

Пространственная изменчивость гена *LZTFL1*, связанного с тяжелым COVID-19, в населении России и мира

Изучена пространственная изменчивость 11 SNP-маркеров гена *LZTFL1*, ассоциированного с двукратным увеличением риска тяжелого течения COVID-19. В анализ включены 28 метапопуляций (97 этносов) Северной Евразии и 34 метапопуляции мира. Создан картографический атлас пространственной изменчивости 11 SNP-маркеров *LZTFL1* в популяциях. В Северной Евразии основные географические векторы изменчивости «север–юг» и «запад–восток» в целом согласуются с антропологическим вектором «европеоидность–монголоидность». Глобальная изменчивость маркеров гена согласуется с распределением населения по частям света и их антропологическим составом. Огромное генетическое разнообразие народов России и сопредельных стран служит мостом, связующим три части света: Европу, Азию и Америку.

Тяжесть заболевания COVID-19 и смертность от него различаются в разных популяциях, что говорит о значимом влиянии генетических факторов. Сбор информации о генетических вариантах, ассоциированных с тяжелым течением болезни, происходит в рамках международного проекта *COVID-19 host genetics initiative* ([COVID-19 HGI](#)). Исследование частоты этих вариантов в разных популяциях мира – задача популяционных генетиков. Коллективом под руководством специалистов Медико-генетического научного центра ранее уже было проведено геногеографическое исследование двух генетических маркеров, связанных с предрасположенностью к тяжелому течению заболевания COVID-19 в населении России и мира (об этой работе [можно прочитать на сайте](#)). В новом исследовании ученые сосредоточились на анализе распространения в популяциях России и мира 11 вариантов гена *LZTFL1*, связанного с двукратным увеличением риска тяжелого течения COVID-19. Статья с результатами этой работы [опубликована в журнале «Вестник РГМУ»](#), руководитель работы — д.б.н., профессор Е.В.Балановская.

Ген *LZTFL1*, как было показано в исследованиях по проекту COVID-19 HG, связан с риском тяжелого протекания заболевания COVID-19. Этот ген экспрессируется в клетках легких человека и кодирует белок, участвующий в транспорте других белков к первичным ресничкам. Таким образом, ген *LZTFL1* критически важен для работы клеток мерцательного эпителия; с ним ассоциированы некоторые нарушения в этих клетках, происходящие у пациентов с COVID-19. Изменения в области хромосомы 3p21.31, включающей ген *LZTFL1*, приводят к двукратному увеличению риска дыхательной недостаточности и более чем двукратному увеличению риска смерти для людей моложе 60 лет.

Целью работы было исследование пространственной изменчивости SNP-маркеров гена *LZTFL1* в коренном населении Северной Евразии и всего мира и создание картографического атласа этой изменчивости.

Исходным материалом послужили данные о коренном населении Северной Евразии, включившие популяции 97 этносов не только России, но и большинства постсоветских стран и Монголии. Образцы ДНК были предоставлены АНО «Биобанк Северной Евразии». Образцы из географически и исторически близких популяций, но с малыми выборками, объединялись в метапопуляции. Итоговый массив данных по Северной Евразии включил 28 метапопуляций ($n = 1980$ индивидуальных образцов) со средним размером выборки 140 хромосом. Массив данных о коренном населении других регионов мира ($n = 1657$) взят из базы «GG-base». Итоговый набор данных о коренном населении мира ($n = 3637$ индивидуальных образцов) включил 34 метапопуляции, в которые были объединены географически и исторически близкие группы населения.

При биоинформатическом анализе обоих массивов данных — коренного населения Северной Евразии и мирового народонаселения — выявлено 10 SNP-полиморфизмов гена *LZTFL1*, репрезентативно представленных в обоих регионах. Из них два SNP-маркера указаны как достоверно ассоциированные с COVID-19: *rs1058961* (3'-нетранслируемая область) и *rs12493471* (2-й интрон). Остальные восемь SNP-маркеров исследовали впервые: *rs11130077* (3-й интрон), *rs17078408*, *rs1860264*, *rs2191031*, *rs2236938*, *rs6441929*, *rs7614952* и *rs9842595* (2-й интрон). Об одиннадцатом маркере — *rs17713054*, ассоциированном с тяжелым течением заболевания COVID-19, — в базе данных «GG-base» отсутствует информация о его изменчивости в популяциях мира, поэтому его анализ проведен только для Северной Евразии. Положение метапопуляций в генетическом пространстве оценивалось двумя независимыми методами — главных компонент и многомерного шкалирования – выявившими сходные закономерности. Для большей сопоставимости результатов анализ метапопуляций Северной Евразии проведен по данным о тех же 10 SNP-маркерах гена *LZTFL1*, что и при анализе населения мира.

