

<https://doi.org/10.29001/2073-8552-2021-36-4-132-143>
УДК 616.98:578.834.1-036.21-039.5:578.5](470+571)

Линии коронавируса SARS-CoV-2 российского происхождения: генетическая характеристика и корреляции с клиническими параметрами, тяжестью коронавирусной инфекции

О.С. Глотов^{1, 2, 4}, А.Н. Чернов⁵, А.И. Коробейников³,
Р.С. Калинин⁴, В.В. Цай⁴, А.Ю. Анисенкова^{1, 3}, С.П. Уразов¹,
А.А. Лапидус³, С.В. Мосенко¹, С.Г. Щербак^{1, 3}

¹Городская больница № 40 Курортного административного района,
197706, Российская Федерация, Санкт-Петербург, Сестрорецк, ул. Борисова, 9

²Научно-исследовательский институт акушерства, гинекологии и репродуктологии имени Д.О. Отта,
199034, Российская Федерация, Санкт-Петербург, Менделеевская линия, 3

³Санкт-Петербургский государственный университет,
199034, Российская Федерация, Санкт-Петербург, Университетская наб., 7

⁴Детский научно-клинический центр инфекционных болезней Федерального медико-биологического агентства,
197022, Российская Федерация, Санкт-Петербург, ул. Профессора Попова, 9

⁵Институт экспериментальной медицины,
197376, Российская Федерация, Санкт-Петербург, ул. Академика Павлова, 12

Аннотация

Во время пандемии COVID-19 актуальным является выявление новых белковых и генных мишеней коронавируса (SARS-CoV-2) и организма человека, которые могут оказаться маркерами тяжести и исхода заболевания.

Цель: провести генетический анализ образцов РНК SARS-CoV-2 и установить корреляции генетических показателей и характера SNP с клиническими данными и степенью тяжести инфекции COVID-19.

Материал и методы. В исследование включены образцы вирусной РНК, выделенной от 56 пациентов с инфекцией COVID-19, находившихся на лечении в ГБУЗ «Городская больница № 40» Санкт-Петербурга в период с 18.04.2020 по 31.01.2021 г. Пациентам проводили объективный осмотр с оценкой параметров гемодинамики, дыхательной системы, оценку по шкале NEWS (National Early Warning Score), компьютерную томографию (КТ) органов грудной клетки, лабораторные исследования: клинический анализ крови, определение уровней ферритина, С-реактивного белка (СРБ), интерлейкина-6 (ИЛ-6), лактатдегидрогеназы (ЛДГ), Д-димера, креатинина, глюкозы. У всех пациентов был положительный тест на РНК SARS-CoV-2, выполненный методом полимеразной цепной реакции (ПЦР). SNP (Single Nucleotide Polymorphism) в вирусной РНК идентифицировали через создание библиотек кДНК таргетным секвенированием (MiSeq Illumina).

Биоинформационный анализ генома вируса был проведен при помощи вычислительной цепочки viralrecon v2, с последующей аннотацией вариантов утилитам Pangolin и Nextstrain. Визуализация собранных геномов была проведена с помощью программы Integrative Genomics Viewer (IGV). Статистическую обработку данных (описательная статистика, графический анализ взаимосвязей данных из разных таблиц) выполняли с помощью прибора GraphPad на платформе Prism 8.01.

Результаты. Проведен сравнительный анализ частот SNP в геноме вируса в образцах у погибших и выписанных пациентов. Идентифицированы SNP, ассоциированные с риском летального исхода ($OR > 1$), нейтральные SNP ($OR = 1$) и протективные SNP ($OR < 1$). Пациенты были инфицированы 14 линиями SARS-CoV-2, пять из которых (B.1.1.129, B.1.1.407, B.1.1.373, B.1.1.397 и B.1.1.152) имели российское происхождение. Установлены положительные корреляционные зависимости между средним количеством SNP, несинонимичными SNP, SNP в S-белке со степенью дыхательной недостаточности, суммой баллов по шкале NEWS, формой заболевания по КТ, длительностью лечения на искусственной вентиляции легких (ИВЛ), исходом заболевания, уровнями ЛДГ, глюкозы, Д-димера, ферритина, количеством РНК в ПЦР тесте. SNP в S-белке отрицательно коррелируют с уровнем лейкоцитов и нейтрофилов.

Глотов Олег Сергеевич, e-mail: olglov@mail.ru.

Выводы. Результаты позволяют предположить влияние генетических факторов SARS-CoV-2 на степень тяжести течения и риск летального исхода заболевания.

Ключевые слова:	SARS-CoV-2, инфекция COVID-19, гены, SNP, степень тяжести, клинические симптомы, биохимические маркеры, корреляции.
Конфликт интересов:	авторы заявляют об отсутствии конфликта интересов.
Прозрачность финансовой деятельности:	финансирование исследования было проведено из средств бюджета СПб ГБУЗ «Городская больница № 40 Курортного административного района» и грантов из средств СПбГУ – ID PURE: 75253103 и ID PURE 73023672.
Для цитирования:	Глотов О.С., Чернов А.Н., Коробейников А.И., Калинин Р.С., Цай В.В., Анисенкова А.Ю., Уразов С.П., Липидус А.Л., Мосенко С.В., Щербак С.Г. Линии коронавируса SARS-CoV-2 российского происхождения: генетическая характеристика и корреляции с клиническими параметрами, тяжестью коронавирусной инфекции. <i>Сибирский журнал клинической и экспериментальной медицины</i> . 2021;36(4):132–143. https://doi.org/10.29001/2073-8552-2021-36-4-132-143 .

The lineage of coronavirus SARS-CoV-2 of Russian origin: Genetic characteristics and correlations with clinical parameters and severity of coronavirus infection

Oleg S. Glotov^{1, 2, 4}, Alexander N. Chernov⁵, Anton I. Korobeynikov³,
Roman S. Kalinin⁴, Viktoria V. Tsai⁴, Anna Yu. Anisenkova^{1, 3},
Stanislav P. Urazov¹, Alla L. Lapidus³, Sergei V. Mosenko¹,
Sergei G. Shcherbak^{1, 3}

¹ City Hospital No. 40 of Kurortny Administrative District,
9, Borisov str., Sestroretsk, Saint Petersburg, 197706, Russian Federation

² D.O. Ott Research Institute for Obstetrics, Gynecology, and Reproduction,
3, Mendeleevskaya liniya, Saint Petersburg, 199034, Russian Federation

³ Saint Petersburg State University,
7, Universitetskaya emb., Saint-Petersburg, 199034, Russian Federation

⁴ Children's Scientific and Clinical Center for Infectious Diseases of the Federal Medical and Biological Agency,
9, Professor Popov str., Saint Petersburg, 197022, Russian Federation

⁵ Institute of Experimental Medicine,
12, Pavlov str., Saint Petersburg, 197376, Russian Federation

Abstract

The identification of new SARS-CoV-2 and human protein and gene targets, which may be markers of the severity and outcome of the disease, are extremely important during the COVID-19 pandemic. The goal of this study was to carry out genetic analysis of SARS-CoV-2 RNA samples to elucidate correlations of genetic parameters (SNPs) with clinical data and severity of COVID-19 infection.

Material and Methods. The study included viral RNA samples isolated from 56 patients with COVID-19 infection who received treatment at the City Hospital No. 40 of St. Petersburg from 04/18/2020 to 04/18/2021. Patients underwent physical examination with the assessments of hemodynamic and respiratory parameters, clinical risk according to National Early Warning Score (NEWS), computed tomography (CT) of the chest, and laboratory studies including clinical blood analysis, assessment of ferritin, C-reactive protein (CRP), interleukin-6 (IL-6), lactate dehydrogenase (LDH), D-dimer, creatinine, and glucose levels. All patients tested positive for SARS-CoV-2 RNA by polymerase chain reaction (PCR). Single nucleotide polymorphisms (SNPs) in viral RNA were identified through the creation of cDNA libraries by targeted sequencing (MiSeq Illumina). Bioinformatic analysis of viral samples was performed using the viralrecon v2 pipeline with the further annotation via Pangolin and Nextlade. Sampled genomes were visualized using the Integrative Genomics Viewer (IGV) software. Statistical

data processing (descriptive statistics and graphical analysis of data relationships from different tables) was performed using a GraphPad device on the Prism 8.01 platform.

