

УДК 578.287

# Генетическое разнообразие и эволюция биологических свойств коронавируса SARS-CoV-2 в условиях глобального распространения

А. А. Никонова\*, Е. Б. Файзулов, А. В. Грачева, И. Ю. Исаков, В. В. Зверев

Научно-исследовательский институт вакцин и сывороток им. И.И. Мечникова, Москва, 105064 Россия

\*E-mail: aa.nikonova@nrcii.ru

Поступила в редакцию 19.01.2021

Принята к печати 13.04.2021

DOI: 10.32607/actanaturae.11337

**РЕФЕРАТ** Новая коронавирусная инфекция (COVID-19) представляет серьезную угрозу для глобального здравоохранения. С начала пандемии число пациентов с подтвержденной инфекцией в мире превысило 117 млн, из них более 2.6 млн скончались. Несмотря на беспрецедентные ответные меры общественного здравоохранения, направленные на сдерживание распространения заболевания, ряд стран столкнулся с кризисом в отделениях интенсивной терапии. В настоящее время предпринимаются большие усилия по мониторингу генетической изменчивости этиологического агента COVID-19 – коронавируса SARS-CoV-2. Секвенирование сотен тысяч геномов вирусов, выделенных в разных регионах мира, позволяет изучать эволюцию SARS-CoV-2 и отслеживать его географическое распространение с течением времени. Понимание эволюции SARS-CoV-2 необходимо не только для ретроспективного анализа причин появления и закономерностей распространения новой коронавирусной инфекции, но и для создания терапевтических и профилактических средств против COVID-19. В этом обзоре представлена общая характеристика SARS-CoV-2 и вызываемого им заболевания, а также проанализированы источники, описывающие генетическое разнообразие вируса и его взаимосвязь с биологическими свойствами SARS-CoV-2, такими, как вирулентность и контагиозность.

**КЛЮЧЕВЫЕ СЛОВА** коронавирусы, SARS-CoV-2, патогенность, вирулентность, контагиозность, эволюция вирусов, вирусный геном.

## ВВЕДЕНИЕ

После появления первых случаев заражения в китайском Ухане в декабре 2019 года новая коронавирусная инфекция COVID-19, вызванная SARS-CoV-2, распространилась во всех странах земного шара и стала первой в истории человечества пандемией, вызванной коронавирусом [1]. По состоянию на март 2021 года COVID-19 диагностирован более чем у 117 млн человек в мире и стал причиной более 2.6 млн смертей [2]. В настоящее время профилактические вакцины доступны не во всех странах и не всем слоям населения, поэтому карантин, социальное дистанцирование и особые санитарные меры предосторожности остаются важнейшими мерами предотвращения распространения инфекции.

В связи с быстрым и широким распространением нового коронавируса неизбежна его дивергенция, появление штаммов с различными биологическими свойствами, наиболее значимым из которых является вирулентность. О фенотипическом разнообразии коронавируса SARS-CoV-2, учитывая короткий пери-

од его изучения, известно очень мало. К сожалению, при описании геномных последовательностей информация о пациенте скудна и ограничивается возрастом и полом, тогда как информация о тяжести, особенностях и исходе заболевания зачастую отсутствует.

Одной из актуальных задач фундаментальной и медицинской вирусологии остается определение природы патогенности и вирулентности вирусов, в том числе животного происхождения. Несмотря на прогресс, достигнутый в понимании эволюции вирусов, вопрос об эволюции вирулентности, возникшей в результате межвидовой передачи, остается дискуссионным. Становится ли вирус более или менее вирулентным в новом хозяине? Каким образом уровень вирулентности модулируется под влиянием естественного отбора и почему? Существуют ли закономерности эволюции вирулентности вирусов в новом хозяине, позволяющие прогнозировать направление этого процесса? Упрощенная интерпретация эволюции вирулентности заключается в том, что естественный отбор будет оптимизировать уровень

вирулентности таким образом, чтобы повысить эффективность передачи вируса, характеризующуюся базовым репродуктивным числом ( $R_0$ ) [3]. Адаптация вируса в новом хозяине формируется сложным набором взаимодействий факторов хозяина-возбудителя. Согласно современным представлениям, при межвидовой передаче первоначальная вирулентность вируса может варьироваться от отсутствия патогенности (бессимптомное носительство) до высокой патогенности, причем прогнозировать направление изменения вирулентности крайне сложно. Встреча человечества с высоковирулентным пандемическим вирусом животного происхождения происходит не очень часто – 1 раз за несколько десятилетий, однако последствия этого тяжелейшие и имеют глобальное значение. В связи с этим крайне важно понимать и прогнозировать, как будут эволюционировать биологические свойства коронавируса SARS-CoV-2. Целью обзора является анализ результатов научных исследований, посвященных поиску взаимосвязи генетических изменений вируса SARS-CoV-2 с его биологическими свойствами, включая патогенность, вирулентность и контагиозность.

Под *патогенностью* вируса понимается его способность вызывать заболевание. Термин «вирулентность» в зависимости от контекста может иметь различные значения. В настоящем обзоре под *вирулентностью* вируса подразумевается мера его патогенности, т.е. способность вызывать более или менее тяжелые заболевания; при этом степень вирулентности определяется коэффициентом летальности. *Контагиозность* (трансмиссивность) вируса – это его способность передаваться от больных организмов здоровым. Мерой контагиозности являются взаимосвязанные показатели: индекс контагиозности (доля лиц из числа восприимчивых, которые заболели после контакта с источником возбудителя) и базовое репродуктивное число  $R_0$  (среднее количество лиц, напрямую инфицированных больных в течение всего заразного периода при условии попадания его в полностью восприимчивую популяцию).

### ОБЩАЯ ХАРАКТЕРИСТИКА SARS-CoV-2

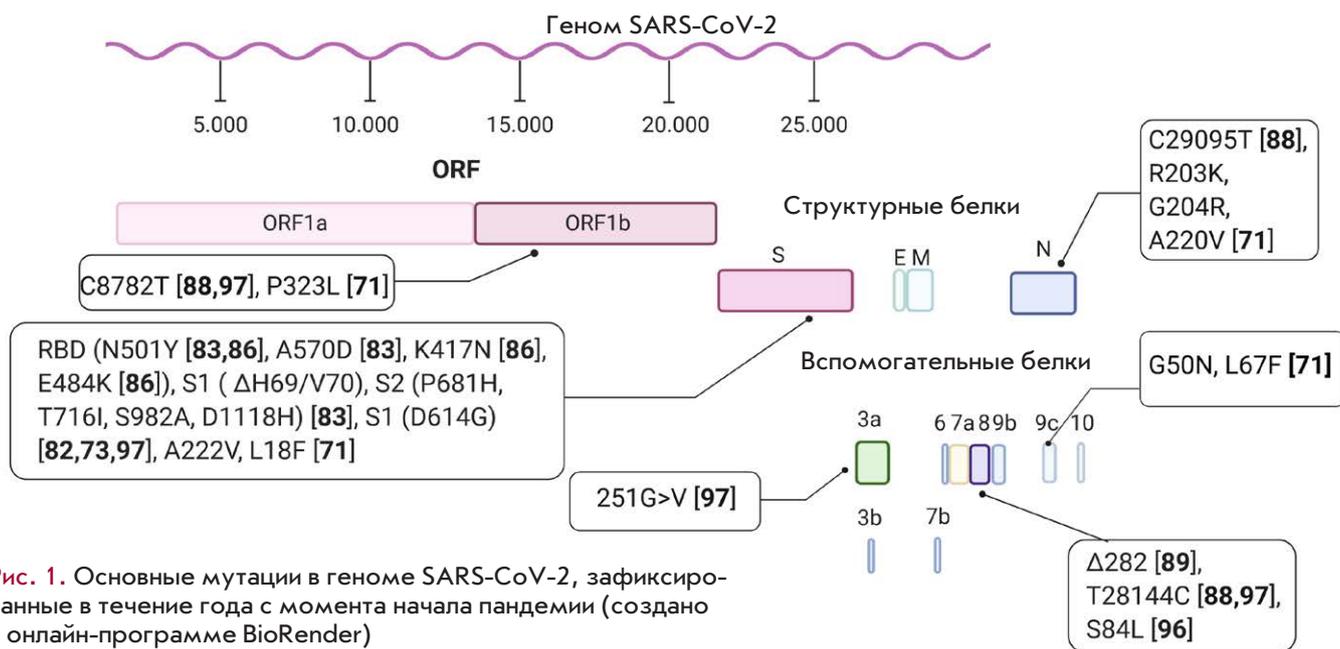
Пандемический SARS-CoV-2, наряду с вирусом SARS-CoV, относится к семейству Coronaviridae, подсемейству Orthocoronavirinae, роду *Betacoronavirus*, подроду *Sarbecovirus*, виду *Severe acute respiratory syndrome-related coronavirus* [4]. Следует отметить, что вместе с перечисленными патогенами к подроду *Sarbecovirus* относятся также коронавирусы, выделенные от летучих мышей, в частности, подковоносов (род *Rhinolophus*) [5]. Установлено, что последовательность генома SARS-CoV-2 на 96.2 и 93.3% совпадает с геномом коронавирусов летучих

мышей RaTG13 [6] и RmYN02 [7] соответственно. Степень сходства нуклеотидных последовательностей и эволюционный анализ лежат в основе гипотезы о том, что именно летучие мыши являются природным резервуаром SARS-CoV-2, который через неизвестных промежуточных хозяев попал в человеческую популяцию [8, 9]. Кроме того, показано, что сходство с геномом коронавирусов, выделенных от панголинов, составляет 85.5–92.4% [10], с геномом SARS-CoV – 80% [6] и MERS-CoV, принадлежащего к подроду *Merbecovirus*, – 50% [11]. Однако степень гомологии генома сильно варьирует в зависимости от генов, а также геномных локусов [5]. При этом основные различия между этими вирусами наблюдаются в последовательности *ORF1a* и гена, кодирующего шиповидный белок S, который играет ключевую роль при взаимодействии вируса с клеткой [12]. Эти особенности организации генома могут быть следствием межвирусной рекомбинации [13].

