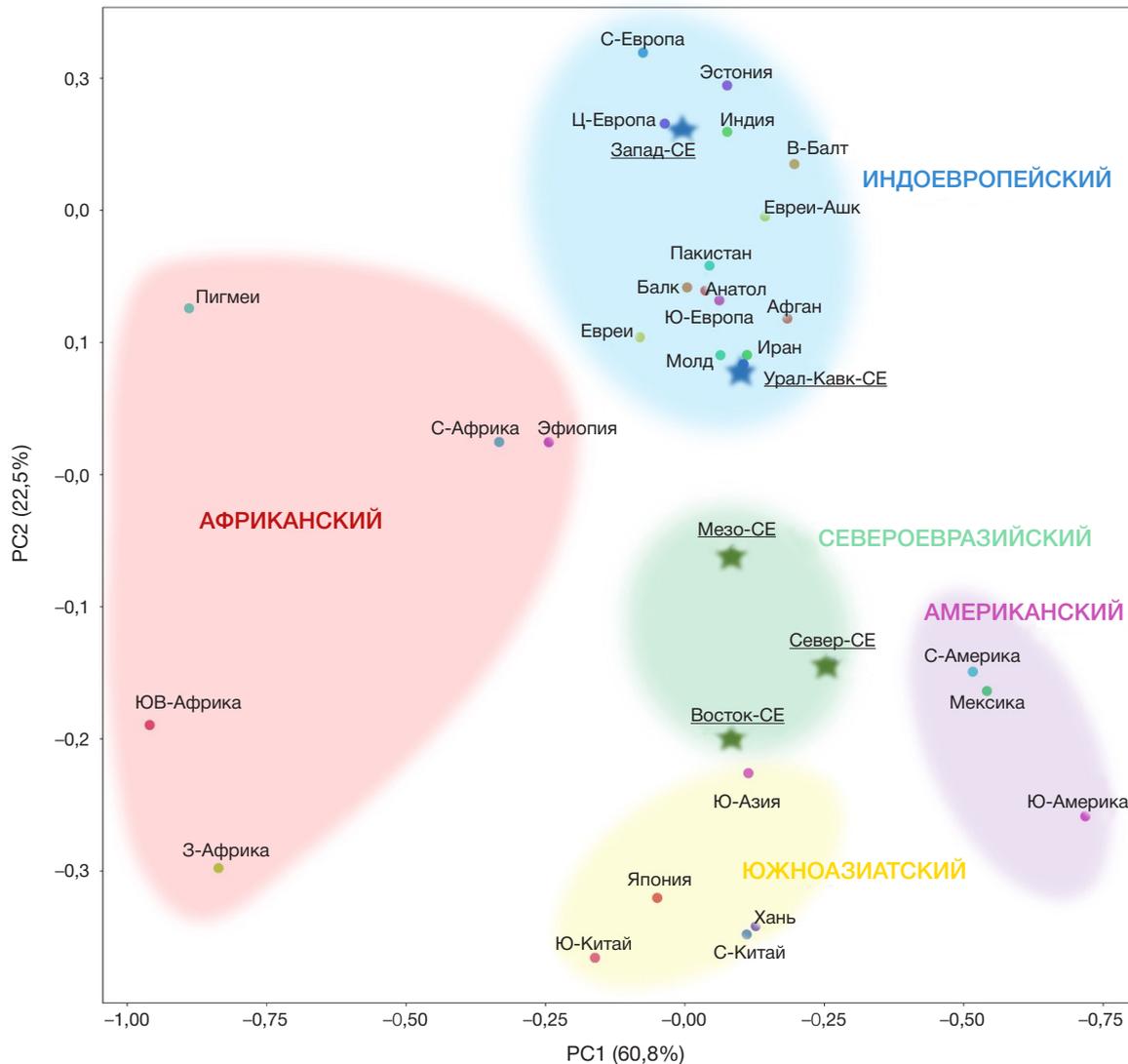
В целом, общий тренд изменчивости SNP-маркеров *LZTFL1* в Северной Евразии согласуется и с географическим вектором «запад–восток», и с антропологическим вектором «европеоидность–монголоидность» (рис. 1).

### Гетерогенность населения мира по SNP-маркерам гена *LZTFL1*

Глобальная изменчивость отдельных SNP-маркеров представлена в табл. 3 приложения, а закономерности по совокупности SNP-маркеров — на графиках главных компонент (приложение, рис. Б). Поскольку полученные обоими методами результаты и структура кластеров близки, рассмотрим график главных компонент (рис. 2).



**Рис. 1.** Положение популяций России и сопредельных стран в генетическом пространстве многомерного шкалирования по данным о 10 SNP-маркерах гена *LZTFL1*. Примечания: график многомерного шкалирования, показатель алиенации 0,1, стресса  $SO = 0,09$ . Синим цветом показан «Северный» кластер; зеленым — «Западный» кластер; желтым — «Урало-Кавказский» кластер; оранжевым — «Мезокластер»; красным — «Восточный» кластер



**Рис. 2.** Расположение популяций коренного населения мира в генетическом пространстве 1 и 2 главных компонент (ГК) изменчивости 10 SNP-маркеров гена *LZTFL1*. Звездочками отмечены пять кластеров, выделенных при многомерном шкалировании популяций России и сопредельных стран). Примечания: доля описываемой дисперсии: ГК1 (61%), ГК2 (22%). Синим цветом показан «Индоевропейский» кластер; красным — «Африканский» кластер; зеленым — «Североевразийский», желтым — «Южноазиатский», фиолетовым — «Американский»

Прежде всего бросается в глаза, насколько высока в мировом масштабе изменчивость популяций России и сопредельных стран — они не вместились ни в один из пяти кластеров в пространстве 1 и 2 главных компонент мирового генофонда (рис. 2). «Западный» кластер России расположился в верхней части «Индоевропейского» мирового кластера между популяциями Северной и Центральной Европы. «Урало-Кавказский» кластер России разместился в противоположной части «Индоевропейского» — среди популяций Передней Азии и Южной Европы. Три «азиатских» кластера Северной Евразии в генетическом пространстве мира образовали свой собственный «Североазиатский» кластер, в котором расположились в соответствии с их происхождением: «Мезокластер» России тяготеет к «Индоевропейскому» кластеру мира; «Восточный» кластер России — к «Южноазиатскому» мира; «Северный» кластер России (включивший народы Чукотки и Камчатки) близок к «Американскому».

