ORIGINAL RESEARCHES

ORIGINAL ARTICLE
DOI: https://doi.org/10.36233/0507-4088-72

Monitoring the spread of the SARS-CoV-2 (*Coronaviridae: Coronavirinae: Betacoronavirus; Sarbecovirus*) variants in the Moscow region using targeted high-throughput sequencing

Nadezhda I. Borisova[1]*, Ivan A. Kotov[1,2]*, Anton A. Kolesnikov[1], Valeria V. Kaptelova[1], Anna S. Speranskaya[1], Larisa Yu. Kondrasheva[1], Elena V. Tivanova[1], Kamil F. Khafizov[1], Vasily G. Akimkin[1]

[1]FSBI «Central Research Institute for Epidemiology» of the Federal Service for Supervision of Consumer Rights Protection and Human Welfare (Rospotrebnadzor), Moscow, 111123, Russia
[2]FSAEI HE «Moscow Institute of Physics and Technology (National Research University)», 141700, Dolgoprudny, Russia

Introduction. Since the outbreak of the COVID-19 pandemic caused by SARS-CoV-2 novel coronavirus, the international community has been concerned about the emergence of mutations altering some biological properties of the pathogen like increasing its infectivity or virulence. Particularly, since the end of 2020, several variants of concern have been identified around the world, including Alpha (B.1.1.7), Beta (B.1.351), Gamma (P.1), and Delta (B.1.617.2). However, the existing mechanism of detecting important mutations are not always effective enough, since only a relatively small part of all pathogen samples can be examined by whole genome sequencing due to its high cost.
Material and methods. In this study, we have designed special primer panel and used it for targeted high-throughput sequencing of several significant S-gene (spike) regions of SARS-CoV-2. The Illumina platform averaged approximately 50,000 paired-end reads with a length of ≥150 bp per sample. This method was used to examine 579 random samples obtained from COVID-19 patients in Moscow and the Moscow Region from February to June 2021.
Results. This study demonstrated the dynamics of distribution of several SARS-CoV-2 strains and its some single mutations. It was found that the Delta strain appeared in the region in May 2021, and became prevalent in June, partially displacing other strains.
Discussion. The obtained results provide an opportunity to assign the viral samples to one of the strains, including the previously mentioned in time- and cost-effective manner. The approach can be used for standardization of the procedure of searching for mutations in individual regions of the SARS-CoV-2 genome. It allows to get a more detailed data about the epidemiological situation in a region.

Key words: *coronavirus; coronavirus infection, SARS-CoV-2; next generation sequencing; COVID-19; genome; strains*

For citation: Borisova N.I., Kotov I.A., Kolesnikov A.A., Kaptelova V.V., Speranskaya A.S., Kondrasheva L.Yu., Tivanova E.V., Khafizov K.F., Akimkin V.G. Monitoring the spread of the SARS-CoV-2 (*Coronaviridae: Coronavirinae: Betacoronavirus; Sarbecovirus*) variants in the Moscow region using targeted high-throughput sequencing. *Problems of Virology* (*Voprosy Virusologii*). 2021; 66(4): 269-278. DOI: https://doi.org/10.36233/0507-4088-72

For correspondence: Kamil F. Khafizov, Ph.D. (Biol.), Head of the Scientific Group for the Development of New Diagnostic Methods, Department of Molecular Diagnostics and Epidemiology, FSBI «Central Research Institute for Epidemiology» of the Federal Service for Supervision of Consumer Rights Protection and Human Welfare (Rospotrebnadzor), Moscow, 111123, Russia. E-mail: kkhafizov@gmail.com

Information about the authors:
Borisova N.I., http://orcid.org/0000-0002-9672-0648
Kotov I.A., http://orcid.org/0000-0003-2416-5689
Kolesnikov A.A., http://orcid.org/0000-0003-3480-953X
Kaptelova V.V., http://orcid.org/0000-0003-0952-0830
Speranskaya A.S., http://orcid.org/0000-0001-6326-1249
Kondrasheva L.Yu., https://orcid.org/0000-0002-0147-4262
Tivanova E.V., http://orcid.org/0000-0003-1286-2612
Khafizov K.F., http://orcid.org/0000-0001-5524-0296
Akimkin V.G., https://orcid.org/0000-0003-4228-9044

* Shared first co-authorship

Contribution: Borisova N.I. – writing of the text, performing experiments; Kotov I.A. – writing of the text, preparing of the illustrations, bioinformatics analysis; Kolesnikov A.A. – performing experiments; Kaptelova V.V. – performing experiments; Speranskaya A.S. – performing experiments; Kondrasheva L.Yu. – organization of the biological material collection, polymerase chain reaction analysis; Tivanova E.V. – organization of the biological material collection; Khafizov K.F. – writing of the text, collection and processing of the materials, preparation of the illustrations, bioinformatics analysis, general edition of the article; Akimkin V.G. – writing of the resume, scientific editing, general edition of the article.

Funding. The research was funded by the State budget.
Conflict of interest. The authors declare no conflict of interest.
Ethics approval. The study was conducted with the informed consent of the patients. The research protocol was approved by the Ethics Committee of FSBI «Central Research Institute for Epidemiology» of the Surveillance of Consumer Rights Protection and Human Wellbeing (Rospotrebnadzor) (Protocol No. 111 dated December 22, 2020).

Received: 1 July 2021
Accepted 14 August 2021
Published 31 August 2021

НАУЧНАЯ СТАТЬЯ

DOI: https://doi.org/10.36233/0507-4088-72

Мониторинг распространения вариантов SARS-CoV-2 (*Coronaviridae: Coronavirinae: Betacoronavirus; Sarbecovirus*) на территории Московского региона с помощью таргетного высокопроизводительного секвенирования

Борисова Н.И.[1]*, Котов И.А.[1,2]*, Колесников А.А.[1], Каптелова В.В.[1], Сперанская А.С.[1], Кондрашева Л.Ю.[1], Тиванова Е.В.[1], Хафизов К.Ф.[1], Акимкин В.Г.[1]

[1]ФБУН «Центральный научно-исследовательский институт эпидемиологии» Федеральной службы по надзору в сфере защиты прав потребителей и благополучия человека (Роспотребнадзор), 111123, Москва, Россия;
[2]ФГАОУ ВО «Московский физико-технический институт (национальный исследовательский университет)», 141700, Долгопрудный, Россия