SARS-CoV-2 представлен вирионами плеоморфной (обычно сферической) формы со средним диаметром  $108 \pm 8$  нм в диапазоне от 84 до 126 нм [14]. На поверхности вирусных частиц расположены выступающие шипы длиной примерно 9–12 нм, что придает вирусу форму короны. Морфология вирионов SARS-CoV-2 имеет сходство с другими представителями семейства Coronaviridae, включая SARS-CoV и MERS-CoV [15].

Геном SARS-CoV-2 представлен несегментированной одноцепочечной РНК позитивной полярности размером 29.9 т.н. и состоит из шести основных открытых рамок считывания (ORF) (рис. 1). Трансляция кодируемой вирусом РНК-зависимой РНК-полимеразы (репликазы) необходима для инициации репликации вируса в клетке и синтеза субгеномных вирусных РНК, которые в свою очередь служат матрицей для синтеза структурных и вспомогательных белков вируса [16]. Размер *ORF1ab*, которая кодирует репликазу, составляет 2/3 от размера всего вирусного генома. За ней следуют гены шиповидного белка (*S*), *ORF3a*, белка оболочки (*E*), белка мембраны (*M*), *ORF6*, *ORF7a*, *ORF7b*, *ORF8*, нуклеокапсида (*N*) и *ORF10*. Кроме того, Nelson и соавт. доказали, что SARS-CoV-2 содержит новый перекрывающийся ген (OLG – overlapping gene) *ORF3d* [17], который имеется и у коронавирусов, выделенных от панголинов в районе Гуанси на юге Китая, но не обнаружен у других коронавирусов, выделенных от панголинов и летучих мышей.

Шиповидный белок S коронавирусов SARS-CoV и SARS-CoV-2 инициирует слияние оболочки вируса с плазматической мембраной клетки-хозяина, а ангиотензинпревращающий фермент 2 (ACE2) служит клеточным рецептором для прикрепления вируса.



**Рис. 1.** Основные мутации в геноме SARS-CoV-2, зафиксированные в течение года с момента начала пандемии (создано в онлайн-программе BioRender)

Рецептором MERS-CoV является hDPP4 (дипептидилпептидаза 4, или CD26 человека) [18]. В S-белке выделяют два домена: S1 и S2. Домен S1 опосредует связывание с ACE2, а S2 – последующее слияние вирусной оболочки с клеточной мембраной [19]. Рецепторсвязывающий домен (RBD) – это ключевой функциональный компонент S1, который отвечает за связывание SARS-CoV-2 с ACE2 [20]. Кроме того, RBD SARS-CoV содержит коровый мотив и рецепторсвязывающий мотив (RBM), который опосредует контакты с ACE2. На поверхности ACE2 расположены две точки, которые SARS-CoV-2 использует для прикрепления [21]. Стадия адсорбции и проникновения SARS-CoV-2 в клетку зависит не только от ACE2-рецептора, но и от трансмембранной сериновой протеазы TMPRSS2 и пропротеинконвертазы фуринов, функция которых состоит в праймировании S-белка SARS-CoV-2 [22, 23]. Таким образом, SARS-CoV-2 может проникать в клетку двумя различными способами (рис. 2): через позднюю эндосому, где S-белок расщепляется катепсинами, или через клеточную мембрану, или через раннюю эндосому, используя трипсиноподобные протеазы для расщепления S-белка [23, 24].

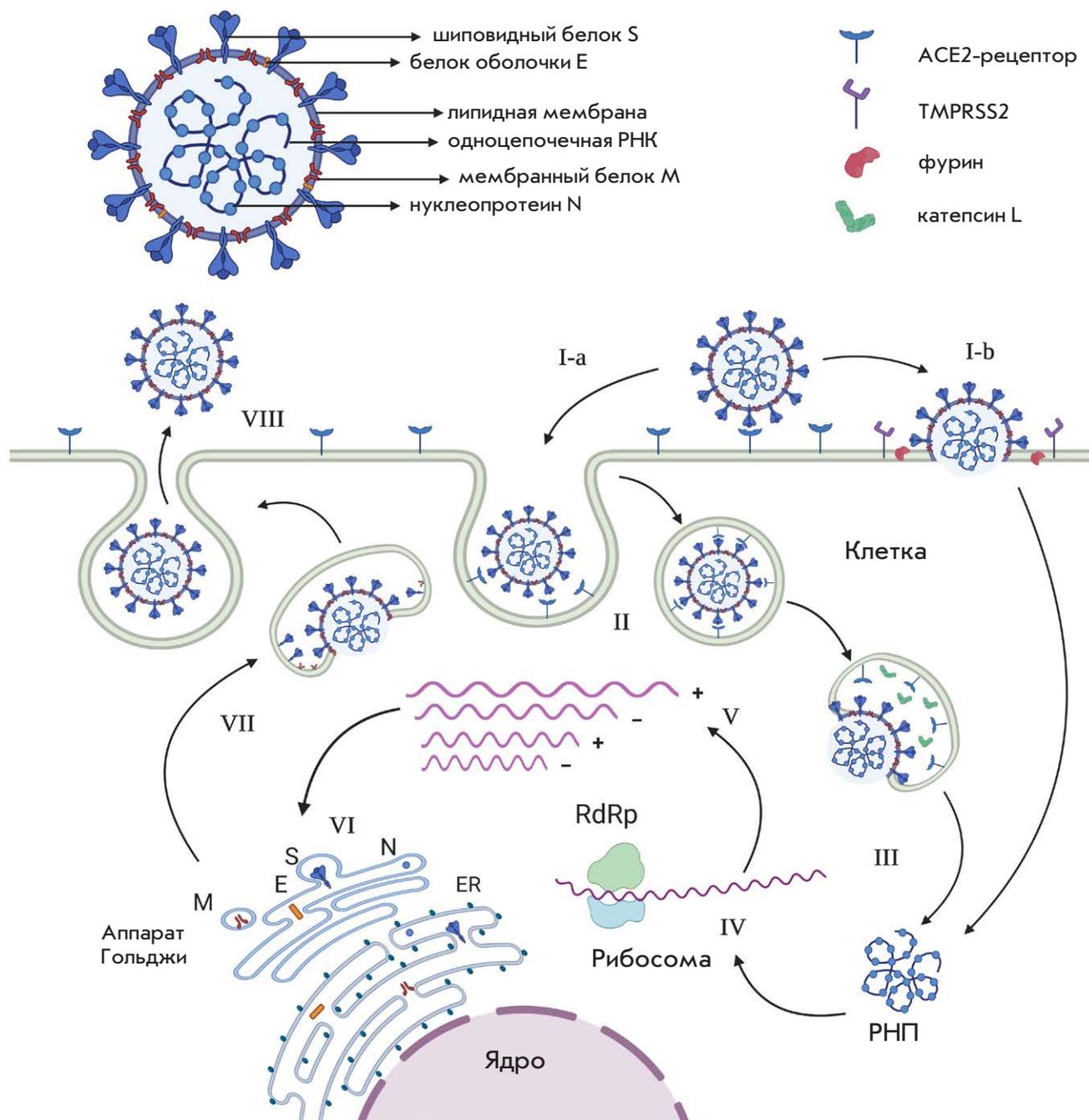
Белок E формирует ионные каналы и регулирует сборку вирионов [25]. Белок M также участвует в сборке вирусных частиц [26], тогда как белок N формирует с вирусной РНК рибонуклеопротеиновый комплекс и выполняет несколько функций, таких, как усиление транскрипции вирусного генома и взаимодействие с вирусным мембранным белком во время сборки вириона [27].

Рецептор клеток-мишеней, который вирус использует для проникновения в клетку, является

фактором, определяющим, какие органы и ткани чувствительны к заражению. Рецептор ACE2 экспрессируется на поверхности эпителиальных клеток альвеол, трахеи, бронхов, бронхиальных желез, а также на альвеолярных макрофагах. Кроме того, ACE2 присутствует на слизистых оболочках, таких, как роговица глаза, бокаловидные и ресничные клетки в носовой полости [28], которые, судя по всему, служат входными воротами инфекции. Патогенетический механизм COVID-19 заключается в том, что вирус проникает в клетку, используя рецептор ACE2, и высвобождает одноцепочечную вирусную РНК, которая соединяется с рибосомой клетки-мишени и инициирует синтез РНК-репликазы, которая в свою очередь воспроизводит копии геномной и субгеномной РНК, а также фрагменты РНК, служащие матрицей для синтеза белков вирусной оболочки. Молекулы геномных вирусных РНК положительной полярности вместе со структурными вирусными белками формируют новые вирионы SARS-CoV-2, которые высвобождаются из клетки и инфицируют интактные клетки-мишени (рис. 2) [29].

### МНОГООБРАЗИЕ КЛИНИЧЕСКИХ ПРОЯВЛЕНИЙ COVID-19

Характер заболевания, вызванного SARS-CoV-2, может существенно различаться [30]. Наблюдается большая вариабельность клинических проявлений COVID-19 даже среди близких контактов инфицированного и членов одной семьи [30]. Спектр симптоматики при COVID-19 колеблется от легкой и умеренной формы до критической и фатальной [31–33].



