

DOI: 10.21055/0370-1069-2020-3-70-82

УДК 616.98:578.2

Я.М. Краснов¹, А.Ю. Попова^{2,3}, В.А. Сафронов¹, А.В. Федоров¹, Д.В. Баданин¹, С.А. Щербакова¹,
В.В. Кутырев¹**АНАЛИЗ ГЕНОМНОГО РАЗНООБРАЗИЯ SARS-CoV-2 И ЭПИДЕМИОЛОГИЧЕСКИХ ПРИЗНАКОВ АДАПТАЦИИ ВОЗБУДИТЕЛЯ COVID-19 К ЧЕЛОВЕЧЕСКОЙ ПОПУЛЯЦИИ (СООБЩЕНИЕ 1)**¹ФКУЗ «Российский научно-исследовательский противочумный институт «Микроб», Саратов, Российская Федерация;²Федеральная служба по надзору в сфере защиты прав потребителей и благополучия человека, Москва, Российская Федерация;³Российская медицинская академия непрерывного профессионального образования, Москва, Российская Федерация

Цель – изучение эволюции генома вируса SARS-CoV-2 при адаптации к организму человека в период текущей пандемии. **Материалы и методы.** Использовалась база данных (GISAID) по нуклеотидным последовательностям геномов вируса SARS-CoV-2, полученных из образцов клинического материала в период с конца декабря 2019 по июль 2020 г. Построение филогенетических деревьев проводили с помощью программного обеспечения BioNumerics 7.6 с использованием алгоритма максимальной экономии (maximum parsimony). **Результаты и обсуждение.** Самым существенным устойчивым изменением в геномах вируса SARS-CoV-2 являются связанные единичные мутации в гене ORF1b (P314L) и гене S (D614G), в результате чего к настоящему времени подавляющее большинство выявляемых изолятов этого вируса имеет данную пару замен. Замена в гене S (D614G) многими исследователями ассоциируется со снижением патогенности содержащих ее штаммов, что, однако, может быть объяснено и улучшением методологии лечения больных в ходе пандемии. Влияние мутации в гене ORF1b (P314L) еще не изучено. Мутации P314L и D614G тесно связаны, и только их совместное присутствие в геноме дало преимущество для распространения данным геновариантам вируса SARS-CoV-2. Анализ обобщенных эпидемиологических данных свидетельствует о том, что распространение новых геновариантов вируса может быть связано с их биологическими свойствами, облегчающими передачу от человека к человеку, при этом ассоциированное снижение летальности может отражать не только прогресс в методах лечения, но и вероятное ослабление вирулентных свойств. Наблюдаемое увеличение потенциала распространения на фоне признаков снижения вирулентности, вероятно, является основной формой адаптации нового коронавируса к человеческой популяции и, по-видимому, будет продолжаться в дальнейшем в виде интеграции SARS-CoV-2 в структуру сезонных возбудителей ОРВИ.

Ключевые слова: COVID-19, SARS-CoV-2, пандемия, передача от человека человеку, вирулентность, эволюция возбудителя.

Корреспондирующий автор: Краснов Ярослав Михайлович, e-mail: rusrapr@microbe.ru.

Для цитирования: Краснов Я.М., Попова А.Ю., Сафронов В.А., Федоров А.В., Баданин Д.В., Щербакова С.А., Кутырев В.В. Анализ геномного разнообразия SARS-CoV-2 и эпидемиологических признаков адаптации возбудителя COVID-19 к человеческой популяции (Сообщение 1). Проблемы особо опасных инфекций. 2020; 3:70–82 DOI: 10.21055/0370-1069-2020-3-70-82

Поступила 16.09.20. Отправлена на доработку 21.09.20. Принята к публ. 23.09.20.

Ya.M. Krasnov¹, A.Yu. Popova^{2,3}, V.A. Safronov¹, A.V. Fedorov¹, D.V. Badanin¹, S.A. Shcherbakova¹,
V.V. Kutyrev¹**Genomic Diversity Analysis of SARS-CoV-2 and Epidemiological Features of Adaptation of COVID-19 Agent to Human Population (Communication 1)**¹Russian Research Anti-Plague Institute “Microbe”, Saratov, Russian Federation;²Federal Service for Surveillance in the Sphere of Consumers Rights Protection and Human Welfare, Moscow, Russian Federation;³Russian Medical Academy of Continuing Professional Education, Moscow, Russian Federation

Abstract. Objective: to study the evolution of SARS-CoV-2 virus in the process of adapting to human organism during the current pandemic. **Materials and methods.** Database (GISAID) on nucleotide sequences of SARS-CoV-2 virus genome, obtained from clinical samples during the period of late December, 2019–July, 2020. Phylogenetic tree diagram construction was carried out applying BioNumerics v.7.6 software using Maximum parsimony algorithm. **Results and discussion.** The most substantial change in the genomes of SARS-CoV-2 virus are associated one-time mutations in ORF1b (P314L) and S (D614G) genes, as a result of which the overwhelming majority of identified isolates of this virus have the stated pair of substitutions to date. Many researches link the substitution in S (D614G) gene to the decreased pathogenicity in the strains that contain it, which may be also explained by enhanced methodology of patient treatment in the course of pandemic. The effect of the mutation in ORF1b (P314L) gene has not yet been investigated. P314L and D614G mutations are closely related and only their combined presence in the genome favored the dissemination of the genovariants of SARS-CoV-2 virus. Analysis of congregated epidemiological data testifies to the fact that the spread of new genovariants may be associated with biological properties facilitating human-to-human transmission. Thereat, associated decrease in lethality may reflect not only advancements in methods of treatment, but possible attenuation of virulent properties. Thus, observed growth in dissemination potential against the background of decrease in virulence is, probably, one of the forms of adaptation of new coronavirus to human population and, apparently, will remain in the future as the integration of SARS-CoV-2 virus into the structure of seasonal ARVI agents.

Key words: COVID-19, SARS-CoV-2, pandemic, human-to-human transmission, virulence, evolution of the agent.

Conflict of interest: The authors declare no conflict of interest.

Corresponding author: Yaroslav M. Krasnov, e-mail: rusrapi@microbe.ru.

Citation: Krasnov Ya.M., Popova A.Yu., Safronov V.A., Fedorov A.V., Badanin D.V., Shcherbakova S.A., Kutyrev V.V. Genomic Diversity Analysis of SARS-CoV-2 and Epidemiological Features of Adaptation of COVID-19 Agent to Human Population (Communication 1). *Problemy Osobo Opasnykh Infektsii [Problems of Particularly Dangerous Infections]*. 2020; 3:70–82. (In Russian). DOI: 10.21055/0370-1069-2020-3-70-82

Received 16.09.20. Revised 21.09.20. Accepted 23.09.20.

Krasnov Ya.M., ORCID: <https://orcid.org/0000-0002-4909-2394>

Popova A.Yu., ORCID: <https://orcid.org/0000-0002-4315-5307>

Safronov V.A., ORCID: <https://orcid.org/0000-0001-9563-2833>

Shcherbakova S.A., ORCID: <https://orcid.org/0000-0003-1143-4069>

Kutyrev V.V. ORCID: <https://orcid.org/0000-0003-3788-3452>

Коронавирусы вызывали эпидемические осложнения задолго до развивающейся в настоящее время пандемии COVID-19. В частности, бетакоронавирусы SARS-CoV и MERS-CoV являлись этиологическими агентами двух крупных вспышек SARS и MERS [1]. Новый бетакоронавирус SARS-CoV-2, вызывающий заболевание у человека – COVID-19, впервые зарегистрирован в декабре 2019 г. в городе Ухань (провинция Хубэй, центральный Китай) [2, 3]. Предполагается, что основными носителями в дикой природе вируса SARS-CoV-2 являются летучие мыши рода подковоносов [3, 4].

С момента первого обнаружения SARS-CoV-2 в декабре 2019 г. накопился значительный массив сведений об изменчивости генома и динамике распространения COVID-19 в странах мира. За период циркуляции в человеческой популяции геноварианты вируса подвергались фильтрации через иммунные системы каждого организма-хозяина, и при каждой передаче происходил некоторый отбор наиболее продуктивных и быстрее передающихся вариантов. На этот отбор влияет и противоэпидемическая деятельность человека, когда тщательно изолируются тяжелобольные, что способствует прекращению распространения наиболее патогенных вариантов. В отличие от упомянутых вспышек SARS и MERS, распространение COVID-19 не удастся взять под контроль за счет использования профилактических и противоэпидемических мероприятий, что делает наиболее вероятным интеграцию этого возбудителя в этиологическую структуру возбудителей ОРВИ в среднесрочной перспективе. В свою очередь внедрение SARS-CoV-2 в ряд возбудителей ОРВИ может сопровождаться снижением его вирулентных свойств по причине селективных преимуществ тех геновариантов, которые легче передаются от человека к человеку. В противовес этому, отрицательной селекции подвергаются штаммы возбудителя, способствующие развитию более тяжелого инфекционного процесса.

