

УДК 575.22/ 616-056.7

Генетическая эпидемиология полногеномных ассоциативных исследований с тяжестью течения COVID-19 (обзор GWAS исследований)

Шкурат Т.П., Ид Муэз, Шкетик А.А., Гутникова Л.В.
Южный федеральный университет, Ростов-на-Дону, Россия

DOI: 10.18522/2308-9709-2024-48-4

Аннотация

Прошло уже более четырех лет с начала эпидемии COVID-19. Клинические и генетические данные о пациентах продолжают накапливаться, но при этом мы только в самом начале нашего понимания генетических механизмов, лежащих в основе инфекции. Цель исследования: Анализ генетических исследований на основе полногеномных ассоциативных исследований (GWAS) проведенных в различных популяциях мира, для выявления ассоциаций наиболее часто встречающихся полиморфных вариантов генов у пациентов с тяжелым течением COVID-19. GWAS исследования проведенные в различных странах мира, выявили более 4 тысяч полиморфных вариантов различных генов ассоциированных с тяжелым течением COVID-19. Нами отмечено низкое совпадением полиморфных маркеров в опубликованных 274 исследованиях ассоциаций (GWAS). Так чаще всех в GWAS исследованиях регистрировали полиморфные маркеры в гене *ZTFL1*, они были отмечены в 53 исследованиях из 274, что составило 19,3%, совпадения полиморфных вариантов гена *SLC6A20* отмечены в 10,4%, а для протекторного полиморфного локуса O (1) гена *ABO* совпадение составило 9% от числа всех GWAS исследований. В патогенез тяжести течения COVID-19 вовлечены не только полиморфные варианты белок кодирующих генов, но и полиморфные варианты функциональных элементов некодирующего и митохондриального генома.

Ключевые слова: GWAS; COVID-19; полиморфные варианты генов, ассоциации.

Genetic Epidemiology of Genome-Wide Association Studies with Severity of COVID-19 (GWAS Review)

Shkurat T.P., Eid M.A., Shketik A. A., Gutnikova L.V.

Southern Federal University, Rostov-on-Don, Russia

tshkurat@sfedu.ru

DOI: 10.18522/2308-9709-2024-48-4

More than four years have passed since the onset of the COVID-19 epidemic. Clinical and genetic data on patients continue to accumulate, but we are only at the very beginning of our understanding of the genetic mechanisms underlying the infection. Objective: To analyse genetic studies based on genome-wide association studies (GWAS) conducted in various populations around the world to identify associations of the most common polymorphic variants of genes in patients with severe COVID-19. GWAS studies conducted in various countries around the world have identified more than 4 thousand polymorphic variants in various genes associated with severe COVID-19. We noted a low concordance of polymorphic markers in 274 published association studies (GWAS). Polymorphic markers in the *ZTFL1* gene were most often registered in GWAS studies; they were noted in 53 out of 274 studies, which constituted 19.3%, while concordance of polymorphic variants of the *SLC6A20* gene were noted in 10.4%, and for the protective polymorphic locus O (1) of the *ABO* gene, the concordance was 9% of all GWAS studies. Not only polymorphic variants of protein-coding genes, but also polymorphic variants of functional elements of the non-coding and mitochondrial genome are involved in the pathogenesis of COVID-19 severity.

Key words: GWAS; COVID-19; polymorphic variants of genes; associations.

Введение

Пандемия COVID-19 является одной из самых серьезных пандемий в современной истории, но также она имеет беспрецедентное число исследований, в том числе и генетических во всех медицинских и биомедицинских центрах мира. По данным ВОЗ на 25 октября 2023 года в мире зарегистрировано 771549718 подтвержденных случаев заболевания и 6974473 пациентов умерло от COVID-19. Пандемия была зарегистрирована в 215 странах, и было применено 13,533 миллиардов доз вакцин (<https://covid19.who.int>).

Эпидемиологические исследования показали, что инфекция SARS-CoV-2 может привести к широкому спектру исходов, от бессимптомной инфекции до тяжелой вирусной пневмонии, представляющей опасность для жизни (Sun et al., 2020). Известно, что у многих людей коронавирус либо не вызывал заболевания, либо оно протекало в легкой форме (<https://covid-stat.com/ru>). Можно предположить, что клиническая картина COVID-19 в каждом конкретном случае обусловлена генетическими особенностями как пациента, так и вируса, что и определяет характер их взаимодействия.

Генетический полиморфизм представителей разных этнических, возрастных групп и разного пола определяет отличия по степени восприимчивости к SARS-CoV-2. Известно, что клинический спектр инфекции SARS-CoV-2 включает, с одной стороны, бессимптомные случаи и пациентов, демонстрирующих спонтанное выздоровление, а с другой - тяжелый острый респираторный синдром (SARS). Во время пандемии все года регистрировали случаи заболевания разной тяжести, несмотря на то что появлялись новые варианты вируса SARS-CoV-2 (Tregoning, 2021; Viana et al., 2022). Только за первые 4 месяца эпидемии COVID-19 было

просеквенировано более 48 тысяч образцов вируса в различных регионах мира и показано 353341 мутационное событие в этих образцах SARS CoV2 по сравнению с первым секвенированным геномом NC_045512.2 Wuhan reference genome (Mercantile, Giorgi, 2020). Однако, эти многочисленные вариации вирусного штамма не объясняют наблюдаемую изменчивость исходов заболевания (от смертельных до бессимптомных), в одном образце чаще всего встречалась только одна мутация, возможно исключение составил только O-микрон. Кроме того, некоторые клинические факторы риска, включая пожилой возраст, повышенный индекс массы тела и мужской пол, ассоциированы с повышенной восприимчивостью к развитию тяжелых форм COVID-19 (Zhou et al., 2020; Cummings et al., 2020; Lopez et al., 2021; Knight et al., 2021). Причины степени тяжести заболевания после инфицирования остаются пока невыясненными до конца, также существуют факторы хозяина, которые связаны с восприимчивостью к заражению вирусом SARS-CoV-2, например, генотипические особенности его врожденного и приобретенного иммунитета, предрасположенность к повышенной свертываемости крови, и др. Неожиданностью стала и широкая распространенность так называемого длительного COVID (long-COVID-19), хронического синдрома COVID (CCS), пост-острых последствий COVID-19 (PASC) или синдрома/состояний после COVID-19, причины которых еще предстоит определить. Были предложены различные патофизиологические механизмы, например, разрушение тканей, длительная инфекция, аутоиммунный процесс и гиперкоагуляция. Скорее всего, задействовано несколько механизмов, и группы пациентов могут иметь очень разные механизмы и траектории заболевания. Учитывая масштабы пандемии, и беспрецедентное число полногеномных исследований во всем мире, а также длительный характер отдаленных последствий данный обзор направлен на проведение анализа полногеномных ассоциативных исследований (GWAS) в

различных популяциях мира, с целью выявления наиболее значимых полиморфных локусов генов у пациентов с тяжелым течением COVID-19.