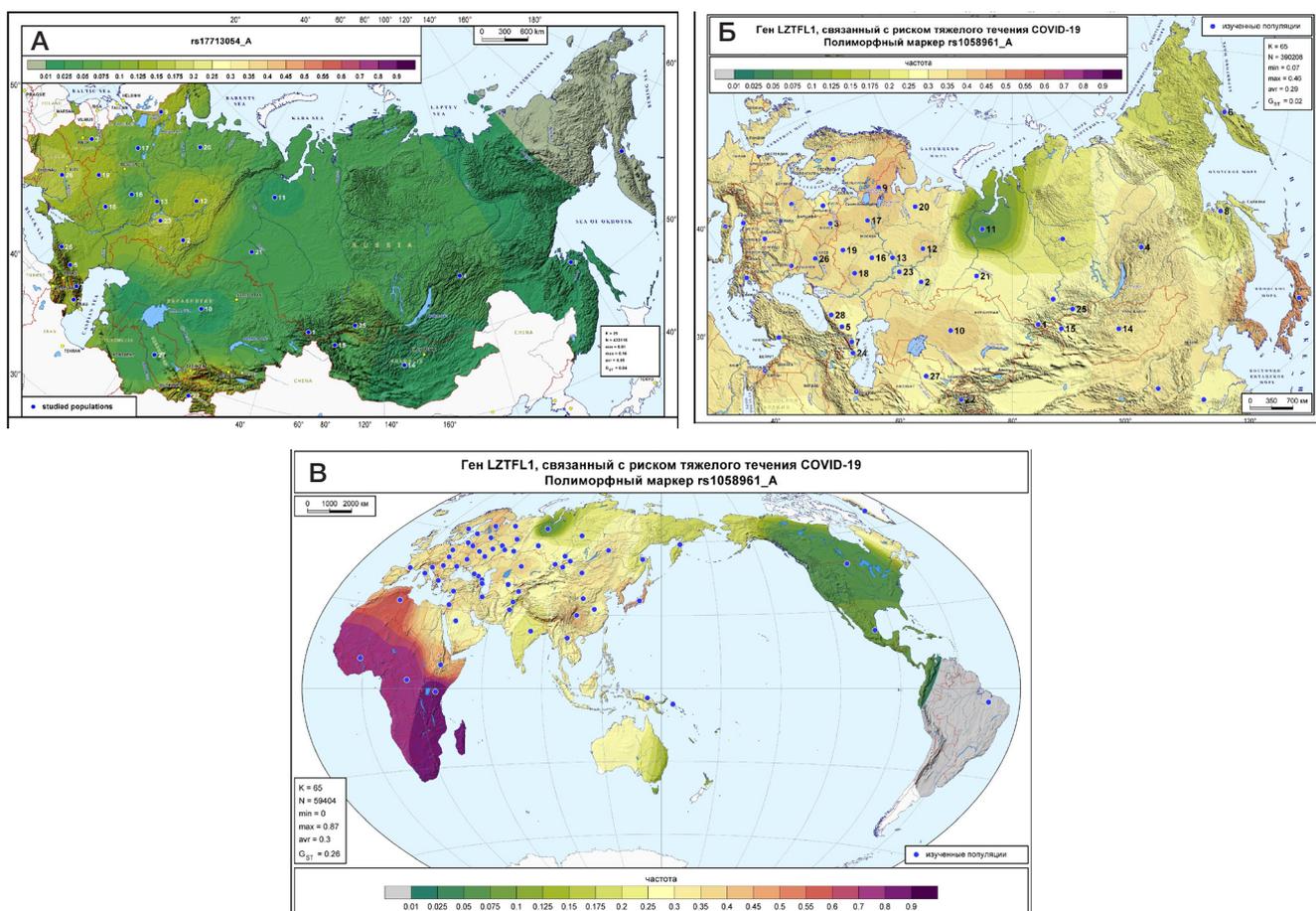
В целом, коренное народонаселение мира распределено по четырем кластерам в соответствии с частями света, но с поправками на антропологическую структуру населения (рис. 2). Три кластера включают популяции только «своей» части света:

Африки, Азии, Америки. Но «Индоевропейский» кластер сочетает географию и историю популяций, включая европеоидные популяции не только Европы, но и Азии (Индии, Пакистана, Афганистана, Ближнего Востока).

Иными словами, основной тренд изменчивости совокупности всех SNP-маркеров гена *LZTFL1* в мире хорошо соответствует антропологическому делению населения мира. Причем все четыре кластера мира отдалены друг от друга, и лишь огромное генетическое разнообразие народов России и сопредельных стран является мостом, связующим три части света — Европу, Азию и Америку.

#### Геногеография 11 SNP-маркеров гена *LZTFL1* в населении России и мира

Карты не являются «иллюстрацией», а добавляют к таблицам еще два измерения географического пространства, становясь эффективным и мощным инструментом анализа. Особое достоинство этого инструментария — возможность быстро охватить огромный объем информации благодаря ее невербальному представлению. Для каждого SNP-маркера *LZTFL1* (кроме



**Рис. 3.** Геноеографическая изменчивость частот SNP-маркеров гена *LZF1* в коренном народонаселении: 3А — изменчивость *rs17713054(A)* в населении Северной Евразии; 3Б — изменчивость *rs1058961(A)* в населении Северной Евразии; 3В — изменчивость *rs1058961(A)* в населении мира. Примечания: номера популяций на карте Северной Евразии соответствуют указанным в табл. 1 приложения, для карты мирового народонаселения — в табл. 2 и на рис. 1 приложения. Для коренного населения Австралии литературные данные не найдены, и на этот регион распространяется интерполяция от Океании

*rs17713054*, для которого нет информации в БД мира) мы создали по две карты его изменчивости: в коренном населении Северной Евразии и мира. Их сравнение позволяет выявлять и глобальные закономерности, и не упускать из вида собственный генетический ландшафт России. Каждая из 28 метапопуляций Северной Евразии отмечена на всех картах номером, позволяющим уточнить и название метапопуляции, и частоту SNP-маркера в ней (приложение, табл. 2). Информация позволяет ориентироваться в метапопуляциях мира (приложение, табл. 3, рис. В). Карты приведены в том же порядке, что и в табл. 1 приложения.

