Распространение генетической предрасположенности к тяжелому COVID-19 в населении мира

Российские исследователи изучили частоту генетических вариантов, которые ассоциированы с тяжелым течением COVID-19, в популяциях мира и отдельно – в популяциях России и сопредельных стран. Один вариант оказался приурочен к популяциям Западной Евразии, второй встречается в повсеместно в Евразии и с особенно высокой частотой – в Африке. В масштабах России обнаружена корреляция между частотой одного из вариантов и долей смертей и выздоровлений среди заболевших COVID-19. В масштабах мира эти корреляции не выявлены, что может быть связано с различиями медицинской статистики между странами.

Как показывают многочисленные клинические данные, тяжесть заболевания COVID-19, вызванного вирусом SARS-CoV-2, значительно различается у разных пациентов – от бессимптомного носительства до летального исхода. Отягощающими факторами являются возраст и сопутствующие хронические болезни, однако они не могут объяснить всей разницы. Эпидемиологические наблюдения указывают на то, что тяжесть заболевания и смертность различаются между странами. Помимо социоэкономических и климатических факторов причины могут лежать и в области генетики.

В ходе ранее проведенного исследования геномных ассоциаций (GWAS) на пациентах Италии и Испании в геноме было найдено два региона, связанных с более тяжелым течением COVID-19. Один регион *rs11385942* на 3-й хромосоме, как было показано в работе Сванте Паабо, появился в геноме современного человека в результате его метисации с неандертальцами. Другой регион *rs657152* на 9-й хромосоме связан с группой крови по системе AB0. Но до сих пор не было известно, как распределены эти генетические варианты в разных регионах и популяциях мира.

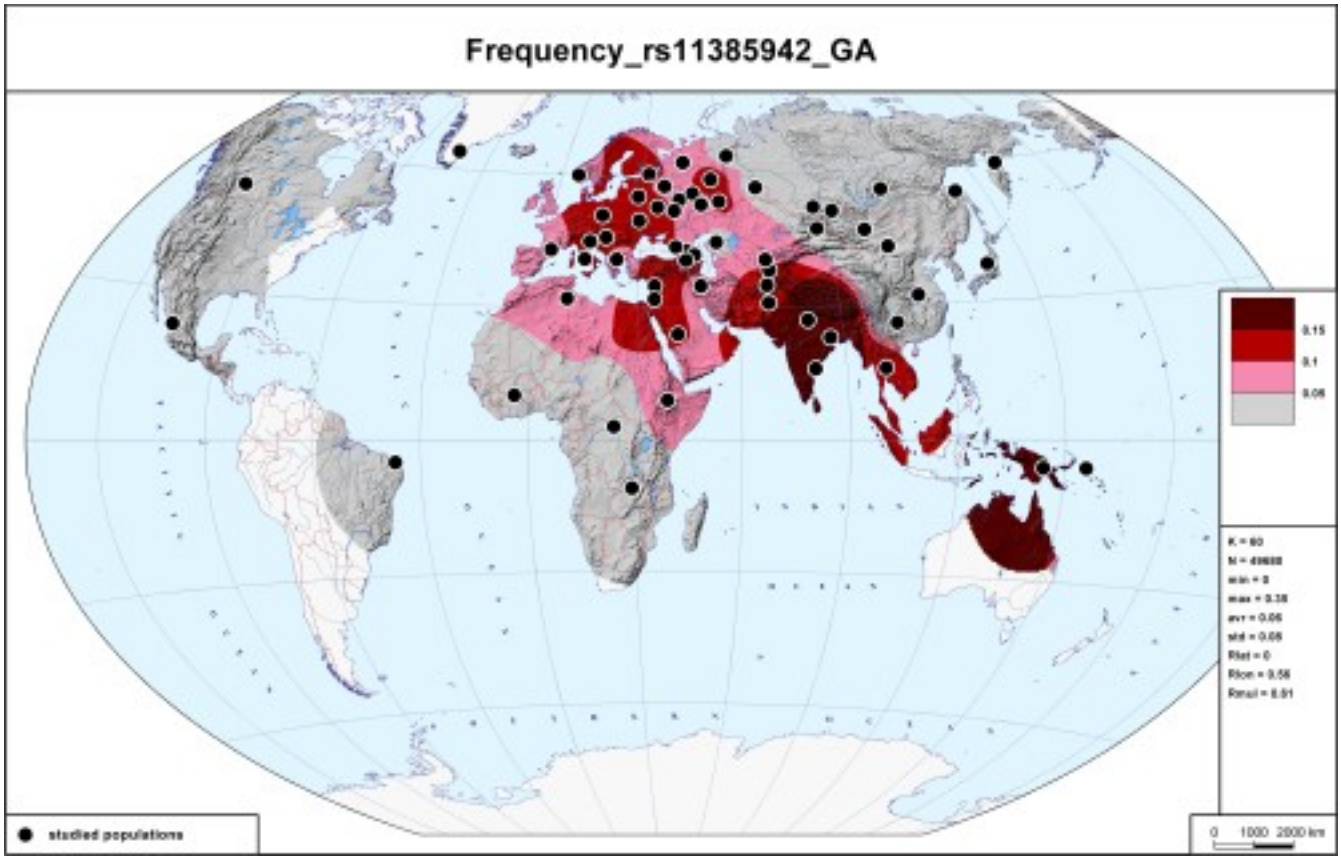
Это и стало основной целью исследования российских популяционных генетиков и медиков, которые опубликовали статью в журнале Pharmacogenomics and Personalized Medicine . Основные авторы представляют Институт общей генетики РАН, Медико-генетический научный центр и Российскую медицинскую академию непрерывного профессионального образования, исследование было поддержано грантом РНФ (проект 21-14-00363).