МАТЕРИАЛ И МЕТОДЫ ИССЛЕДОВАНИЙ

Материалом для исследования послужили данные включенные в базу данных GWAS полных геномов больных COVID -19, которые включали 274 исследования из различных стран мира (<https://ebi.ac.uk/gwas>). Данные из баз данных GWAS COVID-19 v.4 (<https://genome.ucsc.edu>); GenOMICC (<https://covid.genomicc.org>); COVID-19 HGI Browser (covid19hg.org); 23andMe COVID-19 Dataset Access Program (<https://research.23andme.com>). Поиск публикаций по GWAS исследованиям при COVID-19 также был осуществлен в поисковой системе научных публикаций Google scholar (<https://scholar.google.ru>).

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Одно из первых исследований по выяснению роли генетических факторов хозяина в восприимчивости и тяжести пандемии вируса SARS-CoV-2 появились в мае 2020, уже через полгода после начала эпидемии (таблица 1). Эти исследования рассматривали тяжелые случаи течения COVID-19 или случаи со смертельным исходом и описаны в статье «Генетика смертности в интенсивной терапии» (Pairo-Castineira et al.,2021).

Таблица 1 - Полиморфные локусы генома ассоциированные со смертностью при COVID-19 (Pairo-Castineira et al.,2021; Zong, Li, 2021)

SNP	Human GRCh37/hg19 Assembly	Аллель риска	Alternative	Ближайший ген к SNP
rs73064425	chr3:45901089-45901089	T	C	<i>LZTFL1</i>
rs9380142	chr6:29798794-29798794	A	G	<i>HLA-G</i>
rs143334143	chr6:31121426-31121426	A	G	<i>CCHCR1</i>
rs10735079	chr12:113380008-113380008	A	G	<i>OAS3</i>

rs74956615	chr19:10427721-10427721	A	T	<i>TYK2</i>
rs2109069	chr19:4719443-4719443	A	G	<i>DPP9</i>
rs2236757	chr21:34624917-34624917	A	G	<i>IFNAR2</i>

Как видно из результатов представленной в таблице 1, значимый уровень достоверности имели однонуклеотидные замены расположенные в не структуре генов кодирующих белки, а те которые были расположены рядом с этими генами, т.е. в регуляторных областях генома. Мы идентифицировали и воспроизвели следующие новые общегеномные значимые ассоциации: на хромосоме 12q24.13 (rs10735079, $P = 1,65 \times 10^{-8}$) в кластере генов, кодирующих антивирусные активаторы ферментов рестрикции (OAS1, OAS2 и OAS3); на хромосоме 19p13.2 (rs74956615, $P = 2,3 \times 10^{-8}$) рядом с геном, кодирующим тирозинкиназу 2 (TYK2); на хромосоме 19p13.3 (rs2109069, $P = 3,98 \times 10^{-12}$) в составе гена, кодирующего дипептидилпептидазу 9 (DPP9); и на хромосоме 21q22.1 (rs2236757, $P = 4,99 \times 10^{-8}$) в гене рецептора интерферона IFNAR2. Мы определили потенциальные цели для перепрофилирования лицензированных препаратов: используя менделевскую рандомизацию, мы обнаружили доказательства того, что низкая экспрессия IFNAR2 или высокая экспрессия TYK2 связаны с опасными для жизни заболеваниями; и общетранскриптомная ассоциация в легочной ткани показала, что высокая экспрессия моноцитарно-макрофагального хемотаксического рецептора CCR2 связана с тяжелым течением COVID-19

В рамках этих первых исследований было выдвинута инициатива, передачи всех генетических и клинических данные исследований в Европейский архив геномов и фенотипов (EGA) с контролируемым доступом для осуществления координации исследований вирус-хозяин, с учетом изменения нуклеотидной последовательности вируса SARS CoV-2 в различных странах, депонированием вирусных последовательностей, и

оценкой показателей GWAS и других бимолекулярных данных инфицированных пациентов в различных учреждениях по всему миру (covid19hg.org). Благодаря этой и другим инициативам было проведено 274 полногеномных ассоциативных исследований по технологии GWAS в различных странах мира. Эти 274 исследования размещены в GWAS браузере (<https://www.ebi.ac.uk/gwas>) исследования показали широкую гетерогенность полученных данных и выявили 4208 полиморфных локусов ассоциированных с течением COVID-19 у пациентов (данные представлены на 15.06.2024). В таблице 2 представлены наиболее часто регистрируемые SNP в GWAS исследованиях «случай-контроль» у пациентов с разной степенью тяжести течения COVID-19.

Таблица 2 Наиболее часто регистрируемые SNP в GWAS исследованиях «случай-контроль» у пациентов с разной степенью тяжести течения COVID-19 (<https://ebi.ac.uk/gwas>).