Введение. С начала пандемического распространения инфекции COVID-19, вызываемой коронавирусом SARS-CoV-2, международное научное сообщество регулярно фиксирует появление мутаций этого патогена, потенциально способных повысить его контагиозность и/или вирулентность. В частности, с конца 2020 г. в мире обнаружено несколько вызывающих озабоченность вариантов, включая альфа (B.1.1.7), бета (B.1.351), гамма (P.1) и дельта (B.1.617.2). Однако существующие механизмы поиска мутаций и выявления штаммов не всегда бывают достаточно эффективными, поскольку лишь небольшая доля получаемых от пациентов образцов возбудителя может быть исследована на наличие генетических изменений, например методом полногеномного секвенирования из-за его высокой стоимости.
Материал и методы. В исследовании применён способ таргетного высокопроизводительного секвенирования нового (следующего) поколения (next generation sequencing, NGS) наиболее значимых регионов гена, кодирующего S-гликопротеин (шиповидный, spike) вируса SARS-CoV-2, для чего разработана соответствующая праймерная панель. В среднем на платформе Illumina на 1 образец приходилось около 50 тыс. парноконцевых прочтений длиной ≥150 п.н. С помощью описанной методики нами исследованы 579 случайных образцов, полученных у проживающих в Московском регионе пациентов с новой коронавирусной инфекцией с февраля по июнь 2021 г.
Результаты. В работе продемонстрирована динамика представленности в Российской Федерации ряда штаммов нового коронавируса и нескольких его мутаций на протяжении февраля–июня 2021 г. Установлено, что штамм дельта появился на территории Москвы и Московской области в мае текущего года, а в июне стал доминирующим, частично вытеснив другие разновидности вируса.
Обсуждение. Полученные данные представляют возможность определять принадлежность образцов к упомянутым и некоторым другим штаммам, а описанный подход может быть использован для стандартизации процедуры поиска новых и существующих разновидностей SARS-CoV-2. Методика делает возможным изучение большого количества образцов в короткие сроки, позволяя получать более детальное представление об эпидемиологической ситуации в регионе.

Ключевые слова: *коронавирус; коронавирусная инфекция; SARS-CoV-2; секвенирование нового поколения; геном; штаммы*

*Первые авторы статьи

Для цитирования: Борисова Н.И., Котов И.А., Колесников А.А., Каптелова В.В., Сперанская А.С., Кондрашева Л.Ю., Тиванова Е.В., Хафизов К.Ф., Акимкин В.Г. Мониторинг распространения вариантов SARS-CoV-2 (*Coronaviridae: Coronavirinae: Betacoronavirus; Sarbecovirus*) на территории Московского региона с помощью таргетного высокопроизводительного секвенирования. *Вопросы вирусологии*. 2021; 66(4): 269-278. DOI: https://doi.org/10.36233/0507-4088-72

Для корреспонденции: Хафизов Камиль Фаридович, канд. биол. наук, руководитель научной группы разработки новых методов диагностики на основе технологий секвенирования следующего поколения отдела молекулярной диагностики и эпидемиологии, ФБУН «Центральный научно-исследовательский институт эпидемиологии» Федеральной службы по надзору в сфере защиты прав потребителей и благополучия человека (Роспотребнадзор), 111123, Москва, Россия. E-mail: kkhafizov@gmail.com

Участие авторов: Борисова Н.И. – написание текста, выполнение экспериментов; Котов И.А. – написание текста, подготовка иллюстраций, биоинформатический анализ; Колесников А.А. – выполнение экспериментов; Каптелова В.В. – выполнение экспериментов; Сперанская А.С. – выполнение экспериментов; Кондрашева Л.Ю. – организация сбора биологического материала, постановка полимеразной цепной реакции; Тиванова Е.В. – организация сбора биологического материала; Хафизов К.Ф. – написание текста, сбор и обработка материалов, подготовка иллюстраций, биоинформатический анализ, общая редакция статьи; Акимкин В.Г. – написание резюме, научное редактирование и общая редакция статьи.

Финансирование. Исследование выполнено за счет Государственного бюджета.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Этическое утверждение. Исследование проводилось при информированном согласии пациентов. Протокол исследования одобрен Этическим комитетом ФБУН «Центральный научно-исследовательский институт эпидемиологии» Федеральной службы по надзору в сфере защиты прав потребителей и благополучия человека (Роспотребнадзор) (ФБУН ЦНИИ эпидемиологии Роспотребнадзора) (Протокол № 111 от 22 декабря 2020 г.)

Поступила 1.07.2021
Принята к печати 14.08.2021
Опубликована 31.08.2021

Introduction

Since it was first detected in Wuhan (the People's Republic of China (PRC)) in December 2019 [1], the novel SARS-CoV-2 coronavirus has rapidly spread around the world and caused >4 million deaths [2]. Since the beginning of the pandemic, a number of effective therapeutic and preventive measures have been adopted to fight the novel coronavirus infection (COVID-19). They include using of therapeutic agents such as monoclonal antibodies (mAbs) [3, 4] and vaccines [5–8] where the spike (S) protein of the virus acts as an antigen.

At the end of 2020, the global scientific community described several SARS-CoV-2 variants of concern and special attention. They are Alpha (formerly referred to as the British variant, B.1.1.7), Beta (South-African, B.1.351), Gamma (Brazilian, P.1) and Delta (Indian, B.1.617.2) variants. The above variants of the virus sparked the interest of researchers after a number of geographic regions of the world had reported an increase in the frequency of human-to-human transmission of these infectious agents; later, the new variants of the pathogen were detected in many countries. For example, the Alpha variant spread rapidly in southeast England, causing a sharp increase in COVID-19 cases; shortly after, it was detected in the United States (USA Centers for Disease Control and Prevention (CDC)) [9], having become the dominant variant circulating in the country by April 2021. Similarly, the variants from the Republic of South Africa and Brazil caused outbreaks of COVID-19 in these countries. The above variants of SARS-CoV-2 give rise to concern due to the E484K mutation in the S protein, which may reduce the efficacy of some therapeutic mAbs, hinder virus neutralization *in vitro* and cause the virus to evade the immune protection developed following the infection or vaccination [8, 10–14].

In addition, 3 variants (Alpha, Beta and Gamma), in a key region of the S protein (the receptor-binding domain (RBD)), have the N501Y mutation associated with increased affinity for the angiotensin-converting enzyme 2 (ACE2) receptor. In its turn, it can contribute to increased transmissibility of the pathogen [15, 16]. Specialists of the Central Research Institute for Epidemiology (CRIE) of the Federal Service for Supervision of Consumer Rights Protection and Human Welfare (Rospotrebnadzor) designed a reagent kit for fast detection of the above mutation in the viral genome by using loop-mediated isothermal amplification (LAMP) [17] to significantly reduce the number of samples required for whole genome sequencing aimed to detect and monitor new strains containing the N501Y mutation. However, the further emergence of similar strains, including strains with other mutations in the S protein gene, showed that LAMP-associated genomic substitutions at primer binding sites could decrease the effectiveness of the above technique. In addition, the detection and exploration of SARS-CoV-2 variants result in an extended list of mutation-caused changes of interest and subject to monitoring. For example, the Delta strain responsible for a sharp increase in the number of cases in India was detected in Russia in May 2021, and has been rapidly spreading in this country since that time. Furthermore, several local strains have been detected in Russia, including Siberian (B.1.1.397+) and Northwestern (B.1.1.370.1) strains carrying mutations in the S protein gene [18, 19]. At the moment, the strains circulating in Russia are being studied and monitored [20]. For the foregoing reasons, scientists need more efficient and universal tools to identify

multiple significant mutations within a single cycle of the laboratory test. Although whole genome sequencing of viruses is undoubtedly the most detailed technique of the genetic analysis [21], economically speaking, it may not be the best option among the advanced sequencing techniques. It may pose difficulties when used in the setting of steadily increasing COVID-19 cases, including repeat disease cases. In addition, whole genome identification may be impractical for effective epidemiological surveillance, considering that the most significant changes take place in a small part of the viral genome.