**Рис. 2.** Строение вириона и основные этапы цикла репродукции SARS-CoV-2 (создано в онлайн-программе BioRender)

- I. Адсорбция вируса. После адсорбции вирус может проникнуть в клетку двумя путями: через эндосому (I-a) или путем слияния с плазматической мембраной (I-b)
- II. Рецептор-опосредованный эндоцитоз
- III. Разделение вируса в результате слияния оболочки вируса с мембраной эндосомы. Высвобождение рибонуклеопротеидного комплекса (РНП)
- IV. Трансляция вирусного генома. Синтез вирусных белков, ответственных за репликацию и транскрипцию генома, включая РНК-зависимую РНК-полимеразу (RdRp)
- V. Репликация и транскрипция вирусного генома
- VI. Синтез вирусных белков на эндоплазматической сети (ЭПС)
- VII. Сборка и транспорт вириона к мембране клетки
- VIII. Выход из клетки путем экзоцитоза

Также часто фиксируются случаи бессимптомного течения заболевания. Установлено, что доля случаев бессимптомного носительства достигает 40–50%, при этом инфицированный человек является источником заражения на протяжении более 14 дней [34]. Кроме того, бессимптомное течение инфекции может быть ассоциировано с субклиническими нарушениями в легких, которые выявляются при проведении компьютерной томографии [34]. Таким образом, SARS-CoV-2 обладает повышенной вирулентностью с тактическим преимуществом – способностью поддерживать передачу от человека к человеку даже при бессимптомном носительстве [35], что позволяет вирусу быстро распространяться.

Согласно докладу Китайского центра по контролю и профилактике заболеваний [36], при анализе 44500 подтвержденных случаев инфекции с оценкой тяжести заболевания установлено, что умеренная форма течения COVID-19 (без пневмонии или со слабовыраженной пневмонией) наблюдается в 81% случаев. Тяжелая форма (с одышкой, гипоксией или >50% поражением легких) зарегистрирована в 14% случаев. Заболевание в критической форме (с дыхательной недостаточностью, шоком или полиорганной дисфункцией) наблюдается в 5% случаев. При этом общий уровень летальности составил 2.3% (среди некритических случаев не зарегистрировано ни одного случая смерти).

Заболевание COVID-19 в тяжелой форме может наблюдаться у практически здоровых людей любого возраста, но встречается преимущественно у лиц старше 65 лет и/или с сопутствующими заболеваниями (сердечно-сосудистые заболевания, сахарный диабет, гипертония, хронические заболевания легких и почек, рак, ожирение, курение) [32, 36, 37], тогда как у большинства молодых взрослых инфекция протекает в легкой форме без осложнений.

Описано несколько осложнений, вызванных COVID-19. Среди них острый респираторный дистресс-синдром – тип дыхательной недостаточности, при которой необходимо проводить реанимационные действия, включая искусственную вентиляцию легких. В такой помощи нуждаются от 12 до 24% госпитализированных пациентов [38, 39]. Также выявляют сердечно-сосудистые [40], тромбоэмболические осложнения [41], воспалительные реакции [42], суперинфекции [43].

Дети наименее подвержены инфицированию вирусом – в общей структуре случаев COVID-19 доля детей варьирует от 1 до 6.3% [44, 45]. Согласно докладу Китайского центра по контролю и профилактике заболеваний, из 72314 случаев заболевания, зарегистрированных по состоянию на 11 февраля 2020 года, только 2% составляли лица в возрасте до 19 лет [36].

У детей с COVID-19 также описан мультисистемный воспалительный синдром с клиническими признаками, сходными с болезнью Кавасаки и синдромом токсического шока [46]. Мониторинг детской заболеваемости, проведенный Мескиной, показал, что доля бессимптомных форм COVID-19 у детей составляет 62%, в том числе у новорожденных – 73.1%, с низкой частотой тяжелых форм – 0.38% [45].

### ГЕНЕТИЧЕСКАЯ ДЕТЕРМИНАЦИЯ БИОЛОГИЧЕСКИХ СВОЙСТВ SARS-CoV-2

Открытым остается вопрос о причинах таких разнообразных клинических проявлений COVID-19 у разных категорий населения. Можно предположить, что это разнообразие зависит от определенных генетических профилей организма-хозяина. В соответствии с этой гипотезой генетическую основу восприимчивости к инфекции можно объяснить полиморфизмом функциональных рецепторов, необходимых для проникновения вируса в клетки-мишени. В частности, наблюдаемая при COVID-19 полиорганная дисфункция, включая фатальные повреждения легких и миокарда, может быть связана с функциональными характеристиками рецепторов ACE2 в популяции [47–49]. Так, Ноу и соавт. на основе анализа ~ 81000 геномов человека исследовали восприимчивость к COVID-19, обусловленную полиморфизмом генов ACE2 и TMPRSS2 (двух ключевых факторов хозяина SARS-CoV-2). Обнаружено, что полиморфизмы (р.) ACE2 (такие, как р.Arg514Gly в африканской/афроамериканской популяции) связаны с сердечно-сосудистыми и легочными заболеваниями вследствие изменения взаимодействий ангиотензиногена и ACE2. Уникальные, но распространенные полиморфизмы (включая р.Val160Met (rs12329760)) в TMPRSS2 потенциально могут быть причиной дифференциальной генетической восприимчивости к COVID-19 [50].

При помощи полногеномного поиска ассоциаций (GWAS) проанализировано 8582968 однонуклеотидных полиморфизмов (SNP) от 1980 пациентов с тяжелой формой COVID-19 из итальянского и испанского эпицентров пандемии в Европе. Исследование не выявило значимых ассоциаций тяжелой формы заболевания с одним геном, скорее с мультигенным кластером на хромосоме 3 (гены *SLC6A20*, *LZTFL1*, *CCR9*, *FYCO1*, *CXCR6* и *XCR1*) [51].

Используя методы NGS, китайские ученые проанализировали генетические профили 332 пациентов с различной степенью тяжести COVID-19. Результаты полногеномного поиска ассоциаций (GWAS) свидетельствуют о том, что наиболее значимый локус, связанный с тяжестью заболевания, находится в *TMEM189-UBE2V1*, который участвует

в сигнальном пути интерлейкина-1 (IL-1). Частота миссенс-вариантов p.Val197Met гена *TMPRSS2*, влияющих на стабильность белка, у пациентов с тяжелой формой инфекции ниже, чем при легком течении заболевания и в общей популяции. Кроме того, установлено, что аллели HLA-A\*11:01, B\*51:01 и C\*14:02 значительно предрасполагают к тяжелому течению COVID-19 [52].

Селективность в отношении генетических профилей организма-хозяина (как фактор вирулентности SARS-CoV-2) может быть одной из особенностей вируса. У вируса гриппа, ставшего причиной глобальной пандемии в 1918 году, такое свойство не описано. Это может быть связано с тем, что в начале XX века уровень техники и знаний не позволял провести исследований на эту тему. Опубликованы данные о восприимчивости к ВИЧ-1, генетически детерминированной вариациями хемокиновых рецепторов хозяина [53]. Эти данные дают основание полагать, что такая избирательность может определять вирулентность и тканевую специфичность и других вирусов, включая SARS-CoV-2.

В исследованиях, направленных на изучение молекулярных механизмов патогенности и контагиозности коронавируса, наибольшее внимание уделяется детерминантам тропизма коронавируса к клеткам дыхательных путей человека, основанного на рецептор-опосредованном проникновении вируса в клетку. Эти детерминанты находятся в поверхностном S-белке коронавируса. Распространено мнение, что именно мутации в эпитопах белка S, ответственных за связывание с вирусными рецепторами, определяют эффективность межвидовой передачи и адаптации вируса к новому хозяину [54]. Экспериментально обоснована способность коронавируса летучих мышей, S-белок которых модифицирован методами обратной генетики, преодолевать видовой барьер (инфицировать клетки человека) [55]. Однако на сегодняшний день отсутствует экспериментальное подтверждение того, что только один S-белок SARS-CoV-2 опосредует контагиозность или высокую вирулентность вируса у человека. Ранее на примере высокопатогенного вируса гриппа А птиц H5N1 доказали, что способность распознавать вирусные рецепторы клеток эпителия дыхательных путей млекопитающих может достигаться введением в гемагглютинин (HA) от двух до четырех аминокислотных замен, важных для связывания HA с  $\alpha$ -2,6-сиаловыми рецепторами [56, 57]. Тем не менее только этих мутаций оказалось недостаточно для того, чтобы вирус приобрел контагиозность и высокую вирулентность для хорьков [56, 57]. Это свидетельствует о том, что дополнительные детерминанты контагиозности и вирулентности вероятно закоди-

рованы во «внутренних генах» вируса. Патогенность вируса опосредована не только его способностью эффективно проникать в клетки-мишени, но и множеством других вирусных факторов. Примером тому служат вакцинные штаммы, входящие в состав живых аттенуированных вакцин. По данным Klimov и соавт. [58], детерминантами аттенуации холодаадаптированного вакцинного штамма вируса гриппа A/Leningrad/134/47/17 являются мутации в генах белков полимеразного комплекса (PB1, PB2, PA, NP), M-белка и неструктурного белка NS2, но не поверхностных белков нейраминидазы N и гемагглютинаина H. Более 10 примеров изменения вирулентности разных вирусов млекопитающих, обусловленных лишь одной или двумя аминокислотными заменами, приведено в обзоре [3]. Подавляющее большинство этих примеров касается РНК-содержащих вирусов (вирусы гриппа А и В, энтеровирусы, вирус Эбола, ВИЧ, вирус лихорадки Западного Нила, вирус болезни Ньюкасла, вирус репродуктивно-респираторного синдрома свиней и др.).