Ген	Количество GWAS исследований, в которых отмечен хоть один SNP для данного гена	Наиболее часто регистрируемые SNP в различных исследованиях	Число пациентов «случай-контроль» в GWAS исследованиях и происхождение пациентов
LZTFL1	53	rs35482426 rs34549672 rs35280891 rs17713054 rs11385942 rs10490770 rs35081325 rs73064425 rs35731912 rs35508621	682 Африка-случай , 2,540 Африка-контроль; 401 Восточная Азия-случай, 759 Восточная Азия-контроль; 13,199 Европейское происхождение -случай, 745,011 Европейское происхождение – контроль; 1,141 Южная Азия-случай; 5,384 Южная Азия-контроль, 8,779 - случай, 1,001,875 - контроль

ABO	25	rs505922 rs9411378 rs687289 rs550057 rs9411378 rs550057 rs495828 rs579459 rs529565 rs145526382	8,461 Африка-случай , 127,310 Африка-контроль, 10,109 Испания или Латинская Америка - случай , 75,831 Испания или Латинская Америка - контроль, 2,047 Восточная Азия-случай, 19,113 Восточная Азия-контроль, 97,783 Европейское происхождение -случай, 2,295,876 Европейское происхождение - контроль 1,103 Ближний Восток, Северная Африка, Персия -случай, 15,818 Ближний Восток, Северная Африка, Персия – контроль; 6,081 Южная Азия- случай; 41,399 Южная Азия- контроль
DPP9	22	rs12610495 rs7256695 rs2109069 rs2277732	2,884 Европейское происхождение -случай, 95,493 Европейское происхождение - контроль 376 Южная Африка-случай, 6,163 Южная Африка-контроль
SLC6A20	38	rs2271616 rs17713054 rs10490770 rs71325088 rs73062389 rs73062394 rs7620663 rs17078346 rs17713054 rs17279437	72 Восточная Азия-случай, 35 Восточная Азия-контроль, 81 Южная Азия- случай18 Южная Азия- контроль, 210 Европейское происхождение -случай, 92 Европейское происхождение - контроль 90 Американское происхождение-случай12 Американское происхождение- контроль.
TYK2	25	rs11085727 rs11085725 rs34536443 rs144309607 rs149329233 rs280519 rs2304256	2,113 Африка-случай , 121,925 Африка-контроль, 1,709 Испания или Латинская Америка - случай , 150,547 Испания или Латинская Америка - контроль, 1,551 Восточная Азия-случай, 15,825 Восточная Азия-контроль, 18,605 Европейское происхождение -случай, 2,490,909 Европейское происхождение - контроль 584 Ближний Восток, Северная Африка, Персияcases, 14,491 Ближний Восток, Северная Африка, Персия -контроль; 465

			Южная Азия- случай 42,575 Южная Азия- контроль.
HTRA1	13	rs2248799 rs2672598 rs28414499 rs28637696 rs28414499	72 Восточная Азия-случай, 35 Восточная Азия-контроль, 81 Южная Азия- случай18 Южная Азия- контроль, 210 Европейское происхождение - случай, 92 Европейское происхождение - контроль 90 Американское происхождение- случай12 Американское происхождение- контроль
RBFOX1	16	rs200693233 rs151059900 rs7188948 rs7203935 rs148817892 rs192187627 rs10492829 rs2881439 rs11077081 rs1861036	2,884 Европейское происхождение -случай, 456,366 Европейское происхождение - контроль
TRIOBP	16	rs550674936	1,120 Европейское происхождение -случай, 458,130 Европейское происхождение - контроль 124 - случай, 27,946 - контроль.
NOL10	14	rs4997810 rs7566875 rs6432114 rs2110776 rs6432125 rs12104950 rs34595606 rs7355649 rs10183883 rs10199831	235 тяжелых случаев, к 841 легких случаев течения COVID-19

Анализируя функции и биологические процессы в которых участвуют гены, полиморфные варианты которых показали значимую ассоциацию у больных COVID-19, отмечаем что генетический компонент, лежащий в основе риска инфекции, связан с биологическими процессами, происходящими в верхних дыхательных путях, такими как мукоцилиарный клиренс, состояние иммунной системы слизистой оболочки и механизмами проникновения и репликации вируса в клетках хозяина.

Чаще других во многих GWAS исследованиях мира отмечали ассоциацию полиморфных локусов гена *LZTFL1* с тяжестью течения COVID-19. Ген *LZTFL1* (leucine zipper transcription factor like 1; транскрипционный фактор-подобный белок 1 с лейциновой «застежкой-молнией») кодирует цитозольный белок, который связывается с E-кадгерином (эпителиальным маркером) и актиновым цитоскелетом и участвует в переносе многочисленных сигнальных молекул. *LZTFL1* экспрессируется в эпителиальных клетках легких, включая реснитчатые эпителиальные клетки, которые были идентифицированы как одна из основных клеточных мишеней для вируса SARS-CoV-2. (Траспов А., и др. 2022). Геногеография в России и мире однонуклеотидных вариантов гена *LZTFL1*, ассоциированных с тяжелым течением COVID-19 была исследована в Медико-генетическом научном центре им. Н. П. Бочкова (Балановская Е.В. и др, 2022). Показано, что мировые экстремумы полиморфного варианта rs17713054 приходятся на коренное население Африки и Северной Европы, в то время, как полиморфный вариант rs2191031(A) чаще встречался у народов Африки, Америки, но максимум сосредоточен в Океании. В этом исследовании была подтверждена значимая сильная ассоциация rs17713054 с тяжелым течением (OR = 4,56; p = 0,0025) в рецессивной модели (Балановская Е.В. и др, 2022). Для жителей России было предложено делить на группы пациентов с потенциально легким течением и потенциально тяжелым течением

заболевания COVID-19 по факту наличия или отсутствия в генотипе пациента выбранных полиморфных локусов гена *TNF* rs1800629 и гена *LZTFL1* rs17713054 (Траспов А., и др. 2022). На рисунке 2 показано отношение шансов для полиморфного локуса rs73064425 (полиморфный вариант локализован недалеко от гена *LZTFL1*, в некодирующем участке генома) у больных госпитализированных с COVID-19 по данным мета анализа GWAS 60 исследований из 25 стран.

Для полиморфного локуса rs73064425 представлены размеры эффекта и доверительные интервалы для каждого GWAS исследования, вносящего вклад. Исследования сгруппированы по происхождению, при этом суммарные эффекты сообщаются для каждой подгруппы происхождений. Суммарный размер эффекта по всем исследованиям также указан на нижней панели. Как видно из рисунка 2 практически во всех популяциях мира полиморфный локус rs73064425 был ассоциирован с тяжестью течения COVID-19. Таким образом, в патогенез тяжести течения COVID-19 вовлечены не только нарушения в белок кодирующих гена, но и полиморфные варианты функциональных элементов некодирующего генома.