***rs17713054*.** Пространственная изменчивость частоты *rs17713054(A)* в Северной Евразии (рис. 3А) мала, но отражает тренд «запад–восток»: как минимум на Дальнем Востоке, как максимум — в европейской части, где высокие значения частоты есть и на западе (Украина, 16%), и на северо-западе (Карелия, 14%), и в Уральском регионе (16%), и на Кавказе (14%). Поэтому можно рассматривать европейскую часть Северной Евразии как регион наиболее высокой частоты встречаемости этого SNP-маркера. Другой регион повышения частоты намечается в Таджикистане ( $p = 0,15$ ), что может косвенно подтверждать более ранние выводы [3] о высоких частотах *rs17713054* в южных районах Азии.

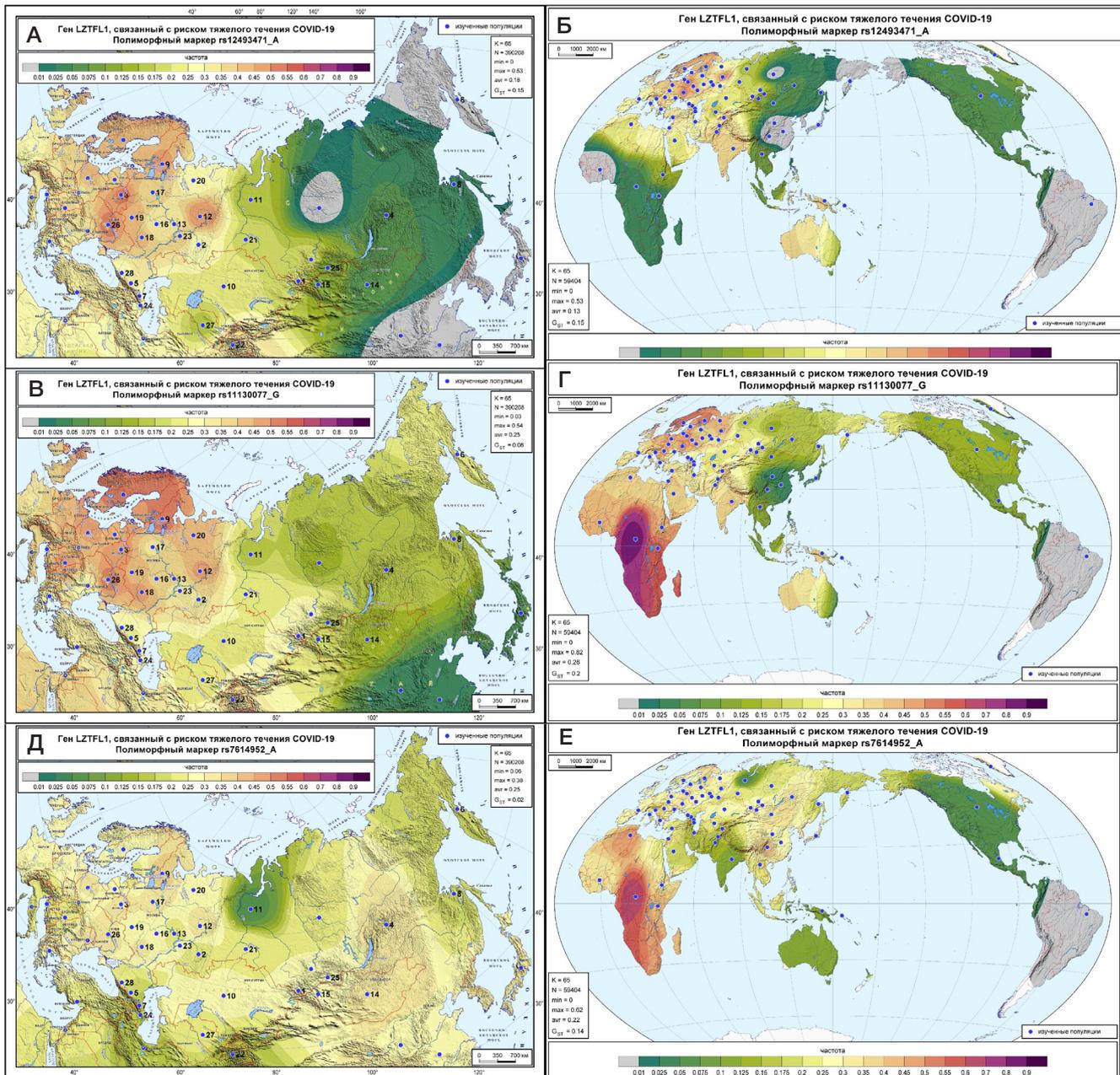
***rs1058961*.** Генетический ландшафт *rs1058961(A)* в Северной Евразии (рис. 3Б) при более высокой средней частоте (30%) отражает схожую, но еще более сглаженную клинальную изменчивость снижения частоты с запада

(43% у карел и вепсов) на северо-восток (20% на Дальнем Востоке). При этом обнаружены локальные минимумы на севере Западной Сибири (8%) и локальные максимумы в Средней Азии (37%).

При сравнении с мировой изменчивостью (рис. 3В) оказывается, что генетический ландшафт Северной Евразии практически полностью вписан в общий паттерн населения мира. Снижение частоты на Дальнем Востоке плавно переходит в снижение частоты на Аляске, снижение в коренном населении Северной Америки и падение до нуля в Южной Америке. Повышенные частоты в Европе постепенно переходят в максимальные частоты (87%) в Африке. Даже локальный максимум в Японии (46%) (рис. 3Б) находит отражение в повышении частоты в Юго-Восточной Азии (42%).