Наша работа направлена на анализ вариаций вируса SARS-CoV-2 и выявление эволюционных изменений его генома с учетом тенденций приобретения новых эпидемически значимых свойств.

Материалы и методы

Для проведения филогенетического анализа использовались данные, представленные на сай-

те организации GISAID (<https://www.gisaid.org>), по нуклеотидным последовательностям геномов вируса SARS-CoV-2, полученным из образцов клинического материала в период с конца декабря 2019 по июль 2020 г. По состоянию на 6 июля 2020 г. в этой базе содержалось 37952 полных последовательности геномов вируса SARS-CoV-2 с высокой степенью покрытия, полученных из всех регионов мира. Позже, 5 августа 2020 г., с сайта GISAID загружена еще 641 последовательность геномов вируса SARS-CoV-2, полученных из образцов клинического материала, взятого в июле 2020 г. В филогенетическом анализе нами использовались только нуклеотидные последовательности геномов вируса с высоким качеством прочтения. Для отбора последовательности геномов выравнивали (использовался сервис <https://mafft.cbrc.jp/alignment/server/>) по референсу – Wuhan-Hu-1 (NC_045512.2). Затем последовательности выровненных геномов обрезали на 106 нуклеотидов от начала генома и на 390 нуклеотидов от его правого края по геному референсной последовательности. В случае наличия в оставшейся части последовательности генома хотя бы одного нераспознанного нуклеотида (N) этот геном удалялся из выборки для анализа. По описанным критериям отобрано 8360 нуклеотидных последовательностей геномов вируса SARS-CoV-2. Построение филогенетических деревьев проводили с помощью программного обеспечения BioNumerics 7.6 (<https://www.applied-maths.com>) с использованием алгоритма максимальной экономии (maximum parsimony). Для статистической обработки частоты встречаемости отдельных мутаций использовали всю базу данных полных геномов вируса SARS-CoV-2 на сайте GISAID (52468 геномов).

Для оценки эпидемиологических тенденций использованы данные о ежедневном числе регистрируемых случаев COVID-19 и летальных исходов в странах мира по данным университета Джонса Хопкинса на 13 июля 2020 г. по 188 странам за 173 дня наблюдений (<https://github.com/CSSEGISandData/COVID-19>).

Результаты и обсуждение

Полученная нами динамика эволюционной изменчивости генома вируса SARS-CoV-2 за период с конца декабря 2019 по июль 2020 г. (рис. 1–3) согласуется с данными, представленными в других работах [5, 6]. Подтверждено наличие ключевых мутаций

в геноме вируса SARS-CoV-2 в процессе его адаптации к организму человека при многократной передаче внутри человеческой популяции. Общее число мутаций, которые в совокупности имеют все секвенированные геномы вируса SARS-CoV-2, составляет

многие тысячи, однако редкие из них зафиксировались и стабильно наследуются. Мы определили только 22 единичные мутации, сохраняющиеся в $\geq 3\%$ от числа исследуемых геномов вируса SARS-CoV-2 (рис. 6, А). Для данного анализа использовали 52468

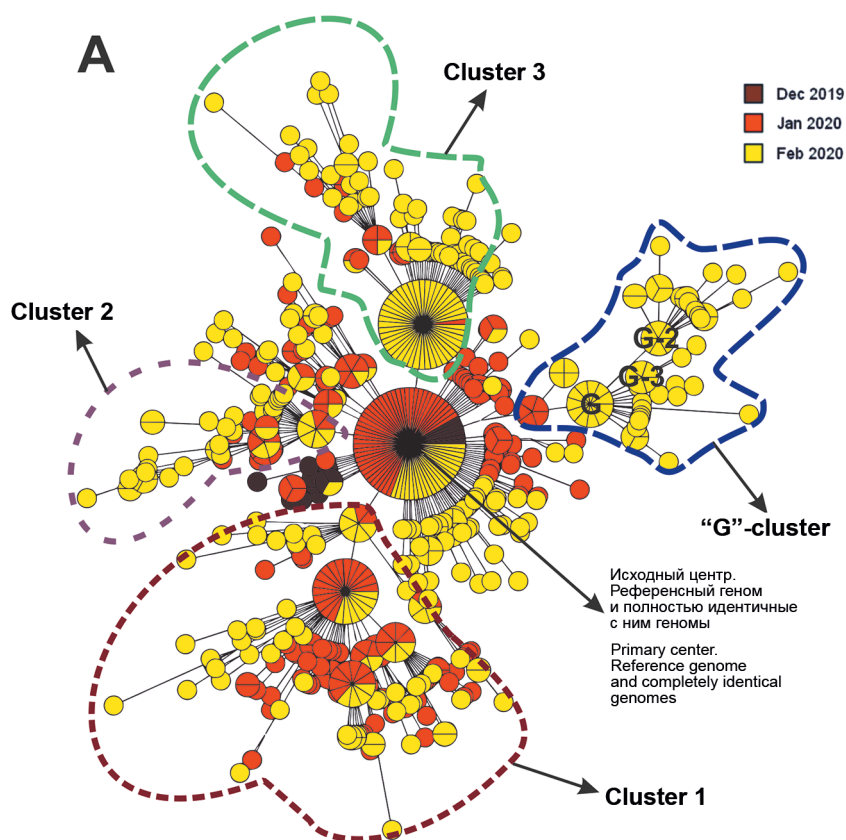
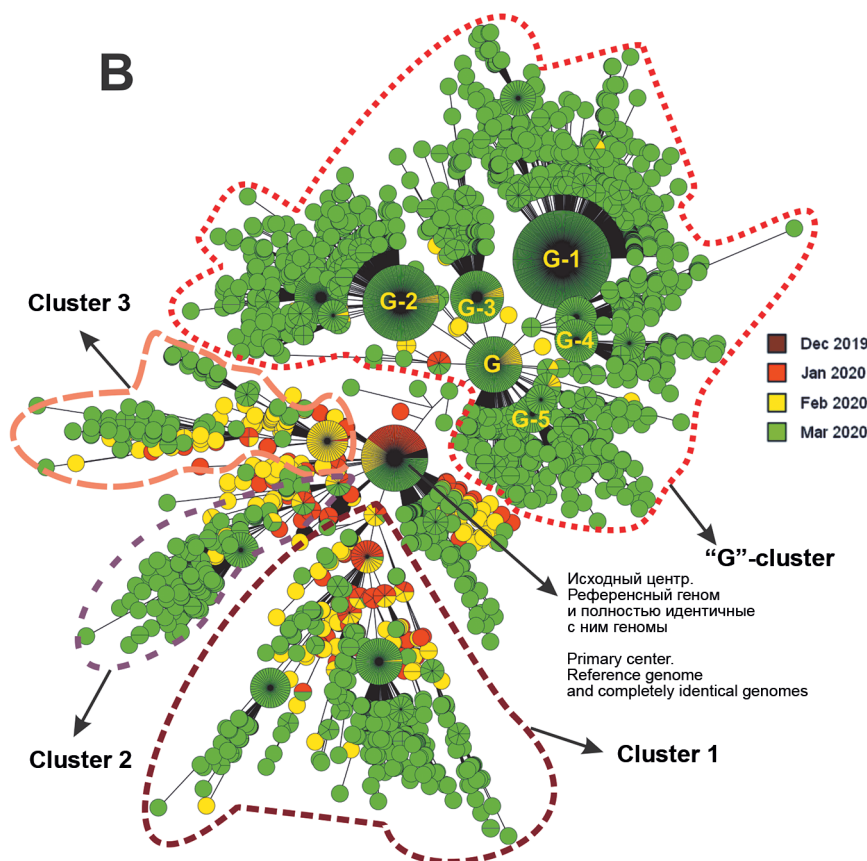


Рис. 1. Филогенетические связи геномов вируса SARS-CoV-2, полученных из клинического материала от пациентов из разных регионов мира (коричневый – декабрь, красный – январь, желтый – февраль, зеленый – март):

A – период с декабря 2019 по февраль 2020 г. *B* – период с декабря 2019 по март 2020 г. В течение марта наблюдается бурный рост числа геномов вируса SARS-CoV-2, имеющих мутации P314L (ORF1b ген) и D614G (S ген)

Fig. 1. Phylogenetic relations of SARS-CoV-2 virus genomes obtained from clinical material of patients from different regions of the world (brown color is for December, red color – January, yellow – February, green – March):