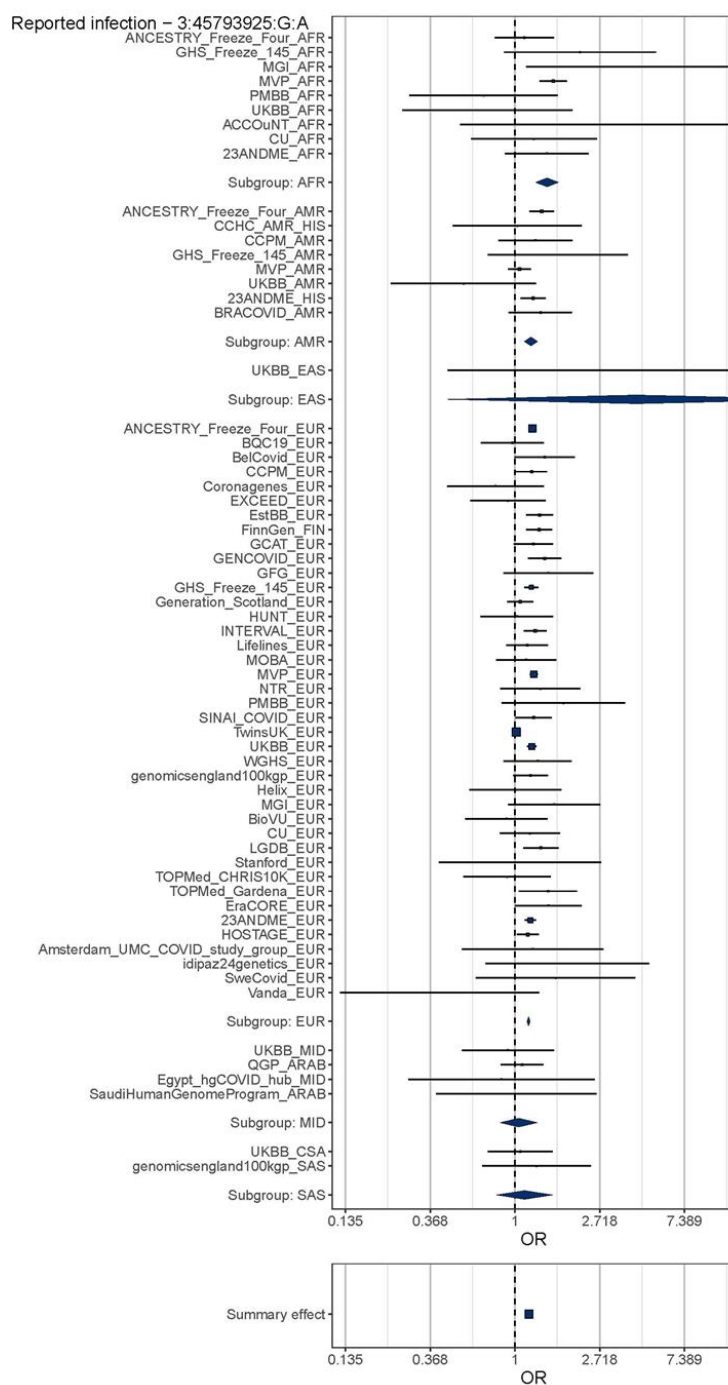


Рисунок 1 - Совокупная оценка отношения шансов локуса rs73064425 3:4579925 G:A (полиморфный локус на 3 хромосоме рядом с геном *LZTFL1*) у больных госпитализированных с COVID-19 по данным мета анализа GWAS 60 исследований из 25 стран (Ganna et al., 2021).

В многих GWAS исследованиях сигнал ассоциации совпадал с локусом группы крови ABO. Исследователи из Китая сообщили, что

аллельные варианты гена ABO статистически тесно ассоциированы не только с тяжелым течением COVID-19, но и с выживаемостью после инфицирования (Zhao et al., 2020). Китайские ученые сравнили распределение ABO у 1775 пациентов с инфекцией SARS-CoV-2 в Китае с 3694 жителей из фоновой здоровой популяции (Zhao et al., 2020). Следует отметить, что их референтная популяция не представляла собой группу неинфицированных пациентов, поступивших в одну и ту же больницу, а скорее представляла собой фоновую референтную популяцию «нормальных людей» из Ухани. Они обнаружили, что группа крови «0» реже встречалась среди инфицированных лиц (25,8%) по сравнению со здоровыми людьми из контрольной группы (33,8%), а группа «А» чаще встречалась среди инфицированных лиц (37,7%) по сравнению со здоровым населением контрольной группы (32,2%, $p < 0,001$). Большинство исследователей признает наличие протекторного эффекта группы крови 0(I) и резус-отрицательного фенотипа при инфицировании SARS-CoV-2, однако в трактовке отдельных наблюдений имеются некоторые разногласия. Имеются данные о том, что антитела анти-А и анти-В у лиц с группой крови 0 (I) могут воздействовать на гликопротеины на поверхности вируса, нарушая процесс проникновения вируса (Pendou et al., 2021). Высказана гипотеза, что защитный эффект группы крови 0(I), осуществляется предположительно за счет отсутствия аглютиногена А на поверхности мембран эритроцитов, являющегося предпочтительным лигандом для вируса SARS-CoV-2 (Зубарева, Донсков, 2022). Неоднозначные результаты корреляции исследователи объясняют неравномерным распределением групп крови в силу этнических, географических и социальных влияний. Например, для жителей Южной Африки значимая ассоциация с тяжестью течения COVID-19 была обнаружена для полиморфного варианта rs10896071 гена *SART1* (spliceosome associated factor 1, фактор 1, ассоциированный со сплайсосомой).

Ни в одном другом из 273 GWAS исследований этой ассоциации не было обнаружено.

Чтобы оценить результаты GWAS COVID-19 с данными об экспрессии тканеспецифических генов, ученые из Roslin Institute, University of Edinburgh, Edinburgh (Великобритания) выполнили исследование ассоциаций в масштабах всего транскриптома (TWAS) и сделали вывод об экспрессии генов с различными полиморфными вариантами. Для этого анализа они изучили экспрессию в двух тканях: образцах цельной крови и легких. Они выбрали гены с $P < 0,05$ в этих тканях и провели комбинированный мета-TWAS-анализ²³, включив данные eQTL из других тканей в GTEx, чтобы оптимизировать мощность для обнаружения различий в прогнозируемой экспрессии в легких или крови. На рисунке 2 представлен график Манхэттена TWAS-анализа (экспрессии генов) в тканях легких и крови.