Results. A comparative analysis of SNP frequencies in the virus genome in samples from deceased and discharged patients was carried out. The SNPs associated with risk of death ($OR > 1$), neutral SNPs ($OR = 1$), and protective SNPs ($OR < 1$) were identified. Patient samples were infected with 14 lines of SARS-CoV-2, five of which (B.1.1.129, B.1.1.407, B.1.1.373, B.1.1.397, and B.1.1.152) were of Russian origin. The SNPs in the samples infected with the strains of non-Russian origin were associated with an increased risk of mortality ($OR = 2.267$, 95% confidence interval 0.1594–8.653) compared to the SNPs in the samples obtained from the group of patients infected with the strains of Russian origin. Positive correlations were identified between the average SNP number, nonsynonymous SNPs, and S-protein SNPs with the degree of respiratory failure, total NEWS score, CT-based form of disease, duration of treatment with mechanical ventilation, disease outcome, levels of LDH, glucose, D-dimer, and ferritin, and RNA amount in the PCR test. S-protein SNPs negatively correlated with the leukocyte and neutrophil counts.

Keywords:	SARS-CoV-2, COVID-19 infection, genes, SNP, severity, clinical symptoms, biochemical markers, correlations.
Conflict of interest:	the authors do not declare a conflict of interest.
Financial disclosure:	the study was financed from the budget of the St. Petersburg State Budgetary Healthcare Institution "City Hospital No. 40 of the Kurortny Administrative District" and by St. Petersburg State University (grants ID PURE: 75253103 and ID PURE: 73023672).
For citation:	Glotov O.S., Chernov A.N., Korobeynikov A.I., Kalinin R.S., Tsai V.V., Anisenkova A.Yu., Urazov S.P., Lapidus A.L., Mosenko S.V., Shcherbak S.G. The lineage of coronavirus SARS-CoV-2 of Russian origin: Genetic characteristics and correlations with clinical parameters and severity of coronavirus infection. <i>The Siberian Journal of Clinical and Experimental Medicine</i> . 2021;36(4):132–143. https://doi.org/10.29001/2073-8552-2021-36-4-132-143 .

Введение

Пандемия коронавирусной инфекции COVID-19, начавшись в конце 2019 г. в китайском Ухане, в 2020–2021 гг. распространилась на 222 страны мира, включая и Россию, и вызвала более 190 млн случаев инфицирования и более 4 млн летальных исходов во всем мире [1]. Высокая скорость распространения и степень тяжести заболевания COVID-19 отчасти обусловлены высокой скоростью мутационного процесса в геноме вируса. Установлено, что геном SARS-CoV-2 состоит из одноцепочечной РНК, содержащей регуляторные последовательности и 10 открытых рамок считывания (ORF), которые транскрибируются с образованием мРНК функциональных белков, из которых мутации в РНК-полимеразе (ORF8), РНК-праймазе (ORF6), S-белке коррелируют с вирулентностью вируса и степенью тяжести коронавирусной инфекции [2].

Многочисленные исследования по секвенированию образцов РНК SARS-CoV-2 установили различные генетические подтипы коронавируса. На основании анализа SNP в образцах из эндемического региона Чанчуань Инь определены четыре генотипа SARS-CoV-2: генотип I (11083G>T), генотип II (26144G>T), генотип III (8782C>T, 28144T>C), генотип IV (241C>T, 3037C>T, 14408C>T, 23403A>G) [3]. Позже, в октябре 2020 г., в образцах от пациентов из штата Рио-де-Жанейро (Бразилия) были обнаружены мутации в S-белке, одна из которых (E484K) оказалась наиболее клинически значимой, вследствие чего данный вариант и получил свое название E484K.V2 [4].

В декабре 2020 г. были зарегистрированы варианты B.1.1.7 и B.1.351 (501Y.V2) коронавируса из Великобритании и Южной Африки, содержащие, соответственно, нуклеотидные замены 69-70del, Y144del и N501Y, K417N,

E484K в S-белке [5]. При анализе последовательностей 405871 образцов из базы GISAID в конце 2020 г. были выявлены варианты B.1.427 и B.1.429 коронавируса из Южной и Северной Калифорнии [6]. По состоянию на 22 января 2021 г. CAL.20C был обнаружен в 26 штатах и других странах [6]. В январе 2021 г. были обнаружены P.1 и P.2 (подлиния B.1.128) варианты, включающие 10 уникальных мутаций (E484K и N501K и др.) в S-белке, в 42% протестированных образцов из Манауса, Бразильского штата Амазонас [7].

В декабре 2020 г. появились данные о распространении линий B.1.1.31 и B.1.1.317 SARS-CoV-2, имеющих российское происхождение, в Великобритании, США, Японии, Сингапуре, Турции, Таиланде, Швейцарии и Бразилии [8]. Однако до настоящего времени не проведен обстоятельный генетический анализ линий B.1.1.31 и B.1.1.317, не установлены корреляции с клиническими симптомами и степенью тяжести коронавирусного заболевания.

Цель исследования: провести генетический анализ образцов РНК SARS-CoV-2 и установить корреляции генетических показателей и SNP с клиническими данными и степенью тяжести инфекции COVID-19.

Материал и методы

Общая и клиническая характеристика пациентов

В исследование включены образцы вирусной РНК, выделенные от 56 пациентов с инфекцией COVID-19, находившихся на лечении в ГБУЗ «Городская больница № 40» Санкт-Петербурга в период с 18.04.2020 по 31.01.2021 г., имевших положительный результат теста на наличие РНК SARS-CoV-2 методом амплификации нуклеиновых кислот в полимеразной цепной реакции (ПЦР). Исследо-

вание было одобрено экспертным советом по этике СПб ГБУЗ «Городская больница № 40» (протокол № 171 от 18 мая 2020 г.).

У 25 пациентов (45,5%) заболевание закончилось летальным исходом (табл. 1).

Таблица 1. Общая характеристика пациентов с COVID-19

Table 1. General characteristics of patients

Параметры Parameters	Исход пациентов Patient Outcome	
	Выписанные Discharged	Умершие Died
Число пациентов Number of patients	31	25
Мужчины Men	17	11
Женщины Women	14	14
Возраст, лет Age, year	64,6 ± 16,6	69,2 ± 14,8
Легкая степень тяжести заболевания Mild disease severity	4	–
Среднетяжелая степень тяжести заболевания Moderate to severe disease severity	14	–
Тяжелая степень тяжести заболевания Severe degree of disease severity	13	25
PHK, нг RNA, ng	26,0 ± 4,9	25,4 ± 5,0

Для всех пациентов собирались данные: эпидемиологического анамнеза, наличие клинических симптомов (кашель, одышка, повышение температуры, лихорадка, слабость, потеря обоняния и вкуса, тяжесть в груди), степень дыхательной недостаточности. Учитывалось наличие сопутствующих заболеваний (ишемическая болезнь сердца (ИБС), острый коронарный синдром (ОКС), цереброваскулярная болезнь (ЦВБ), состояния после инсульта или инфаркта (ПИКС), сахарный диабет, ревматоидный артрит, хроническая болезнь почек, хроническая обструктивная болезнь легких (ХОБЛ), бронхиальная астма, хронический бронхит, гидроторакс. Выяснялись степень тяжести заболевания: легкая, средняя, тяжелая и крайне тяжелая, информация о дне заболевания при обращении в больницу.