A – the period between December, 2019 and February, 2020. *B* – the period from December, 2019 to March, 2020. Throughout the March, an explosive growth in the number of genomes of SARS-CoV-2 virus that have P314L (ORF1b gene) and D614G (S gene) mutations is observed



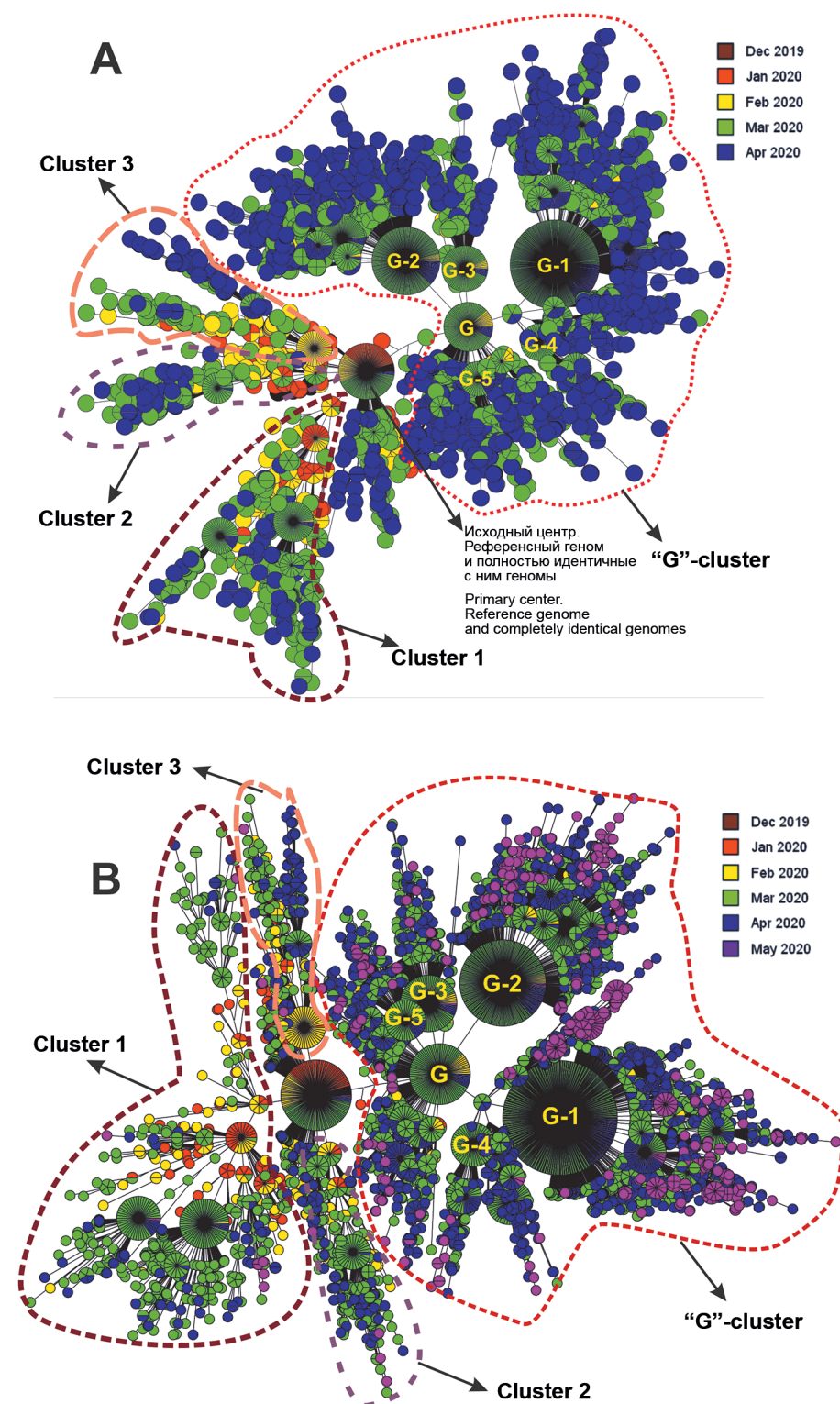


Рис. 2. Филогенетические связи геномов вируса SARS-CoV-2, полученных из клинического материала от пациентов из разных регионов мира:

A – в период с декабря 2019 по апрель 2020 г. (апрель выделен синим). *B* – в период с декабря 2019 по май 2020 г. (май выделен фиолетовым). Наблюдается дальнейший бурный рост числа геномов вируса, относящихся к «G»-кластеру. Число новых секвенированных геномов вируса SARS-CoV-2, не имеющих мутации P314L (ORF1b ген) и D614G (S ген), к маю 2020 г. не превышает 16 %, а к июню – 9 %

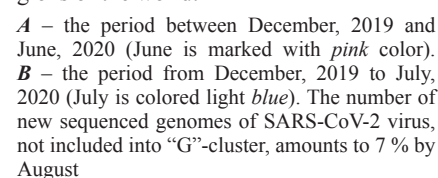
Fig. 2. Phylogenetic relations of SARS-CoV-2 virus genomes obtained from clinical material of patients from different regions of the world:

A – the period between December, 2019 and April, 2020 (April is marked with blue color). *B* – the period from December, 2019 to May, 2020 (May is colored violet). Further explosive growth in the number of genomes of SARS-CoV-2 virus appertaining to “G”-cluster is observed. The number of new sequenced genomes of SARS-CoV-2 virus without P314L (ORF1b gene) and D614G (S gene) mutations is less than 16 % towards May, while in June – only 9 %

полных геномов вируса, секвенированных из клинического материала, взятого с конца декабря 2019 по июль 2020 г. Стоит отметить, что все эти мутации зафиксированы с начала января по начало марта 2020 г., а позже в геноме вируса не фиксируется новых значительно представленных мутаций (рис. 6, B).

На ранних этапах развития пандемии COVID-19 (к середине января 2020 г.) в геноме вируса SARS-CoV-2 стабильно зафиксировались три основные группы мутаций, сформировавшие, соответствен-

но, три ветви эволюционного развития от одного исходного генома вируса (кластеры 1–3, рис. 1, A). Ключевые мутации, сформировавшие эти кластеры, следующие: кластер 1 (C8782T, T28144C); кластер 2 (G11083T, G26144T); кластер 3 (G1397A, G11083T, T28688C, или C6312A, G11083T, C13730T, C23929T, C28311T). Основное число мутаций, сформировавших три данных кластера, относится к гену ORF1ab, большинство из которых приводит к замене аминокислот (подчеркнутые позиции G1397A,



и N (C28311T, кластер 3). Развитие эволюционных направлений данных трех кластеров замедляется к концу марта 2020 г., когда в основном начинают об-

наруживаться варианты штаммов вируса с мутацией в гене ORF1b (позиция P314L) и гене S (D614G).

Появление единичной мутации в гене ORF1b (P314L или C14408T) и гене S (D614G или A23403G) является самым существенным устойчивым изменением в геноме вируса SARS-CoV-2, в результате чего к настоящему времени подавляющее большинство (более 93 %) выявляемых изолятов этого вируса имеет данную пару замен. Замена в гене S (D614G) многими исследователями ассоциируется с повышением способности вируса к передаче и снижением его патогенности, последнее, однако, может быть объяснено и улучшением методологии лечения больных в ходе пандемии. Влияние мутации в гене ORF1b (P314L) еще не изучено. Появление и рост числа встречаемости данных мутаций в геномах вируса SARS-CoV-2, а также всеобщее распространение данных геновариантов вируса отмечено в ряде публикаций [5, 6, 13, 14]. Из анализа используемых нами данных установлено, что первые варианты вируса с этими двумя мутациями в геноме зафиксированы в клинических образцах от 24 января 2020 г. на территории Китая и позднее в Австралии, а в феврале уже в большинстве стран Западной Европы, а также в Саудовской Аравии, США, Канаде, Мексике, Бразилии, Марокко и Сенегале. Появление этих новых геновариантов вируса SARS-CoV-2 в дальнейшем привело к их стремительному распространению почти по всем странам мира (развитие «G»-кластера на рис. 1–3) и исчезновению предшествующих им, менее адаптированных к организму человека, вариантов вируса SARS-CoV-2 (кластеры 1–3 на рис. 1–3). Изначально от генома референсного штамма (Wuhan-Hu-1, NC_045512.2) новые геноварианты («G»-кластер) отличались 4 единичными мутациями: C241T, C3037T, C14408T, A23403G. Проведенный нами анализ показывает, что мутации C14408T и A23403G тесно связаны и только совместное их присутствие в геноме дало преимущество для распространения данным геновариантам вируса SARS-CoV-2, а наличие каждой из этих мутаций в отдельности (A23403G в отсутствие C14408T, или наоборот), независимо от сочетания других присутствующих единичных замен, не позволило получить распространение таким геновариантам вируса. Стремительное распространение новых геновариантов привело к тому, что геномы штаммов вируса SARS-CoV-2, не содержащие мутации C14408T и A23403G, уже к апрелю 2020 г. встречались только в 25 % случаев (кластер 1–3 на рис. 1, B). На территории России также наблюдается наличие вариантов штаммов вируса с данной парой мутаций в большинстве представленных геномов. Из 331 генома вируса SARS-CoV-2, секвенированного из клинического материала на территории России, 327 геномов (98,8 %) содержат данные мутации (рис. 4). Детальное сравнение геномов вируса SARS-CoV-2 на территории России и зарубежных стран подтверждает, что именно геноварианты вируса SARS-CoV-2, циркулирующие на территории Западной Европы, в

марте и апреле 2020 г. массово завезены в Россию. В Западной Европе к тому времени большинство геновариантов вируса SARS-CoV-2 уже имели мутацию C14408T (ген ORF1b) и A23403G (ген S).