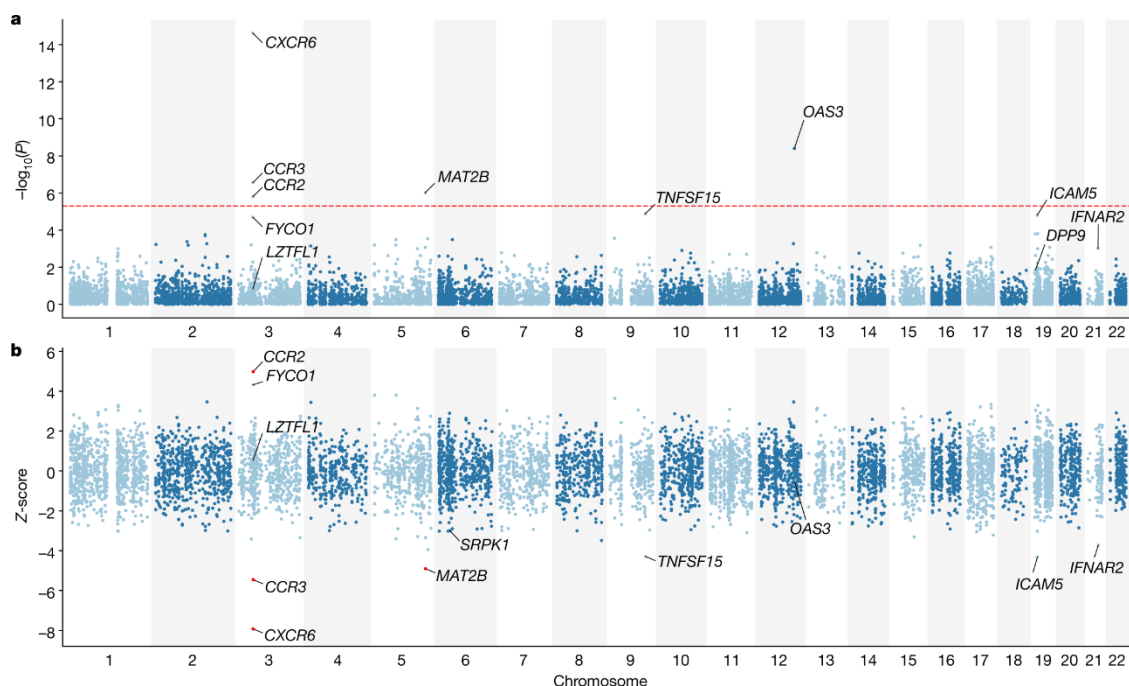


Рисунок 2 - График Манхэттена на уровне генов, мета-TWAS-анализ тканей (Pairo-Castineira et al., 2021)

Красная горизонтальная линия показывает значимость на уровне гена для всего генома при $-\log_{10}(5 \times 10^{-6})$. Z-показатели, предполагаемая по генотипу экспрессия транскриптов, которые кодируют гены, кодирующие белок, в легочной ткани.

Также было показано, что полиморфные варианты генов: *SLC22A31*, *SFTPD4* и *TKY2* оказывают значительное влияние на дифференциальную экспрессию этих генов у тяжелобольных COVID-19. В этих исследованиях показано, что низкая экспрессия *IFNAR2* или высокая экспрессия *TYK2* ассоциированы с тяжестью COVID-19. Транскриптомные исследования в легочной ткани показали, что высокая экспрессия хемотаксического рецептора моноцитов-макрофагов *CCR2* ассоциирована с тяжелым течением COVID-19 (Pairo-Castineira et al., 2021).

Zeberg H, Pääbo S. (2020) показали, что геномный сегмент на 3 хромосоме размером около 50 килобаз, на котором регистрируется самое большое число полиморфных локусов ассоциированных с тяжелым течением COVID-19, был унаследован от неандертальцев и сохранился примерно у 50% людей в Южной Азии и примерно 16% людей в Европе (рис.3).

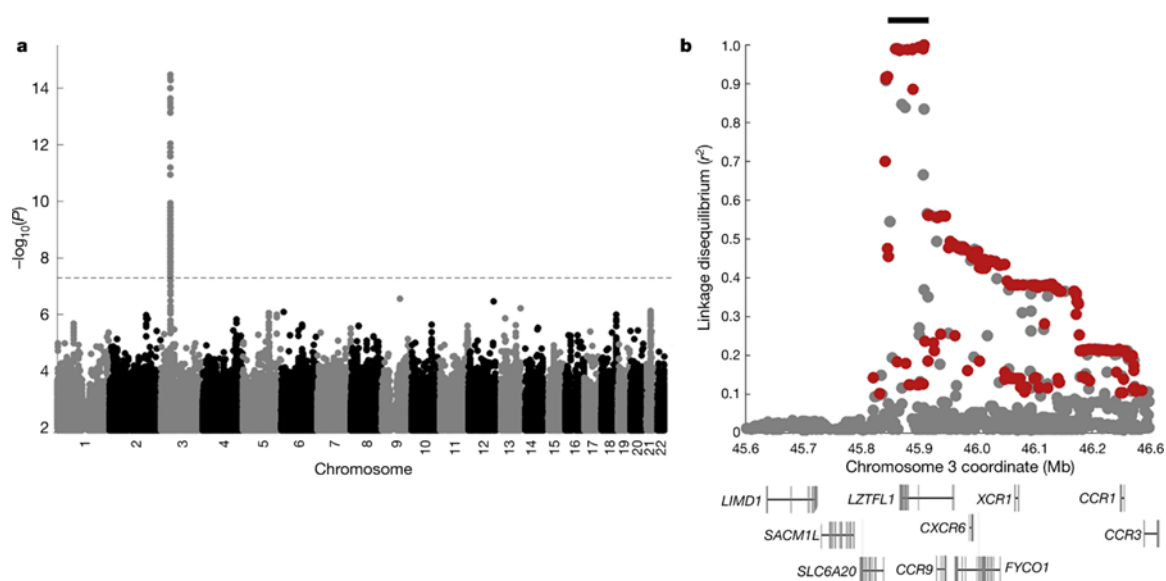


Рис. 3 Манхэттенский график исследования ассоциаций по всему геному 3199 госпитализированных пациентов с COVID-19 и 897 488 контрольных популяций. Пунктирная линия указывает на значимость по всему геному ($P = 5 \times 10^{-8}$). Данные были изменены из COVID-19 Host Genetics Initiative2 (<https://www.covid19hg.org/>).