Вирусы, относящиеся к одному биологическому виду, могут значительно различаться вирулентностью, что связано с дивергенцией в ходе эволюции. Смертность при заражении сезонными вирусами гриппа А (вид *Influenza A virus* семейства Orthomyxoviridae) серотипов H3N2, H1N1 составляет 0.04–1.0%, тогда как смертность от заболеваний, вызванных отдельными штаммами вируса птичьего гриппа А, в том числе H5N1, H7N7, H9N2, H7N3 и H7N9, достигала 60% [59, 60]. Коронавирусы человека не являются исключением. Так называемые сезонные коронавирусы (HCoV-NL63, -229E, -OC43, -HKU1) ассоциированы в основном с легкими и среднетяжелыми формами ОРВИ, тогда как коронавирусы животного происхождения – SARS-CoV, MERS-CoV и SARS-CoV-2 – с развитием тяжелого острого респираторного синдрома и более высоким риском летальности (по разным оценкам от 1 до 40% от числа лабораторно подтвержденных случаев).

Группа Кунина недавно выявила возможные генетические детерминанты, определяющие повышенную летальность в результате заражения высоковирулентными коронавирусами – SARS-CoV, MERS-CoV и SARS-CoV-2 – по сравнению с низковирулентными сезонными HCoV-NL63, -229E, -OC43 и -HKU1 [61]. В геноме высокопатогенных коронавирусов SARS-CoV, MERS-CoV и SARS-CoV-2 в результате анализа более 3000 коронавирусных геномов выявили четыре области (три в гене нуклеопротеина N и одну в гене S-белка), значимо отличающиеся по аминокислотным последовательностям от сезонных коронавирусов. Отличия в гене N предположительно определяют усиление сигналов ядер-

ной локализации и экспорта этого белка. Отличия в гене *S* находятся в области узнавания рецепторов и слияния вирусной оболочки с клеточной мембраной и гипотетически отвечают за усиление этапа прикрепления и входа вируса в клетку. Полученные результаты проливают свет на вероятные детерминанты вирулентности коронавирусов, но не имеют пока эмпирического подтверждения, поскольку работа выполнена методами компьютерного анализа.

В геноме любого вируса постоянно появляются новые мутации, часть которых способна изменить биологические свойства вируса, включая степень контагиозности, способность ускользать от иммунного ответа организма-хозяина, вирулентность. РНК-содержащий SARS-CoV-2 характеризуется высокой (но ниже, чем у других РНК-вирусов) частотой мутаций [62].

На сегодняшний день известны сотни тысяч геномных последовательностей коронавируса SARS-CoV-2. Результаты исследований пополняют базу данных геномных последовательностей GISAID, которая на январь 2021 года включает информацию о более чем 323493 последовательностей. В GISAID, помимо SARS-CoV-2, представлены геномные последовательности коронавирусов, выделенных от летучих мышей и панголинов. Основываясь на данных о вирусных последовательностях, а также информации о географическом происхождении образцов из GISAID, другой информационный ресурс Nextstrain (<https://nextstrain.org>) [63] размещает филогенетический, географический и геномный анализ SARS-CoV-2. Благодаря базе данных GISAID и ресурсу Nextstrain за эволюцией вируса можно наблюдать в режиме реального времени. Анализ Nextstrain прогнозирует появление примерно 26 замен в геноме SARS-CoV-2 в год. Учитывая размеры генома SARS-CoV-2 (29.9 т.н.), расчетная скорость эволюции составляет примерно  $0.90 \times 10^{-3}$  замен/сайт/год [5]. Значение этого показателя сопоставимо с показателями, зарегистрированными для SARS-CoV ( $0.80-2.38 \times 10^{-3}$ ) [64], MERS-CoV ( $0.63-1.12 \times 10^{-3}$ ) [65] и HCoV-OC43 ( $0.43 \times 10^{-3}$ ) [66]. Частота мутаций (количество однонуклеотидных замен на сайт за цикл репликации) у SARS-CoV-2 скорее всего будет ниже, чем у других РНК-содержащих вирусов, таких, как, например, вирусы гриппа. Это связано с тем, что геном коронавирусов кодирует 3'-5'-экзорибонуклеазу (nsp14-ExoN), обладающую редактирующей активностью [67]. Этим обусловлена высокая стабильность генома коронавирусов, включая SARS-CoV-2. Анализ нуклеотидных последовательностей 48635 изолятов вируса подтверждает низкую частоту мутаций, которая составила в среднем 7.23 мутации на образец по от-

ношению к эталонным последовательностям генома SARS-CoV-2 (NC\_045512.2) [68].

Кроме того, показано, что геном SARS-CoV-2 имеет гораздо более низкую частоту мутаций и генетическое разнообразие по сравнению с вирусом SARS-CoV, который вызвал вспышку атипичной пневмонии в 2002–2003 годах [69]. Следует также отметить, что RBD-домен S-белка (~90 аминокислот) SARS-CoV-2, который непосредственно реагирует с рецептором ACE2 на поверхности клеток-мишеней, значительно отличается от RBD SARS-CoV, особенно в двух областях, которые взаимодействуют с ACE2, и является самой вариабельной частью SARS-CoV-2 [70]. Последнее предполагает участие нескольких механизмов, задействованных в процессе проникновения вируса в клетку. Определены шесть аминокислотных остатков RBD S-белка (Leu455, Phe486, Gln493, Ser494, Asn501 и Tyr505), которые играют ключевую роль в связывании с ACE2. Причем пять из них отличаются от RBD SARS-CoV, что следует учитывать при разработке специфических противовирусных препаратов, блокирующих вход вируса в клетку [70].

Следует отметить, что многочисленные элементы генома вируса в разной степени подвержены мутациям. Так, на основе анализа около 223000 полноразмерных последовательностей протеома SARS-CoV-2 была рассчитана частота мутаций каждого вирусного белка. Самой высокой частота мутаций была у белков S, NSP12, NS9c и N [71].

Анализ нуклеотидных последовательностей изолятов SARS-CoV-2 выявил несколько областей генома с повышенной частотой мутаций [72–81]. Среди них следует выделить одну – D614G, в С-концевой области домена S1 [72–74, 77, 81]. Показано, что в Европе доминирует мутантный вирус с заменой D614G в S1-домене [68]. Сравнение функциональных свойств S-белка с аспарагиновой кислотой в положении 614 (S<sup>D614</sup>) и глицином (S<sup>G614</sup>) показало, что псевдовirusы, несущие S<sup>G614</sup>, проникают в клетки, экспрессирующие ACE2, более эффективно, чем вирусы с S<sup>D614</sup> [82]. Хотя накопление данных все еще продолжается, однако растущая доля вируса с заменой D614G предполагает, что вирусы с этой мутацией имеют преимущество в виде более эффективной передачи от человека к человеку. Интересно, что эта мутация, по-видимому, не влияет существенно на тяжесть течения заболевания [73, 79].

В декабре 2019 года появились сообщения о выделении в Великобритании нового штамма SARS-CoV-2, характеризующегося повышенным уровнем контагиозности [83]. По данным филогенетического анализа этот штамм формирует отдельный филогенетический кластер (линия B.1.1.7) [84]. В S-белке этого вируса идентифицированы семь характер-

Таблица 1. Наиболее значимые генетические варианты SARS-CoV-2

| Генетический вариант SARS-CoV-2 | Регион, где впервые был зафиксирован | Характерные мутации  | Характерные особенности   |
|---------------------------------|--------------------------------------|--|---|
| V.1.1.7                         | Великобритания                       | S-белок: RBD (N501Y, A570D), S1 ( $\Delta$ H69/V70), S2 (P681H, T716I, S982A и D1118H) | Высокая контагиозность  |
| V.1.351 (N501Y.V2)              | ЮАР                                  | S-белок: RBD (K417N, E484K и N501Y)  | Некоторые вакцины менее эффективны в отношении этого варианта, высокая контагиозность |
| P1-потомок варианта V.1.1.28    | Бразилия                             | S-белок: RBD (E484K, K417T и N501Y).   | Высокая контагиозность  |
| Fin-796H                        | Финляндия                            | S-белок: RBD (E484K, K417T и N501Y)  | Не детектируется в ПЦР  |

ных мутаций: RBD (N501Y, A570D), S1 ( $\Delta$ H69/V70) и S2 (P681H, T716I, S982A и D1118H) [83]. Мутация N501Y в рецепторсвязывающем домене (RBD) обеспечивает повышенную аффинность по отношению с ACE2 человека и мыши [85]. Делеция  $\Delta$ H69/V70 в S1 способствует ускользанию вируса от иммунного ответа. Мутация P681H непосредственно примыкает к участку расщепления фурином между S1 и S2 в S-белке. Кроме того, представлены данные, указывающие на существование нескольких независимых линий SARS-CoV-2, для которых характерна делеция  $\Delta$ H69/V70 в S-белке, и рост циркуляции этих вирусов в некоторых странах Европы, наблюдаемый с августа 2020 года [83].

В январе 2021 года в Южной Африке зафиксировано появление новой линии SARS-CoV-2 (501Y. V2), которая быстро распространилась и стала доминирующей в нескольких регионах страны. Выявлено восемь характерных для этой линии мутаций в S-белке, включая три в RBD (K417N, E484K и N501Y), которые могут иметь функциональное значение. Две из них (E484K и N501Y) расположены в рецепторсвязывающем мотиве (RBM), который непосредственно взаимодействует с ACE2 [86]. Мутация N501Y также характерна для линии V.1.1.7, выявленной в Великобритании. Возможно, эта мутация определяет уровень контагиозности SARS-CoV-2.

Также появились сообщения о новой линии P.1. SARS-CoV-2 в Бразилии [87]. Следует отметить появление конвергентных мутаций, общих для линий P1, V.1.1.7 и V.1.351 (табл. 1). Это мутация N501Y в S-белке и делеция в ORF1b (del11288–11296 (3675–3677 SGF)) общие для P.1 и британского V.1.1.7, а также мутации в RBD (K417N/T, E484K, N501Y), характерные как для P.1, так и для южноафриканского V.1.351.