Также проводился объективный осмотр с оценкой параметров гемодинамики, дыхательной системы (частота дыхательных движений (ЧДД), частота сердечных сокращений (ЧСС), артериальное давление (АД), SpO₂, степень дыхательной недостаточности). Осуществлялась оценка по шкале NEWS (National Early Warning Score), рекомендованной к использованию для пациентов с COVID-19 [9]. Проводилась компьютерная томография (КТ) органов грудной клетки с оценкой формы заболевания по 4-значной шкале без внутривенного контрастного усиления (КТ-1, КТ-2, КТ-3, КТ-4). Выполнялись лабораторные исследования (клинический анализ крови, биохимический минимум, определение уровней ферретина, С-реактивного белка (СРБ), IL-6, лактатдегидрогеназы (ЛДГ), Д-димера, креатинина, глюкозы). Проводилась электрокардиография, при необходимости использовались дополнительные инструментальные методики.

Терапия пациентов с коронавирусной инфекцией

Лечение инфекции COVID-19 и ее осложнений включало противовирусные препараты, профилактику гиперкоагуляции и синдрома диссеминированного внутрисосудистого свертывания (ДВС-синдрома). У пациентов с прогрессирующим течением заболевания для предупреждения или лечения цитокинового шторма (ЦШ) стандартная терапия была дополнена назначением патогенредуцированной плазмы реконвалесценто, антицитокиновых препаратов: ингибиторов рецептора интерлейкина-6 (IL-6) (тоцилизумаба, олокизумаба, левалимаба), IL-1 (канакинумаба, RH104), JAK-киназ (тофацитиниба, руксолитиниба, барицитиниба), тирозинкиназы Bcr-Abl (радотиниба), в ряде случаев – глюкокортикостероидов. По показаниям проводили этапную респираторную терапию, антибактериальную терапию, лечение сепсиса и септического шока, экстракорпоральную детоксикацию и гемокоррекцию, экстракорпоральную мембранную оксигенацию.

Выделение РНК

Выделение РНК из мазков носоглотки проводили согласно инструкции к набору реагентов для экстракции GeneJET RNA Purification Kit (Thermo Scientific) [10]. Концентрацию РНК измеряли на флуориметре Quantus с использованием набора QuantiFluor RNA System (Promega). Оценку качества РНК выполняли при помощи прибора 4200 TapeStation System и набора High Sensitivity RNA ScreenTape Analysis. Отбор положительных проб (с Cq < 25) для дальнейшего исследования проводили с помощью Real-Time PCR с одним из наборов для диагностики вируса SARS-CoV-2.

RT-PCR анализ

RT-PCR анализ проводили с помощью наборов реагентов для выявления РНК коронавируса SARS-CoV-2 в клиническом материале (производство НИИ эпидемиологии и микробиологии имени Пастера) и GeneFinderTM COVID-19 plus RealAmp (OSANG Healthcare Co., Ltd., Korea) на приборах CFX96 Real-Time PCR Detection System (Biorad, USA).

Секвенирование

Для исследования вируса методом таргетного секвенирования выделяли кДНК с использованием следующих наборов: Mint-2 cDNA synthesis kit (Евроген, Россия), Maxima H Minus Double-Stranded cDNA Synthesis kit for RT-qPCR (ThermoFisher, USA). Оценка качества полученной кДНК проводилась с помощью прибора 4200 TapeStation System и набора D1000 ScreenTape и High Sensitivity D1000 ScreenTape, концентрацию измеряли на флуориметре Quantus с использованием набора QuantiFluor dsDNA System (Promega, USA). После получения подходящих по качеству образцов кДНК проводили подготовку библиотек для секвенирования по протоколу с использованием набора Illumina DNA Prep with Enrichment Kit для MiSeq Illumina [11]. Исследование полного транскриптома выполняли с помощью набора для подготовки библиотек MGIEasy RNA Library Prep Set User Manual на приборе MGISEQ-2000.

Биоинформатический анализ

Биоинформатический анализ включал в себя сборку с использованием существующего референсного генома вируса при помощи вычислительной цепочки viralrecon

v2 [12] с последующей аннотацией вариантов утилитам Pangolin и Nextstrain. Также обработку данных после секвенирования на приборе MiSeq выполняли при помощи сервиса basespace.illumina с использованием приложения «DRAGEN RNA Pathogen Detection». Визуализация генома производилась в программе The Integrative Genomics Viewer (IGV).

Статистический анализ

Статистическую обработку данных (описательная статистика, графический анализ взаимосвязей данных из разных таблиц) выполняли с помощью прибора GraphPad на платформе Prism 8.01. Взаимосвязи категориальных показателей (пола, степени и формы патологических процессов, жалоб) выявлялись с помощью χ^2 -критерия Пирсона и точного критерия Фишера. Пороговые уровни

для возраста, индекса NEWS и лабораторных данных определяли с помощью метода иерархической кластеризации. Отношение шансов OR (odds ratio, OR) и границы 95% доверительного интервала (CI) выполняли с помощью калькулятора на портале www.medstatistic.ru [13].

Результаты

Нами было выявлено всего 389 вариантов в РНК SARS-CoV-2, из них 263 (67,6%) были связаны с аминокислотной заменой в белках, 126 синонимичных мутаций (32,4%), причем 139 (35,7%) в образцах пациентов с летальным исходом. Вначале изучали корреляцию общего количества, количества несинонимичных, молчащих мутаций, и мутаций в S-белке SARS-CoV-2 со степенью тяжести коронавирусного заболевания (табл. 2).

Таблица 2. Количество мутаций в SARS-CoV-2 в зависимости от степени тяжести заболевания

Table 2. The number of mutations in SARS-CoV-2 depending on the severity of the disease

Степень тяжести COVID-19 Severity of COVID-19	Среднее количество мутаций Average number of mutations	Среднее количество несинонимичных мутаций Average number of nonsynonymous mutations	Среднее количество молчащих мутаций Average number of silent mutations	Среднее количество мутаций в S-белке Average number of mutations in S-protein
I (n = 4)	11,0 ± 0,2	7,0 ± 1,4	4,0 ± 1,4	1,0
II (n = 14)	11,4 ± 1,9	8,3 ± 1,9	3,3 ± 1,6	2,0 ± 1,0
III (n = 13)	13,0 ± 3,6	9,0 ± 2,4	4,0 ± 2,1	2,2 ± 1,0***
Умершие, III Died, 3 (n = 25)	17,1 ± 5,0	10,6 ± 3,0	6,6 ± 2,2	1,6 ± 0,9

Примечание: *** – статистическая значимость при $p < 0,001$. Легкая степень тяжести обозначена I, среднетяжелая степень – II, тяжелая степень – III, все умершие пациенты имели тяжелую III степень тяжести инфекции COVID-19.

Note: *** statistical significance at $p < 0,001$. Mild, moderate, and high severity degrees are designated by I, II, and III, respectively. All patients who died had degree III severity of COVID-19 infection.

Из данных таблицы 2 следует, что только количество мутаций в S-белке в образцах пациентов с III степенью тяжести было достоверно выше ($p < 0,001$), чем число мутаций в образцах пациентов с I степенью тяжести заболевания. Также отмечалась тенденция к росту общего числа и несинонимичных мутаций у умерших лиц по сравнению

с пациентами с I степенью тяжести заболевания. Далее для обнаруженных мутаций оценивали их потенциальную опасность, т. е. ассоциацию со степенью тяжести заболевания, рассчитывая частоты и OR в группах выписанных и умерших пациентов. Частоты и значения OR удалось оценить только для 38 (9,7%) мутаций (табл. 3).