В работе было поставлено две цели. Первая состояла в определении частот вариантов *rs11385942* и *rs657152* в популяциях на глобальном уровне и в пределах России. Вторая цель – определить, есть ли корреляция между частотой этих вариантов в популяциях и числом выздоровлений и смертей от COVID-19.

Образцы ДНК популяций России и сопредельных стран Северной Евразии (страны бывшего СССР и Монголия) были получены из коллекций Биобанка Северная Евразия, они включали большинство этнических групп этого региона. В общей сложности авторы генотипировали 1883 образца из 91 популяций из Армении, Азербайджана, Белоруссии, Грузии, Казахстана, Киргизии, Литвы, Монголии, России, Таджикистана, Украины и Узбекистана. Для данного исследования отобрали генотипы по маркеру *rs657152* и данные по SNP, позволяющие однозначно определить генотип *rs11385942*. Не все популяции были представлены достаточным числом образцов, и в этом случае географически и генетически близкие популяции объединяли в группы. Первая полученная база данных состояла из 28 объединенных популяций, каждая группа включала в среднем 128 хромосом (64 индивида), она носила условное название «российской» базы данных.

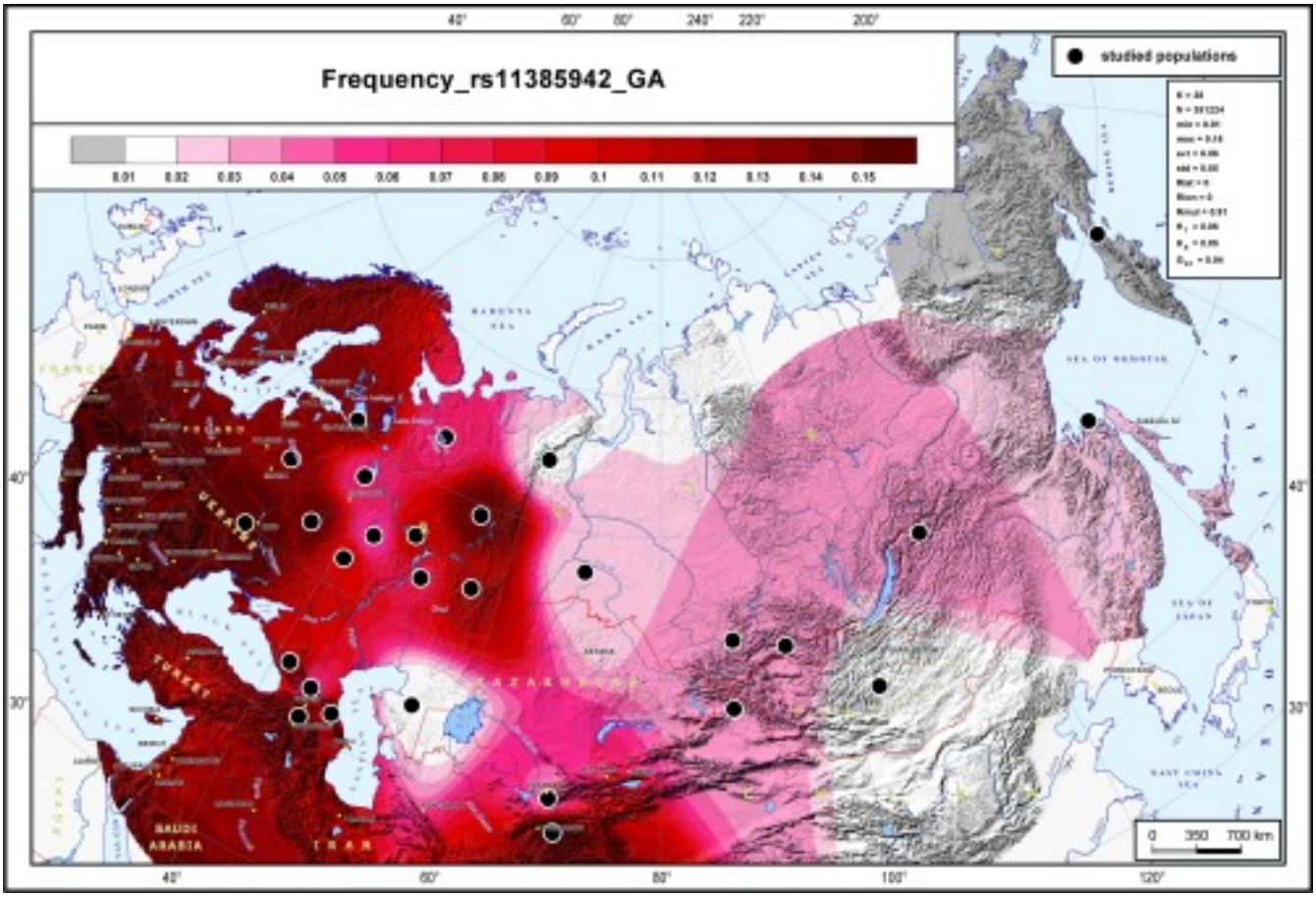