Набор мутаций/делеций, характерных для линий P.1, V.1.1.7 и V.1.351, возник, вероятно, совершенно независимо. Кроме того, мутации, общие для P.1 и V.1.351, по-видимому, связаны с быстрым увеличением случаев инфекции в местах, где ранее уже был зафиксирован высокий уровень заболеваемости.

Поэтому крайне важно выяснить, существует ли повышенная вероятность повторного заражения людей, перенесших COVID-19 [87]. В средствах массовой информации появилась информация о выделении нового штамма SARS-CoV-2 (Fin-796H), который имеет сходство как с британским, так и с южноафриканским вариантом вируса, но при этом выявление этого варианта при ПЦР-анализе может быть затруднено.

Следует отметить, что мутации в S-гене представляют особый интерес для исследователей. На ресурсе GISAID регулярно обновляются данные о вариантах гена S-белка вируса SARS-CoV-2. Самые распространенные из них на январь 2021 года указаны на рис. 1.

Анализ 95 полноразмерных последовательностей генома SARS-CoV-2, доступных в GenBank на период с декабря 2019 по апрель 2020 года, позволил обнаружить 116 мутаций, из которых наиболее часто встречались 8782C > T в гене *ORF1ab*, 28144T > C в гене *ORF8* и 29095C > T в гене *N*. Предполагается, что выявленные мутации могут влиять на вирулентность и контагиозность SARS-CoV-2 [88].

Еще одна попытка исследовать взаимосвязь определенных мутаций в геноме SARS-CoV-2 с вирулентностью вируса сделана Young и соавт. [89]. В частности, изучено, каким образом делеция нуклеотида в положении 382 ( $\Delta$ 382) в области *ORF8* генома SARS-CoV-2 влияла на клинические особенности инфекции. Установлено, что вариант  $\Delta$ 382 SARS-CoV-2, по-видимому, связан с более легким течением инфекции.

В настоящее время продолжается сбор и анализ данных о взаимосвязи мутаций в геноме SARS-CoV-2 с вирулентностью и контагиозностью вируса. Основные мутации, зафиксированные за год циркуляции пандемического вируса, обозначены на рис. 1. Очевидно, что значительная доля мутаций, которые влияли на трансмиссивность вируса, содержится в гене, кодирующем S-белок. Это очень важное наблюдение должно быть принято во внимание разработчиками вакцин против SARS-CoV-2, подавляющее число которых основано на S-белке [90]. Показано, что сы-

воротки от 20 человек, вакцинированных BNT162b2 (РНК-вакцина, кодирующая S-белок), нейтрализовали псевдовirus SARS-CoV-2 с мутациями N501 и Y501 [91]. Возможно, что при разработке вакцин следует обратить внимание на другие белки вируса, включая белок нуклеокапсида N. Так, например, 90% эпитопов при Т-клеточном ответе расположено в ORF1ab гена белка нуклеокапсида SARS-CoV-2 [92].

### СИСТЕМАТИЗАЦИЯ И ГЕОГРАФИЧЕСКАЯ РАСПРОСТРАНЕННОСТЬ ГЕНЕТИЧЕСКИХ ВАРИАНТОВ SARS-CoV-2

Молекулярно-генетический мониторинг новой коронавирусной инфекции и филогенетический анализ позволили выделить разные генетические варианты SARS-CoV-2, отличающиеся своей географической распространенностью. Существует несколько подходов к сравнительному геномному анализу вариантов SARS-CoV-2. В одном из них, предложенном Forster P. и соавт., выделяют три основных варианта SARS-CoV-2 (A, B, C), которые различаются аминокислотными заменами. При проведении филогенетического анализа близкородственный коронавирус летучих мышей BatCoVraTG13, выделенный в провинции Юньнань [93], определен как предковый и помещен в основу филогенетического дерева (кластер A) [94]. Выделяют два подкластера A, которые различаются синонимичной мутацией T29095C. Вариант B произошел от A в результате двух мутаций: синонимичной T8782C и несинонимичной C28144T, изменяющей лейцин на серин. При этом тип C отличается от своего родительского типа наличием несинонимичной мутации G26144T, которая приводит к замене глицина на валин [94]. Типы A и C встречаются в основном в Европе и Америке. Напротив, тип B наиболее распространен в Восточной Азии, а его предковый геном, видимо, не вышел за пределы Восточной Азии, что предполагает существование иммунологической или экологической устойчивости к этому типу за пределами Азии [94]. Эти исследования дополнены группой ученых из Гонконга [95], которые провели филогенетический и филодинамический анализ 247 геномных последовательностей SARS-CoV-2, доступных в базе данных GISAID на 5 марта 2020 года. Среди них идентифицировали четыре генетических кластера вирусов, названных «суперраспространителями» (super-spreaders, SS), т.е. вирусы, которые вызвали крупные вспышки заболевания COVID-19 в разных странах мира. Так, SS1 был широко распространен в Азии и США и в основном ответствен за вспышки в штатах Вашингтон и Калифорния, а также в Южной Корее, тогда как SS4 способствовал развитию пандемии в Европе. Используя сигнатурные мутации каждого SS в качестве маркеров, допол-

нительно проанализировали 1539 последовательностей генома SARS-CoV-2, зарегистрированных после 29 февраля 2020 года, и обнаружили, что 90% этих геномов были «суперраспространителями», причем доминировал SS4 [95]. Проводя параллели с работой [94], следует отметить, что вирус, идентифицированный как SS1, эквивалентен типу B, SS2 – типу C, а тип A – предковый вариант. Результаты географического распространения разных типов вирусов совпадают в обоих исследованиях.

Популяционный генетический анализ 103 геномов SARS-CoV-2 показал [96], что вирусы можно разделить на два основных типа (L и S), которые различаются двумя точечными мутациями в аминокислотной последовательности сайта 84 (S84L) гена ORF8. Несмотря на то что тип L (~70%) более распространен, чем тип S (~30%), результаты эволюционного анализа предполагают, что тип S, скорее всего, является предковой версией SARS-CoV-2. Кроме того, предполагается, что тип L более агрессивен и распространяется быстрее, чем S, а вмешательство человека, возможно, изменило соотношение L- и S-типов вскоре после первой вспышки SARS-CoV-2. Однако в настоящее время неясно, произошел ли L-тип в результате эволюции S-типа коронавируса человека или промежуточных хозяев. Также неясно, является ли L-тип более вирулентным, чем S-тип [96].

Чтобы оценить связь генетических мутаций с уровнем вирулентности вируса, Zhang и соавт. изучили клинические, молекулярные и иммунологические характеристики 326 пациентов с подтвержденной инфекцией SARS-CoV-2 в Шанхае [97]. Они идентифицировали два основных клайда. Клайд I включал несколько подгрупп, характеризующихся различиями в ORF3a – p.251G> V (субклайд V), или S – p.614D> G (субклайд G). Клайд II отличается от клайда I двумя связанными мутациями в ORF8: p.84L> S (28144T> C) и ORF1ab: p.2839S (8782C> T). Такая классификация не согласуется с S/L-типированием [96], несмотря на то, что она основана на тех же двух связанных полиморфизмах. Кроме того, авторы не выявили существенных различий в частоте мутаций и трансмиссивности у вирусов, относящихся к клайду I или II, а также в клинических особенностях вызываемых ими заболеваний.

Еще один подход к систематизации генетических вариантов SARS-CoV-2 представлен в препринте [98]. При сравнении полноразмерных вирусных геномов рассчитывали коэффициент сходства Жаккара. При этом в анализ не включали информацию о географическом происхождении образцов и не пытались моделировать эволюционные взаимоотношения различных геномов SARS-CoV-2 с помощью филогенетического анализа. Несмотря на это, результаты

Таблица 2. Современные подходы к субвидовой классификации SARS-CoV-2

| Клайды по GISAID [101] и характерные мутации |   | Линии по Rambaut [103] | Клайды по Hodcroft [105] и характерные мутации |                               |
|--|---|------------------------|--|-------------------------------|
| <b>S</b>                                     | C8782T, T28144C, включая NS8-L84S   | <b>A</b>               | <b>19A</b><br><b>19B</b>                       | C8782T<br>T28144C             |
| <b>L</b>                                     | C241, C3037, A23403, C8782, G11083, G26144, T28144 (референсная последовательность - штамм WIV04, GISAID: hCoV-19/Wuhan/WIV04/2019) | <b>B.2</b>             | <b>20A</b>                                     | C3037T, C14408T<br>и A23403G  |
| <b>V</b>                                     | G11083T, G26144T NSP6-L37F + NS3-G251V  | <b>B.1</b>             |  |                               |
| <b>G</b>                                     | C241T, C3037T, A23403G, включая S-D614G   | <b>B.1*</b>            |  |                               |
| <b>GH</b>                                    | C241T, C3037T, A23403G, G25563T, включая S-D614G + NS3-Q57H   | <b>B.1.1.</b>          | <b>20C</b>                                     | C1059T и G25563T              |
| <b>GR</b>                                    | C241T, C3037T, A23403G, G28882A, включая S-D614G + N-G204R  |                        | <b>20B</b>                                     | G28881A, G28882A<br>и G28883C |
| <b>GV</b>                                    | C241T, C3037T, A23403G, C22227T, включая S-D614G + S-A222V  |                        |  |                               |

проведенного анализа отражают хронологическое распространение SARS-CoV-2 по всему земному шару, начиная с первых случаев заболевания, наблюдавшихся в Китае, и заканчивая современными вспышками заболевания в Европе и Северной Америке. Кроме того, использование данных о нуклеотидной последовательности 7 640 геномов SARS-CoV-2, представленных в базе данных GISAID, позволило сделать вывод о том, что вирусы группируются в четыре отдельные генетические подгруппы [98].