Таблица 3. Сравнение частот мутаций в генах вируса SARS-CoV-2 со степенью тяжести заболевания

Table 3. Association of the frequencies of mutations in the genes of SARS-CoV-2 virus with the severity of the disease

Ген Gene	Нуклеотидная позиция Nucleotide position	Аминокислотная замена Amino acid substitution	Белок Protein	Выписанные Discharged Частота Frequency	Умершие Died Частота Frequency	Отношение шансов OR	Доверительный интервал 95% CI
5'UTR	21C>T	–	–	–	0,095	3,522	0,342–36,216
5'UTR	241C>T	–	–	0,928	0,954	1,615	0,137–19,068
ORF1a	346C>T	–	Nsp1	–	0,095	3,522	0,342–36,216
ORF1a	872G>A	Asp203Asn	Nsp2	0,071	–	0,439	0,043–4,516
ORF1a	1648C>T	–		0,036	0,095	2,947	0,249–34,85
ORF1a	1681G>T	Glu472Asp		0,071	0,046	0,619	0,052–7,307
ORF1a	3037C>T	Phe924Phe	Папаин-подобная протеаза (PL2pro), Nsp3 Papain-like protease (PL2pro), Nsp3	0,758	0,864	1,727	0,379–7,864
ORF1a	3058A>G	–		0,036	0,095	2,700	0,229–31,894
ORF1a	3737C>T	Pro1158Ser		0,107	0,095	0,912	0,139–6,005
ORF1a	4965C>T	Thr1567Ile		0,107	0,046	0,433	0,042–4,485
ORF1a	5040>G	–		0,143	0,136	1,042	0,207–5,237
ORF1a	7296C>T	Ala2344Val		0,071	–	0,439	0,043–4,516

Окончание табл. 3
End of table 3

Ген Gene	Нуклеотидная позиция Nucleotide position	Аминокислотная замена Amino acid substitution	Белок Protein	Выписанные Discharged Частота Frequency	Умершие Died Частота Frequency	Отношение шансов OR	Доверительный интервал 95% CI
ORF1a	8944T>C	Gly2893Gly	Nsp4	0,036	0,095	2,700	0,229–31,894
ORF1a	9967C>T	Leu3234Leu		0,036	0,095	2,700	0,229–31,894
ORF1a	10046G>T	Val 3261Phe		–	0,095	3,522	0,342–36,216
ORF1a	11916C>T	Ser3884Leu	Nsp7	0,143	0,046	0,313	0,032–3,021
ORF1a	13554C>T	–	РНК-полимераза, Nsp12 RNA polymerase, Nsp12	0,036	0,095	2,700	0,229–31,894
ORF1b	14408C>T	Pro323Leu		0,892	1,0	1,167	0,980–1,389
ORF1b	15738 C>T	–		0,071	0,046	0,619	0,052–7,307
ORF1b	16775A>G	Tyr103Cys	Геликаза, Nsp13 Helicase, Nsp13	0,071	0,227	3,824	0,664–22,005
ORF1b	17082A>G	–	–	0,036	0,095	2,700	0,229–31,894
ORF1b	18479A>G	Lys1671Arg	Эндонуклеаза Endonuclease	0,071	0,046	0,619	0,052–7,307
ORF1b	19839T>C	–	ЭндоРНКазы, Nsp15 EndoRNase, Nsp15	0,107	0,095	0,600	0,099–3,623
ORF1b	20028T>C	–	–	0,143	0,095	0,600	0,099–3,623
S	22020T>C	Met153Thr	S-белок S-protein	0,107	0,046	0,433	0,042–4,485
S	23403A>G	Asp614Gly		0,928	0,954	1,615	0,137–19,068
S	23586A>G	Gln675Arg		0,107	0,046	0,433	0,042–4,485
S	23604C>A	Pro681His		0,036	0,095	2,700	0,229–31,894
S	24095G>T	Ala845Ser		0,071	–	0,439	0,043–4,516
S	24383A>G	Thr941Ala		0,036	0,095	2,700	0,229–31,894
S	25046C>T	Phe1162Ser		–	0,095	3,522	0,342–36,216
ORF3a	25910A>G	Asp173Gly	Orf3a	–	0,095	3,522	0,342–36,216
N	28881G>A	Arg203Lys	N-белок нуклеокапсида N-nucleocapsid protein	1,0	1,0	1,000	–
N	28882G>A	Arg203Arg	–	1,0	1,0	1,000	–
N	28883G>A	Gly204Arg	–	1,0	1,0	1,000	–
N	28905C>T	Ala211Val	–	0,250	18,1	1,400	0,405–4,836
N	28975G>T	Met 234 Ile	–	0,071	–	0,439	0,043–4,516
N	29422G>A	–	–	–	0,095	3,522	0,342–36,216
3'UTR	29708C>T	–	–	–	0,095	3,522	0,342–36,216

Как следует из данных из таблицы 3, мутации можно сгруппировать в 3 кластера. В первую группу включены мутации, потенциально ассоциированные с тяжестью коронавирусной инфекции и летальным исходом ($OR > 1$). Они локализованы в генах, кодирующих лидерную 5'-концевую сар-область и белки Nsp1, Nsp2, Nsp3, Nsp4, РНК-зависимую РНК-полимеразу (Nsp12), геликазу (Nsp13), S-белок, Orf3a и в 3'-концевую poly-A-последовательность. Ко второй группе относились нейтральные

мутации ($OR = 1$), локализованные в генах, кодирующих Nsp3 и N-белки. Третий кластер образовали потенциально протекторные мутации ($OR < 1$), локализованные в генах, кодирующих белки: Nsp2, Nsp3, Nsp7, эндонуклеазу (Nsp14), эндоРНКазу (Nsp15), S- и N-белки.

В результате анализа мутаций с помощью утилит Nextclade и Pangolin, а также программы IGV мы установили типы линий вируса, инфицирующие каждый образец (табл. 4).

Таблица 4. Линии SARS-CoV-2, инфицирующие образцы пациентов из Санкт-Петербурга

Table 4. SARS-CoV-2 lines, infecting patient samples from St. Petersburg

Линия вируса Virus lineage	Часто встречаемые страны в родословной Common countries in pedigree	Количество образцов Number of samples	Исход (выписан) Outcome Discharged	Исход (умер) Outcome Died
Линии нероссийского генеза Lines of non-Russian genesis				
B.1	UK 33,0%, USA 13,0%, Japan 9,0%	3	1	2
B.1.1	UK 33,0%, USA 13,0%, Japan 9,0%	19	8	11
B.1.1.29	USA 100,0%	2	1	1
B.1.1.174	USA 74,0%, Ireland 4,0%, Turkey 3,0%, UK 3,0%, India 2,0%	1	0	1

Окончание табл. 4
End of table 4

Линия вируса Virus lineage	Часто встречаемые страны в родословной Common countries in pedigree	Количество образцов Number of samples	Исход (выписан) Outcome Discharged	Исход (умер) Outcome Died
B.1.1.57	South Africa 84,0%, India 5,0%, Russia 4,0%, Zimbabwe 3,0%, Japan 1,0%	1	1	0
B.1.610	UK 50,0%, Spain 20,0%, Turkey 10,0%, USA 5,0%, Japan 3,0%	1	0	1
B.1.1.155	United Kingdom 100,0%	2	1	1
Общее количество образцов Total number of samples		29	12	17
Отношение шансов (Odds ratio OR)		2,267	95% CI	0,594–8,653
Линии российского генеза Lines of Russian genesis				
B.1.1.129	Russia 90,0%, Germany 8,0%, Mexico 2,0%	2	1	1
B.1.1.407	Russia 81,0%, Germany 5,0%, Indonesia 5,0%, USA 5,0%, Finland 5,0%	4	2	2
B.1.1.373	Russia 81,0%, Estonia 5,0%, UK 5,0%, Hong Kong 2,0%, Indonesia 2,0%	5	4	1
B.1.1.397	Russia 46,0%, Estonia 14,0%, Latvia 10,0%, Israel 10,0%, Finland 4,0%	2	2	0
B.1.1.152	Russia 82,0%, Germany 14,0%, Finland 5,0%	1		1
Общее количество образцов Total number of samples		14	8	5
Отношение шансов (Odds ratio OR)		0,230	95%CI	0,05–0,889
Линии смешанного генеза Lines of mixed genesis				
B.1.1.317	Germany 46,0%, Russia 25,0%, Estonia 5,0%, Finland 4,0%, USA 3,0%	4	3	1
B.1.1.141	Latvia 53,0%, Russia 23,0%, Germany 11,0%, Sweden 3,0%, Georgia 2,0%	1	1	0
Общее количество образцов Total number of samples		5	4	1

Примечание: для 9 образцов происхождение вируса невозможно было определить.