In this article, we describe the identification of isolates belonging to different SARS-CoV-2 variants (including Alpha, Beta and Delta strains) by using targeted high-throughput next-generation sequencing. We designed a primer panel intended for fast and efficient targeted amplification of genomic fragments, bypassing ligation of adapters when preparing samples for sequencing due to modification of oligonucleotides during the synthesis. The amplified regions include the most epidemiologically significant mutations corresponding to K417T, L452R, T478K, E484K, S494P, N501Y, A570D, P681H and other amino acid substitutions as well as to HV69-70 and Y144 deletions **(Fig. 1)**. Therefore, not only the cost of sample preparation is significantly reduced, but also the amount of generated data is substantially reduced, making it possible to study a larger number of virus samples within a short period. The latter point is especially important in the situation when the frequency of different strains should be promptly assessed for the region or country of interest. Using our primer panel, we studied 579 random virus samples collected in Moscow and Moscow Region in February–June 2021, and found changes in frequencies of some mutations and strains in the above region.

Materials and methods

In our study, we used biological materials from nasopharyngeal swabs from patients with symptoms of novel coronavirus infection. The presence of SARS-CoV-2 in these patients was confirmed with a real-time reverse-transcription polymerase chain reaction (rRT-PCR) assay by using an AmpliSens Cov-Bat-FL reagent kit (AmpliSens, Russia). The study was conducted with the informed consent of the patients; the research protocol was approved by the CRIE Ethics Committee (Protocol No. 111 of December 22, 2020). The samples were placed in the transport medium (CRIE, Russia). The RNA was extracted from the clinical material with the help of a RIBO-prep reagent kit (AmpliSens); the reverse transcription was performed by using a REVERTA-L kit (AmpliSens). Only the clinical samples, in which the cycle threshold (Ct) did not exceed 20 in the PCR test, were selected.

For amplification in a T100 Thermal Cycler (BioRad, USA), we used PCR-mix-2 blue (AmpliSens) containing *Taq*-polymerase. Then, the PCR-products were cleaned from the reaction mixture by using AMPureXP beads (Beckman Coulter, USA). The temperature profile for the amplification was as follows: 1) denaturation at 95 ºC for 30 seconds; 2) 38 amplification cycles: 95 ºC – 30 sec, 60 ºC – 20 sec, 72 ºC – 60 sec; 3) final elongation at 72ºC for 3 min. Indexing was performed by using PCR-mix-2 blue and EvaGreen (Biotium, USA) as a dye and a QuantStudio 5 Real-Time PCR System (ThermoFisher Scientific, USA). The indexing was conducted with index primers compatible with a Nextera XT Index Kit v2 (N7xx – 26 possible variants and S5xx – 18 variants; for more information see the Illumina Adapter Sequences, Document # 1000000002694 v16, April 2021, https://support-docs.illumina.com/SHARE/AdapterSeq/illumina-adapter-sequences.pdf). The double indexing improves the accuracy of sample identification and makes it possible to test a large number of samples at a time. The temperature profile of the indexing was: 1) 98 ºC – 30 sec; 2) 15 cycles: 98 ºC – 10 sec, 65 ºC – 1 min 15 sec. Then, the samples went through another clean-up from the reaction mixture. The DNA concentration was measured with a Qubit 4.0 fluorometer (ThermoFisher Scientific). Based on the concentration, the samples were pooled and the pool quality was evaluated by using an

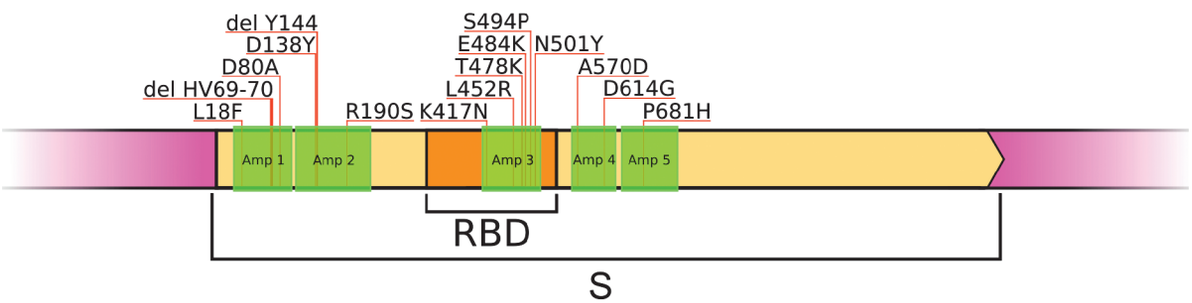

Fig. 1. Localization of the amplicons obtained using the primer panel relative to the SARS-CoV-2 virus genome.

Note. Several amino acid substitutions and deletions covered by the panel are indicated.
Amp 1–5, amplicons; S, S-protein; RBD, receptor binding domain.

Рис. 1. Локализация ампликонов, получаемых с помощью разработанной панели праймеров, относительно генома SARS-CoV-2.

Примечание. Указан ряд аминокислотных замен и делеций, покрываемых панелью.
Amp 1–5 – ампликоны; S – S-белок; RBD – рецептор-связывающий домен.

Agilent 2100 bioanalyzer (Agilent Technologies, USA). Denaturation and estimation of the loading volume were performed according to the manufacturer's instructions.

The sequencing was conducted on the Illumina MiSeq platform (Illumina, USA) by using a MiSeq Reagent Kit v2 (PE 150 + 150 or PE 250 + 250 cycles) or a MiSeq Reagent Kit v3 (PE 300 + 300 cycles) as well as on the Illumina HiSeq platform by using a v2 Rapid SBS kit (PE 250 + 250 cycles). There were around 50 thousand reads per sample, which is equivalent to approximately 0.3% of the information obtained when using a MiSeq Reagent Kit v2 (PE 150 + 150) or approximately 0.1% – when using a MiSeq Reagent Kit v3 (PE 300 + 300 cycles) during sequencing. The above approach makes it possible to avoid excessive expenses associated with testing numerous samples, while the average coverage level for most of the samples is ≥2,000.

The primers for the panel for targeted amplification of fragments of the S protein gene were selected manually based on the available information about the known epidemiologically significant mutations and the information about conserved regions of the genome. The melting temperature of oligonucleotides and the level of interaction between them were estimated with a Multiple Primer Analyzer (ThermoFisher Scientific). The BLASTn program [22] was used to evaluate the specificity of each obtained sequence in reference to all known organisms (first of all, the human organism, the genetic material of which is best represented in the sample), thus eliminating any non-specific interaction between the primer and DNA regions of human and other organisms. A total of 5 pairs of oligonuc-leotides were obtained; they also contained additional adapter sequences required for time and cost reduction during sample preparation. The synthesis was performed at Syntol Research and Production Company, LLC (Russia). The structure of primers is shown in the **Table**. The lengths of amplicons were selected to provide overall coverage of the target regions during their high-throughput sequencing on Illumina MiSeq platforms with the help of MiSeq v2 reagent kits (300 cycles), v2 (500 cycles) and v3 (600 cycles) and on Illumina HiSeq platforms by using a v2 Rapid SBS kit (500 cycles).