***rs12493471*.** Та же клинальная изменчивость «запад–восток» в Северной Евразии обнаружена и для *rs12493471(A)* (рис. 4А), сцепленного с *rs17713054(A)* (рис. 3А; приложение, табл. 1). Но градиент падения частоты с запада на восток выражен намного ярче и четче: максимумы ( $\approx 50\%$ ), охватывая европейскую часть России, почти не заходят за Урал. Падение частоты на востоке захватывает все регионы и снижается до нуля на Дальнем Востоке, в Японии и Китае. При этом пик частоты оказывается в Восточной Европе и Фенноскандии и снижается в Западной и Южной Европе.

Карта глобальной изменчивости (рис. 4Б) показывает, что максимум частоты в Восточной и Северной Европе является мировым максимумом — от него снижение



**Рис. 4.** Геногеографическая изменчивость частот SNP-маркеров гена *LZF1* в коренном народонаселении: 4А — изменчивость *rs12493471(A)* в населении Северной Евразии; 4Б — изменчивость *rs12493471(A)* в населении мира; 4В — изменчивость *rs1130077(G)* в населении Северной Евразии; 4Г — изменчивость *rs1130077(G)* в населении мира; 4Д — изменчивость *rs7614952(A)* в населении Северной Евразии; 4Е — изменчивость *rs7614952(A)* в населении мира. Примечания: номера популяций на карте Северной Евразии соответствуют указанным в табл. 1 приложения, для карты мирового народонаселения — в табл. 2 и на рис.1 приложения. Для коренного населения Австралии литературные данные не найдены, и на этот регион распространяется интерполяция от Океании

частоты идет во всех направлениях с локальными максимумами (40%) в Индостане (влияние которого доходит до Памира) и в Океании.

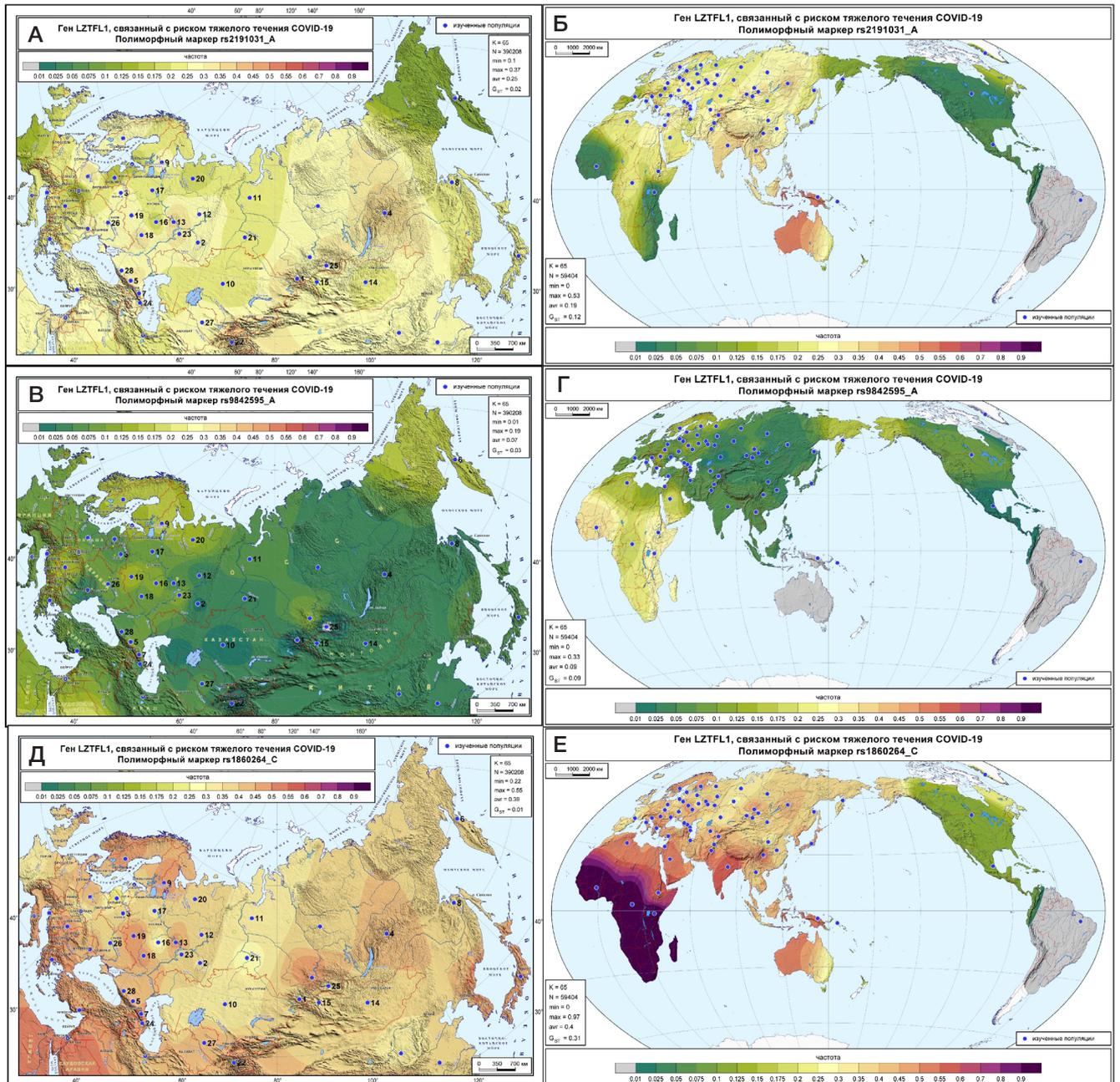
**rs1130077.** В Северной Евразии клинальная изменчивость «запад-восток» характерна и для *rs1130077(G)* (рис. 4В). Незначительные отличия связаны со смещением максимумов в Фенноскандию (54%), но они вновь не преодолевают Урал и не заходят на Кавказ. Однако в азиатской части картина не столь однозначна, как на предыдущих картах. При снижении частоты в Сибири до 16–17% на Дальнем Востоке намечается некоторое повышение частоты до 20%, а в Южной Сибири — до 27%.