Анализ десятков тысяч геномов SARS-CoV-2, проведенный группой ученых из Темпльского университета, позволил выявить его прародительский штамм (препринт [99], опубликованный на сайте bioRxiv.org). Со временем мутации в геноме предкового вируса дали начало семи доминантным линиям, которые распространились по разным континентам. Использование технологии молекулярного баркодирования позволило установить, что геномные последовательности североамериканских коронавирусов отличаются от последовательностей коронавирусов, циркулирующих в то же время в Европе и Азии [99].

На основе анализа 75 полноразмерных геномов выделены шесть кластеров, названных Ухань, Бриллиантовая принцесса, Азиатский, Европейский, США и Пекин [100]. Обнаруженные в образцах из Южной Кореи, Индии, Греции, Испании, Австралии, Швеции и Юньнани мутации в гене, кодирующем шиповидный гликопротеин S, возможно, предполагают преобладание мутировавших штаммов с разной степенью вирулентности.

Несмотря на многообразие подходов к классификации SARS-CoV-2, консорциум GISAID сформировал собственную, обобщенную систему классификации [101], согласно которой выделяют семь основных клайдов (на основе характерных наборов мутаций): S, L, V, G, GH, GR, GV (табл. 2).

Согласно [68], клайды G и GR преобладают в Европе, в то время как S и GH – в основном

в Северной и Южной Америке. Референсный клайд L представлен преимущественно последовательностями из Азии. В настоящее время представители клайда G и его производных – GH и GR, наиболее распространены среди секвенированных геномов SARS-CoV-2, на долю которых в глобальном масштабе приходится 74% всех известных последовательностей. В частности, клайд GR, несущий комбинацию мутаций D614G S-белка и RG203KR N-белка, в настоящее время является самым распространенным представителем SARS-CoV-2 в мире. Исходный вирусный штамм, представленный клайдом L, по-прежнему составляет 7% от числа секвенированных геномов [68].

Проведен анализ 1566 последовательностей генома SARS-CoV-2, выделенных в 10 азиатских странах [102]. Последовательности сравнивали с эталонной последовательностью штамма WIV04 (Accession no MN996528.1) для идентификации мутаций в различных областях генома. Анализ *in silico* показал, что изоляты из 10 азиатских стран формируют клайды G, GH, GR, L, S, O и V. При этом самый высокий уровень мутаций зафиксирован в клайдах GH и GR [102].

Классификация GISAID дополняется более детальной динамической системой номенклатуры, предложенной Rambaut и соавт. [103]. Согласно этой системе, выделяют 81 линию SARS-CoV-2 и большая их часть относится к линиям A, B и B.1. Выделены шесть линий, производных от линии A (A.1–A.6), и две сублинии-потомки A.1 (A.1.1 и A.3). Также описано 16 линий, производных от линии B. Линия B.1, включающая на апрель 2020 года 70 сублиний, является преобладающей. Линия B.2 имеет шесть сублинейных потомков. Согласно этой классификации, клайды S, V, G, GH, GR, GV соответствуют линиям A, B.2, B.1, B.1\*, B.1.1. (табл. 2) [68]. На основе этой системы разработано программное обеспечение pangolin [104], которое позволяет автоматически классифицировать новые геномы.

Еще один подход к систематизации описан в работе Hodcroft с соавт. [105]. Авторы предлагают указывать в названии год выявления штаммов, входящих в определенный клайд. При этом клайд формируется из штаммов, которые циркулируют в течение нескольких месяцев и имеют характерное географическое распространение. Согласно этой классификации, в настоящее время выделяют следующие клайды: 19A, 19B, 20A, 20B, 20C (табл. 1). Клайды 19A и 19B преобладали в Азии в начале пандемии, тогда как 20A был детектирован в Европе в начале 2020 года. Клайд 20B также распространился в Европе, тогда как 20C в основном в Северной Америке.

Таким образом, в настоящее время продолжается формирование удобной и понятной системы классификации пандемического SARS-CoV-2. Следует отметить, что на момент написания обзора не было опубликовано официальных рекомендаций МКТВ (ICTV) по субвидовой таксономии SARS-CoV-2.

В конце января 2020 года в России зафиксировали первые случаи заражения SARS-CoV-2, а с мая 2020 года Россия входит в число четырех стран с наибольшим количеством подтвержденных случаев заболевания COVID-19. На март 2021 года в России зафиксировано 4,3 млн случаев COVID-19 и 87000 смертельных исходов. Однако вспышка в России началась позже, чем во многих соседних европейских странах, возможно, благодаря принятым мерам по ограничению транспортного сообщения с Китаем. Филогенетический анализ изолятов SARS-CoV-2, выделенных в России, показал, что большинство образцов соответствуют линиям B.1, B.1.1 и B.1\* (номенклатура PANGOLIN) или клайдам G, GR и GH (номенклатура GISAID), широко распространенным в Европе [106]. При этом самой распространенной генетической линией является GR / 20B / B.1.1 (номенклатура GISAID, Nextstrain, Pangolin соответственно) [107]. В результате филогенетического анализа российских штаммов установлено, что, как и повсеместно, российские изоляты SARS-CoV-2 характеризовались низкой частотой мутаций. Однако обнаружена высокая частота несинонимичных мутаций, приводящих к неконсервативным заменам.

Большинство несинонимичных замен найдены в нуклеотидных последовательностях, кодирующих нуклеопротеин N. Предполагается, что это наблюдение может служить косвенным доказательством интенсивной циркуляции вируса в человеческой популяции и его адаптации к новым носителям [108].

## ЗАКЛЮЧЕНИЕ

При глобальном распространении SARS-CoV-2 с отчетливым порогом начала пандемии новой для иммунной системы человека вирусной инфекции возникли условия, когда возможно получение достаточно убедительных данных о том, зависит ли структура клинических форм COVID-19 от динамических изменений генетически обусловленных биологических свойств вируса, или же она детерминирована только характеристиками хозяина. Этот вопрос принципиально важен для разработки вакцин и планирования ресурсов общественного здравоохранения. По истечении 12 месяцев от начала распространения нового коронавируса в человеческой популяции все меньше остается сомнений в дивергенции SARS-CoV-2, т.е. в появлении штаммов, отличающихся по биологическим свойствам, что обусловлено высокой пластичностью геномов РНК-вирусов и благоприятными условиями для их эволюции.

Любые изменения вирусного генома, нарушающие взаимодействие с клеткой хозяина, изменяющие условия для репродукции коронавируса, экспрессию генов хозяина или противодействие иммунитету хозяина, могут изменить степень его контагиозности и вирулентности. Более того, биологические свойства вируса может изменить одна или несколько точечных мутаций, как показано в ряде исследований. При этом взаимодействие между коронавирусом и хозяином является ключом к патогенезу коронавирусных заболеваний и, в конечном счете, определяет исход инфекции. ●

*Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 20-04-60079.*

## СПИСОК ЛИТЕРАТУРЫ

- Liu Y.C., Kuo R.L., Shih S.R. // *Biomed. J.* 2020. V. 43. № 4. P. 328–333.
- COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU). <https://gisanddata.maps.arcgis.com/apps/opsdashboard/index.html#/bda7594740fd40299423467b48e9ecf6>.
- Geoghegan J.L., Holmes E.C. // *Nat. Rev. Genet.* 2018. V. 19. № 12. P. 756–769.
- ICTV (International Committee on Taxonomy of Viruses). [https://talk.ictvonline.org/ictv-reports/ictv\\_9th\\_report/positive-sense-rna-viruses-2011/w/posrna\\_viruses/222/coronaviridae](https://talk.ictvonline.org/ictv-reports/ictv_9th_report/positive-sense-rna-viruses-2011/w/posrna_viruses/222/coronaviridae).
- Nakagawa S., Miyazawa T. // *Inflamm. Regen.* 2020. V. 40. № 1. P. 17.
- Zhou P., Yang X.L., Wang X.G., Hu B., Zhang L., Zhang W., Si H.R., Zhu Y., Li B., Huang C.L., et al. // *Nature.* 2020. V. 579. № 7798. P. 270–273.
- Zhou H., Chen X., Hu T., Li J., Song H., Liu Y., Wang P., Liu D., Yang J., Holmes E.C., et al. // *Curr. Biol.* 2020. V. 30. № 11. P. 2196–2203 e2193.
- Guo Y.R., Cao Q.D., Hong Z.S., Tan Y.Y., Chen S.D., Jin H.J., Tan K.S., Wang D.Y., Yan Y. // *Mil. Med. Res.* 2020. V. 7. № 1. P. 11.
- Yuan S., Jiang S.C., Li Z.L. // *Front. Vet. Sci.* 2020. V. 7. P. 379.
- Lam T.T., Jia N., Zhang Y.W., Shum M.H., Jiang J.F., Zhu H.C., Tong Y.G., Shi Y.X., Ni X.B., Liao Y.S., et al. // *Nature.* 2020.