To analyze the sequencing data, the obtained reads were aligned to the reference genome of the SARS-CoV-2 virus by using the BWA software [23]. The BBTools program [24] was used to trim the adapter sequences in the reads. The coverage of 5 reads was accepted as the minimum coverage. The GATK program was used to find genetic variants [25]. The obtained sequences were uploaded to the VGARus nucleotide sequence database for SARS-CoV-2 and its mutations (https://genome.crie.ru/).

The VMD software was used for visualization of the S protein molecule and for the subsequent drawings [26]. We used a structural model of S protein (the Protein Data Bank (PDB) ID: 7CAB), which was received by using cryoelectron microscopy [27].

Results

The period, during which the study was conducted in the Moscow Region, coincided with the time when SARS-CoV-2 strains of concern started spreading all over the world, including Russia, contributing to the risk of new waves of disease cases in a number of countries.

While only a relatively small (~2%) proportion of the Alpha strain was found in the isolates obtained in February 2021, its frequency increased to ~20% in March, supporting the data on high contagiousness of the strain [28]. However, no further widespread occurrence of the strain was recorded; the frequency rate gradually decreased, reaching almost zero in the middle of June. Most likely, this decrease was caused by the Delta strain that spread to Russia in May 2021 and previously may have caused an increase in the number of cases and deaths among the population of India. By mid-June 2021, the prevalence of this strain increased to 70% in Moscow, and, according to the data not included in this study, it has been steadily increasing to >90%. In addition, at the end of June 2021, there were cases of infection with the Delta Plus strain containing the additional K417N substitution [29], which was previously found in the Beta strain and which is located in one of the SARS-CoV-2 genome regions amplifiable by using the new primer panel **(Fig. 2)**.

Notably, no significant spread of the Beta variant has been recorded in Moscow, though in April 2021, its proportion quickly increased to 13%, causing concern among the population. The data suggest that the above strain of the novel coronavirus can partially evade the neutralizing effect of antibodies (Abs) developed following the coronavirus infection or vaccination [30].

In addition, the B.1.1.523 strain should be mentioned; its frequency had significantly increased by May 2021. The fact that it has the E484K mutation shows that it (like the Beta strain) can be more resistant to Abs. Besides, this variant has alterations in the S protein – the S494P substitution and deletion of 3 amino acids. On the other hand, its proportion also sharply decreased in June 2021, when the Delta variant became the dominant variant. At the same time, we have not detected any cases of infection caused by the Gamma (Brazilian) strain.

Fig. 2 shows changes in the frequency of virus variants with N501Y or E484K mutations and not assigned to any of the specified strains due to the absence of other relevant changes in the genome.

Charts in **Fig. 3** show changes in the prevalence rate for some mutations, which were previously identified as mutations altering the virus properties and causing increased transmissibility or evasion of the protective effect of Abs. For example, in February 2021, almost 15% of all isolates had the E484K mutation, though the Beta strain carrying this substitution had not been detected in Russia by that time. This fact proves that although such alterations in the genome can give certain advantages to the pathogen, their presence not always results in the widespread occurrence of the pathogen, if there are no other significant mutations, the combinations of which are frequently unclear. The subsequent months were characterized by an increase in the prevalence of several mutations, including E484K, N501Y, and S494P. In April 2021 when the Beta strain was de-

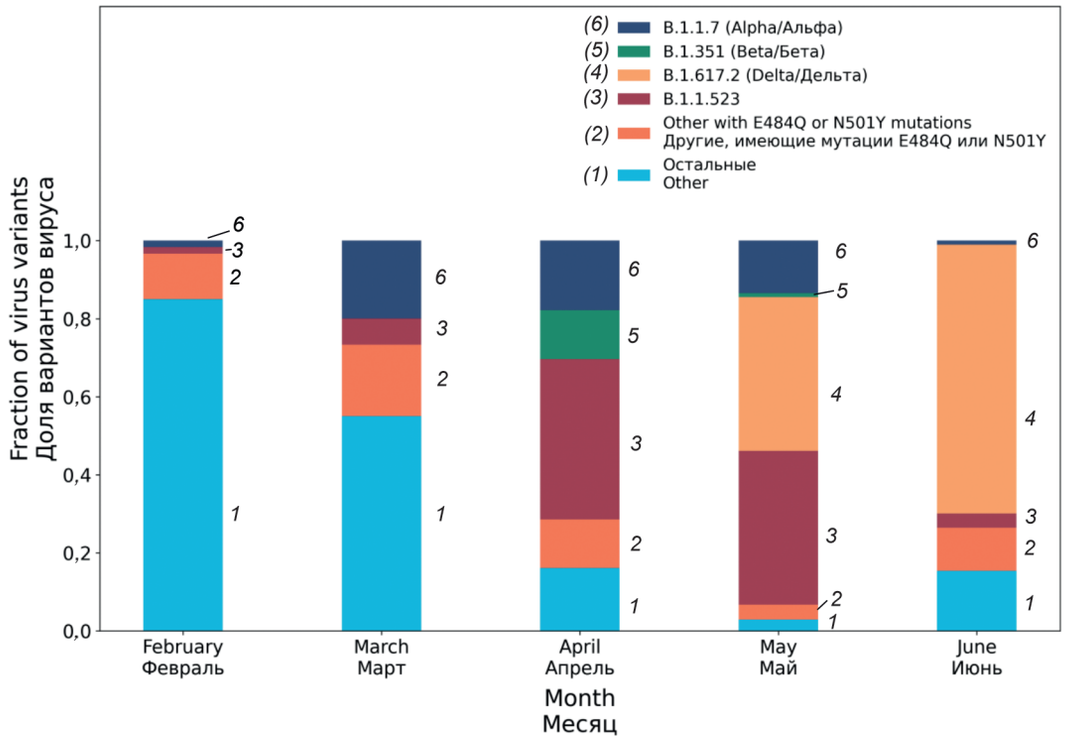

Fig. 2. Representation of various variants of SARS-CoV-2 from February to June 2021 in Moscow and the Moscow Region.
Рис. 2. Представленность различных вариантов вируса SARS-CoV-2 с февраля по июнь 2021 г. на территории Москвы и Московской области.