В генетическом пространстве метапопуляции Северной Евразии распределились строго по осям «север–юг» и «запад–восток»: «северный» кластер, объединивший северные популяции Западной Сибири (ненцев, манси, хантов) и Дальнего Востока (ительменов, коряков, чукчей), противостоит трем остальным кластерам, расположившимся строго по оси «запад–восток». Среди них «западный» кластер включил восточных славян (белорусов, русских, украинцев) и финноязычные народы России (бесермян, карел коми-пермяков, мордву, удмуртов). «Урало-кавказский» кластер объединил все народы Кавказа, Закавказья,

Таджикистана, тюрок Урала (башкир, поволжских татар, чувашей), а также финноязычных марийцев, поскольку они включены в одну метапопуляцию с чувашами. В «мезокластер» — переходный от «западного» кластера к «восточному» — вошли 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 (бурятов, калмыков, народы Монголии), тувинцев (до середины XX века входивших в состав Монголии) и малочисленные народы Дальнего Востока (нанайцев, нивхов, ульчей, эвенов). В целом, общий тренд изменчивости SNP-маркеров *LZTFL1* в Северной Евразии согласуется не только с географическим вектором «запад–восток», но и с антропологическим вектором «европеоидность–монголоидность» (рис. 1).



Рис. 1. Положение популяций России и сопредельных стран в генетическом пространстве многомерного шкалирования.

В генетическом пространстве метапопуляций мира выявлено пять основных кластеров (рис. 2). Бросается в глаза огромная изменчивость — даже в мировом масштабе метапопуляций России и сопредельных стран — они не вместились ни в один из мировых кластеров. «Западный» кластер России (выявленный на рис. 1) 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 антропологию населения. Три кластера включают популяции только «своей» части света: Африки, Азии, Америки. Но «индоевропейский» кластер, в соответствии с географией и историей популяций, включает европеоидные популяции не только Европы, но и Азии (Индии, Пакистана, Афганистана, Ближнего Востока).