Распространение по всему миру геновариантов вируса SARS-CoV-2, имеющих мутации D614G (ген S) и P314L (ген ORF1b), видимо, обусловлено ускоренным процессом их связывания с клеткой хозяина и проникновением в нее. На основе обобщения эпидемиологических данных видно, что распространение штаммов новых геновариантов вируса («G»-кластер) ассоциировано не только с ростом заболеваемости, но и со снижением летальности, как показано на рис. 5. Уже с марта 2020 г. наблюдается устойчивая тенденция к росту ежедневного числа регистрируемых больных на фоне распространения штаммов, относящихся к «G»-кластеру – с 16 % к началу марта до 97 % в конце июня, что косвенно подтверждает наличие более высокого потенциала передачи вируса от человека к человеку. При этом с середины апреля 2020 г. наблюдается динамическое снижение летальности с 10 до 2 %. Представленные данные дополнительно свидетельствуют в пользу того, что распространение штаммов вируса, относящихся к «G»-кластеру, может быть связано с его биологическими свойствами, облегчающими передачу

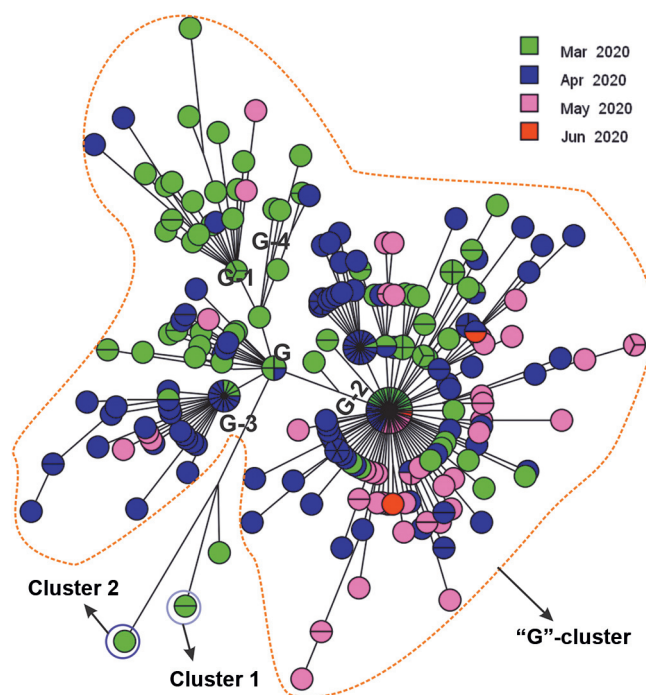


Рис. 4. Филогенетические связи 331 генома вируса SARS-CoV-2, полученного из клинического материала от пациентов на территории России в марте–июне 2020 г. Наиболее представлен подкластер G-2 (229 геномов) и подкластер G-3 (39 геномов), а также исходные варианты «G»-кластера (24 генома). Кластеры 1 и 2 представлены двумя и одним геномом вируса SARS-CoV-2 соответственно

Fig. 4. Phylogenetic relations of 331 genomes of SARS-CoV-2 virus obtained from clinical material of patients in the territory of Russia in March–June, 2020. The most represented sub-cluster is G-2 one (229 genomes), and G-3 sub-cluster (39 genomes), as well as initial variants of «G»-cluster (24 genomes). Clusters 1 and 2 are represented by two and one genome of SARS-CoV-2 virus, respectively

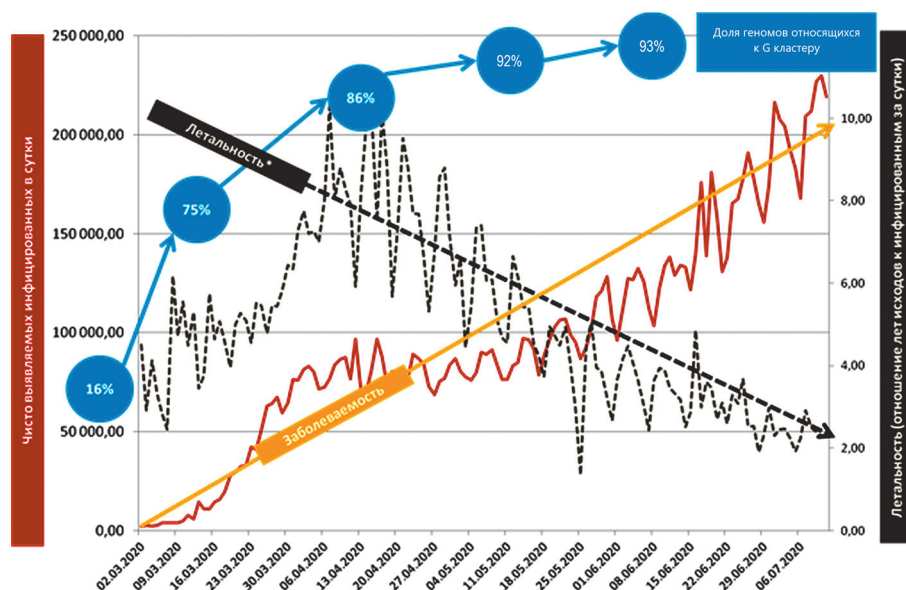


Рис. 5. Сводный график заболеваемости (красная линия), соотношения суточного числа летальных исходов и числа инфицированных (черная пунктирная линия, обозначенная как летальность) и распространенность геномов в «G»-кластере (светло-синяя линия)

Fig. 5. Summary graph of incidence (red line), the ratio of daily number of lethal cases and the number of infected persons (black dotted line, marked as lethality), and prevalence of genomes in "G"-cluster (light-blue line)

от человека к человеку, при этом ассоциированное снижение летальности может отражать не только прогресс в методах лечения, но и вероятное ослабление вирулентных свойств.

Формирование иммунной прослойки у населения против COVID-19 будет в основном зависеть от напряженности и длительности формируемого постинфекционного иммунитета, а также от эффективности разрабатываемых вакцин и порядка их применения. В описанных условиях SARS-CoV-2 будет встраиваться в этиологическую структуру возбудителей сезонных острых респираторных инфекционных болезней.

Скорость мутаций в геноме вируса SARS-CoV-2 на ранних этапах эпидемии (середина января 2020 г.) оценивалась в интервалах от $1,8 \cdot 10^{-4}$ до $4 \cdot 10^{-3}$ и от $5,2 \cdot 10^{-3}$ до $8,1 \cdot 10^{-3}$ на нуклеотид в год [7, 8]. В обеих статьях использовались небольшие количества полных геномов (33 и 54 соответственно). Позже в другом исследовании, где анализировалась выборка из 7666 геномов вируса SARS-CoV-2, результаты показали скорость его эволюции около $6 \cdot 10^{-4}$ на нуклеотид в год [9].

Для вируса SARS-CoV, послужившего причиной вспышки атипичной пневмонии в Китае в 2003 г., скорость мутации оценивалась от $0,8 \cdot 10^{-3}$ до $2,38 \cdot 10^{-3}$ на нуклеотид в год [10]. По-видимому, новый вариант вируса SARS-CoV-2 значительно уступает в скорости мутации близкородственному вирусу SARS-CoV.