Изменчивость *rs1130077(G)* в населении мира (рис. 4Г) в целом вторит предыдущей карте за одним резким исключением: мировой максимум приходится не на

Европу (53%), а на население Африки, в которой различия по этому маркеру велики: от 40% в Северной Африке до 82% у пигмеев Африки.

**rs7614952.** В отличие от предыдущих карт, генетический ландшафт *rs7614952(A)* (рис. 4Д) в Северной Евразии не отличается четкой закономерностью. Хотя максимальные значения по-прежнему находятся на западе региона и в направлении к Дальнему Востоку снижаются, но минимум частоты приходится на север Западной Сибири, а локальный максимум обнаружен не только в Европе, но и в Забайкалье.

В контексте мировой изменчивости (рис. 4Е) мы видим, что байкальский максимум является частью повышенных частот во всем регионе Восточной Азии. В этом наиболее яркое отличие от предыдущей карты:



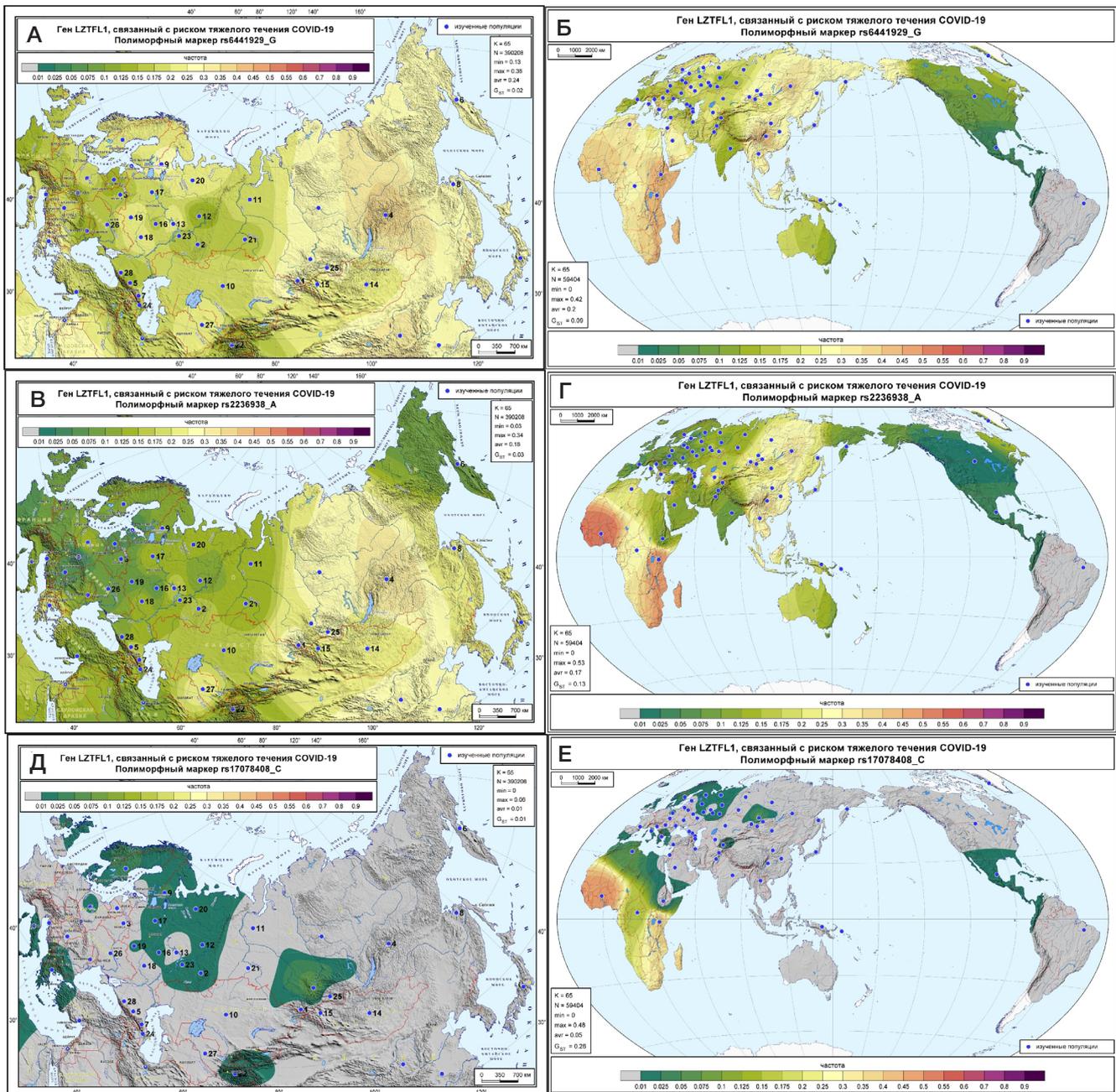
**Рис. 5.** Геногеографическая изменчивость частот SNP-маркеров гена *LZTF1* в коренном народонаселении: 5А — изменчивость *rs2191031(A)* в населении Северной Евразии; 5Б — изменчивость *rs2191031(A)* в населении мира; 5В — изменчивость *rs9842595(A)* в населении Северной Евразии; 5Г — изменчивость *rs9842595(A)* в населении мира; 5Д — изменчивость *rs1860264(C)* в населении Северной Евразии; 5Е — изменчивость *rs1860264(C)* в населении мира. Примечания: номера популяций на карте Северной Евразии соответствуют указанным в табл. 1 приложения, для карты мирового народонаселения — в табл. 2 и на рис.1 приложения. Для коренного населения Австралии литературные данные не найдены, и на этот регион распространяется интерполяция от Океании

в Восточной Азии частоты *rs11130077(G)* минимальны (рис. 4Г), но по *rs7614952(A)* (рис. 4Е) они высоки. И напротив: вместо высоких частот в Индостане и Передней Азии мы видим резкое падение частоты *rs7614952(A)*. Однако вновь мировой максимум находится в Африке, а мировой минимум — в Америке.