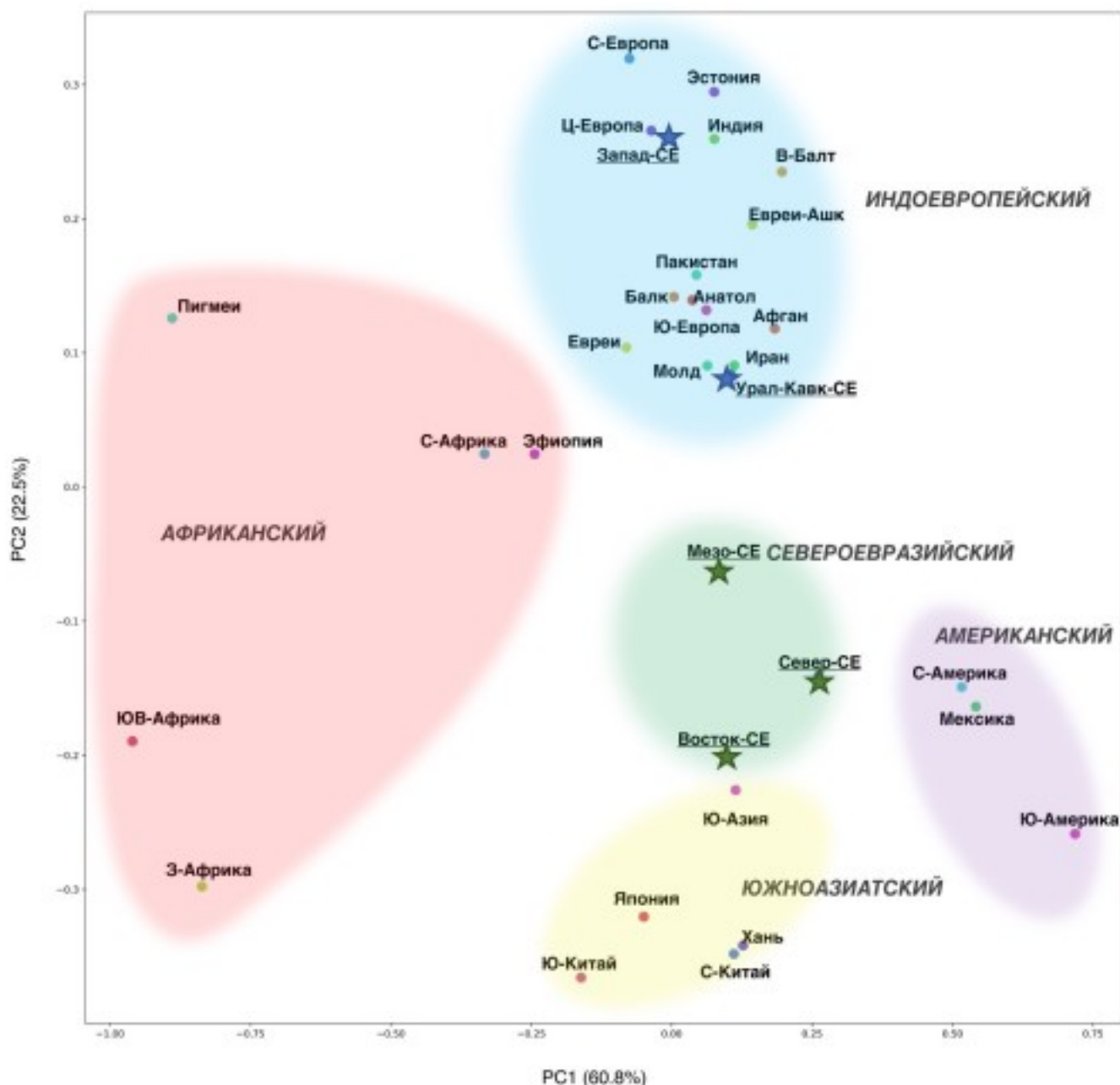


Рис.2. Расположение популяций коренного населения мира в генетическом пространстве главных компонент изменчивости 10 SNP-маркеров гена LZTFL1.

Для каждого SNP-маркера *LZTFL1* (кроме *rs17713054*, для которого нет информации в мировой базе данных) авторы создали по две карты его изменчивости: в коренном населении Северной Евразии и мира.