- V. 583. № 7815. P. 282–285.
11. Wu C., Liu Y., Yang Y., Zhang P., Zhong W., Wang Y., Wang Q., Xu Y., Li M., Li X., et al. // *Acta Pharm. Sin B*. 2020. V. 10. № 5. P. 766–788.
  12. Chen Y., Liu Q., Guo D. // *J. Med. Virol.* 2020. V. 92. № 4. P. 418–423.
  13. Li X., Giorgi E.E., Marichanegowda M.H., Foley B., Xiao C., Kong X.P., Chen Y., Gnanakaran S., Korber B., Gao F. // *Sci. Adv.* 2020. V. 6. № 27. P. eabb9153.
  14. Liu C., Mendonca L., Yang Y., Gao Y., Shen C., Liu J., Ni T., Ju B., Liu C., Tang X., et al. // *Structure*. 2020. V. 28. № 11. P. 1218–1224 e1214.
  15. Zhu N., Zhang D., Wang W., Li X., Yang B., Song J., Zhao X., Huang B., Shi W., Lu R., et al. // *N. Engl. J. Med.* 2020. V. 382. № 8. P. 727–733.
  16. Lim Y.X., Ng Y.L., Tam J.P., Liu D.X. // *Diseases*. 2016. V. 4. № 3. P. 26.
  17. Nelson C.W., Ardern Z., Goldberg T.L., Meng C., Kuo C.-H., Ludwig C., Kolokotronis S.-O., Wei X. // *eLife*. 2020. V. 9. P. e59633.
  18. Davidson A.M., Wysocki J., Batlle D. // *Hypertension*. 2020. V. 76. № 5. P. 1339–1349.
  19. Li W., Choe H., Farzan M. // *Nidoviruses*. 2006. doi: 10.1007/978-0-387-33012-9-36
  20. Walls A.C., Park Y.J., Tortorici M.A., Wall A., McGuire A.T., Veesler D. // *Cell*. 2020. V. 181. № 2. P. 281–292 e6.
  21. Shang J., Ye G., Shi K., Wan Y., Luo C., Aihara H., Geng Q., Auerbach A., Li F. // *Nature*. 2020. V. 581. № 7807. P. 221–224.
  22. Hoffmann M., Kleine-Weber H., Schroeder S., Kruger N., Herrler T., Erichsen S., Schiergens T.S., Herrler G., Wu N.H., Nitsche A., et al. // *Cell*. 2020. V. 181. № 2. P. 271–280 e8.
  23. Bestle D., Heindl M.R., Limburg H., Van Lam van T., Pilgram O., Moulton H., Stein D.A., Hards K., Eickmann M., Dolnik O., et al. // *Life Sci. Alliance*. 2020. V. 3. № 9. P. e202000786.
  24. Tang T., Bidon M., Jaimes J.A., Whittaker G.R., Daniel S. // *Antiviral Res.* 2020. V. 178. P. 104792.
  25. Ruch T.R., Machamer C.E. // *Viruses*. 2012. V. 4. № 3. P. 363–382.
  26. Neuman B.W., Kiss G., Kunding A.H., Bhella D., Baksh M.F., Connelly S., Droese B., Klaus J.P., Makino S., Sawicki S.G., et al. // *J. Struct. Biol.* 2011. V. 174. № 1. P. 11–22.
  27. McBride R., van Zyl M., Fielding B.C. // *Viruses*. 2014. V. 6. № 8. P. 2991–3018.
  28. Sungnak W., Huang N., Becavin C., Berg M., Queen R., Litvinukova M., Talavera-Lopez C., Maatz H., Reichart D., Sampaziotis F., et al. // *Nat. Med.* 2020. V. 26. № 5. P. 681–687.
  29. Zhang X.Y., Huang H.J., Zhuang D.L., Nasser M.I., Yang M.H., Zhu P., Zhao M.Y. // *Infect. Dis. Poverty*. 2020. V. 9. № 1. P. 99.
  30. Khan F.A. // *New Microbes New Infect.* 2020. V. 36. P. 100697.
  31. Russell C.D., Millar J.E., Baillie J.K. // *Lancet*. 2020. V. 395. № 10223. P. 473–475.
  32. Yang X., Yu Y., Xu J., Shu H., Xia J.a., Liu H., Wu Y., Zhang L., Yu Z., Fang M., et al. // *Lancet Respiratory Med.* 2020. V. 8. № 5. P. 475–481.
  33. Liu K., Fang Y.Y., Deng Y., Liu W., Wang M.F., Ma J.P., Xiao W., Wang Y.N., Zhong M.H., Li C.H., et al. // *Chin. Med. J. (Engl.)*. 2020. V. 133. № 9. P. 1025–1031.
  34. Oran D.P., Topol E.J. // *Ann. Intern. Med.* 2020. V. 173. № 5. P. 362–367.
  35. Abduljalil J.M., Abduljalil B.M. // *New Microbes New Infect.* 2020. V. 35. P. 100672.
  36. Wu Z., McGoogan J.M. // *J. Am. Med. Ass.* 2020. V. 323. № 13. P. 1239–1242.
  37. Harrison S.L., Fazio-Eynullayeva E., Lane D.A., Underhill P., Lip G.Y.H. // *PLoS Med.* 2020. V. 17. № 9. P. e1003321.
  38. Richardson S., Hirsch J.S., Narasimhan M., Crawford J.M., McGinn T., Davidson K.W., the Northwell C.-R.C., Barnaby D.P., Becker L.B., Chelico J.D., et al. // *J. Am. Med. Ass.* 2020. V. 323. № 20. P. 2052–2059.
  39. Petrilli C.M., Jones S.A., Yang J., Rajagopalan H., O'Donnell L., Chernyak Y., Tobin K.A., Cerfolio R.J., Francois F., Horwitz L.I. // *Brit. Med. J.* 2020. V. 369. P. m1966.
  40. Wang D., Hu B., Hu C., Zhu F., Liu X., Zhang J., Wang B., Xiang H., Cheng Z., Xiong Y., et al. // *J. Am. Med. Ass.* 2020. V. 323. № 11. P. 1061–1069.
  41. Zhang Y., Xiao M., Zhang S., Xia P., Cao W., Jiang W., Chen H., Ding X., Zhao H., Zhang H., et al. // *N. Engl. J. Med.* 2020. V. 382. № 17. P. e38.
  42. Mehta P., McAuley D.F., Brown M., Sanchez E., Tattersall R.S., Manson J.J., Hlh Across Speciality Collaboration U.K. // *Lancet*. 2020. V. 395. № 10229. P. 1033–1034.
  43. Rawson T.M., Moore L.S.P., Zhu N., Ranganathan N., Skolimowska K., Gilchrist M., Satta G., Cooke G., Holmes A. // *Clin. Infect. Dis.* 2020. V. 71. № 9. P. 2459–2468.
  44. Ludvigsson J.F. // *Acta Paediatrica*. 2020. V. 109. № 6. P. 1088–1095.
  45. Meskina E.R. // *J. Microbiol. Epidemiol. Immunobiol.* 2020. V. 97. № 3. P. 202–215.
  46. Ahmed M., Advani S., Moreira A., Zoretic S., Martinez J., Chorath K., Acosta S., Naqvi R., Burmeister-Morton F., Burmeister F., et al. // *Eclin. Med.* 2020. V. 26. P. 100527.
  47. Hamming I., Timens W., Bulthuis M.L., Lely A.T., Navis G., van Goor H. // *J. Pathol.* 2004. V. 203. № 2. P. 631–637.
  48. LoPresti M., Beck D.B., Duggal P., Cummings D.A.T., Solomon B.D. // *Am. J. Hum. Genet.* 2020. V. 107. № 3. P. 381–402.
  49. Fujikura K., Uesaka K. // *J. Clin. Pathol.* 2021. V. 74. № 5. P. 307–313.
  50. Hou Y., Zhao J., Martin W., Kallianpur A., Chung M.K., Jehi L., Sharifi N., Erzurum S., Eng C., Cheng F. // *BMC Med.* 2020. V. 18. № 1. P. 216.
  51. Ellinghaus D., Degenhardt F., Bujanda L., Buti M., Albillos A., Invernizzi P., Fernandez J., Prati D., Baselli G., Asselta R., et al. // *N. Engl. J. Med.* 2020. V. 383. № 16. P. 1522–1534.
  52. Wang F., Huang S., Gao R., Zhou Y., Lai C., Li Z., Xian W., Qian X., Li Z., Huang Y., et al. // *Cell Discov.* 2020. V. 6. № 1. P. 83.
  53. Zimmerman P.A., Buckler-White A., Alkhatib G., Spalding T., Kubofcik J., Combadiere C., Weissman D., Cohen O., Rubbert A., Lam G., et al. // *Mol. Med.* 1997. V. 3. № 1. P. 23–36.
  54. de Wilde A.H., Snijder E.J., Kikkert M., van Hemert M.J. // *Curr. Top. Microb. Immunol.* 2018. V. 419. P. 1–42.
  55. Menachery V.D., Yount B.L., Debbink K., Agnihothram S., Gralinski L.E., Plante J.A., Graham R.L., Scobey T., Ge X.-Y., Donaldson E.F., et al. // *Nat. Med.* 2015. V. 21. № 12. P. 1508–1513.
  56. Imai M., Watanabe T., Hatta M., Das S.C., Ozawa M., Shinya K., Zhong G., Hanson A., Katsura H., Watanabe S., et al. // *Nature*. 2012. V. 486. № 7403. P. 420–428.
  57. Maines T.R., Chen L.-M., Van Hoesven N., Tumpey T.M., Blixt O., Belser J.A., Gustin K.M., Pearce M.B., Pappas C., Stevens J., et al. // *Virology*. 2011. V. 413. № 1. P. 139–147.
  58. Klimov A.I., Cox N.J., Yotov W.V., Rocha E., Alexandrova G.I., Kendal A.P. // *Virology*. 1992. V. 186. № 2. P. 795–797.
  59. Lin J.-H., Wu H.-S. // *BioMed Research International*. 2015. V. 2015. P. 805306.
  60. Jernigan D.B., Cox N.J. // *Annu. Rev. Med.* 2015. V. 66. № 1. P. 361–371.
  61. Gussow A.B., Auslander N., Faure G., Wolf Y.I., Zhang F., Koonin E.V. // *Proc. Natl. Acad. Sci. USA*. 2020. V. 117. № 26. P. 15193–15199.