**rs2191031.** Изменчивость *rs2191031(A)* в Северной Евразии невыразительна (рис. 5А) из-за чересполосицы локальных максимумов и минимумов. Минимум вновь находится на Дальнем Востоке (10%), а также на западе региона (18%) и в Западной Сибири (22%). Высокие частоты тянутся от Забайкалья (36%) через Южную Сибирь (34%) к Центральной Азии (36%), но характерны и для Поволжья (37%), и для Кавказа (34%).

Этот паттерн, однако, полностью вписан в ландшафт мировой изменчивости (рис. 5Б): снижение частоты в Западной Европе (16%) переходит в минимум в Западной и Восточной Африке (3%), а снижение частоты на Дальнем Востоке — в минимумы коренного населения Америки (0–4%). Максимумы юга Сибири и Центральной Азии являются частью области высоких частот Юго-Восточной и Южной Азии (35–40%), достигая мирового максимума в Океании (53%).

**rs9842595.** Еще менее выразителен генетический ландшафт *rs9842595(A)* (рис. 5В): повышенные частоты разбросаны по Северной Евразии: на Дальнем Востоке (19%), на севере (17%) и юге (15%) Европы, в Приуралье (12%) и на Кавказе (11%).



**Рис. 6.** Геногеографическая изменчивость частот SNP-маркеров гена *LZF1* в коренном народонаселении: 6А — изменчивость *rs6441929(G)* в населении Северной Евразии; 6Б — изменчивость *rs6441929(G)* в населении мира; 6В — изменчивость *rs2236938(A)* в населении Северной Евразии; 6Г — изменчивость *rs2236938(A)* в населении мира; 6Д — изменчивость *rs17078408(C)* в населении Северной Евразии; 6Е — изменчивость *rs17078408(C)* в населении мира. Примечания: номера популяций на карте Северной Евразии соответствуют указанным в табл. 1 приложения, для карты мирового народонаселения — в табл. 2 и на рис.1 приложения. Для коренного населения Австралии литературные данные не найдены, и на этот регион распространяется интерполяция от Океании

Ему вторит генетический ландшафт мира (рис. 5Г), практически весь охваченный областью низких частот за исключением Африки южнее Сахары, где частота маркера повышается до 33%.

*rs1860264*. Частота маркера *rs1860264(C)*, напротив, нигде в Северной Евразии (рис. 5Д) не опускается ниже 22% и в целом следует общему тренду «запад–восток», хотя локальные минимумы и максимумы разбросаны в разных регионах. Так, полоса минимальных частот тянется через всю Западную Сибирь к Казахстану, но обнаруживается также в Приуралье и Прибалтике. А повышенные частоты выявлены не только на западе региона (50%), но и в Центральной Азии, Южной Сибири и в Прибайкалье (40–42%).

Карта глобального генетического ландшафта (рис. 5Е) показывает, что Евразия представляет собой постепенный переход от африканского максимума (97%) к американскому минимуму (0%).

*rs6441929*. Закономерность изменчивости *rs6441929(G)* вновь следует тренду «запад–восток» (рис. 6А), но наибольшие значения находятся на востоке с максимумом в Забайкалье (38%) и с минимальными частотами на Урале, Кавказе и в Прибалтике.

Такой ландшафт полностью вписан в мировой (рис. 6Б): деление Евразии на западную с низкими значениями и восточную с более высокими значениями длится на юг вплоть до границы между Индостаном и Юго-Восточной Азией. При этом Евразия в мировом масштабе не

содержит экстремумов — они вновь обнаруживаются в Африке (максимумы) и в Америке (минимумы).

**rs2236938**. Генетический ландшафт *rs2236938(A)* вторит предыдущему (рис. 6В). Лишь западная зона минимумов выражена более ярко и четко, а в восточной зоне высоких частот появилась зона резкого снижения частоты на северо-востоке, тяготеющая к американскому мировому минимуму.

Мировой ландшафт *rs2236938(A)*, тоже вторя предыдущему, более контрастен (рис. 6Г). В Африке частота повышается до 53%, но при этом Аравия, северная и северо-восточная Африка присоединились к западной Евразийской зоне низких частот.

**rs17078408**. Завершает панораму маркер, практически отсутствующий на всем мировом пространстве (рис. 6Д, Е), кроме Африки, где частота *rs17078408(C)* достигает 48%.

## ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Обобщая геногеографические ландшафты всех рассмотренных маркеров гена *LZTFL1*, связанного с риском тяжелого течения COVID-19, укажем на основные закономерности: 1) мировые экстремумы наиболее характерны для коренного населения Африки и Америки и, как правило, альтернативны; 2) Евразия обычно составляет переходную зону между этими двумя экстремумами, но обладает собственными закономерностями и огромным размахом изменчивости в мировом масштабе; 3) генетический ландшафт России органично вписан в Евразийский ландшафт.

Эти основные закономерности выполняются далеко не всегда.

Из первой закономерности есть два исключения. Минимальные частоты встречаемости *rs12493471(A)* сосредоточены и в Африке, и в Америке, и в Восточной Азии, а максимумы — в Европе и Индостане. Аналогично минимумы *rs2191031(A)* встречены и в Африке, и в Америке, вся Евразия занята высокими частотами, но максимум сосредоточен в Океании. Отметим, что коренное население Америки всегда подчинено единой закономерности — низкие частоты всех рассмотренных маркеров снижаются до нуля в Южной Америке.