***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 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.

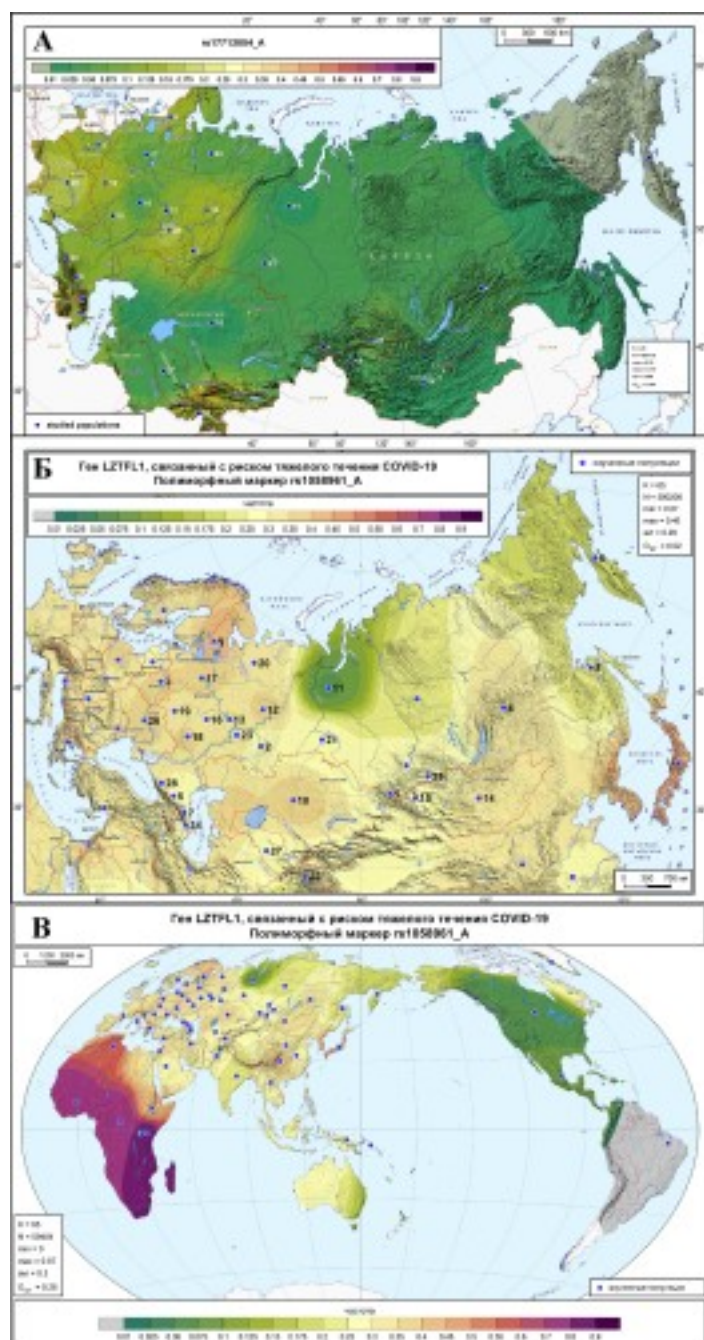


Рис. 3. Геногеографическая изменчивость частот SNP-маркеров гена LZTFL1 в коренном народонаселении: 3А – изменчивость rs17713054(A) в населении Северной Евразии; 3Б – изменчивость rs1058961(A) в населении Северной Евразии; 3В – изменчивость rs1058961(A) в населении мира.

rs12493471. Та же клинальная изменчивость «запад–восток» в Северной Евразии обнаружена и для rs12493471(A) (рис. 4А), сцепленного с rs17713054(A) (рис. 3А). 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 Финноскандии 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 частоты идет во всех направлениях с локальными максимумами (40%) в Индостане и в Океании.

rs11130077. В Северной Евразии клинальная изменчивость «запад-восток» характерна и для rs11130077(G) (рис. 4В). Незначительные отличия связаны со смещением максимумов в Финноскандию (54%). Однако в азиатской части картина не столь однозначна. При снижении частоты в Сибири до 16–17% на Дальнем Востоке намечается некоторое повышение частоты

до 20%, а в Южной Сибири — до 27%. Изменчивость *rs11130077(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) видно, что байкальский максимум является частью повышенных частот во всем регионе Восточной Азии. Вновь мировой максимум находится в Африке, а мировой минимум — в Америке.