Проведенный сравнительный анализ частоты мутаций на выборке из 8360 геномов показывает, что большинство вновь секвенированных геномов вируса SARS-CoV-2, включая варианты, полученные из клинического материала за июль 2020 г., до сих пор не отличаются от генома штамма референса (Wuhan-Hu-1, NC_045512.2) более чем на 10–12 единичных мутаций и редко имеют более 15 единичных мутаций. Исходя из полученных данных, можно утверждать, что однородность генома вируса SARS-CoV-2

в течение 7 мес. развития пандемии сохраняется на уровне $\geq 99,95\%$. На примере сделанной нами выборки установлено, что в совокупности анализируемые геномы содержат 5273 единичные мутации, отличающие их от исходного (референсного) штамма Wuhan-Hu-1 (NC_045512.2). Однако в значительном количестве геномов вируса SARS-CoV-2 ($\geq 3\%$ от общего числа геномов) в исследуемой выборке сохраняются всего 22 единичные мутации. Остальные единичные мутации – уникальные или редко встречающиеся. После проведения аналогичного поиска значительно представленных мутаций среди всех доступных 52468 геномов вируса на сайте GISAID выявлено также 22 единичные мутации, присутствующие в $\geq 3\%$ от числа исследуемых геномов вируса SARS-CoV-2 (рис. 6, А), однако три из них были другими. Столбцами красного цвета на рис. 6, А обозначены несинонимичные единичные замены, столбцы синего цвета соответствуют синонимичным мутациям, серым цветом обозначена мутация за пределами открытой рамки считывания. Распределение отобранных 22 единичных мутаций по первой дате их обнаружения отображено на рис. 6, В. Все они зафиксированы в период с начала января по начало марта 2020 г. Позже, включая июль, в глобальном масштабе в геноме вируса нами не зафиксированы значительно представленные мутации. К тому же 4 единичные мутации (выделены серым цветом) перестали фиксироваться в геномах вируса SARS-CoV-2 с конца июня 2020 г. (рис. 6, В), три из них относятся к геновариантам из кластера 1 и одна мутация принадлежит кластеру 2 (рис. 1–3).

Вышеуказанные 22 единичные мутации располагаются в следующих генах: ORF1ab (12), S (2), ORF3a (2), ORF8 (1) и N (4), мутация в позиции C241T относится к некодирующей области. Остальные гены в большинстве геновариантов вируса остались без изменений (M, E, ORF6, ORF7ab и ORF10). Функциональные характеристики указанных выше генов, несущих закрепившиеся мутации,

можно разделить на две основные группы: гены, ответственные за синтез неструктурных белков вируса – ORF1ab, ORF3a и ORF8, и гены, синтезирующие соответствующие структурные белки вириона – S и N. Белок S образует гомотримерные шипы на поверхности вирусной частицы и опосредует вход вируса в клетку хозяина. Белок N связывает вирусную РНК и требуется для упаковки вирусной РНК в вирусную частицу во время ее сборки. В то же время белки М и Е не имеют закрепившихся мутаций, хотя являются важными в структуре вириона – белок М придает ему форму и имеет решающее значение вместе с белком Е в организации сборки вируса и формировании зрелых вирусных частиц [11, 12]. Таким образом, важные гены вируса SARS-CoV-2, необходимые для его взаимодействия с клеткой хозяина и связывания вирусной РНК, имеют мутации, закрепившиеся в процессе эволюционной адаптации к организму человека.

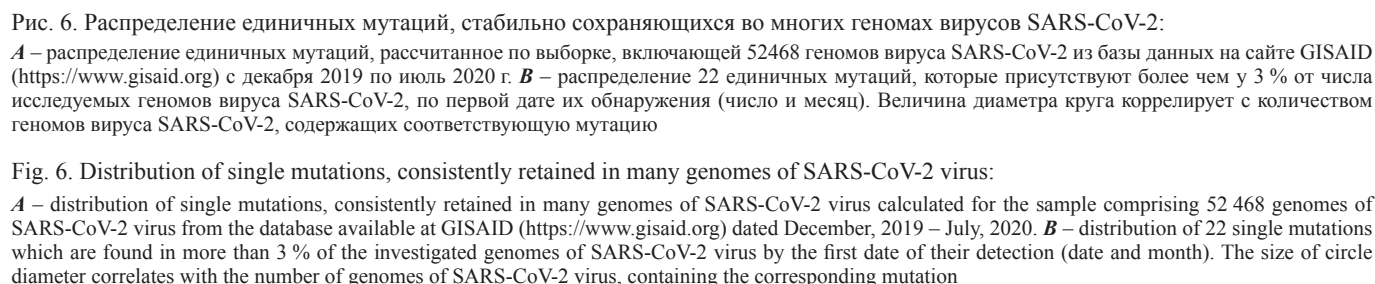
На рис. 1–3 представлены филогенетические связи между геномами вируса, распределенные по месяцам, каждый месяц маркирован определенным цветом. Исходя из филогенетических связей, представленных на рис. 1–3, видно, что геном вируса SARS-CoV-2 в январе 2020 г. уже имел ряд ключевых мутаций, которые в дальнейшем сформировали четыре основных кластера его геновариантов от одного исходного генома (геном штамма-референса и группа полностью идентичных ему геномов). Геномы вируса, составляющие «G»-кластер, содержат пару мутаций – D614G (ген S) и P314L (ген ORF1b), а в кластерах 1–3 данные мутации отсутствуют.

Ключевые мутации, сформировавшие три первых кластера (рис. 1–3), следующие: кластер 1 (C8782T, T28144C); кластер 2 (G11083T, G26144T); кластер 3 (G1397A/C6312A, G11083T, C28311T/T28688C). При этом единичные мутации, приводящие к замене аминокислот (подчеркнутая позиция), присутствуют в четырех генах: ORF1ab (G1397A, C6312A, G11083T – кластеры 2 и 3), ORF8 (T28144C – кластер 1), ORF3a (G26144T – кластер 2) и N (C28311T – кластер 3). Все эти гены, за исключением гена N (нуклеокапсид), кодируют аминокислотные последовательности неструктурных белков вируса [11, 12].

Первые геномы вируса, содержащие базовые мутации, формирующие кластер 1 (состоит из 797 геномов, рис. 3, В), секвенированы из клинического материала в 2020 г. на территории Китая (05.01.20), позднее они зафиксированы в США (19.01.20), Австралии (24.01.20), на Тайване (24.01.20), а также во Вьетнаме (24.01.20), Таиланде (25.01.20), Гонконге (27.01.20), Южной Корее (30.01.20) и Индии (31.01.20). В феврале данные по секвенированию геномов вируса, относящихся к кластеру 1, имелись только в Испании (28.02.20). Динамика развития кластера 1 в январе и феврале показана на рис. 1, А (кластер 1: красным цветом выделены геномы вируса, циркулирующие в январе, желтым – варианты генома за февраль). В марте штаммы

вируса из кластера 1 также зафиксированы в Чили (02.03.20), Мексике (03.03.20), Бразилии (05.03.20), Грузии (08.03.20), Нидерландах (09.03.20), Канаде (09.03.20), Шотландии (15.03.20), Омане (15.03.20), Люксембурге (17.03.20), Польше (18.03.20), Греции (18.03.20), Саудовской Аравии (23.03.20), Португалии (23.03.20), России (23.03.20), Казахстане (25.03.20) и Нигерии (29.03.20). Динамика развития кластера 1 в марте отображена на рис. 1, В (кластер 1: геномы выделены зеленым цветом). На изображении видно, что варианты штаммов вируса SARS-CoV-2 из кластера 1 к апрелю 2020 г. распространились по множеству стран, расположенных на всех континентах. Однако в апреле, мае, июне и июле распространение штаммов вируса из кластера 1 на новых территориях не зафиксировано. Частота их встречаемости убывает, как и количество стран, где они продолжают фиксироваться (рис. 2–3: геномы выделены синим, фиолетовым, розовым и голубым цветом). В мае штаммы вируса этих геновариантов зафиксированы в Индии, Шотландии, США и Казахстане, а в июне геномы вируса из кластера 1 фиксируются только в Индии (2 генома). Добавляются также новые страны, на территории которых обнаружены данные варианты вируса: Египет, Пакистан, Великобритания (представлено по 1 геному). В июле вирусы этих геновариантов не зафиксированы ни в одной из стран.