Клинические проявления COVID-19 варьируемы и варьируются от полного отсутствия симптомов до тяжелой дыхательной недостаточности и летального исхода. Эта клиническая изменчивость предполагает, что генетика хозяина, его полиморфные варианты генома, играет важную роль в восприимчивости и влиянии проявления COVID-19.

В китайской популяции ученые выявили общие варианты SNP, характерными для других популяций (*ABO* и *FOXP4-AS1*), которые в значительной степени ассоциированы с тяжестью течения COVID-19 (таблица 3). В этом исследовании обнаружен и вариант редкого риска в *MEF2B* (*myocyte enhancer factor 2B*, *миоцит фактор-энхансер-2*) специфичен для выходцев из Восточной Азии и обеспечивает примерно восьмикратное увеличение риска тяжелого течения COVID-19 среди его носителей (Wu et al., 2021).

Таблица 3 - Значимые локусы, ассоциированные с тяжестью COVID -19 по данным Chinese (GWAS) и HGI и 23andMe (Wu et al., 2021)

Locus	Dataset	Sample size	Lead variant ^a	OR (95 % CI) ^c	P
3p21.31	Chinese (GWAS)	598/2260	rs35044562	–	–
<i>LZTFL1</i>	HGI (B2_release3)	3199/897,488	chr3:45867532	1.60 (1.42–1.79)	3.11×10^{-15}
	Chinese (WGS)	474/1615	A/G, Intronic	–	–
6p21.1	Chinese (GWAS)	598/2260	rs1853837	1.30 (1.13–1.50)	3.24×10^{-4}
<i>FOXP4-AS1</i>	HGI (B2_release3)	3199/897,488	chr6:41529297	1.28 (1.15–1.42)	5.24×10^{-6}
<i>FOXP4</i>	Chinese (WGS)	474/1615	C/A	1.27 (1.07–1.51)	7.06×10^{-3}
	Meta-analysis	4271/901,363	Intronic	1.28 (1.19–1.39)	2.51×10^{-10}
9q34.2	Chinese (GWAS)	598/2260	rs8176719	1.28 (1.12–1.46)	3.19×10^{-4}

Locus	Dataset	Sample size	Lead variant ^a	OR (95 % CI) ^c	P
<i>ABO</i>	HGI (B2_release3)	3199/897,488	chr9:133257521	1.17 (1.09–1.26)	1.27×10^{-5}
	Chinese (WGS)	474/1615	T/TC	1.17 (0.98–1.38)	8.03×10^{-2}
	Meta-analysis	4271/901,363	Exonic (frameshift)	1.19 (1.12–1.26)	8.98×10^{-9}
19q13.11	Chinese (GWAS)	598/2260	rs74490654	–	–
<i>MEF2B</i>	HGI (B2_release3)	3199/897,488	chr19:19163581	–	–
	Chinese (WGS)	474/1615	C/G, Intronic	8.73 (4.14–18.41)	1.22×10^{-8}

Обзор вариаций в митохондриальной ДНК (мтДНК), проведенный Wu соавторами [Wu Y, e.a., 2021], указывает на различия в генах ND1, ND2, ND3 и в области D-петли, как связанные с повышенным риском тяжелого течения COVID-19.

Полногеномное ассоциативное исследование (GWAS) тяжести COVID-19 двух групп европейских пациентов госпитализированных в больницах Италии и Испании первоначально выделили 8 582 968 SNP. Исследователи провели GWAS у 835 итальянских пациентов и 1255 членов контрольной группы, а также у 775 испанских пациентов и 950 членов контрольной группы. После проведения SRM анализа ассоциированных SNP осталось 1054. Затем экспериментальные результаты были подвергнуты метаанализу для получения окончательных SNP рисков. Большинство SNP не повлияли на риск заражения COVID-19, значимые полиморфные варианты представлены (таблица 4). Среди идентифицированных полиморфных вариантов генов всего было шесть белок-кодирующих генов (UQCRH, PPA2, PAPSS1, ABO, PANO1 и ARSA) и пять полиморфных вариантов были локализованы в некодирующих участках генома (Zong, 2021).

Таблица 4- GWAS исследования геномов пациентов из Испании и Италии ассоциированные с тяжестью COVID-19 с использованием метода SMR (Zong, 2021)

SNP	Ген	Хромосома	Позиция	P_GWAS	P_eQTL
rs41292543	<i>UQCRH</i>	1	46309111	0.001553	2.65E-16
rs35258888	<i>PPA2</i>	4	105355205	0.001035	2.61E-10
rs70947091	<i>PAPSS1</i>	4	107694523	0.001899	2.28E-56
rs8176719	<i>ABO</i>	9	133257521	8.76E-07	1.27E-34
rs10678686	<i>AP006621.5</i>	11	780321	3.96E-05	1.28E-100
rs10678686	<i>CMB9-55F22.1</i>	11	780321	3.96E-05	9.62E-143
rs10678686	<i>AP006621.6</i>	11	780321	3.96E-05	1.12E-44
rs7104929	<i>PANO1</i>	11	784340	3.73E-05	0.0111629
rs10407383	<i>CTD-2027H9.2</i>	19	24134099	0.000412	4.61E-09
rs6122883	<i>LINC01273</i>	20	50172836	0.000494	3.43E-34
rs6151429	<i>ARSA</i>	22	50625049	0.000522	4.70E-50

Анализ гена *ACE2*, который расположен в X-хромосоме и характеризуется высокой степенью полиморфизма в его кодирующих областях, показал, что разные варианты по-разному способствуют инфицированию COVID-19. Также было показано, что *TMPRSS2* вместе с

клеточным рецептором ACE2 и экзопептидазой DPP4 являются мембраносвязанными белками, вовлеченными в инфекцию COVID-19 и проникновение вируса в клетку-хозяина, и, следовательно, способствуют тяжести заболевания (Irham LM, e.a., 2020; Senapati S, e.a., 2020; Singh H, 2021). Показана значительная ассоциация полиморфного варианта rs4880 *SOD2* с тяжестью COVID-19. Носители комбинации аллелей *SOD1 7958G * SOD2 47T * CAT 262C* имели более высокий риск тяжелой формы COVID-19 (Eid M. A. et al. , 2024).