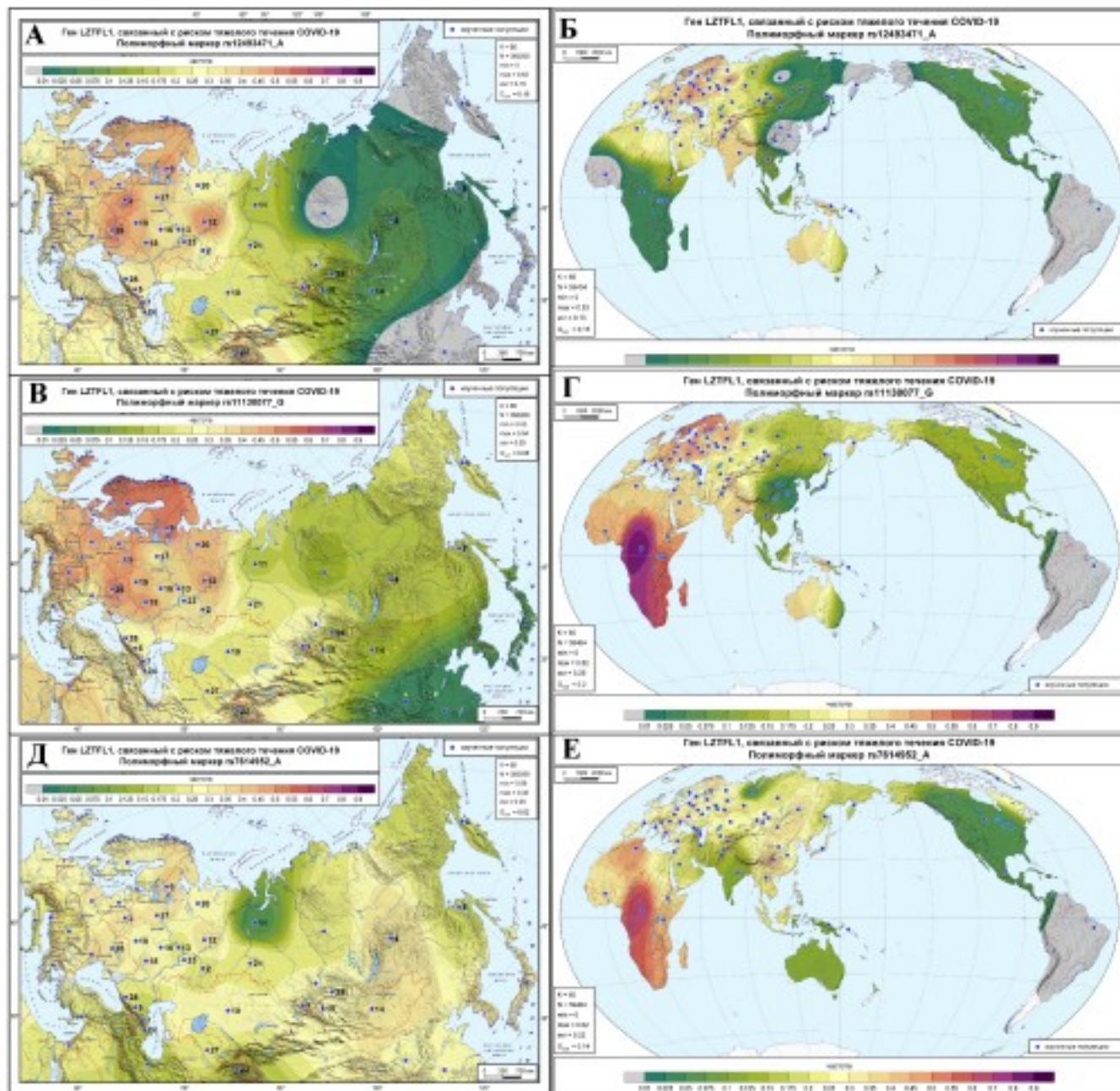


Рис. 4. Геногеографическая изменчивость частот SNP-маркеров гена LZTFL1 в коренном народонаселении: 4А – изменчивость *rs12493471(A)* в населении Северной Евразии; 4Б – изменчивость *rs12493471(A)* в населении мира; 4В – изменчивость *rs11130077(G)* в населении Северной Евразии; 4Г – изменчивость *rs11130077(G)* в населении мира; 4Д – изменчивость *rs7614952(A)* в населении Северной Евразии; 4Е – изменчивость *rs7614952(A)* в населении мира.

***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).

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

rs1860264. Частота маркера *rs1860264*(C) 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%).