Первые геномы вируса SARS-CoV-2 в базе GISAID, содержащие базовые мутации, формирующие кластер 2 (состоит из 342 геномов, рис. 3, В), получены из клинического материала, взятого 17.01.2020 г. на территории Китая и Испании, США (21.01.20), Тайваня (23.01.20), Франции (23.01.20), Таиланда (23.01.20), Австралии (25.01.20), Сингапура (27.01.20), Италии (29.01.20) и Гонконга (30.01.20). В феврале зафиксировано распространение геновариантов вируса из этого кластера на территорию ряда других стран: Южная Корея (06.02.20), Швеция (07.02.20) и Испания (28.02.20). Динамика развития кластера 2 в январе и феврале показана на рис. 1, А. В марте геноварианты вируса из кластера 2 также зафиксированы в следующих странах: Норвегия (04.03.20), Нидерланды (06.03.20), Бельгия (07.03.20), Греция (07.03.20), Ямайка (09.03.20), Грузия (10.03.20), Португалия (11.03.20), Шотландия (12.03.20), Бразилия (18.03.20), Оман (18.03.20), Конго (20.03.20), Россия (24.03.20), Германия (28.03.20), Великобритания (29.03.20), Нигерия (29.03.20), Польша (31.03.20) и Шри-Ланка (31.03.20). Динамика развития кластера 2 в марте отображена на рис. 1, В (кластер 2: геномы выделены зеленым цветом). В итоге к апрелю 2020 г. вирус SARS-CoV-2 с вариантами генома из кластера 2 также распространился по всем континентам. Далее распространение штаммов вируса из кластера 2 на новых территориях не зафиксировано. Наблюдается уменьшение частоты их встречаемости и количества стран, в которых геноварианты из кластера 2 продолжают фиксироваться (рис. 2–3, кластер 2). Так,



в мае они зарегистрированы только в Шотландии и Нидерландах, а в июне и июле уже отсутствуют данные секвенирования, подтверждающие их присутствие в какой-либо из стран.

Первые геномы вируса SARS-CoV-2, содержащие базовые мутации, формирующие кластер 3 (состоит из 349 геномов, рис. 3, В), получены из клинического материала, взятого в конце января на территории Китая (20.01.20), Гонконга (24.01.20) и Франции (29.01.20). В феврале зафиксировано распространение геновариантов вируса из этого кластера на территорию ряда других стран: Сингапур (01.02.20), Япония (круизный лайнер, 15.02.20), США (круизный лайнер, 18.02.20), Канада (20.02.20), Оман (23.02.20), Израиль (23.02.20), ОАЭ (25.02.20), Тайвань (25.02.20), Швеция (27.02.20) и Австралия (28.02.20). Динамика развития кластера 3 в январе и феврале показана на рис. 1, А. В марте геноварианты вируса из кластера 3 фиксируются на территории еще ряда новых стран: Германия (02.03.20), Кувейт (02.03.20), Нидерланды (03.03.20), Австралия (03.03.20), Пакистан (04.03.20), Саудовская Аравия (07.03.20), Индия (10.03.20), Бразилия (10.03.20), Малазия (16.03.20), Шри-Ланка (16.03.20), Турция (19.03.20), Бруней (21.03.20), Иран (26.03.20), Тунис (28.03.20), Словения (29.03.20), Индонезия (30.03.20), Уганда (30.03.20). Динамика развития кластера 3 в марте отображена на рис. 1, В (кластер 3: геномы выделены зеленым цветом). В апреле зафиксировано распространение геновариантов вируса из этого кластера на территорию Казахстана (26.04.20). В дальнейшем распространение штаммов вируса из кластера 3 на новых территориях не зафиксировано. Убывает частота их встречаемости и количество стран, в которых данные геноварианты фиксируются (рис. 2–3, кластер 3). В июне они зарегистрированы в Сингапуре, Индии и Австралии, а в июле продолжали фиксироваться только в Сингапуре.

Таким образом, в ходе анализа дат забора клинического материала и территорий, на которых зафиксированы геноварианты из трех вышеописанных кластеров, становится очевидным, что к апрелю 2020 г. варианты штаммов вируса SARS-CoV-2 из кластеров 1–3 распространились по большинству стран и регионов, расположенных на всех континентах. Однако уже к лету данные геноварианты встречаются редко. Нами проведен поиск геновариантов вируса, еще не имеющих мутации C14408T и A23403G (маркерные метки «G»-кластера), в доступной базе, представленной на сайте GISAID за июнь (3618 геномов) и июль (1260 геномов) по состоянию на 6 августа 2020 г. Результат поиска показал, что в клинических образцах за июнь геноварианты вируса, не имеющие данных мутаций, преобладают в Сингапуре и редко встречаются в США, Великобритании, Индии, Пакистане и Австралии. В клинических образцах за июль геноварианты вируса, не имеющие данных мутаций, по-прежнему преобладают в Сингапуре, один геновариант зафиксирован в Австралии.

Обобщая вышеизложенное, можно утверждать, что геноварианты вируса SARS-CoV-2 из кластеров 1–3 на сегодняшний день имеют слабое влияние на эпидемиологическую ситуацию. Число образцов клинического материала, содержащего такие геноварианты вируса, на территории абсолютного большинства стран составляет около 7 % и менее. Причиной этому послужило появление и широкое распространение геновариантов из «G»-кластера (рис. 1–3), более адаптированных к организму человека, которые конкурентно вытеснили первичные геноварианты, уже циркулирующие на территории всех континентов.

Первый геномом вируса SARS-CoV-2, содержащий мутации C14408T и A23403G, зафиксирован в Китае (Zhejiang_HZ103_2020_EPI_ISL_422425, получен из клинического материала от 24.01.2020 г.), а позднее в Австралии (Australia_NSW2153_2020_EPI_ISL_509505, клинический материал от 25.01.2020 г.). В первых числах февраля такие геноварианты зафиксированы в Великобритании и Саудовской Аравии. К концу февраля мутации C14408T и A23403G в геномах вируса уже зафиксированы на территории большинства стран Западной Европы, Северной и Южной Америки, а также Африки. Следует отметить, что мутации C14408T и A23403G тесно связаны и только их совместное присутствие в геноме дало преимущество для распространения данным геновариантам вируса SARS-CoV-2. Профиль нуклеотидных замен относительно референсного штамма (Wuhan-Hu-1, NC_045512.2) у исходных вариантов генома вируса, сформировавших «G»-кластер и получивших на сегодняшний день абсолютное большинство, был следующим: C241T, C3037T, C14408T, A23403G. Наличие в геноме вируса мутации A23403G в отсутствие C14408T имеют небольшое число геновариантов вируса и их количество уменьшается. Нами не найдены геноварианты вируса, имеющие в геноме только мутации C14408T и A23403G без C241T и C3037T, а также имеющие мутацию C14408T без мутации A23403G. Таким образом, только комплексное наличие мутаций (как минимум C14408T и A23403G) позволило геновариантам вируса получить преимущество в распространении среди человеческой популяции. Например, в базе данных GISAID среди геномов вируса, относящихся к маю, июню и июлю, нами найдено соответственно 51, 54 и 24 генома, имеющих мутации C241T, C3037T и A23403G, но не имеющих замену C14408T, что однозначно указывает на отсутствие преимуществ к распространению у данных геновариантов, несмотря на их циркуляцию во многих странах мира. Тем не менее в экспериментальной работе [15] представлены доказательства того, что геноварианты вируса с мутацией в S-белке (D614G) демонстрируют повышенную инфекционную активность (в 5 раз) в клетках, экспрессирующих ACE2, без участия мутации P314L в гене ORF1b при тестировании в псевдотипированных лентивирусных векторах. Этот результат, однако, не

согласуется с малым количеством геномов вируса, имеющих мутацию D614G и не имеющих мутацию P314L, чего не должно быть в случае нейтральности замены P314L.

В нуклеотидной последовательности генома вируса мутация A23403G находится в S-гене, в аминокислотной последовательности S-белка это привело к замене в позиции 614 аспарагиновой кислоты (D) на глицин (G). Белок, кодируемый S-геном, – гликопротеин (spike glycoprotein) – имеет функции распознавания клетки хозяина, связывания с ней и обеспечивает проникновение вируса в клетку [16]. Мутация C14408T в гене ORF1b привела к замене в позиции 314 пролина (P) на лейцин (L). Функция гена ORF1b связана с образованием ряда неструктурных белков, мутация P314L находится в структуре белка RdRp (nsp12), который непосредственно участвует в репликации вируса, и его точность влияет на вероятность появления мутаций в геноме вируса SARS-CoV-2 [13]. Данные замены оказались настолько важными (устойчивыми), что дальнейшая эволюция генома вируса, по всей вероятности, будет развиваться только на основе геновариантов вируса, содержащих эту пару замен. Появление и рост числа встречаемости данных мутаций уже отмечены в ряде публикаций [5, 6, 13, 14], а также представлены на сайте GISAID.