В дополнение к ней авторы использовали глобальную базу данных по разным регионам, собранную из 16 опубликованных статей, в которых выборки из популяций были генотипированы по панели Illumina. Эта база включает 3088 образцов, генотипированных по маркеру *rs657152* и SNP маркера *rs11385942*. Так же, как и в случае «российской» базы данных, популяции с недостаточным числом образцов объединили в группы. В итоге, в нее вошли 32 объединенных популяции, в каждую было включено в среднем 104 хромосомы, эту базу данных назвали «мировой».

Итоговый анализ был проведен на объединенной базе данных по 60 популяциям. Максимальная частота *rs11385942* найдена в популяциях Южной Азии (20-30%), за ними по частоте идут популяции Ближнего Востока и Европы (5-15%). Этот маркер очень редок или вообще отсутствует в Восточной Азии, в Сибири, у американских индейцев и в африканских популяциях южнее Сахары, хотя он есть в Северной Африке. Данные по Юго-Восточной Азии и Папуа Новая Гвинея недостаточны, но показывают, что он может там достигать значительной частоты. В целом этот вариант может быть охарактеризован как западноевразийский, поскольку известно, что население Южной Азии генетически ближе к Западной Евразии, чем к Восточной. Отсутствие этого аллеля в Африке южнее Сахары согласуется с тем, что он был получен от метисации с неандертальцами, в то время как его отсутствие в Восточной Азии можно объяснить генетическим дрейфом после разделения западноевразийской и восточноевразийской ветвей.