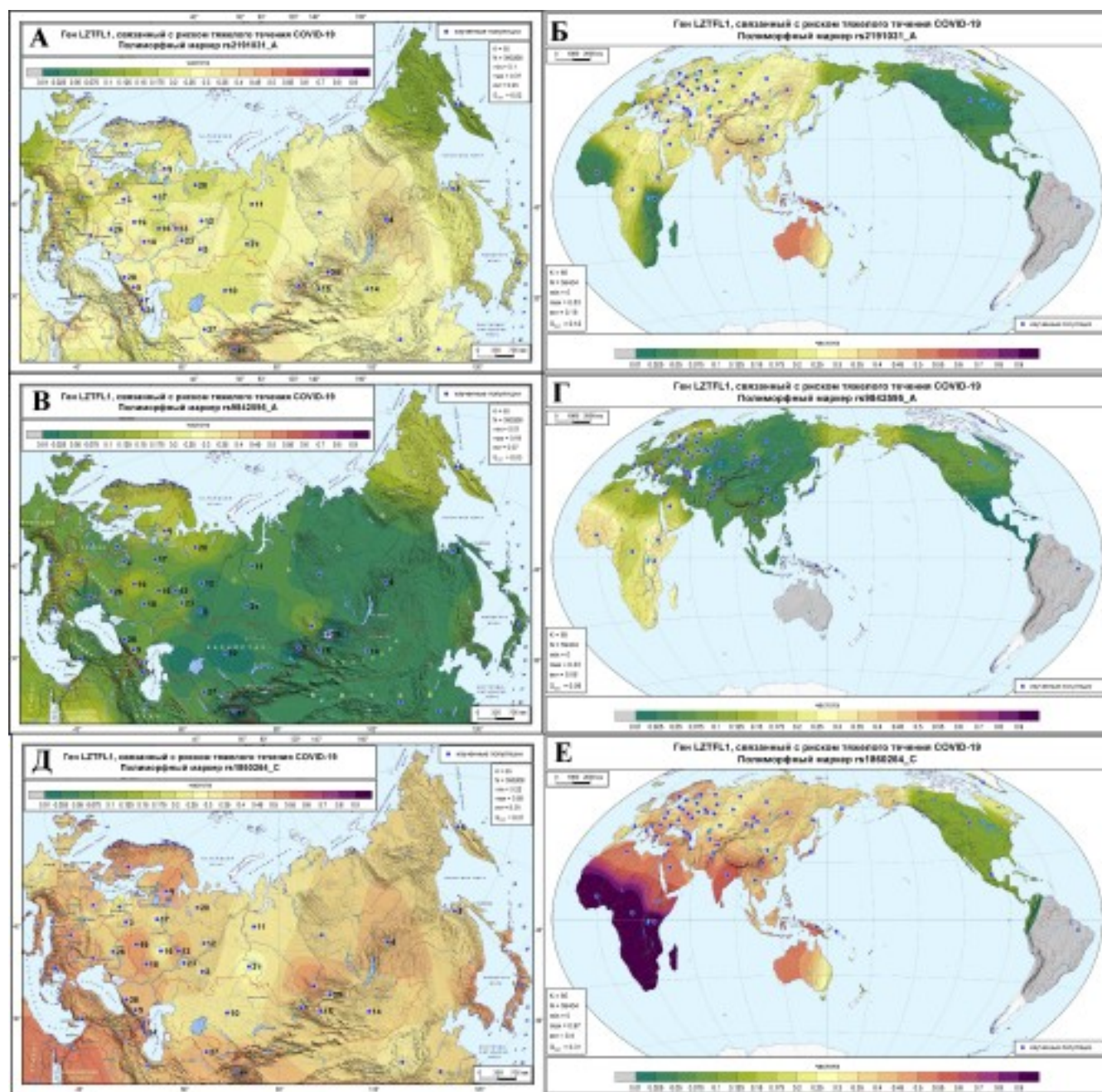


Рис.. 5. Геногеографическая изменчивость частот SNP-маркеров гена LZTFL1 в коренном народонаселении: 5А – изменчивость *rs2191031*(A) в населении Северной Евразии; 5Б – изменчивость *rs2191031*(A) в населении мира; 5В – изменчивость *rs9842595*(A) в населении Северной Евразии; 5Г – изменчивость *rs9842595*(A) в населении мира; 5Д – изменчивость *rs1860264*(C) в населении Северной Евразии; 5Е – изменчивость *rs1860264*(C) в населении мира.

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)* 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%.