Помимо мутации D614G в S-гене и P314L в гене ORF1b, в феврале и начале марта 2020 г. появляется ряд других единичных замен в геноме вируса SARS-CoV-2, которые разделяют «G»-кластер на подкластеры (рис. 1–3). Профиль основных генетических мутаций в данном случае представлен следующими единичными мутациями: исходный «G»-кластер (C241T, C3037T, C14408T, A23403G); подкластер G-1 (C241T, C1059T, C3037T, C14408T, A23403G, G25563T), известный как GH-кластер; подкластер G-2 (C241T, C3037T, C14408T, A23403G, комплексная мутация трех нуклеотидов подряд – GGG28881-28883AAC), известный как GR-кластер; подкластер G-3 (C241T, C3037T, C14408T, A20268G, A23403G); подкластер G-4 (C241T, C2416T, C3037T, C18877T, C14408T, A23403G, G25563T, C26735T) и подкластер G-5 (C241T, C3037T, C14408T, C15324T, A23403G). Подкластер G-1 сформировался благодаря дополнительным мутациям в ORF1ab гене (C1059T) и гене ORF3a (G25563T); геномы вируса из подкластера G-2, помимо исходного профиля мутаций «G»-кластера, имеют комплексную мутацию трех нуклеотидов подряд в последовательности гена N – GGG28881-28883AAC. Подкластеры G-3 (A20268G, ген ORF1ab), G-4 (C2416T, C18877T, ген ORF1ab; G25563T, ген ORF3a; C26735T, ген M) и G-5 (C15324T, ген ORF1ab) формируются за счет дополнительных синонимичных и несинонимичных мутаций, указанных в скобках. Как можно заметить, кроме базовых мутаций C241T, C3037T, C14408T и A23403G, геномы вируса из «G»-кластера имеют на сегодняшний день ряд других закрепившихся мута-

ций, отличных от геномов вируса из кластеров 1–3.

Частота встречаемости вариантов вируса SARS-CoV-2 с геномами, образующими «G»-кластер, в образцах клинического материала составляет на сегодняшний день для большинства регионов мира в среднем около 93 % (рис. 3, B).

По данным секвенирования, в конце февраля и первых числах марта в «G»-кластере появляются пять основных быстро развивающихся ответвлений (подкластеров): G-1, G-2, G-3, G-4 и G-5, – маркерные мутации которых описаны выше. Первые геноварианты из подкластера G-1 сначала фиксируются на территории Сингапура (16.02.20), Франции (21.02.20), Сенегала (28.02.20), Бельгии (29.02.20) и США (29.02.20). В марте данные геноварианты фиксируются также в Канаде, Швеции, Нидерландах, Норвегии и России. В итоге геноварианты из подкластера G-1 (1908 геномов, рис. 3, B) становятся наиболее характерными для территории Северной Америки.

Первые геноварианты из подкластера G-2 фиксируются в Великобритании (16.02.20) и позднее в Италии (24.02.20), Австрии (24.02.20), Германии (25.02.20), Дании (26.02.20), Швейцарии (27.02.20), Мексике (27.02.20), Шотландии (28.02.20), США (28.02.20) и Испании (29.02.20). В итоге геноварианты из подкластера G-2 (1861 геном, рис. 3, B) становятся наиболее характерными для территории Западной Европы.

Первые геноварианты из подкластера G-3 фиксируются в Великобритании (24.02.20), Италии (26.02.20), Швейцарии (27.02.20) и США (28.02.20). В итоге геноварианты из подкластера G-3 (369 геномов, рис. 3, B) становятся типичными для территории Западной Европы.

Первые геноварианты из подкластера G-4 фиксируются в Саудовской Аравии (03.02.20), затем в Канаде (27.02.20), США (28.02.20) и Франции (29.02.20). В итоге геноварианты из подкластера G-4 (676 геномов, рис. 3, B) встречаются как на территории Северной Америки, так и Западной Европы.

Геноварианты из подкластера G-5 сначала зафиксированы на территории Китая (01.02.20), а в конце февраля – в Германии и Объединенных Арабских Эмиратах. В марте они фиксируются в Швейцарии, Бельгии и Франции. В итоге геноварианты из подкластера G-5 (261 геном, рис. 3, B) встречаются на территории Западной Европы, редко в Северной Америке и Индии.

На территорию России вирус SARS-CoV-2 завезен массово в марте и начале апреля 2020 г. гражданами, прибывшими из-за рубежа. На рис. 4 представлено филогенетическое дерево из 331 генома вируса SARS-CoV-2, полученного на территории России. Только 4 генома (1,2 %) вируса SARS-CoV-2 из данной выборки не относятся к «G»-кластеру. Филогенетический анализ показал, что основное количество геномов (81 %) вируса SARS-CoV-2, циркулирующего на территории России, относится

к подкластерам G-2 и G-3 (229 и 39 геномов соответственно). Варианты данных геномов вируса SARS-CoV-2 типичны для территории Западной Европы уже с конца февраля 2020 г. Близкие им варианты генома только в небольших количествах секвенированы из клинического материала от пациентов из Северной и Южной Америки, Саудовской Аравии и Индии. Таким образом, почти все генетическое разнообразие вариантов генома вируса SARS-CoV-2, распространенных на территории России, относится к «G»-кластеру и в пределах этого кластера разделяется на 5 групп: G, G-1, G-2, G-3 и G-4.

Полученные нами данные подтверждают, что основным регионом, из которого осуществлен завоз штаммов вируса SARS-CoV-2 на территорию России, является Западная Европа. В это время в Западной Европе в большинстве секвенированных клинических образцов уже содержались геноварианты вируса из «G»-кластера [13].

На сегодняшний день мы наблюдаем повсеместное распространение геновариантов вируса SARS-CoV-2, имеющих мутацию D614G (ген S) и P314L (ген ORF1b), в отличие от геновариантов вируса SARS-CoV-2, не имеющих данных мутаций, что, видимо, обусловлено ускоренным процессом связывания вируса с клеткой хозяина и проникновением в нее. Характерно, что указанный тренд наблюдается как в мире в целом, так и, при независимом анализе, в разрезе отдельных регионов, что подтверждает наличие эволюционного преимущества новых геновариантов. В литературе описано две гипотезы, объясняющие, как мутация D614G связана с повышенной передачей вируса SARS-CoV-2. Первая гипотеза основана на структурном положении мутации D614G [14, 17] и связана с уменьшением взаимодействия между субъединицами S1 и S2 в S-белке, которое способствует их облегченному расщеплению и, соответственно, дальнейшему слиянию вируса со стенкой клетки хозяина и последующему проникновению в клетку. Вторая гипотеза состоит в том, что эта мутация может косвенно усилить взаимодействие рецептор-связывающего домена на поверхности субъединицы S1 белка S вируса с ангиотензин-превращающим ферментом 2 (связь RBD-ACE2), что также способствует ускоренному взаимодействию частицы вируса с клеткой [17, 18].

Следует отметить, что распространение штаммов «G»-кластера ассоциировано не только с ростом заболеваемости (до 200 и более тысяч ежедневно выявляемых инфицированных), но и со снижением летальности, как показано на рис. 5. Уже с марта 2020 г. наблюдается устойчивая тенденция к росту ежедневного числа регистрируемых больных на фоне распространения штаммов, относящихся к «G»-кластеру – с 16 % к началу марта до 93 % в конце июня, что косвенно подтверждает наличие более высокого потенциала у данных геновариантов вируса к передаче от человека к человеку. При этом с середины апреля 2020 г. наблюдается динамиче-

ское снижение летальности с 10 до 2 % (оцененной в данном исследовании как отношение суточного числа летальных исходов к суточному числу зарегистрированных больных в процентном выражении). Представленные данные дополнительно свидетельствуют в пользу того, что распространение штаммов вируса, относящихся к «G»-кластеру, может быть связано с его биологическими свойствами, облегчающими передачу от человека к человеку, при этом ассоциированное снижение летальности может отражать не только прогресс в методах лечения, но и вероятное ослабление вирулентных свойств.

Таким образом, одним из актуальных вопросов оценки глобального риска, связанного с пандемией COVID-19, остается интерпретация наблюдаемых изменений в геноме, что имеет значение для долгосрочного прогноза развития эпидемического процесса. До тех пор, пока отсутствует общедоступная и эффективная вакцина, развитие эпидемического процесса обуславливается, с одной стороны, биологическими свойствами возбудителя, а с другой – комплексом неспецифических профилактических и противоэпидемических мероприятий. При сохранении текущих темпов распространения COVID-19 формирование иммунной прослойки будет в основном зависеть от напряженности и длительности формируемого постинфекционного иммунитета, а также от эффективности разрабатываемых вакцин и порядка их применения. В описанных условиях SARS-CoV-2 будет встраиваться в этиологическую структуру возбудителей сезонных острых респираторных инфекционных болезней.