Note: The origin of the virus could not be determined in nine samples.

Из данных таблицы 4 следует, что пациенты из Санкт-Петербурга были инфицированы 14 линиями SARS-CoV-2. Причем 14 образцов (29,8%) из 47, представленных в таблице, были инфицированы 5 линиями вируса российского происхождения. В целом OR для 5 линий вируса, имеющих российское происхождение, составил 0,441 (95% CI = 0,116–1,684). Это свидетельствует о том, что мутации в линиях SARS-CoV-2 нероссийско-

го генеза ассоциированы с повышенным риском (OR = 2,267, 95% CI = 0,1594–8,653) летального исхода в группе российских пациентов.

Также мы оценили корреляции общего количества, несинонимичных и мутаций в S-белке линий SARS-CoV-2 с клиническими данными, сопутствующими заболеваниями, биохимическими маркерами и исходом заболевания (табл. 5).

Таблица 5. Корреляции общего количества, несинонимичных и мутаций в S-белке линий SARS-CoV-2 с клиническими симптомами, сопутствующими заболеваниями, биохимическими маркерами

Table 5. Correlations of total number, nonsynonymous mutations, and mutations in the S-protein of SARS-CoV-2 lines with clinical symptoms, concomitant diseases, and biochemical markers

Линия вируса Virus line	Коэффициент корреляции r , уровень значимости p Correlation coefficient, r , significance level p			Параметр Parameter
	среднее количество SNP Average number of SNPs	Среднее количество несинонимичных SNP Average number of nonsynony- mous SNPs	Среднее количество SNP в S-белке Average number of SNPs in S-protein	
	0,62 ($p = 0,03$)	0,65 ($p = 0,02$)	–	Слабость Weakness
	0,71 ($p = 0,009$)	–	–	Степень дыхательной недостаточности при выписке Grade of respiratory failure at dis- charge

Окончание табл. 5
End of table 5

Линия вируса Virus line	Коэффициент корреляции r , уровень значимости p Correlation coefficient, r , significance level p			Параметр Parameter
	среднее количество SNP Average number of SNPs	Среднее количество несинонимичных SNP Average number of nonsynonymous SNPs	Среднее количество SNP в S-белке Average number of SNPs in S-protein	
Российского генеза Lines of Russian genesis	0,67 ($p = 0,01$)	–	–	NEWS при выписке (сумма баллов) NEWS score at discharge
	0,75 ($p = 0,008$)	0,66 ($p = 0,02$)	–	Форма забол. по КТ 1–4 к началу терапии Disease form according to CT 1–4 at the beginning of therapy
	0,82 ($p = 0,006$)	0,76 ($p = 0,02$)	–	Форма заболевания по КТ 1–4 при выписке Disease form according to CT 1–4 at discharge
	0,73 ($p = 0,009$)	–	–	Количество суток на ИВЛ Number of days on mechanical ventilation
	0,62 ($p = 0,03$)	–	–	Исход заболевания Outcome of disease
	–	–0,52 ($p = 0,05$)	–0,57 ($p = 0,05$)	Лейкоциты Leukocytes
	–	–0,52 ($p = 0,05$)	–0,54 ($p = 0,05$)	Нейтрофилы Neutrophils
	0,83 ($p = 0,009$)	–	–	ЛДГ LDH
	–	0,68 ($p = 0,02$)	0,81 ($p = 0,002$)	Глюкоза Glucose
Нероссийского генеза Lines of non-Russian genesis	–	–0,44 ($p = 0,03$)	–	Одышка Dyspnea
	0,49 ($p = 0,02$)	0,50 ($p = 0,02$)	0,46 ($p = 0,02$)	Слабость Weakness
	0,43 ($p = 0,04$)	–	–	Применение системных кортикостероидов Systemic corticosteroid use
	–	–	0,41 ($p = 0,05$)	Количество РНК RNA amount
	0,42 ($p = 0,04$)	–	–	Койко-дни Length of hospital stay (days in hospital)
	0,49 ($p = 0,05$)	0,50 ($p = 0,05$)	–	Д-димер D-dimer
	–	–	0,85 ($p = 0,03$)	Ферритин Ferritin

Данные таблицы 5 свидетельствуют, что среднее количество SNP линий SARS-CoV-2 российского происхождения коррелирует со слабостью, степенью дыхательной недостаточности при выписке, суммой баллов NEWS при выписке, формой заболевания по КТ к началу терапии и при выписке, количеством суток на искусственной вентиляции легких (ИВЛ), исходом заболевания и уровнем ЛДГ. Среднее количество несинонимичных SNP и SNP в S-белке коррелируют с уровнем глюкозы и отрицательно коррелируют с уровнем лейкоцитов и нейтрофилов. Кроме того, среднее количество несинонимичных SNP также коррелирует со слабостью и формой заболевания по КТ к началу терапии и при выписке.

Среднее количество SNP, несинонимичных и SNP в S-белке линий SARS-CoV-2 нероссийского происхождения положительно коррелирует со слабостью ($p = 0,02$).

Среднее количество SNP и несинонимичных SNP положительно коррелирует с уровнем Д-димера ($p = 0,05$). Среднее количество SNP также коррелирует с применением кортикостероидов и количеством койко-дней. Среднее количество несинонимичных SNP отрицательно коррелирует с одышкой. Число SNP в S-белке коррелирует с количеством РНК и уровнем ферритина.

Обсуждение

В данном исследовании мы изучали корреляцию общего количества, несинонимичных мутаций, молчащих-мутаций и мутаций в S-белке SARS-CoV-2 со степенью тяжести вирусной инфекции и установили, что количество мутаций в S-белке было выше ($p < 0,001$) в образцах пациентов с III степенью тяжести, чем в образцах от пациентов со II степенью тяжести коронави

русной инфекции. Коллективом авторов из университета Кейо (Токио, Япония) при проведении секвенирования геномной РНК SARS-CoV-2 из образцов от 90 пациентов с инфекцией COVID-19 и при изучении ассоциации вирусных гаплотипов с тяжестью заболевания путем подсчета количества мутаций, обнаружено, что количество несинонимичных мутаций (особенно Pro108Ser в 3 химо-трипсиноподобной протеазе (3CLpro) и Pro151Leu в N-белке нуклеокапсида) обратно коррелирует ($OR = 0,24$; 95% CI = 0,07–0,88; $p = 0,032$) со степенью тяжести инфекции COVID-19 и потребностью в кислородной терапии [14].

Нами также оценены частоты и OR SNP для степени тяжести инфекции COVID-19. При этом OR мутаций рассчитывали, когда мутация присутствовала в 2 и более образцах (частота > 5%) в обеих группах пациентов (выписанные и умершие) (см. табл. 3). Интересно, что 75,8% образцов наших пациентов были инфицированы IV генотипом (241C>T, Nsp3: 3037C>T; Nsp12, РНК-полимераза: 14408C>T, Pro323Leu; S-белок: 23403A>G, Asp614Gly) SARS-CoV-2, установленным Чанчуань Инем [6]. Гаплотип 3037C>T, 14408C>T, 23403A>G и SNP 241C>T наиболее часто встречаются в образцах из Европы, отражая европейское происхождение вируса. Например, в исследовании, проведенном международным коллективом авторов из Вьетнама, Австралии, Великобритании и США, данный мутантный гаплотип обнаружен в 90% (40 из 44) образцов пациентов из Великобритании ($n = 15$), России ($n = 6$), Германии ($n = 5$), Франции ($n = 4$), Италии ($n = 2$), Испании ($n = 2$), Нидерландов ($n = 1$), Вьетнама ($n = 6$) и стран Азии ($n = 3$). Также у данных пациентов вместе с этой группой мутаций в 75% (33 из 44) геномов наблюдались SNP в N-белке 28881G>A, 28882G>A, 28883G>C [16].