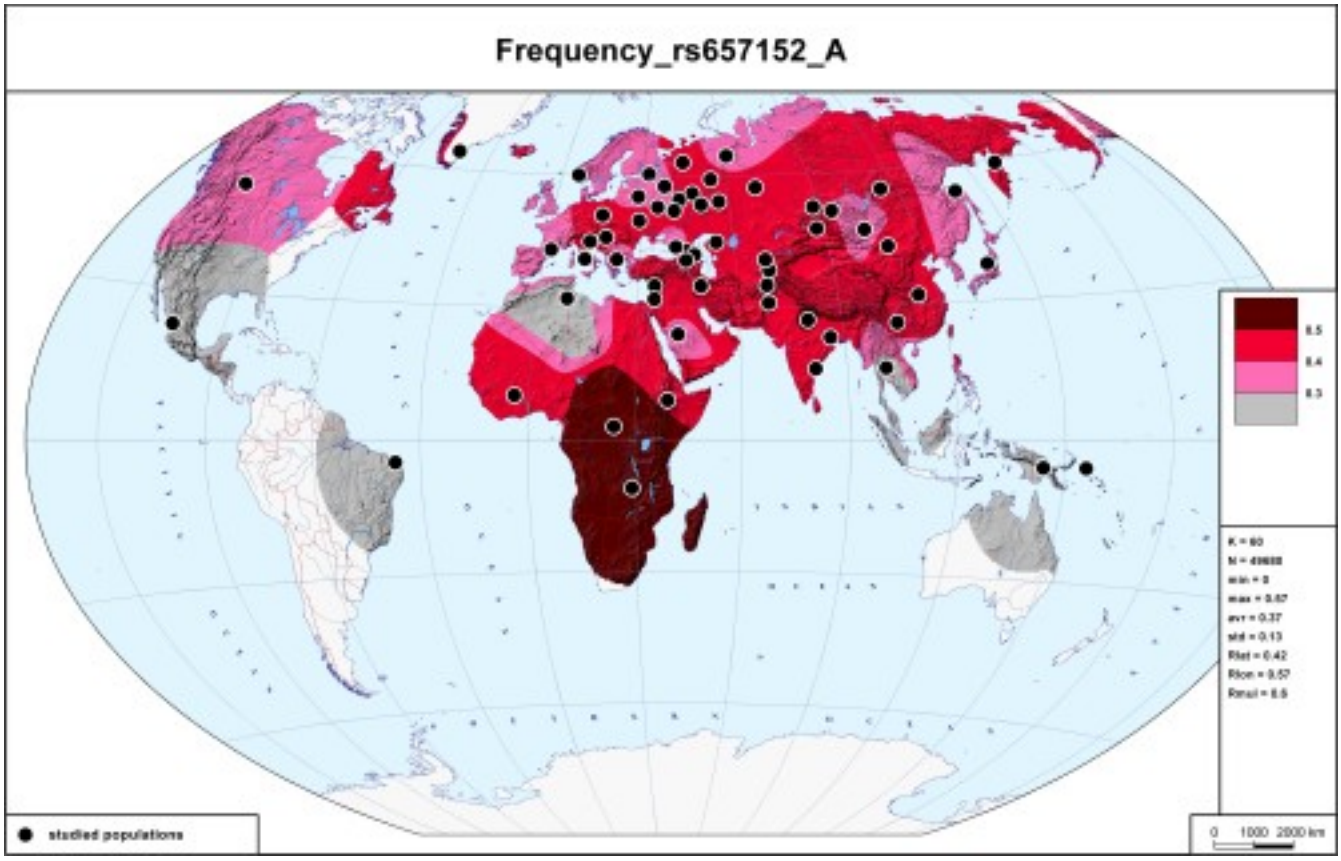
Частота аллеля rs11385942_GA в популяциях мира (более высокая частота соответствует более темному цвету).

Более детальный анализ распределения маркера *rs11385942* авторы провели в популяциях России и сопредельных стран. Карта показывает, что различия в его частоте между Европой и Сибирью не скачкообразны, а, скорее, отражают клинальную изменчивость. Происходит постепенное изменение частоты маркера от 13-16% в Украине, Белоруссии и более западных русских популяциях до 0% в коренном населении Камчатки и Чукотки. В среднем в популяциях Европейской России его частота составляет 11%, в Сибири 3%. В странах Центральной Азии его частота низка (1-4%) за исключением Таджикистана, где она составляет 14%. Но надо отметить, что население Таджикистана географически и генетически близко к популяциям Южной Азии, у которых при анализе в мировом масштабе были обнаружены максимальные частоты.



Частота аллеля rs11385942_GA в популяциях России и сопредельных стран (более высокая частота соответствует более темному цвету).