62. Islam M.R., Hoque M.N., Rahman M.S., Alam A., Akther M., Puspo J.A., Akter S., Sultana M., Crandall K.A., Hossain M.A. // *Sci. Rep.* 2020. V. 10. № 1. P. 14004.
63. Hadfield J., Megill C., Bell S.M., Huddleston J., Potter B., Callender C., Sagulenko P., Bedford T., Neher R.A. // *Bioinformatics.* 2018. V. 34. № 23. P. 4121–4123.
64. Zhao Z., Li H., Wu X., Zhong Y., Zhang K., Zhang Y.P., Boerwinkle E., Fu Y.X. // *BMC Evol. Biol.* 2004. V. 4. P. 21.
65. Cotten M., Watson S.J., Zumla A.I., Makhdoom H.Q., Palser A.L., Ong S.H., Al Rabeeah A.A., Alhakeem R.F., Assiri A., Al-Tawfiq J.A., et al. // *mBio.* 2014. V. 5. № 1. P. e01062–01013.
66. Vijgen L., Keyaerts E., Moës E., Thoelen I., Wollants E., Lemey P., Vandamme A.-M., van Ranst M. // *J. Virol.* 2005. V. 79. № 3. P. 1595–1604.
67. Smith E.C., Blanc H., Surdel M.C., Vignuzzi M., Denison M.R. // *PLoS Pathog.* 2013. V. 9. № 8. P. e1003565.
68. Mercatelli D., Giorgi F.M. // *Front. Microbiol.* 2020. V. 11. P. 1800.
69. Jia Y., Shen G., Zhang Y., Huang K.-S., Ho H.-Y., Hor W.-S., Yang C.-H., Li C., Wang W.-L. // *bioRxiv.* 2020. doi: 10.1101/2020.04.09.034942.
70. Naqvi A.A.T., Fatima K., Mohammad T., Fatima U., Singh I.K., Singh A., Atif S.M., Hariprasad G., Hasan G.M., Hassan M.I. // *Biochim. Biophys. Acta Mol. Basis. Dis.* 2020. V. 1866. № 10. P. 165878.
71. Vilar S., Isom D.G. // *bioRxiv.* 2020. doi: 10.1101/2020.12.16.423071.
72. Morais I.J., Poveiro R.C., Souza G.M., Bortolin D.I., Sassaki F.T., Lima A.T.M. // *Sci. Rept.* 2020. V. 10. № 1. P. 18289.
73. Korber B., Fischer W.M., Gnanakaran S., Yoon H., Theiler J., Abfalterer W., Hengartner N., Giorgi E.E., Bhattacharya T., Foley B., et al. // *Cell.* 2020. V. 182. № 4. P. 812–827 e819.
74. Laha S., Chakraborty J., Das S., Manna S.K., Biswas S., Chatterjee R. // *Infection, Genetics and Evolution.* 2020. V. 85. P. 104445.
75. Alouane T., Laamarti M., Essabbar A., Hakmi M., Bouricha E.M., Chemaou-Elfihri M.W., Kartti S., Boumajdi N., Bendani H., Laamarti R., et al. // *Pathogens.* 2020. V. 9. № 10. P. 829.
76. Lokman S.M., Rasheduzzaman M., Salauddin A., Barua R., Tanzina A.Y., Rumi M.H., Hossain M.I., Siddiki A.M.A.M.Z., Mannan A., Hasan M.M. // *Infect. Genet. Evol.* 2020. V. 84. P. 104389.
77. Comandatore F., Chiodi A., Gabrieli P., Biffignandi G.B., Perini M., Ricagno S., Mascolo E., Petazzoni G., Ramazzotti M., Rimoldi S.G., et al. // *bioRxiv.* 2020. doi: 10.1101/2020.04.30.071027.
78. Zhang L., Wang S., Ren Q., Yang J., Lu Y., Zhang L., Gai Z. // *medRxiv.* 2020. doi: 10.1101/2020.04.27.20081349.
79. van Dorp L., Richard D., Tan C.C., Shaw L.P., Acman M., Balloux F. // *bioRxiv.* 2020. doi: 10.1101/2020.05.21.108506.
80. Bhowmik D., Pal S., Lahiri A., Talukdar A., Paul S. // *bioRxiv.* 2020. doi: 10.1101/2020.04.26.062471.
81. Cortey M., Li Y., Díaz I., Clilverd H., Darwich L., Mateu E. // *bioRxiv.* 2020. doi: 10.1101/2020.05.16.099499.
82. Zhang L., Jackson C.B., Mou H., Ojha A., Peng H., Quinlan B.D., Rangarajan E.S., Pan A., Vanderheiden A., Suthar M.S., et al. // *Nat. Commun.* 2020. V. 11. № 1. P. 6013.
83. Kemp S., Harvey W., Datir R., Collier D., Ferreira I., Carabelli A., Robertson D., Gupta R. // *bioRxiv.* 2020. doi: 10.1101/2020.12.14.422555.
84. Preliminary genomic characterisation of an emergent SARS-CoV-2 lineage in the UK defined by a novel set of spike mutations. 2020. <https://virological.org/t/preliminary-genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-the-uk-defined-by-a-novel-set-of-spike-mutations/563>.
85. Starr T.N., Greaney A.J., Hilton S.K., Crawford K.H.D., Navarro M.J., Bowen J.E., Tortorici M.A., Walls A.C., Veasley D., Bloom J.D. // *bioRxiv.* 2020. doi: 10.1101/2020.06.17.157982.
86. Tegally H., Wilkinson E., Giovanetti M., Iranzadeh A., Fonseca V., Giandhari J., Doolabh D., Pillay S., San E.J., Msomi N., et al. // *medRxiv.* 2020. doi: 10.1101/2020.12.21.20248640.
87. Genomic characterisation of an emergent SARS-CoV-2 lineage in Manaus: preliminary findings. 2021. <https://virological.org/t/genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-manaus-preliminary-findings/586>.
88. Khailany R.A., Safdar M., Ozaslan M. // *Gene Rep.* 2020. V. 19. P. 100682.
89. Young B.E., Fong S.-W., Chan Y.-H., Mak T.-M., Ang L.W., Anderson D.E., Lee C.Y.-P., Amrun S.N., Lee B., Goh Y.S., et al. // *Lancet.* 2020. V. 396. № 10251. P. 603–611.
90. Krammer F. // *Nature.* 2020. V. 586. № 7830. P. 516–527.
91. Xie X., Zou J., Fontes-Garfias C.R., Xia H., Swanson K.A., Cutler M., Cooper D., Menachery V.D., Weaver S., Dormitzer P.R., et al. // *bioRxiv.* 2021. doi: 10.1101/2021.01.07.425740.
92. Ferretti A.P., Kula T., Wang Y., Nguyen D.M.V., Weinheimer A., Dunlap G.S., Xu Q., Nabilsi N., Perullo C.R., Cristofaro A.W., et al. // *Immunity.* 2020. V. 53. № 5. P. 1095–1107 e1093.
93. Wahba L., Jain N., Fire A.Z., Shoura M.J., Artiles K.L., McCoy M.J., Jeong D.E. // *mSphere.* 2020. V. 5. № 3. P. e00160–20.
94. Forster P., Forster L., Renfrew C., Forster M. // *Proc. Natl. Acad. Sci. USA.* 2020. V. 117. № 17. P. 9241–9243.
95. Yang X., Dong N., Chan E.W., Chen S. // *Emerg. Microbes Infect.* 2020. V. 9. № 1. P. 1287–1299.
96. Tang X., Wu C., Li X., Song Y., Yao X., Wu X., Duan Y., Zhang H., Wang Y., Qian Z., et al. // *Nat. Sci. Rev.* 2020. V. 7. № 6. P. 1012–1023.
97. Zhang X., Tan Y., Ling Y., Lu G., Liu F., Yi Z., Jia X., Wu M., Shi B., Xu S., et al. // *Nature.* 2020. V. 583. № 7816. P. 437–440.
98. Hahn G., Lee S., Weiss S.T., Lange C. // *bioRxiv.* 2020. doi: 10.1101/2020.05.05.079061.
99. Kumar S., Tao Q., Weaver S., Sanderford M., Caraballo-Ortiz M.A., Sharma S., Pond S.L.K., Miura S. // *bioRxiv.* 2020. doi: 10.1101/2020.09.24.311845.
100. Sundru Manjulata D., Annapurna P., Balakuntalam K., Kumar S. // *Res. Square.* 2021. doi: 10.21203/rs.3.rs-29557/v1.
101. GISAID: Clade and lineage nomenclature, July 4, 2020. <https://www.gisaid.org/references/statements-clarifications/clade-and-lineage-nomenclature-aids-in-genomic-epidemiology-of-active-hcov-19-viruses/>
102. Sengupta A., Hassan S.S., Choudhury P.P. // *bioRxiv.* 2020. doi: 10.1101/2020.11.30.402487.
103. Rambaut A., Holmes E.C., Hill V., O’Toole Á., McCrone J., Ruis C., du Plessis L., Pybus O.G. // *bioRxiv.* 2020. doi: 10.1101/2020.04.17.046086.
104. Phylogenetic Assignment of Named Global Outbreak LINEages (PANGOLIN) updated 2020. <https://github.com/cov-lineages/pangolin>.
105. Hodcroft E.B., Hadfield J., Neher R.A., Bedford T. Year-letter genetic clade naming for SARS-CoV-2 on nextstrain.org. 2020. <https://nextstrain.org/blog/2020-06-02-SARSCoV2-clade-naming>
106. Komissarov A.B., Safina K.R., Garushyants S.K., Fadeev A.V., Sergeeva M.V., Ivanova A.A., Danilenko D.M., Lioznov D., Shneider O.V., Shvyrev N., et al. // *Nat. Commun.* 2021. V. 12. № 1. P. 649.
107. Shchetinin A.M., Tsyganova E.V., Protsenko D.N. // *Cureus.* 2021. V. 13. № 3. P. e13733.
108. Kozlovskaya L., Pinaeva A., Ignatyev G., Selivanov A., Shishova A., Kovpak A., Gordeyuk I., Ivin Y., Berestovskaya A., Prokhortchouk E., et al. // *Int. J. Infect. Dis.* 2020. V. 99. P. 40–46.