В наших образцах SNP 28881G>A, 28882G>A, 28883G>C присутствовали в 100% образцов, что указывает на их нейтральный характер и на происхождение вируса, относящегося в кладе 20В линии В1.1 [14]. Данные мутации локализованы в небольшом участке (194–204) аминокислотной последовательности N-белка. Поскольку N-белок участвует в упаковке вирусной РНК в спиральный рибонуклеокапсид, мутации 28881G>A, 28882G>A, 28883G>C могут повышать эффективность транскрипции субгеномных РНК, способствуя выживанию клеток, репликации и персистенции вируса [15].

В 47 из 50 (94%) наших образцов также наблюдался SNP 241C>T, локализованный в 5'-нетранслируемой области (5'UTR) вирусной РНК. Эта замена имеет одни из самых высоких частот 0,758 (9673/12754) и 0,809 (36786/45494) соответственно в США и мире, влияет на регуляцию транскрипции и экспрессии генов вируса и его репликацию [16]. SNP 14408C>T (Pro323Leu) и мутация 13554C>T, ассоциированные с тяжелым течением и летальным исходом, также обнаружены в РНК-зависимой РНК-полимеразе (белок RdRp или Nsp12) с частотами 0,940 (47 из 50) и 0,06 (3 из 50) в образцах наших пациентов. Первая из этих мутаций также является наиболее распространенной в образцах из США (AF 0,464, 5918/12754) [16]. О биологической роли этой мутации до сих пор ведется дискуссия. R. Wang и соавт. считают, что поскольку обе аминокислоты пролин (Pro) и лейцин (Leu) являются неполярными и алифатическими, в связи с этим Pro323Leu может не оказывать влияния на функцию белка Nsp12 [16].

В 94% (47/50) образцов пациентов встречается SNP 23403A>G (Asp614Gly) в гене S-белка, связанный с замедленной аспарагиновой кислоты в положении 614 на глицин и повышенным риском тяжелого течения или летального исхода пациентов. Как уже упоминалось, этот SNP часто сочетается с тремя другими SNP (241C>T, 3037C>T и 14408C>T) и наследуется как гаплотип [17]. Исследователи из Гарвардского, Вашингтонского (США) и Шеффилдского университетов (Великобритания), изучив клинические данные и генетические последовательности SARS-CoV-2 у 999 пациентов с COVID-19, установили, что вариант Arg614Gly коррелирует с высоким уровнем вирусной РНК в верхних дыхательных путях и скоростью трансмиссии у пациентов по сравнению с диким типом. Это указывает на более высокую инфицированность Gly614. Однако B. Korber и соавт. не смогли установить ассоциации между D614G и продолжительностью госпитализации, т. е. степенью тяжести заболевания, что свидетельствует о повышенной инфекционности D614G и об отсутствии его влияния на тяжесть клинического исхода болезни [17].

Поскольку нами показано (см. табл. 1), что среднее количество SNP имеет тенденцию к увеличению при повышении степени тяжести инфекции COVID-19, а сумма баллов по шкале NEWS используется для определения степени тяжести состояния пациентов, то между данными показателями может наблюдаться положительная корреляция.

Ученые из госпиталя Королевского колледжа (Лондон, Великобритания) исследовали маркеры крови и физиологические параметры у 1276 пациентов с тяжелым исходом инфекции COVID-19 через 14 сут после госпитализации для улучшения применения шкалы NEWS2. Разработанная авторами модель NEWS2 и возраст для тяжелой инфекции COVID-19 имела площадь под кривой 0,700. Эти данные мы использовали при своей оценке риска [18]. В исследовании норвежских ученых, проведенном на 66 пациентах, 23% из которых имели тяжелую и крайне тяжелую степень заболевания, показано, что оценка NEWS2 ≥ 6 при поступлении характеризует тяжелое течение инфекции COVID-19 с чувствительностью 80,0% и специфичностью 84,3% [19]. В нашей выборке вследствие большого разброса не было статистически значимых различий по среднему индексу NEWS при поступлении у пациентов, инфицированных линиями коронавируса нероссийского генеза, – 4,7 (диапазон 0–12) и у пациентов, инфицированных линиями российского происхождения, – 4,2 (диапазон 0–9). Средний индекс NEWS при поступлении у пациентов, в дальнейшем имевших тяжелое течение заболевания, составил $5,35 \pm 0,39$ баллов, против $2,23 \pm 0,50$ баллов у пациентов легкого и среднетяжелого течения.

Нами выявлены корреляции среднего количества SNP в линиях вируса российского генеза со степенью дыхательной недостаточности при выписке и продолжительностью пребывания пациента на ИВЛ (см. табл. 5). В силу того, что патогенез коронавируса SARS-CoV-2 включает инфицирование бронхолегочного древа и развитие острой респираторной недостаточности, степень дыхательной недостаточности и длительность нахождения пациента на ИВЛ может отражать тяжесть заболевания. Например, американские ученые из больницы Иова Фэйрфакса (Вирджиния, США) изучили клиниче-

ские параметры и исходы 1023 пациентов с инфекцией COVID-19, из которых 164 (16,0%) проводилась инвазивная ИВЛ. У 70 из 164 (42,7%) пациентов терапия закончилась летальным исходом. Умершие пациенты были старше, чем выписанные (средний возраст – 66 против 55 лет, $p < 0,0001$) [20].

В настоящем исследовании нами также установлены корреляционные зависимости между общим количеством, несинонимичных и SNP в S-белке линий SARS-CoV-2 и одышкой, риском летального исхода, уровнем ферритина, Д-димера и глюкозы (см. табл. 5). Американские исследователи из Темплавского университета (Филадельфия, США), оценив биохимические показатели (уровень лейкоцитов, АЛТ, АСТ, билирубин, креатинин, СРБ, ЛДГ, Д-димер, тропонин I) у 513 пациентов с инфекцией COVID-19, из которых 64 были с ЦШ, установили зависимость тяжести заболевания от снижения количества лимфоцитов ($p < 0,001$), увеличения нейтрофилов ($p < 0,001$), СРБ ($p = 0,016$), Д-димера ($p = 0,002$), ЛДГ ($p < 0,001$), тропонина I ($p = 0,045$) и АСТ ($p = 0,028$) [21]. В другой работе на 164 пациентах также установлена зависимость тяжести инфекции с повышенными уровнями Д-димера ($p = 0,005$) и ферритина ($p = 0,016$) [20].

Нами также установлены положительные корреляции между средним количеством SNP в образцах и формой заболевания по КТ к началу терапии и при выписке (см. табл. 5). Поскольку наблюдается увеличение среднего количества SNP по мере тяжести инфекции COVID-19 (см. табл. 1), а форма заболевания по КТ отражает степень поражения легких при данном заболевании, то могут существовать корреляции между количеством SNP и формой заболевания по КТ и степенью тяжести инфекции COVID-19. Например, в исследовании, проведенном турецкими авторами на 130 пациентах с инфекцией COVID-19, выявлена положительная корреляция между степенью поражения легких на КТ и уровнем ферритина в сыворотке ($p = 0,530$; $p = 0,0001$), Д-димера ($p = 0,375$; $p = 0,0001$). Таким образом, в группе КТ-положительных пациентов установлена положительная корреляция уровней ферритина и Д-димера в сыворотке с тяжестью и генерализацией заболевания по данным КТ [22]. Эти данные также согласуются с нашими результатами, свидетельствующими об увеличении уровней Д-димера и ферритина у пациентов с III степенью тяжести инфекции.