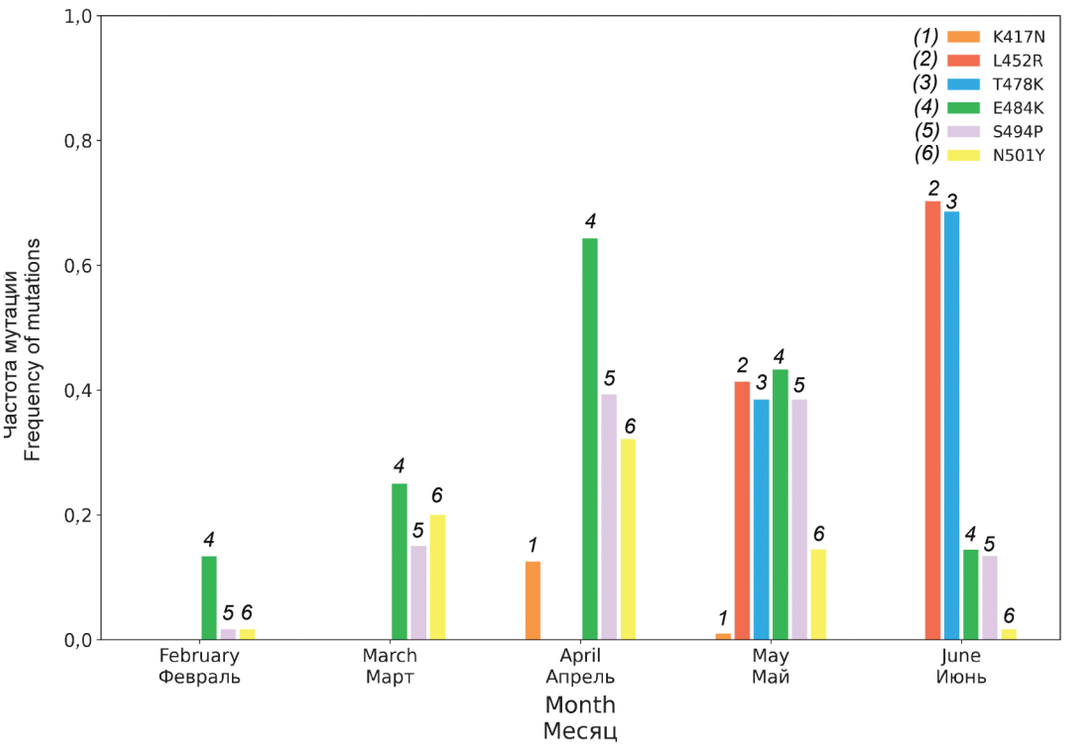

Рис. 3. Частота отдельных мутаций вируса SARS-CoV-2 в разные месяцы 2021 г. на территории Москвы и Московской области.
Fig. 3. Frequency of individual SARS-CoV-2 mutations in different months of 2021 in Moscow and the Moscow Region.

tected in Russia, there was also an increase in the frequency of the K417N mutation, which is also present in the Delta Plus variant that was first detected in the Russian Federation only at the end of June 2021. Finally, L452R and T478K mutations belonging to the Delta strain occurred in May 2021 and are currently present in most of the SARS-CoV-2 genomes.

The obtained data show that by using a small primer panel for amplification of genomic fragments of the novel coronavirus and through the subsequent targeted sequencing, we can detect nearly all known alterations in the coronavirus S protein gene and identify strains of the pathogen, making it possible to monitor their frequency (prevalence) during specific periods. At the same time, it should be remembered that, by no means always, some mutations can result in emergence of a more contagious and epidemiologically dangerous variant of the virus; only in their combination, mutation changes can give new

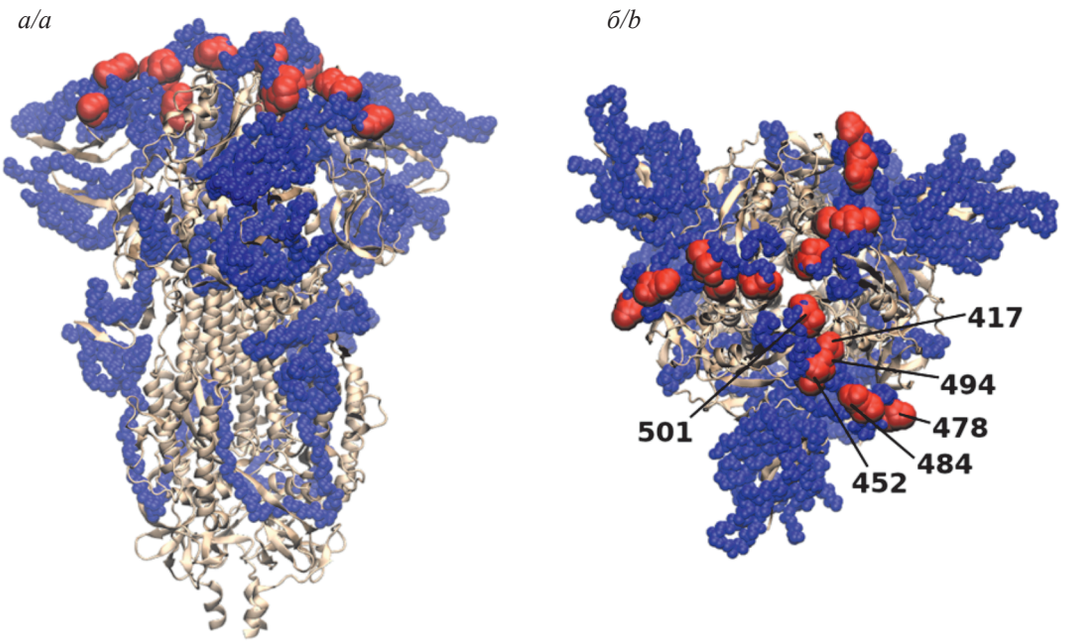

Fig. 4. Structural model of S-protein obtained using cryo-electron microscopy (PDB ID: 7CAB). View from the side (*a*) and from above (*b*).

Note. The numbers from 417 to 501 indicate the amino acid residues in the corresponding positions. The regions included in the primer panel are highlighted in blue and shown as small spheres. Significant mutations (K417T, L452R, T478K, E484K, S494P, N501Y) are shown in red as larger spheres and are indicated in figure (*b*).

Рис. 4. Структурная модель S-белка, полученная с использованием криоэлектронной микроскопии (PDB ID: 7CAB).
а) – вид сбоку, *б)* – вид сверху.

Примечание. Числами от 417 до 501 обозначены аминокислотные остатки в соответствующих позициях. Регионы, вошедшие в праймерную панель, подсвечены синим цветом и показаны в виде небольших сфер. Значимые мутации (K417T, L452R, T478K, E484K, S494P, N501Y) показаны красным цветом как сферы большего диаметра и указаны на рисунке (*б*).

Sequences of oligonucleotides in the primer panel

Последовательности олигонуклеотидов в праймерной панели

Oligonucleotide ID Идентификатор олигонуклеотида	Oligonucleotide sequence Последовательность олигонуклеотида
F1	5’-TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG-TGT TTT TCT TGT TTT ATT GCC ACT AGT CTC-3’
R1	5’-GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA G-TC TTA TGT TAG ACT TCT CAG TGG AAG CA-3’
F2	5’-TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG-GCT GGA TTT TTG GTA CTA CTT TAG ATT CG-3’
R2	5’-GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA G-AA TCT ACC AAT GGT TCT AAA GCC GAA AAA C-3’
F3	5’-TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG-GCT CCA GGG CAA ACT GGA AA-3’
R3	5’-GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA G-CT GTA TGG TTG GTA ACC AAC ACC AT-3’
F4	5’-TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG-CCA ACA ATT TGG CAG AGA CAT TG-3’
R4	5’-GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA G-CGC CAA GTA GGA GTA AGT TGA TCT G-3’
F5	5’-TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG-AAA CAC GTG CAG GCT GTT TAA TAG G-3’
R5	5’-GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA G-CT ACT GAT GTC TTG GTC ATA GAC ACT GG-3’

Note. Specific fragments are separated from adapters with a «-» symbol.