Из второй закономерности тоже есть исключения. Например, для *rs11130077(G)* мировой максимум характерен не только для Африки, но и для Северной Европы. Евразия является таким же сосредоточением низких частот встречаемости *rs9842595(A)*, как и Северная Америка. А евразийский ландшафт *rs6441929(G)* хотя и промежуточен между африканским и американским экстремумами, но резко противоречит понятию «переходной» зоны: максимальные частоты африканского континента граничат с минимальными частотами в Европе и Индостане, а мировой

минимум Америки граничит с высокими значениями частоты в Восточной Азии. Такая же евразийская «чересполосица» обнаружена и для *rs2236938(A)*. Наиболее характерным для Евразии оказывается деление на «запад» и «восток» по 80-му меридиану, разделяющему Индостан и Юго-Восточную Азию на юге и, постепенно размываясь, доходящему до Северного Ледовитого океана. В эту закономерность не укладываются: *rs1058961(A)* (практически вся Евразия занята высокими частотами); *rs9842595(A)* и *rs17078408(C)* (вся Евразия гомогенна по низким частотам этого маркера); *rs2191031(A)* (частота нарастает к югу, переходя в океанийский мировой максимум); слабо выражена для *rs9842595(A)*. Но для половины маркеров *LZTFL1* эта закономерность прослеживается четко.

Из третьей закономерности также есть исключения. Как правило, они связаны с локальными экстремумами — на севере Западной Сибири, на юге Средней Сибири и в Приуралье. Но в целом они не отменяют, а лишь оттеняют общую встроенность Российского генетического ландшафта в Евразийский.

## ВЫВОДЫ

1. В пространственной изменчивости изученных SNP-маркеров гена *LZTFL1*, ассоциированных с тяжелым течением COVID-19, выявлены закономерности, характерные для коренного народонаселения России и мира.
2. В генетическом пространстве Северной Евразии основной закономерностью является следование географической изменчивости по двум осям: «север–юг» и «запад–восток». Этот тренд хорошо согласуется с антропологическим вектором «европеоидность–монголоидность».
3. Основной вектор глобальной изменчивости четко соответствует антропологическому делению населения мира. Кластеры коренного населения Африки, Азии и Америки включают популяции только «своей» части света. «Индоевропейский» кластер сочетает географию и историю популяций, включая европеоидные популяции и Европы, и Азии.
4. Все четыре кластера коренного населения мира отдалены друг от друга, и только огромное генетическое разнообразие народов России и сопредельных стран является мостом, связующим генофонды трех частей света — Европы, Азии и Америки.
5. Создан картографический атлас пространственной изменчивости 11 SNP-маркеров *LZTFL1* в населении Северной Евразии и мира, выявивший основные закономерности в их генетических ландшафтах: а) мировые экстремумы приходятся на коренное население Африки и Америки; б) Евразия является переходной зоной между этими двумя экстремумами, но обладает собственными закономерностями; в) генетический ландшафт России органично вписан в Евразийский ландшафт.

## Литература

1. The COVID-19 Host Genetics Initiative. Mapping the human genetic architecture of COVID-19. *Nature*. 2021; 600: 472–7. DOI: 10.1038/s41586-021-03767-x.
2. COVID19-hg GWAS meta-analyses round 6. The COVID-19 Host Genetics Initiative. [cited 2022 Sep 13]. Available from: <https://www.covid19hg.org/results/r6/>.
3. Downes DJ, Cross AR, Hua P, Roberts N, Schwessinger R, Cutler AJ, et al. Identification of LZTFL1 as a candidate effector gene at a COVID-19 risk locus. *Nat Genet*. 2021; 53: 1606–15. DOI: 10.1038/s41588-021-00955-3.
4. Воложанин Д. А., Голота А. С., Камилова Т. А., Шнейдер О. В., Щербак С. Г. Генетика COVID-19. Клиническая практика. 2021; 12 (1): 41–52. DOI: 10.17816/clinpract64972.
5. Bardet-Biedl syndrome-17; BBS17. Online Mendelian Inheritance in Man — OMIM. [cited 2021 Nov 25]. Available from: <https://omim.org/entry/615994>.

6. Waters AM, Beales PL. Ciliopathies: an expanding disease spectrum. *Pediatr Nephrol.* 2011; 26: 1039–56. DOI: 10.1007/s00467-010-1731-7.
7. Потрохова Е. А., Бабаян М. Л., Балева Л. С., Сафонова М. П., Сипягина А. Е. Синдром Барде–Бидля. *Российский вестник перинатологии и педиатрии.* 2020; 65 (6): 76–83. DOI: 10.21508/1027-4065-2020-65-6-76-83.
8. Seo S, Zhang Q, Bugge K, Breslow DK, Searby CC, Nachury MV, et al. A novel protein LZTFL1 regulates ciliary trafficking of the BBSome and Smoothened. *PLoS Genet.* 2011; 7 (11): e1002358. DOI: 10.1371/journal.pgen.1002358.
9. GeneCards: The Human Gene Database [Internet]. Rehovot, Israel: Weizmann Institute of Science. c1996-2022 — LZTFL1 Gene — Leucine Zipper Transcription Factor Like 1; [cited 2022 Sep 12]. Available from: <https://www.genecards.org/cgi-bin/carddisp.pl?gene=LZTFL1>.
10. Klink BU, Gatsogiannis C, Hofnagel O, Wittinghofer A, Raunser S. Structure of the human BBSome core complex. *eLife.* 2020; 9: e53910. DOI: 10.7554/eLife.53910.
11. Nakanishi T, Pigazzini S, Degenhardt F, Cordioli M, Butler-Laporte G, Maya-Miles D, et al. Age-dependent impact of the major common genetic risk factor for COVID-19 on severity and mortality. *J Clin Invest.* 2021; 131 (23): e152386. DOI: 10.1172/JCI152386.
12. Ellinghaus D, Degenhardt F, Bujanda L, Buti M, Albillos A, Invernizzi P, et al. Genomewide Association Study of Severe Covid-19 with Respiratory Failure. *Engl J Med.* 2020; 383 (16): 1522–34. DOI: 10.1056/NEJMoa2020283.
13. Pairo-Castineira E, Clohisey S, Klaric L, Bretherick AD, Rawlik K, Pasko D, et al. Genetic mechanisms of critical illness in COVID-19. *Nature.* 2021; 591 (7848): 92–8. DOI: 10.1038/s41586-020-03065-y.
14. Balanovsky O, Petrusenko V, Mirzaev K, Abdullaev S, Gorin I, Chernevskiy D, et al. The variation of genome sites associated with severe COVID-19 across populations the worldwide and national patterns. *Pharmgenomics Pers Med.* 2021; 14: 1391–402. DOI: 10.2147/PGPM.S320609.
15. Secolin R, de Araujo TK, Gonsales MC, Rocha CS, Naslavsky M, Marco L, et al. Genetic variability in COVID-19-related genes in the Brazilian population. *Hum Genome Var.* 2021; 8 (15). DOI: 10.1038/s41439-021-00146-w.
16. Балановская Е. В., Жабалин М. К., Агджоян А. Т., Чухряева М. И., Маркина Н. В., Балаганская О. А. и др. Популяционные биобанки: принципы организации и перспективы применения в геногеографии и персонализированной медицине. *Генетика.* 2016; 52 (12): 1371–87. DOI: 10.7868/S001667581612002X.
17. Горин И. О., Петрушенко В. С., Записецкая Ю. С., Кошель С. М., Балановский О. П. 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
18. GG-base [cited 2022 Sep 10]. Available from: <https://gg-base.org/>.