Несколько негенетических факторов также были ассоциированы с тяжелым течением COVID-19, включая возраст, мужской пол и сопутствующие заболевания. Показана ассоциация тяжести течения COVID-19 в зависимости от возраста, мужского пола, наличия коморбидных состояний, в частности, гипертонии и сахарным диабетом 2 типа (СД2). Как гипертония, так и СД2 были определены как независимые факторы риска повышенной смертности у пациентов с COVID-19 (Mancia G, e.a., 2020; Liang X, e.a., 2020; Gao C, e.a., 2020; Holman N, e.a., 2020). Механизм, посредством которого гипертония является фактором риска, был приписан гиперактивации пути ренин-ангиотензиновой системы, что усиливает воспалительную реакцию, цитокиновый шторм, ремоделирование миокарда, острое повреждение легких и повреждение эндотелия. Аналогичным образом, было высказано предположение, что СД2 способствует тромбоемболическим осложнениям и повреждению органов через глюкотоксичность, окислительный стресс и повышенную продукцию цитокинов (Saad F. M. et al. 2024).

ЗАКЛЮЧЕНИЕ

Прошло уже более четырех лет с начала эпидемии COVID-19 клинические и генетические данные о пациентах продолжают накапливаться, но при этом мы только в самом начале нашего понимания генетических

механизмов, лежащих в основе инфекции. Имея огромные базы данных, созданные во время пандемии, теперь пришло время для проведения более глубоких исследований, чтобы понять различные исходы инфекции COVID-19 и понять их роль в патогенезе инфекции. Это обусловлено и относительно низким совпадением полиморфных маркеров в опубликованных 274 исследованиях ассоциаций (GWAS). Так чаще всех в GWAS исследованиях регистрировали полиморфные маркеры в гене *ZTFL1*, они были отмечены в 53 исследованиях из 274, что составило 19,3%, для протекторной роли полиморфного локуса O (1) гена ABO совпадение составило 9% от числа всех исследований, для остальных локусов еще меньше. Низкое совпадение полиморфных локусов ассоциированных с COVID-19 в GWAS исследованиях объясняется не только различными генофондами в различных этнических группах и популяциях, но и разными подходами к оценке тяжести COVID-19. Например, одни исследователи оценивали только летальные случаи, другие госпитализированных пациентов, третьи проводили исследования с помощью анкетирования, при этом чаще всего маркеры сравнивали с данными регистрации аллельных вариантов в популяции, которые были получены ранее, без учета инфицирования вирусом SARS CoV-2. Одним из наиболее ограничивающих факторов при применении популяционного GWAS к большим геномам является размер выборки. С точки зрения статистики недостаточный размер выборки приводит к повышенным ошибкам. Для геномов, состоящих из миллиардов пар оснований, большой размер выборки ограничен затратами на секвенирование. GWAS исследования используют для больших выборок микрочипы, которые могут оценить достаточно большое, но ограниченное число полиморфных локусов.

Тем не менее, на основе данных GWAS и публикаций исследований по одиночным полиморфным локусам на сегодняшний день было выявлено несколько генетических вариантов, которые ассоциированы с

восприимчивостью и устойчивостью к COVID-19. Они в основном ассоциированы с проникновением вируса, рецепторами и иммунологическими процессами. GWAS исследования позволили выявить практически во всех популяциях мира наиболее значимые полиморфные локусы следующих генов: *ZTFL1*, *TYK2*, *PAPSS1*, *ABO*, *FOXP4*, *DPP9*, *SLC6A2*, *HTRA1*, которые наиболее часто встречались практически во всех странах мира и были ассоциированы с COVID-19. Только для китайской популяции обнаружен вариант редкого риска в гене *MEF2B*, и только для жителей Южной Африки отмечена значимая ассоциация с тяжестью течения COVID-19 для полиморфного варианта rs10896071 гена *SART1*. Выявлен ряд значимых полиморфных локусов расположенных в некодирующих участках генома, роль которых в патогенезе COVID-19 только еще предстоит выяснить.

Источник финансирования. Работа выполнена при финансовой поддержке Министерства науки и высшего образования РФ в рамках государственного задания в сфере научной деятельности № FENW-2023-0018.

Список использованных источников:

Балановская, Е. В., Горин, И. О., Петрушенко, В. С., Черневский, Д. К., Кошель, С. М., Темирбулатов, И. И., ... & Агджоян, А. Т. (2022). Геногеография в России и мире SNP-маркеров гена *LZTFL1*, ассоциированных с тяжелым течением COVID-19. *Вестник Российского государственного медицинского университета*, (5), 31-40.

Зубарева Л. М., Донсков С. И. Группы крови и острая респираторная вирусная инфекция COVID-19. *Гематология и трансфузиология*. 2022; 67(1):122–130.

А. А. Траспов, М. М. Минашкин, С. В. Поярков, А. Г. Комаров, И. А. Штинова, Г. И. Спешилов, И. А. Карбышев, Н. В. Позднякова. Полиморфизмы rs17713054 и rs1800629 генов LZTFL1 и TNF ассоциированы с тяжестью течения COVID-19. «Вестник РГМУ», 6/2022. Опубликовано online 28.12.2022

COVID-19 Host Genetics Initiative aganna@broadinstitute.org. (2020). The COVID-19 host genetics initiative, a global initiative to elucidate the role of host genetic factors in susceptibility and severity of the SARS-CoV-2 virus pandemic. *European Journal of Human Genetics*, 28(6), 715-718.