Выводы

В нашем исследовании образцы пациентов из Санкт-Петербурга были инфицированы 14 линиями SARS-CoV-2, из которых 5 имели российское происхождение. Выявлено, что мутации в линиях SARS-CoV-2 нероссийского генеза ассоциированы с повышенным риском летального исхода в группе российских пациентов.

Количество SNP в S-белке в образцах пациентов с III степенью тяжести было выше ($p < 0,001$), чем число SNP в образцах пациентов со II степенью тяжести инфекции COVID-19.

Сравнительный анализ частот для каждого SNP в группах образцов пациентов с риском летального исхода позволил сгруппировать SNP в 3 кластера: I) инфекции, ассоциированные с тяжестью COVID-19 и риском летального исхода ($OR > 1$), локализованные в 5'-концевой сар-области и генах, кодирующих Nsp1, Nsp2, Nsp3, Nsp4 белки, РНК-зависимую-РНК-полимеразу (Nsp12), геликазу (Nsp13), S, Orf3a белки и 3'-концевую poly-A-последовательность. II) нейтральные SNP ($OR = 1$), локализованные в генах, кодирующих Nsp3 и N-белки. III) потенциально протекторные SNP ($OR < 1$), локализованные в генах, кодирующих Nsp2, Nsp3, Nsp7 белки, эндонуклеазу (Nsp14), эндоРНКазу (Nsp15), S- и N-белки.

Установлены положительные корреляционные зависимости между средним количеством SNP линий SARS-CoV-2 российского происхождения со слабостью, степенью дыхательной недостаточности при выписке, суммой баллов NEWS при выписке, формой заболевания по КТ к началу терапии и при выписке, количеством суток на ИВЛ, исходом заболевания и уровнем ЛДГ. Среднее количество несинонимичных SNP и SNP в S-белке положительно коррелируют с уровнем глюкозы и отрицательно с уровнем лейкоцитов и нейтрофилов. Среднее количество несинонимичных SNP также положительно коррелирует со слабостью и формой заболевания по КТ к началу терапии и при выписке.

Среднее количество SNP, несинонимичных и SNP в S-белке линий SARS-CoV-2 нероссийского происхождения положительно коррелируют со слабостью ($p = 0,02$). Среднее количество SNP и несинонимичных SNP положительно коррелирует с уровнем Д-димера ($p = 0,05$). Среднее количество SNP коррелирует с применением кортикостероидов и количеством койко-дней. Среднее количество несинонимичных SNP отрицательно коррелирует с одышкой. Число SNP в S-белке коррелирует с количеством РНК и уровнем ферритина.

Литература / References

- Reported Cases and Deaths by Country or Territory. URL: <https://www.worldometers.info/coronavirus/#countries>.
- Nguyen T.T., Pathirana P.N., Nguyen T., Nguyen Q.V.H., Bhatti A., Nguyen D.C. et al. Genomic mutations and changes in protein secondary structure and solvent accessibility of SARS-CoV-2 (COVID-19 virus). *Sci. Rep.* 2021;11(1):3487. DOI: 10.1038/s41598-021-83105-3.
- Yin C. Genotyping coronavirus SARS-CoV-2: Methods and implications. *Genomics*. 2020;112(5):3588–3596. DOI: 10.1016/j.ygeno.2020.04.016.
- Toovey O.T.R., Harvey K.N., Bird P.W., Tang J.W.W. Introduction of Brazilian SARS-CoV-2 484K.V2 related variants into the UK. *J. Infect.* 2021;82(5):e23–e24. DOI: 10.1016/j.jinf.2021.01.025.
- World Health Organization. SARS-CoV-2 Variants. URL: <https://www.who.int/csr/don/31-december-2020-sars-cov-2-variants/en/>
- Zhang W., Davis B.D., Chen S.S., Sincuir Martinez J.M., Plummer J.T., Vail E. Emergence of a novel SARS-CoV-2 variant in southern California. *JAMA*. 2021;325(13):1324–1326. DOI: 10.1001/jama.2021.1612.
- Sabino E.C., Buss L.F., Carvalho M.P.S., Prete C.A. Jr., Crispim M.A.E., Fraiji N.A. et al. Resurgence of COVID-19 in Manaus, Brazil, despite high seroprevalence. *Lancet*. 2021;397(10273):452–455. DOI: 10.1016/S0140-6736(21)00183-5.
- PANGO lineages. URL: https://cov-lineages.org/lineage_designation.html
- Carr E., Bendayan R., Bean D., Stammers M., Wang W., Zhang H. et al. Evaluation and improvement of the National Early Warning Score (NEWS2) for COVID-19: A multi-hospital study. *BMC Med.* 2021;19(1):23. DOI: 10.1186/s12916-020-01893-3.
- Gene JET RNA Purification Kit. URL: <https://www.thermofisher.com/order/catalog/product/K0731>
- Illumina DNA Prep with Enrichment. Reference Guide. URL: https://support.illumina.com/content/dam/illumina-support/documents/documentation/chemistry_documentation/illumina_prep/illumina-dna-prep-with-enrichment-reference-1000000048041-06.pdf
- Patel H., Varona S., Monzón S., Espinosa-Carrasco J., Heuer M.L., Gabernet G. et al. nf-core/viralrecon: nf-core/viralrecon. *Zenodo*. 2021. DOI: 10.5281/zenodo.514625.
- Медицинская статистика. URL: <https://medstatistic.ru/calculators/calccdds.html>
- Abe K., Kabe Y., Uchiyama S., Iwasaki Y.W., Ishizu H., Uwamino Y. et al. Pro108Ser mutant of SARS-CoV-2 3CLpro reduces the enzy-

- matic activity and ameliorates COVID-19 severity in Japan. *medRxiv*. 2020;11(24):20235952. DOI: 10.1101/2020.11.24.20235952.
15. Caccuri F., Zani A., Messali S., Giovanetti M., Bugatti A., Campisi G. et al. A persistently replicating SARS-CoV-2 variant derived from an asymptomatic individual. *J. Transl. Med.* 2020;18(1):362. DOI: 10.1186/s12967-020-02535-1.
16. Wang R., Chen J., Gao K., Hozumi Y., Yin C., Wei G.W. Analysis of SARS-CoV-2 mutations in the United States suggests presence of four substrains and novel variants. *Commun. Biol.* 2021;4(1):228. DOI: 10.1038/s42003-021-01754-6.
17. Korber B., Fischer W.M., Gnanakaran S., Yoon H., Theiler J., Abfalterer W. et al. Tracking changes in SARS-CoV-2 spike: Evidence that D614G increases infectivity of the COVID-19 virus. *Cell*. 2020;182(4):812–827.e19.
18. Shcherbak S.G., Anisenkova A.Y., Mosenko S.V., Glotov O.S., Chernov A.N., Apalko S.V. et al. Basic predictive risk factors for cytokine storms in COVID-19 patients. *Front. Immunol.* 2021;12:745515. DOI: 10.3389/fimmu.2021.745515.
19. Myrstad M., Ihle-Hansen H., Tveita A.A., Andersen E.L., Nygård S., Tveit A. et al. National Early Warning Score 2 (NEWS2) on admission predicts severe disease and in-hospital mortality from Covid-19 – a prospective cohort study. *Scand. J. Trauma Resusc. Emerg. Med.* 2020;28(1):66. DOI: 10.1186/s13049-020-00764-3.
20. King C.S., Sahjwani D., Brown A.W., Feroz S., Cameron P., Osborn E. et al. Outcomes of mechanically ventilated patients with COVID-19 associated respiratory failure. *PLoS One*. 2020;15(11):e0242651. DOI: 10.1371/journal.pone.0242651.
21. Caricchio R., Gallucci M., Dass C., Zhang X., Gallucci S., Fleece D. et al. Preliminary predictive criteria for COVID-19 cytokine storm. *Ann. Rheum. Dis.* 2020;0:1–8. DOI: 10.1136/annrheumdis-2020-218323.
22. Yilmaz A., Sabirli R., Seyit M., Ozen M., Oskay A., Cakmak V. et al. Association between laboratory parameters and CT severity in patients infected with Covid-19: A retrospective, observational study. *Am. J. Emerg. Med.* 2021;42:110–114. DOI: 10.1016/j.ajem.2021.01.040.