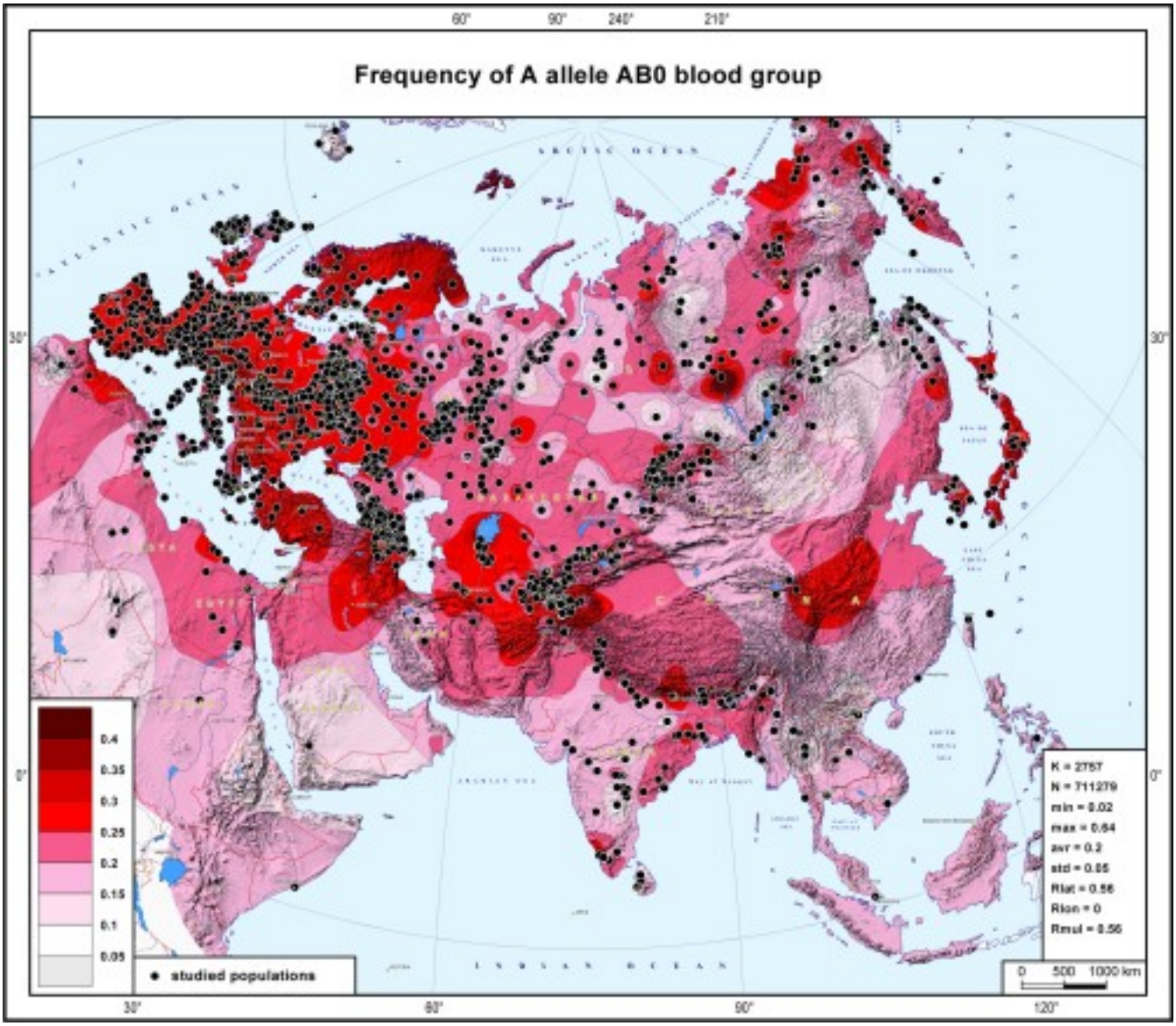
Частота другого маркера — *rs657152* - более равномерно распределена по миру и в целом довольно высока в странах Старого Света. Максимальная его частота (свыше 50%) отмечается в Африке южнее Сахары, в то время как в большинстве популяций Евразии частота составляет от 40 до 50%. На периферии Евразии (Атлантическое побережье Европы, Дальний Восток, Юго-Восточная Азия) частота падает ниже 40%, а у американских индейцев и австралийских аборигенов этот маркер почти отсутствует.