Примечание. Специфические фрагменты отделены от адаптеров символом «-».

properties to the pathogen, sometimes causing its widespread occurrence.

Discussion

In this methodological work, we discussed the possibility to detect a number of SARS-CoV-2 strains, including Alpha (British, B.1.1.7), Beta (South-African, B.1.351), Gamma (Brazilian, P.1), Delta (Indian, B.1.617.2) variants, by using targeted high-throughput next-generation sequencing. For the above purpose, we designed a primer panel (**Table**) intended for efficient targeted amplification of genomic fragments of the novel coronavirus. The amplified regions include a large number of known significant mutations altering the properties of the virus, thus making it possible to decrease substantially the sequencing costs and to increase the number of samples under study. The latter advantage is highly important when significant changes in the S protein gene must be detected to identify the virus isolates belonging to different strains. To show how the primer panel is used, we studied 579 random SARS-CoV-2 samples collected from patients in Moscow and Moscow Region in February–June 2021. Note that most of the significant changes in the S protein gene are located in its small fragments primarily encoding amino acids found on the surface **(Fig. 4)**. Our study also demonstrates the importance of promotion of population, genomic and epidemiological studies aimed at detection, prevalence tracking and monitoring of new variants of viral pathogens.

We have shown a rapid change in the proportion of different genetic variants of the COVID-19 pathogen within the specified period, including the occurrence and swift spread of the Delta strain in Moscow and Moscow Region during May–June 2021; the above strain is partially responsible for the new wave of disease cases in Moscow this summer. In the meantime, the population's inobservance of social distancing and personal protection measures as well as a low vaccination level are important factors boosting the spread of infection. Thus, the conducted study also demonstrates the importance of promotion of population, genomic and epidemiological studies aimed at detection, prevalence tracking and monitoring of the new SARS-CoV-2 variants and variants of other viral pathogens characterized by high variability.

Although we should assume that the new significant genomic variants may affect other important fragments of the viral genetic material, the using of a panel with few primers provides a substantial reduction in expenses on detection of currently circulating strains, while the simple design makes it possible to change promptly the structure of oligonucleotides, bringing in regular updates on genomes. It should be noted that targeted sequencing cannot completely replace whole genome one, which is essential for detecting all genomic alterations and a detailed phylogenetic analysis. For a detailed bioinformation analysis, specialists can use the recently built VGARus coronavirus genome database that as of the end of June 2021 had ~15,000 sequences, including ~8,000 complete genomes obtained through sequencing of isolates from different regions of Russia.

REFERENCES

1. Zhou P., Yang X.L., Wang X.G., Hu B., Zhang L., Zhang W., et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*. 2020; 579(7798): 270–3. https://doi.org/10.1038/s41586-020-2012-7
2. COVID-19 data in motion. Available at: https://coronavirus.jhu.edu (accessed July 24, 2021).
3. Chen P., Nirula A., Heller B., Gottlieb R.L., Boscia J., Morris J., et al. SARS-CoV-2 neutralizing antibody LY-CoV555 in outpatients with Covid-19. *N. Engl. J. Med*. 2021; 384(3): 229–37. https://doi.org/10.1056/nejmoa2029849
4. Weinreich D.M., Sivapalasingam S., Norton T., Ali S., Gao H., Bhore R., et al. REGN-COV2, a neutralizing antibody cocktail, in outpatients with Covid-19. *N. Engl. J. Med*. 2021; 384(3): 238–51. https://doi.org/10.1056/nejmoa2035002
5. Baden L.R., El Sahly H.M., Essink B., Kotloff K., Frey S., Novak R., et al. Efficacy and safety of the mRNA-1273 SARS-CoV-2 vaccine. *N. Engl. J. Med*. 2021; 384(5): 403–16. https://doi.org/10.1056/nejmoa2035389
6. Polack F.P., Thomas S.J., Kitchin N., Absalon J., Gurtman A., Lockhart S., et al. Safety and efficacy of the BNT162b2 mRNA Covid-19 Vaccine. *N. Engl. J. Med*. 2020; 383(27): 2603–15. https://doi.org/10.1056/nejmoa2034577
7. Jones I., Roy P. Sputnik V COVID-19 vaccine candidate appears safe and effective. *Lancet*. 2021; 397(10275): 642–3. https://doi.org/10.1016/s0140-6736(21)00191-4
8. Ryzhikov A.B., Ryzhikov E.A., Bogryantseva M.P., Usova S.V., Danilenko E.D., Nechaeva E.A., et al. A single blind, placebo-controlled randomized study of the safety, reactogenicity and immunogenicity of the "EpiVacCorona" vaccine for the prevention of COVID-19, in volunteers aged 18–60 years (phase I–II) *[Prostoe slepoe platsebo-kontroliruemoe randomizirovannoe issledovanie bezopasnosti, reaktogennosti i immunogennosti vaktsiny «EpiVakKorona» dlya profilaktiki COVID-19 na dobrovol'tsakh v vozraste 18–60 let (faza I–II)]*. *Infektsiya i immunitet*. 2021; 11(2): 283–96. https://doi.org/10.15789/2220-7619-ASB-1699 (in Russian)
9. About Variants of the Virus that Causes COVID-19. Available at: https://www.cdc.gov/coronavirus/2019-ncov/Transmission/variant.html (accessed July 26, 2021).
10. Wang W.B., Liang Y., Jin Y.Q., Zhang J., Su J.G., Li Q.M. E484K mutation in SARS-CoV-2 RBD enhances binding affinity with hACE2 but reduces interactions with neutralizing antibodies and nanobodies: binding free energy calculation studies. *bioRxiv*. 2021; Preprint. https://doi.org/10.1101/2021.02.17.431566
11. Garcia-Beltran W.F., Lam E.C., St. Denis K., Nitido A.D., Garcia Z.H., Hauser B.M., et al. Circulating SARS-CoV-2 variants escape neutralization by vaccine-induced humoral immunity. *medRxiv*. 2021; Preprint. https://doi.org/10.1101/2021.02.14.21251704
12. Liu H., Wei P., Zhang Q., Chen Z., Aviszus K., Downing W., et al. 501Y.V2 and 501Y.V3 variants of SARS-CoV-2 lose binding to bamlanivimab *in vitro*. *MAbs*. 2021; 13(1): 1919285. https://doi.org/10.1080/19420862.2021.1919285
13. Yuan M., Huang D., Lee C.D., Wu N.C., Jackson A.M., Zhu X., et al. Structural and functional ramifications of antigenic drift in recent SARS-CoV-2 variants. *Science*. 2021; eabh1139. https://doi.org/10.1126/science.abh1139
14. Ikegame S., Siddiquey M.N.A., Hung C.-T., Haas G., Brambilla L., Oguntuyo K.Y., et al. Neutralizing activity of Sputnik V vaccine sera against SARS-CoV-2 variants. *medRxiv.* 2021.03.31.21254660. doi: https://doi.org/10.1101/2021.03.31.21254660
15. Gard N., Buzko O., Spilman P., Niazi K., Rabizadeh S., Soon-Shiong P. Molecular dynamic simulation reveals E484K mutation enhances spike RBD-ACE2 affinity and the combination of E484K, K417N and N501Y mutations (501Y.V2 variant) induces conformational change greater than N501Y mutant alone, potentially result-