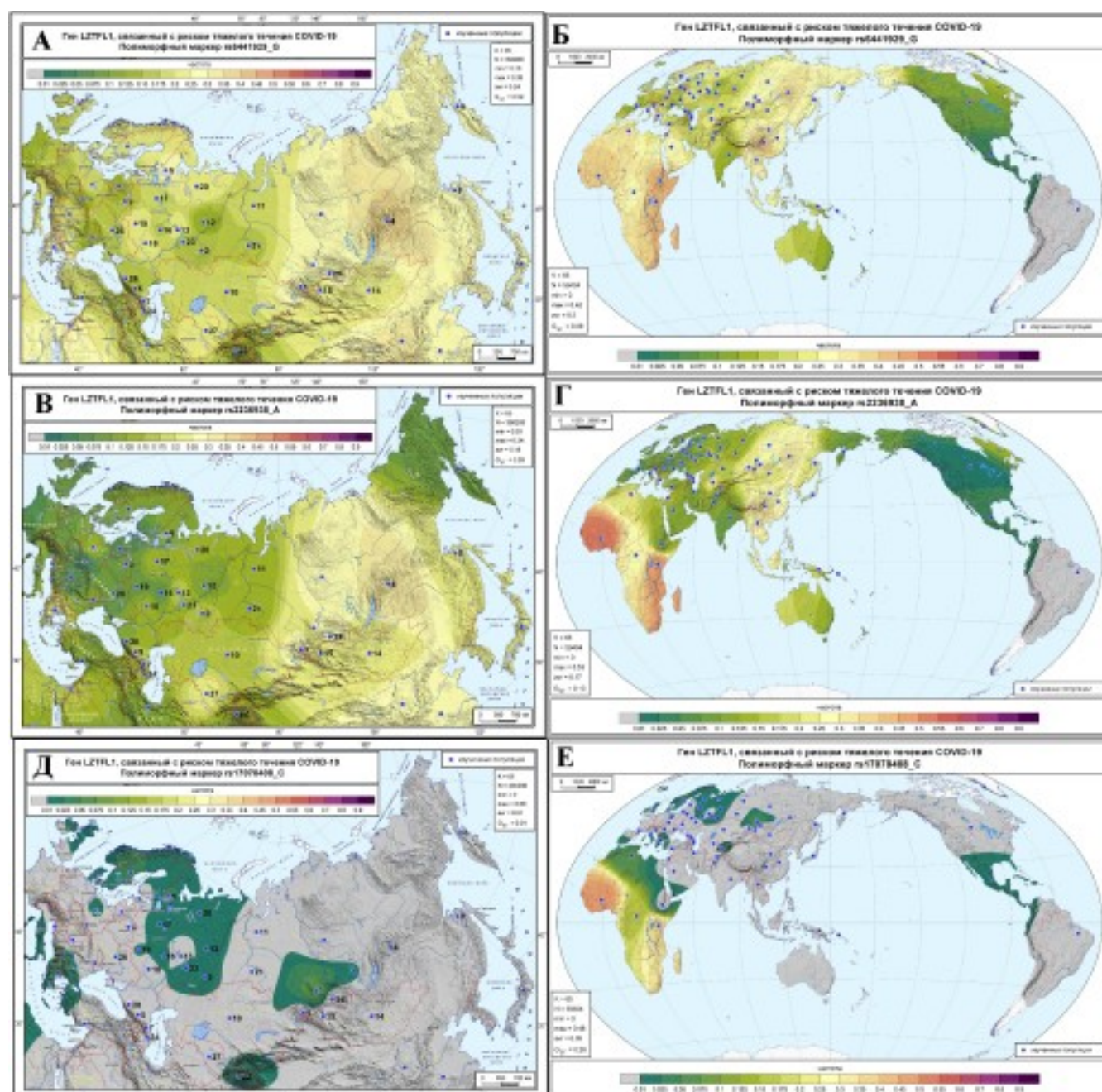


Рис. 6. Геногеографическая изменчивость частот SNP-маркеров гена LZTFL1 в коренном народонаселении: 6А – изменчивость *rs6441929(G)* в населении Северной Евразии; 6Б – изменчивость *rs6441929(G)* в населении мира; 6В – изменчивость *rs2236938(A)* в населении Северной Евразии; 6Г – изменчивость *rs2236938(A)* в населении мира; 6Д –

изменчивость rs17078408(C) в населении Северной Евразии; 6E – изменчивость rs17078408(C) в населении мира.

В геногеографических ландшафтах изученных маркеров гена *LZTFL1* авторы отмечают основные закономерности и по результатам исследования приходят к нескольким выводам.

- 1) Основной вектор глобальной изменчивости маркеров гена *LZTFL1* четко соответствует антропологическому делению населения мира. Мировые экстремумы частот маркеров наиболее характерны для коренного населения Африки и Америки и, как правило, альтернативны.
- 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) Генетический ландшафт России органично вписан в евразийский ландшафт. Огромное генетическое разнообразие народов России и сопредельных стран является мостом, связующим генофонды трех частей света — Европы, Азии и Америки.

Источник:

Балановская Е.В., Горин И.О., Петрушенко В.С., Черневский Д.К., М. Кошель С.М., Темирбулатов И.И., Пылёв В. Ю., Агджоян А.Т. Геногеография в России и мире SNP-маркеров гена LZTFL1, ассоциированных с тяжелым течением COVID-19 // Вестник РГМУ, № 5. DOI: 10.24075/vrgmu.2022.047

<https://vestnik.rsmu.press/archive/2022/5/4/abstract?lang=ru>