## References

1. The COVID-19 Host Genetics Initiative. Mapping the human genetic architecture of COVID-19. *Nature.* 2021; 600: 472–7. DOI: 10.1038/s41586-021-03767-x.
2. COVID19-hg GWAS meta-analyses round 6. The COVID-19 Host Genetics Initiative. [cited 2022 Sep 13]. Available from: <https://www.covid19hg.org/results/r6/>.
3. Downes DJ, Cross AR, Hua P, Roberts N, Schwessinger R, Cutler AJ, et al. Identification of LZTFL1 as a candidate effector gene at a COVID-19 risk locus. *Nat Genet.* 2021; 53: 1606–15. DOI: 10.1038/s41588-021-00955-3.
4. Vologzhanin DA, Golota AS, Kamilova TA, Shneider OV, Sherbak SG. Genetics of COVID-19. *Journal of Clinical Practice.* 2021; 12 (1): 41–52. DOI: 10.17816/clinpract64972. Russian.
5. Bardet-Biedl syndrome-17; BBS17. Online Mendelian Inheritance in Man — OMIM. [cited 2021 Nov 25]. Available from: <https://omim.org/entry/615994>.
6. Waters AM, Beales PL. Ciliopathies: an expanding disease spectrum. *Pediatr Nephrol.* 2011; 26: 1039–56. DOI: 10.1007/s00467-010-1731-7.
7. Potrokhova EA, Babayan ML, Baleva LS, Safonova MP, Sipyagina AE. Bardet–Biedl Syndrome. *Rossiyskiy Vestnik Perinatologii i Pediatrii (Russian Bulletin of Perinatology and Pediatrics).* 2020; 65 (6): 76–83. DOI: 10.21508/1027-4065-2020-65-6-76-83. Russian.
8. Seo S, Zhang Q, Bugge K, Breslow DK, Searby CC, Nachury MV, et al. A novel protein LZTFL1 regulates ciliary trafficking of the BBSome and Smoothened. *PLoS Genet.* 2011; 7 (11): e1002358. DOI: 10.1371/journal.pgen.1002358.
9. GeneCards: The Human Gene Database [Internet]. Rehovot, Israel: Weizmann Institute of Science. c1996-2022 — LZTFL1 Gene — Leucine Zipper Transcription Factor Like 1; [cited 2022 Sep 12]. Available from: <https://www.genecards.org/cgi-bin/carddisp.pl?gene=LZTFL1>.
10. Klink BU, Gatsogiannis C, Hofnagel O, Wittinghofer A, Raunser S. Structure of the human BBSome core complex. *eLife.* 2020; 9: e53910. DOI: 10.7554/eLife.53910.
11. Nakanishi T, Pigazzini S, Degenhardt F, Cordioli M, Butler-Laporte G, Maya-Miles D, et al. Age-dependent impact of the major common genetic risk factor for COVID-19 on severity and mortality. *J Clin Invest.* 2021; 131 (23): e152386. DOI: 10.1172/JCI152386.
12. Ellinghaus D, Degenhardt F, Bujanda L, Buti M, Albillos A, Invernizzi P, et al. Genomewide Association Study of Severe Covid-19 with Respiratory Failure. *Engl J Med.* 2020; 383 (16): 1522–34. DOI: 10.1056/NEJMoa2020283.
13. Pairo-Castineira E, Clohisey S, Klaric L, Bretherick AD, Rawlik K, Pasko D, et al. Genetic mechanisms of critical illness in COVID-19. *Nature.* 2021; 591 (7848): 92–8. DOI: 10.1038/s41586-020-03065-y.
14. Balanovsky O, Petrusenko V, Mirzaev K, Abdullaev S, Gorin I, Chernevskiy D, et al. The variation of genome sites associated with severe COVID-19 across populations the worldwide and national patterns. *Pharmgenomics Pers Med.* 2021; 14: 1391–402. DOI: 10.2147/PGPM.S320609.
15. Secolin R, de Araujo TK, Gonsales MC, Rocha CS, Naslavsky M, Marco L, et al. Genetic variability in COVID-19-related genes in the Brazilian population. *Hum Genome Var.* 2021; 8 (15). DOI: 10.1038/s41439-021-00146-w.
16. Balanovska EV, Zhabagin MK, Agdjoyan AT, Chuhryaeva MI, Markina NV, Balaganskaya OA, et al. Population biobanks: Organizational models and prospects of application in gene geography and personalized medicine. *Russ J Genet.* 2016; 52 (12): 1227–43. DOI: 10.1134/S1022795416120024.
17. Gorin IO, Petrusenko VS, Zapisetskaya YS, Koshel SM, Balanovsky OP. 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. DOI: 10.15829/1728-8800-2020-2732. Russian.
18. GG-base [cited 2022 Sep 10]. Available from: <https://gg-base.org/>.