Следует подчеркнуть, что эволюция SARS-CoV-2 продолжается и требует непрерывной научной оценки с использованием передовых методов анализа генетических последовательностей. При этом представленные данные свидетельствуют о том, что распространение штаммов вируса из «G»-кластера может быть связано с их биологическими свойствами, облегчающими передачу от человека к человеку, в то время как снижение летальности может отражать ослабление вирулентных свойств. Отдельного внимания заслуживает изучение биологических свойств геновариантов SARS-CoV-2.

Конфликт интересов. Авторы подтверждают отсутствие конфликта финансовых/нефинансовых интересов, связанных с написанием статьи.

Благодарность. Авторский коллектив выражает глубокую признательность члену-корреспонденту РАН, доктору биологических наук Нетёсову Сергею Викторовичу, за объективные комментарии, сделанные по содержанию статьи, учет которых позволил сделать работу существенно сильнее и убедительнее.

References / Список литературы

1. Cui J., Li F., Shi Z.-L. Origin and evolution of pathogenic coronaviruses. *Nat. Rev. Microbiol.* 2019; 17(3):181–92. DOI: 10.1038/s41579-018-0118-9.
2. Wu F., Zhao S., Yu B., Chen Y.-M., Wang W., Song Z.-G., Hu Y., Tao Z.-W., Tian J.-H., Pei Y.-Y., Yuan M.-L., Zhang Y.-L., Dai F.-H., Liu Y., Wang Q.-M., Zheng J.-J., Xu L., Holmes E.C., Zhang

- Y.-Z. A new coronavirus associated with human respiratory disease in China. *Nature*. 2020; 579(7798):265–9. DOI: 10.1038/s41586-020-2008-3.
3. Zhou P., Yang X.-L., Wang X.-G., Hu B., Zhang L., Zhang W., Si H.-R., Zhu Y., Li B., Huang C.-L., Chen H.-D., Chen J., Luo Y., Guo H., Jiang R.-D., Liu M.-Q., Chen Y., Shen X.-R., Wang X., Zheng X.-S., Zhao K., Chen Q.-J., Deng F., Liu L.-L., Yan B., Zhan F.-X., Wang Y.-Y., Xiao G.-F., Shi Z.-L. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*. 2020; 579(7798):270–3. DOI: 10.1038/s41586-020-2012-7.
4. Lu R., Zhao X., Li J., Niu P., Yang B., Wu H., Wang W., Song H., Huang B., Zhu N., Bi Y., Ma X., Zhan F., Wang L., Hu T., Zhou H., Hu Z., Zhou W., Zhao L., Chen J., Meng Y., Wang J., Lin Y., Yuan J., Xie Z., Ma J., Liu W.J., Wang D.Y., Xu W., Holmes E.C., Gao G.F., Wu G., Chen W., Shi W., Tan W. Genomic characterization and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *The Lancet*. 2020; 395(10224):565–74. DOI: 10.1016/S0140-6736(20)30251-8.
5. Kupferschmidt K. The pandemic virus is slowly mutating. But does it matter? *Science*. 2020; 369(6501):238–9. DOI: 10.1126/science.369.6501.238.
6. Bhattacharyya Ch., Das Ch., Ghosh A., Singh A.K., Mukherjee S., Majumder P.P., Basu A., Biswas N.K. Global spread of SARS-CoV-2 subtype with spike protein mutation D614G is shaped by human genomic variations that regulate expression of *TMPRSS2* and *MX1* genes. bioRxiv. 2020.05.04.075911. DOI: 10.1101/2020.05.04.075911.
7. Li X., Wang W., Zhao X., Zai J., Zhao Q., Li Y., Chaillon A. Transmission dynamics and evolutionary history of 2019-nCoV. *J. Med. Virol.* 2020; 92(5):501–11. DOI:10.1002/jmv.25701.
8. Benvenuto D., Giovanetti M., Salemi M., Prosperi M., De Flora C., Alcantara L.C.J., Angeletti S., Ciccozzi M. The global spread of 2019-nCoV: a molecular evolutionary analysis. *Pathog. Glob. Health*. 2020; 114(2):64–7. DOI: 10.1080/20477724.2020.1725339.
9. Van Dorp L., Acman M., Richard D., Shaw L.P., Ford C.E., Ormond L., Owena C.J., Pang J., Tan C.C.S., Boshier F.A.T., Ortiz A.T., Balloux F. Emergence of genomic diversity and recurrent mutations in SARS-CoV-2. *Infect. Genet. Evol.* 2020; 83:104351. DOI: 10.1016/j.meegid.2020.104351.
10. Zhao Z., Li H., Wu X., Zhong Y., Zhang K., Zhang Ya-P., Boerwinkle E., Fu Y.-X. Moderate mutation rate in the SARS coronavirus genome and its implications. *BMC Evol. Biol.* 2004; 4:21. DOI: 10.1186/1471-2148-4-21.
11. Mousavizadeh L., Ghasemi S. Genotype and phenotype of COVID-19: Their roles in pathogenesis. *J. Microbiol. Immunol. Infect.* 2020. DOI: 10.1016/j.jmii.2020.03.022.
12. Ashour H.M., Elkhateb W.F., Rahman Md.M., Elshabrawy H.A. Insights into the Recent 2019 Novel Coronavirus (SARS-CoV-2) in Light of Past Human Coronavirus Outbreaks. *Pathogens*. 2020; 9(3):186. DOI: 10.3390/pathogens9030186.
13. Pachetti M., Marini B., Benedetti F., Giudici F., Mauro E., Storici P., Masciovecchio C., Angeletti S., Ciccozzi M., Gallo R.C., Zella D., Ippodrino R. Emerging SARS-CoV-2 mutation hot spots include a novel RNA-dependent-RNA polymerase variant. *J. Transl. Med.* 2020; 18:179. DOI: 10.1186/s12967-020-02344-6.
14. Korber B., Fischer W.M., Gnanakaran S., Yoon H., Theiler J., Abfalterer W., Foley B., Giorgi E.E., Bhattacharya T., Parker M.D., Partridge D.G., Evans C.M., de Silva T.I., LaBranche C.C., Montefiori D.C. Spike mutation pipeline reveals the emergence of a more transmissible form of SARS-CoV-2. bioRxiv. DOI: 10.1101/2020.04.29.069054.
15. Ogawa J., Zhu W., Tonnu N., Singer O., Hunter T., Ryan (Firth) A.L., Pao G.M. The D614G mutation in the SARS-CoV2 Spike protein increases infectivity in an ACE2 receptor dependent manner. bioRxiv. DOI: 10.1101/2020.07.21.214932.
16. Chen W.H., Hotez P.J., Bottazzi M.E. Potential for developing a SARS-CoV receptor-binding domain (RBD) recombinant protein as a heterologous human vaccine against coronavirus infectious disease (COVID)-19. *Hum. Vaccin. Immunother.* 2020; 16(6):1239–42. DOI: 10.1080/21645515.2020.1740560.
17. Wrapp D., Wang N., Corbett K.S., Goldsmith J.A., Hsieh C.-L., Abiona O., Graham B.S., McLellan J.S. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science*. 2020; 367(6483):1260–3. DOI: 10.1126/science.abb2507.
18. Walls A.C., Park Y.-J., Tortorici M.A., Wall A., McGuire A.T., Veesler D. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell*. 2020; 181(2):281–292.e6. DOI: 10.1016/j.cell.2020.02.058.

Authors:

Krasnov Ya.M., Safronov V.A., Fedorov A.V., Badanin D.V., Shcherbakova S.A., Kutyrev V.V. Russian Research Anti-Plague Institute "Microbe". 46, Universitetskaya St., Saratov, 410005, Russian Federation. E-mail: rusrap@microbe.ru.

Popova A.Yu. Federal Service for Surveillance in the Sphere of Consumers Rights Protection and Human Welfare; 18, Bld. 5 and 7, Vadkovsky Pereulok, Moscow, 127994, Russian Federation. Russian Medical Academy of Continuing Professional Education; 2/1, Barrikadnaya St., Moscow, 125993, Russian Federation.

Об авторах:

Краснов Я.М., Сафронов В.А., Федоров А.В., Баданин Д.В., Щербакова С.А., Кутырев В.В. Российский научно-исследовательский противочумный институт «Микроб». Российская Федерация, 410005, Саратов, ул. Университетская, 46. E-mail: rusrap@microbe.ru.

Попова А.Ю. Федеральная служба по надзору в сфере защиты прав потребителей и благополучия человека; Российская Федерация, 127994, Москва, Вадковский переулок, дом 18, строение 5 и 7. Российская медицинская академия непрерывного профессионального образования; Российская Федерация, 125993, Москва, ул. Баррикадная, 2/1.