Частота аллеля rs657152_A в популяциях мира (более высокая частота соответствует более темному цвету).

Маркер *rs657152* расположен в гене групп крови AB0, но сам по себе не определяет ни одну из основных групп крови. Но при этом ранее было показано, что люди с группой крови A имеют более высокий риск дыхательной недостаточности при COVID-19, чем люди с группой 0.

Поэтому дополнительно с картированием частоты *rs657152* авторы проанализировали карту распределения «рисковой» группы крови A в 2757 популяциях. Эта карта была построена несколько лет назад по «классическим» генетическим маркерам, то есть на данных, полученных не прямым анализом ДНК, а иммунологическим анализом антигенов. Хотя данные по классическим маркерам сейчас воспринимаются как устаревшие, за счет их обилия карта на порядок более точна и подробна, чем аналогичные карты по первым двум маркерам, построенные по широкогеномным массивам данных. Эта карта демонстрирует, что высокие частоты (30% и выше) группы крови A отмечаются в Западной Европе, несколько ниже в Волго-Уральском регионе Восточной Европы и еще ниже (10-20%) в Сибири, Восточной Азии и Африке южнее Сахары. В большинстве популяций Южной Азии частота группы A средняя (15-20%).



Частота аллеля группы крови A(II) в популяциях.

Корреляция между частотой аллелей и показателями COVID-19

Авторы сравнили частоту изученных маркеров в популяциях России и мира с официальной статистикой по COVID-19: количества случаев заболевания, выздоровлений и смертей, нормализованные к численности населения, а также вычисляли показатель смертности как отношение числа смертей к числу случаев COVID-19, для которых известен исход (выздоровление или летальный). Положительная корреляция обнаружена между частотой аллеля *rs657152* в популяциях России и смертностью (0,63, p=0,01)**,** для частоты аллеля *rs11385942* корреляция была положительной, но не достигала уровня достоверности. По объединенной базе данных (включая Россию и мир) достоверных корреляций частоты аллелей ни с каким из эпидемиологических показателей не обнаружено.

Как было показано, один аллель риска (*rs11385942_GA)* с высокой частотой встречается в Южной Азии, с несколько более низкой — в Европе и на Ближнем Востоке и с низкой — в других регионах мира. Для другого аллеля риска (*rs657152_A)* наблюдается высокая частота в Африке и средняя в Евразии. Наконец, группа крови A (II) наиболее часта в Европе и имеет среднюю частоту в остальных регионах Евразии. Получается, что в популяциях Европы и Южной Азии высокая/средняя частота рискового аллеля сочетается с высокой/средней частотой группы крови A (II) по сравнению с популяциями в других регионах.

Что касается корреляций, их отсутствие между частотой рисковых аллелей и числом случаев COVID-19 было ожидаемо, так как эти аллели не связаны с чувствительностью к вирусу. Вместе с тем, они могут оказывать влияние на то, что происходит с уже заразившимся человеком, и достоверная корреляция с исходом болезни действительно обнаружена – правда, только для одного из двух маркеров и только по «российской» базе данных, но не по объединенной. Авторы подчеркивают, что эпидемиологические показатели могут по-разному учитываться в разных странах, и это сильно влияет на статистику, делая

невозможным выявить корреляцию с генетическими факторами. В пределах одной страны учет ведется более единообразно, и корреляции становятся видимыми.

Ученые полагают, что генетическое разнообразие вносит небольшой, но реальный вклад в гетерогенность пандемии COVID-19. Они рассчитывают, что в будущих исследованиях будут уточнены корреляции между частотой разных исходов болезни и генетическими особенностями популяций.

текст Надежда Маркина, Олег Балановский

Источник:

Balanovsky O., Petrushenko V., Mirzaev K., Abdullaev S., Gorin I., Chernevskiy D., Agdzhoyan A., Balanovska E., Sychev D. The variation of genome sites associated with severe COVID-19 across populations the worldwide and national patterns //Pharmgenomoc Pers Med. 2021 Nov 4;14:1391-1402. doi: 10.2147/PGPM.S320609. eCollection 2021.