Информация о вкладе авторов

Щербак С.Г., Анисенкова А.Ю., Глотов О.С. предложили концепцию исследования и разработали его протокол, участвовали в заключительной редакции рукописи.

Глотов О.С., Чернов А.Н., Мосенко С.В. написали рукопись статьи, провели статистический анализ, анализировали и интерпретировали данные исследований.

Калинин Р.С., Цай В.В. провели выделение РНК, ДНК из образцов, ПЦР, секвенирование и биоинформатический анализ данных, а также участвовали в обсуждении результатов.

Анисенкова А.Ю. осуществляла лечение и обследование пациентов, составляла клиническую базу пациентов, проводила клиническую интерпретацию результатов.

Мосенко С.В., Уразов С.П. проводили литературный поиск, составляли клиническую базу пациентов, проводили лечение пациентов.

Коробейников А.И., Липидус А.Л. провели биоинформатический анализ данных, предложили алгоритмы анализа данных секвенирования, а также участвовали в обсуждении результатов.

Все авторы дали окончательное согласие на подачу рукописи и согласились нести ответственность за все аспекты работы, ручаясь за их точность и безупречность.

Information on author contributions

Shcherbak S.G., Anisenkova A.Yu., and Glotov O.S. proposed a research concept and developed its protocol, participated in the final edition of the manuscript.

Glotov O.S., Chernov A.N., and Mosenko S.V. wrote the manuscript, conducted statistical analysis, analyzed and interpreted the research data.

Kalinin R.S. and Tsai V.V. performed the extraction of RNA, DNA from the samples, PCR and sequencing and bioinformatics analysis of the data, and also participated in the discussion of the results.

Anisenkova A.Yu. performed treatment and examination of patients, contributed to compilation of clinical database, and provided clinical interpretation of results.

Mosenko S.V. and Urazov S.P. contributed to literature search, compilation of clinical database, and treatment of patients.

Korobeynikov A.I. and Lapidus A.L. conducted bioinformatic data analysis, proposed sequencing data analysis algorithms, and participated in the discussion of the results.

All authors have given their final consent to the submission of the manuscript and agreed to be responsible for all aspects of the work, vouching for their accuracy and flawlessness.

Сведения об авторах

Глотов Олег Сергеевич, канд. биол. наук, заведующий отделом вирусологических и молекулярно-генетических методов диагностики, Детский научно-клинический центр инфекционных болезней Федерального медико-биологического агентства. ORCID 0000-0002-0091-2224.

E-mail: olglotov@mail.ru.

Чернов Александр Николаевич, научный сотрудник отдела Института экспериментальной медицины. ORCID 0000-0003-2464-7370.

E-mail: al.chernov@mail.ru.

Коробейников Антон Иванович, канд. физ.-мат. наук, доцент кафедры статистического моделирования, Санкт-Петербургский государственный университет. ORCID 0000-0002-2937-9259.

E-mail: a.korobeynikov@spbu.ru.

Калинин Роман Сергеевич, младший научный сотрудник отдела вирусологических и молекулярно-генетических методов диагностики, Детский научно-клинический центр инфекционных болезней Федерального медико-биологического агентства. ORCID 0000-0003-1791-7045.

E-mail: pancu43@gmail.com.

Цай Виктория Викторовна, младший научный сотрудник отдела вирусологических и молекулярно-генетических методов диагностики, Детский научно-клинический центр инфекционных болезней Федерального медико-биологического агентства. ORCID 0000-0001-6488-8369.

E-mail: viktoriya14054@gmail.com.

Анисенкова Анна Юрьевна, канд. мед. наук, доцент, врач-терапевт, кардиолог, заведующий терапевтическим отделением, Городская боль-

Information about the authors

Oleg S. Glotov, Cand. Sci. (Biol.), Head of Department of Virological and Molecular Genetic Methods of Diagnostics, Children's Scientific and Clinical Center for Infectious Diseases of the Federal Medical and Biological Agency. ORCID 0000-0002-7465-4504.

E-mail: olglotov@mail.ru.

Alexander N. Chernov, Research Scientist, Institute of Experimental Medicine. ORCID 0000-0003-2464-7370.

E-mail: al.chernov@mail.ru.

Anton I. Korobeynikov, Cand. Sci. (Phys.-Math.), Associate Professor, Department of Statistical Modelling, Saint Petersburg State University. ORCID 0000-0002-2937-9259.

E-mail: a.korobeynikov@spbu.ru.

Roman S. Kalinin, Junior Research Scientist, Department of Virological and Molecular Genetic Methods of Diagnostics, Children's Scientific and Clinical Center for Infectious Diseases, Federal Medical and Biological Agency. ORCID 0000-0003-1791-7045.

E-mail: pancu43@gmail.com.

Viktoria V. Tsai, Junior Research Scientist, Department of Virological and Molecular Genetic Methods of Diagnostics, Children's Scientific and Clinical Center for Infectious Diseases, Federal Medical and Biological Agency. ORCID 0000-0001-6488-8369.

E-mail: viktoriya14054@gmail.com.

Anna Yu. Anisenkova, Cand. Sci. (Med.), Associate Professor, Internist, Cardiologist, Head of Internal Medicine Department, City Hospital No. 40 of

ница № 40 Курортного административного района. ORCID 0000-0001-5642-621X.

E-mail: anna_anisenkova@list.ru.

Уразов Станислав Петрович, врач-кардиолог, Городская больница № 40 Курортного административного района. ORCID 0000-0002-5441-2911.

E-mail: urasta@list.ru.

Лapidus Алла Львовна, канд. биол. наук, директор Центра биоинформатики и алгоритмической биологии, Санкт-Петербургский государственный университет. ORCID 0000-0003-0427-8731.

E-mail: a.lapidus@spbu.ru.

Мосенко Сергей Викторович, канд. мед. наук, врач-невролог высшей категории, Городская больница № 40 Курортного административного района. ORCID 0000-0002-1357-4324.

E-mail: neurologist@mail.ru.

Щербак Сергей Григорьевич, д-р мед. наук, профессор, главный врач Городской больницы № 40 Курортного административного района. ORCID 0000-0001-5047-2792.

E-mail: sgsherbak@mail.ru.

 **Глотов Олег Сергеевич**, e-mail: olglotov@mail.ru.

Kurortny Administrative District. ORCID 0000-0001-5642-621X.

E-mail: anna_anisenkova@list.ru.

Stanislav P. Urazov, Cardiologist, City Hospital No. 40 of Kurortny Administrative District. ORCID 0000-0002-5441-2911.

E-mail: urasta@list.ru.

Alla L. Lapidus, Cand. Sci. (Biol.), Director of Center for Bioinformatics and Algorithmic Biotechnology, Saint Petersburg State University. ORCID 0000-0003-0427-8731.

E-mail: a.lapidus@spbu.ru.

Sergei V. Mosenko, Cand. Sci. (Med.), City Hospital No. 40 of Kurortny Administrative District. ORCID 0000-0002-1357-4324.

E-mail: neurologist@mail.ru.

Sergei G. Shcherbak, Dr. Sci. (Med.), Professor, Chief Physician, City Hospital No. 40 of Kurortny Administrative District. ORCID 0000-0001-5047-2792.

E-mail: sgsherbak@mail.ru.

 **Oleg S. Glotov**, e-mail: olglotov@mail.ru.

Received September 30, 2021

Поступила 30.09.2021