Eid, M. A., Aleksandrova, A. A., Kosenko, P. O., & Shkurat, T. P. (2024). Association of superoxide dismutase and catalase genetic variants and their gene-gene interactions with the severity of COVID-19. *Ecological genetics*.

Everitt, A. R., Clare, S., Pertel, T., John, S. P., Wash, R. S., Smith, S. E., ... & Wellcome Trust Sanger Institute, UK Kellam P. 24. (2012). IFITM3 restricts the morbidity and mortality associated with influenza. *Nature*, 484(7395), 519-523.

Gao C, Cai Y, Zhang K, Zhou L, Zhang Y, Zhang X, et al. Association of hypertension and antihypertensive treatment with COVID-19 mortality: a retrospective observational study. *Eur Heart J*. (2020) 41:2058–66. doi: 10.1093/eurheartj/ehaa433

Ganna A. COVID-19 Host Genetics Initiative, Mapping the human genetic architecture of COVID-19: an update //MedRxiv. – 2021. – С. 2021.11.08.21265944.

Holman N, Knighton P, Kar P, O'Keefe J, Curley M, Weaver A, et al. Risk factors for COVID-19-related mortality in people with type 1 and type 2 diabetes in England: a population-based cohort study. *Lancet Diabetes Endocrinol*. (2020) 8:823–33. doi: 10.1016/S2213-8587(20)30271-0

Irham LM, Chou W-H, Calkins MJ, Adikusuma W, Hsieh S-L, Chang W-C. Genetic variants that influence SARS-CoV-2 receptor TMPRSS2 expression among population cohorts from multiple continents. *Biochem Biophys Res Commun.* 2020;529: 263–269. pmid:32703421

Knight, J. S., Caricchio, R., Casanova, J. L., Combes, A. J., Diamond, B., Fox, S. E., ... & McCune, W. J. (2021). The intersection of COVID-19 and autoimmunity. *The Journal of clinical investigation*, 131(24).

Liang X, Shi L, Wang Y, Xiao W, Duan G, Yang H, et al. The association of hypertension with the severity and mortality of COVID-19 patients: evidence based on adjusted effect estimates. *J Infect.* (2020) 81:e44–7. doi: 10.1016/j.jinf.2020.06.060

Lopez-Leon, S., Wegman-Ostrosky, T., Perelman, C., Sepulveda, R., Rebolledo, P. A., Cuapio, A., & Villapol, S. (2021). More than 50 long-term effects of COVID-19: a systematic review and meta-analysis. *Scientific reports*, 11(1), 1-12.

Mancia G, Rea F, Ludergnani M, Apolone G, Corrao G. Renin–Angiotensin–Aldosterone System Blockers and the Risk of Covid-19. *N Engl J Med.* (2020) 382:2431–40. doi: 10.1056/NEJMoa200692

Mercatelli D. and Giorgi F.M., Geographic and genomic distribution of SARS-CoV-2 mutations, *Front. Microbiol.*, 11 (2020), p. 1800.

Pairo-Castineira, E., Clohisey, S., Klaric, L., Bretherick, A. D., Rawlik, K., Pasko, D., ... & Princess of Wales Hospital, Llantrisant, UK Sathe S. 133 Davies E. 133 Roche L. 133. (2021). Genetic mechanisms of critical illness in COVID-19. *Nature*, 591(7848), 92-98.

Pairo-Castineira, E., Rawlik, K., Bretherick, A. D., Qi, T., Wu, Y., Nassiri, I., ... & Great Western Hospital, Swindon, UK Prout Rachel 164 Langton Helen 164 Watters Malcolm 164 Hunt Charlotte 164 Novis Catherine 164. (2023). GWAS and meta-analysis identifies 49 genetic variants underlying critical COVID-19. *Nature*, 617(7962), 764-768.

Saad, F. M., Daif, A., Sakr, M., & Ismail, H. (2024). Gene Expression Analysis of ACE2, TMPRSS2, and LZTFL1 in COVID-19 Patients: a pilot study. *Research Journal of Applied Biotechnology*, 10(1), 1-14.

Senapati S, Kumar S, Singh AK, Banerjee P, Bhagavatula S. Assessment of risk conferred by coding and regulatory variations of TMPRSS2 and CD26 in susceptibility to SARS-CoV-2 infection in human. *J Genet.* 2020;99: 53. pmid:32661206

Singh H, Choudhari R, Nema V, Khan AA. ACE2 and TMPRSS2 polymorphisms in various diseases with special reference to its impact on COVID-19 disease. *Microb Pathog.* 2021;150: 104621. pmid:33278516

Tregoning, J. S., Flight, K. E., Higham, S. L., Wang, Z., & Pierce, B. F. (2021). Progress of the COVID-19 vaccine effort: viruses, vaccines and variants versus efficacy, effectiveness and escape. *Nature reviews immunology*, 21(10), 626-636.

Viana, M. V., Pantet, O., Charrière, M., Favre, D., Piquilloud, L., Schneider, A. G., ... & Berger, M. M. (2022). Specific nutrition and metabolic characteristics of critically ill patients with persistent COVID-19. *Journal of Parenteral and Enteral Nutrition*, 46(5), 1149-1159.

Wu Y, Wang X-H, Li X-H, Song L-Y, Yu S-L, Fang Z-C, et al. Common mtDNA variations at C5178a and A249d/T6392C/G10310A decrease the risk of severe COVID-19 in a Han Chinese population from Central China. *Mil Med Res.* 2021;8: 57.

Zhao, J., Yang, Y., Huang, H., Li, D., Gu, D., Lu, X., ... & Wang, P. G. (2021). Relationship between the ABO blood group and the coronavirus disease 2019 (COVID-19) susceptibility. *Clinical Infectious Diseases*, 73(2), 328-331.

Zeberg H, Pääbo S. The major genetic risk factor for severe COVID-19 is inherited from Neanderthals. *Nature.* 2020;587: 610–612. pmid:32998156

Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *The Lancet*. 2020;395: 1054–1062. pmid:32171076

Статья поступила в редакцию 28 июня 2024 г.

Поступила после доработки 25 июля 2024 г.

Принята к печати 30 июля 2024 г.

Received 28, June, 2024

Revised 25, July, 2024

Accepted 30, July, 2024