ing in an escape mutant *bioRxiv.* 2021.01.13.426558. doi: https://doi.org/10.1101/2021.01.13.426558
16. Tian F., Tong B., Sun L., Shi S., Zheng B., Wang Z., et al. Mutation N501Y in RBD of spike protein strengthens the interaction between COVID-19 and its receptor ACE2. *bioRxiv*. 2021; Preprint. https://doi.org/10.1101/2021.02.14.431117
17. Khafizov K.F., Petrov V.V., Krasovitov K.V., Zolkina M.V., Akimkin V.G. Rapid diagnostics of novel coronavirus infection by loop-mediated isothermal amplification *[Ekspress-diagnostika novoy koronavirusnoy infektsii s pomoshch'yu reaktsii petlevoy izotermicheskoy amplifikatsii]*. *Voprosy virusologii*. 2021; 66(1): 17–28. https://doi.org/10.36233/0507-4088-42 (in Russian)
18. Gladkikh A., Dolgova A., Dedkov V., Sbarzaglia V., Kanaeva O., Popova A., et al. Characterization of a novel SARS-CoV-2 genetic variant with distinct spike protein mutations. *Viruses*. 2021; 13(6): 1029. https://doi.org/10.3390/v13061029
19. Klink G.V., Safina K.R., Garushyants S.K., Moldovan M., Nabieva E., Komissarov A.B., et al. Spread of endemic SARS-CoV-2 lineages in Russia. *medRxiv*. 2021; Preprint. https://doi.org/10.1101/2021.05.25.21257695
20. Komissarov A.B., Safina K.R., Garushyants S.K., Fadeev A.V., Sergeeva M.V., Ivanova A.A., et al. Genomic epidemiology of the early stages of the SARS-CoV-2 outbreak in Russia. *Nat. Commun*. 2021; 12(1): 649. https://doi.org/10.1038/s41467-020-20880-z
21. Long S.W., Olsen R.J., Christensen P.A., Subedi S., Olson R., Davis J.J., et al. Sequence Analysis of 20,453 Severe Acute Respiratory Syndrome Coronavirus 2 Genomes from the Houston Metropolitan Area Identifies the Emergence and Widespread Distribution of Multiple Isolates of All Major Variants of Concern. *Am. J. Pathol.* 2021; 191(6): 983–92. https://doi.org/10.1016/j.ajpath.2021.03.004
22. Altschul S.F., Gish W., Miller W., Myers E.W., Lipman D.J. Basic local alignment search tool. *J. Mol. Biol.* 1990; 215(3): 403–10. https://doi.org/10.1016/s0022-2836(05)80360-2
23. Li H., Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*. 2009; 25(14): 1754–60. https://doi.org/10.1093/bioinformatics/btp324
24. Bushnell B., Rood J., Singer E. BBMerge – Accurate paired shotgun read merging via overlap. *PLoS One*. 2017; 12(10): e0185056. https://doi.org/10.1371/journal.pone.0185056
25. Poplin R., Ruano-Rubio V., DePristo M.A., Fennell T.J., Carneiro M.O., Van der Auwera G.A., et al. Scaling accurate genetic variant discovery to tens of thousands of samples. *bioRxiv.* 2018; 201178. doi: https://doi.org/10.1101/201178
26. Humphrey W., Dalke A., Schulten K. VMD: visual molecular dynamics. *J. Mol. Graph*. 1996; 14(1): 33–8, 27–8. https://doi.org/10.1016/0263-7855(96)00018-5
27. Lv Z., Deng Y.Q., Ye Q., Cao L., Sun C.Y., Fan C., et al. Structural basis for neutralization of SARS-CoV-2 and SARS-CoV by a potent therapeutic antibody. *Science*. 2020; 369(6510): 1505–9. https://doi.org/10.1126/science.abc5881
28. Davies N.G., Abbott S., Barnard R.C., Jarvis C.I., Kucharski A.J., Munday J.D., et al. Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England. *Science*. 2021; 372(6538): eabg3055. https://doi.org/10.1126/science.abg3055
29. Expert comment on the 'Delta plus' variant (B.1.617.2 with the addition of K417N mutation). Available at: https://www.sciencemediacentre.org/expert-comment-on-the-delta-plus-variant-b-1-617-2-with-the-addition-of-k417n-mutation/ (accessed July 26, 2021).
30. Weekly epidemiological update on COVID-19 – 22 June 2021. Available at: https://www.who.int/publications/m/item/weekly-epidemiological-update-on-covid-19---22-june-2021 (accessed July 24, 2021).

ЛИТЕРАТУРА

1. Zhou P., Yang X.L., Wang X.G., Hu B., Zhang L., Zhang W., et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*. 2020; 579(7798): 270–3. https://doi.org/10.1038/s41586-020-2012-7
2. COVID-19 data in motion. Available at: https://coronavirus.jhu.edu (accessed July 24, 2021).
3. Chen P., Nirula A., Heller B., Gottlieb R.L., Boscia J., Morris J., et al. SARS-CoV-2 neutralizing antibody LY-CoV555 in outpatients with Covid-19. *N. Engl. J. Med*. 2021; 384(3): 229–37. https://doi.org/10.1056/nejmoa2029849
4. Weinreich D.M., Sivapalasingam S., Norton T., Ali S., Gao H., Bhore R., et al. REGN-COV2, a neutralizing antibody cocktail, in outpatients with Covid-19. *N. Engl. J. Med.* 2021; 384(3): 238–51. https://doi.org/10.1056/nejmoa2035002
5. Baden L.R., El Sahly H.M., Essink B., Kotloff K., Frey S., Novak R., et al. Efficacy and safety of the mRNA-1273 SARS-CoV-2 vaccine. *N. Engl. J. Med*. 2021; 384(5): 403–16. https://doi.org/10.1056/nejmoa2035389
6. Polack F.P., Thomas S.J., Kitchin N., Absalon J., Gurtman A., Lockhart S., et al. Safety and efficacy of the BNT162b2 mRNA Covid-19 Vaccine. *N. Engl. J. Med*. 2020; 383(27): 2603–15. https://doi.org/10.1056/nejmoa2034577
7. Jones I., Roy P. Sputnik V COVID-19 vaccine candidate appears safe and effective. *Lancet*. 2021; 397(10275): 642–3. https://doi.org/10.1016/s0140-6736(21)00191-4
8. Рыжиков А.Б., Рыжиков Е.А., Богрянцева М.П., Усова С.В., Даниленко Е.Д., Нечаева Е.А., и др. Простое слепое плацебо-контролируемое рандомизированное исследование безопасности, реактогенности и иммуногенности вакцины «ЭпиВакКорона» для профилактики COVID-19 на добровольцах в возрасте 18–60 лет (фаза I–II). *Инфекция и иммунитет*. 2021; 11(2): 283–96. https://doi.org/10.15789/2220-7619-ASB-1699
9. About Variants of the Virus that Causes COVID-19. Available at: https://www.cdc.gov/coronavirus/2019-ncov/Transmission/variant.html (accessed July 26, 2021).
10. Wang W.B., Liang Y., Jin Y.Q., Zhang J., Su J.G., Li Q.M. E484K mutation in SARS-CoV-2 RBD enhances binding affinity with hACE2 but reduces interactions with neutralizing antibodies and nanobodies: binding free energy calculation studies. *bioRxiv*. 2021; Preprint. https://doi.org/10.1101/2021.02.17.431566
11. Garcia-Beltran W.F., Lam E.C., St. Denis K., Nitido A.D., Garcia Z.H., Hauser B.M., et al. Circulating SARS-CoV-2 variants escape neutralization by vaccine-induced humoral immunity. *medRxiv*. 2021; Preprint. https://doi.org/10.1101/2021.02.14.21251704
12. Liu H., Wei P., Zhang Q., Chen Z., Aviszus K., Downing W., et al. 501Y.V2 and 501Y.V3 variants of SARS-CoV-2 lose binding to bamlanivimab *in vitro*. *MAbs*. 2021; 13(1): 1919285. https://doi.org/10.1080/19420862.2021.1919285
13. Yuan M., Huang D., Lee C.D., Wu N.C., Jackson A.M., Zhu X., et al. Structural and functional ramifications of antigenic drift in recent SARS-CoV-2 variants. *Science*. 2021; eabh1139. https://doi.org/10.1126/science.abh1139
14. Ikegame S., Siddiquey M.N.A., Hung C.-T., Haas G., Brambilla L., Oguntuyo K.Y., et al. Neutralizing activity of Sputnik V vaccine sera against SARS-CoV-2 variants. *medRxiv.* 2021.03.31.21254660. doi: https://doi.org/10.1101/2021.03.31.21254660
15. Gard N., Buzko O., Spilman P., Niazi K., Rabizadeh S., Soon-Shiong P. Molecular dynamic simulation reveals E484K mutation enhances spike RBD-ACE2 affinity and the combination of E484K, K417N and N501Y mutations (501Y.V2 variant) induces conformational change greater than N501Y mutant alone, potentially resulting in an escape mutant *bioRxiv.* 2021.01.13.426558. doi: https://doi.org/10.1101/2021.01.13.426558
16. Tian F., Tong B., Sun L., Shi S., Zheng B., Wang Z., et al. Mutation N501Y in RBD of spike protein strengthens the interaction between COVID-19 and its receptor ACE2. *bioRxiv*. 2021; Preprint. https://doi.org/10.1101/2021.02.14.431117
17. Хафизов К.Ф., Петров В.В., Красовитов К.В., Золкина М.В., Акимкин В.Г. Экспресс-диагностика новой коронавирусной инфекции с помощью реакции петлевой изотермической амплификации. *Вопросы вирусологии*. 2021; 66(1): 17–28. https://doi.org/10.36233/0507-4088-42

18. Gladkikh A., Dolgova A., Dedkov V., Sbarzaglia V., Kanaeva O., Popova A., et al. Characterization of a novel SARS-CoV-2 genetic variant with distinct spike protein mutations. *Viruses*. 2021; 13(6): 1029. https://doi.org/10.3390/v13061029
19. Klink G.V., Safina K.R., Garushyants S.K., Moldovan M., Nabieva E., Komissarov A.B., et al. Spread of endemic SARS-CoV-2 lineages in Russia. *medRxiv*. 2021; Preprint. https://doi.org/10.1101/2021.05.25.21257695
20. Komissarov A.B., Safina K.R., Garushyants S.K., Fadeev A.V., Sergeeva M.V., Ivanova A.A., et al. Genomic epidemiology of the early stages of the SARS-CoV-2 outbreak in Russia. *Nat. Commun*. 2021; 12(1): 649. https://doi.org/10.1038/s41467-020-20880-z
21. Long S.W., Olsen R.J., Christensen P.A., Subedi S., Olson R., Davis J.J., et al. Sequence Analysis of 20,453 Severe Acute Respiratory Syndrome Coronavirus 2 Genomes from the Houston Metropolitan Area Identifies the Emergence and Widespread Distribution of Multiple Isolates of All Major Variants of Concern. *Am. J. Pathol*. 2021; 191(6): 983–92. https://doi.org/10.1016/j.ajpath.2021.03.004
22. Altschul S.F., Gish W., Miller W., Myers E.W., Lipman D.J. Basic local alignment search tool. *J. Mol. Biol.* 1990; 215(3): 403–10. https://doi.org/10.1016/s0022-2836(05)80360-2
23. Li H., Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*. 2009; 25(14): 1754–60. https://doi.org/10.1093/bioinformatics/btp324
24. Bushnell B., Rood J., Singer E. BBMerge – Accurate paired shotgun read merging via overlap. *PLoS One*. 2017; 12(10): e0185056. https://doi.org/10.1371/journal.pone.0185056
25. Poplin R., Ruano-Rubio V., DePristo M.A., Fennell T.J., Carneiro M.O., Van der Auwera G.A., et al. Scaling accurate genetic variant discovery to tens of thousands of samples. *bioRxiv.* 2018; 201178. doi: https://doi.org/10.1101/201178
26. Humphrey W., Dalke A., Schulten K. VMD: visual molecular dynamics. *J. Mol. Graph*. 1996; 14(1): 33–8. https://doi.org/10.1016/0263-7855(96)00018-5
27. Lv Z., Deng Y.Q., Ye Q., Cao L., Sun C.Y., Fan C., et al. Structural basis for neutralization of SARS-CoV-2 and SARS-CoV by a potent therapeutic antibody. *Science*. 2020; 369(6510): 1505–9. https://doi.org/10.1126/science.abc5881
28. Davies N.G., Abbott S., Barnard R.C., Jarvis C.I., Kucharski A.J., Munday J.D., et al. Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England. *Science*. 2021; 372(6538): eabg3055. https://doi.org/10.1126/science.abg3055
29. Expert comment on the 'Delta plus' variant (B.1.617.2 with the addition of K417N mutation). Available at: https://www.sciencemediacentre.org/expert-comment-on-the-delta-plus-variant-b-1-617-2-with-the-addition-of-k417n-mutation/ (accessed July 26, 2021).
30. Weekly epidemiological update on COVID-19 – 22 June 2021. Available at: https://www.who.int/publications/m/item/weekly-epidemiological-update-on-covid-19---22-june-2021 (accessed July 24, 2021).