



ИДЕНТИФИКАЦИЯ РЕДКИХ ОДНОНУКЛЕОТИДНЫХ ГЕНЕТИЧЕСКИХ ПОЛИМОРФИЗМОВ В ГЕНОМЕ ГОСПИТАЛИЗИРОВАННЫХ БОЛЬНЫХ COVID-19

А.А. Передерий¹, А.Б. Комиссаров¹, Д.М. Даниленко¹, Е.В. Венев^{1,2,3}, С.А. Коноплева⁴, М. Коржанова¹, М.М. Писарева¹, Д.А. Гусев², Д.А. Лиознов^{1,3}

¹ Научно-исследовательский институт гриппа им. А.А. Смородинцева, Санкт-Петербург, Россия

² Клиническая инфекционная больница им. С.П. Боткина, Санкт-Петербург, Россия

³ Первый Санкт-Петербургский государственный медицинский университет им. академика И.П. Павлова, Санкт-Петербург, Россия

⁴ Санкт-Петербургский государственный университет, Санкт-Петербург, Россия

Identification of rare single nucleotide polymorphisms in the genome of hospitalized COVID-19 patients

A.A. Perederiy¹, A.B. Komissarov¹, D.M. Danilenko¹, E.V. Venev^{1,2,3}, S.A. Konopleva⁴, M. Korzhanova¹, M.M. Pisareva¹, D.A. Gusev², D.A. Lioznov^{1,3}

¹ Research Institute of Influenza named after A.A. Smorodintsev, Saint-Petersburg, Russia

² Clinical Infectious Diseases Hospital named after S.P. Botkin, Saint-Petersburg, Russia

³ First Saint-Petersburg State Medical University named after academician I.P. Pavlov, Saint-Petersburg, Russia

⁴ Saint-Petersburg State University, Saint-Petersburg, Russia

Резюме

Целью данного контролируемого исследования было выявление полиморфизмов в геноме больных COVID-19, ассоциированных с частотой госпитализации.

Материалы и методы: сформировано 2 группы пациентов: основная — 56 больных COVID-19, госпитализированных не менее двух раз за период исследования, и контрольная — 107 пациентов, для которых был подтверждена однократная госпитализация с COVID-19 за период исследования. Проведено полноэкзомное секвенирование материала из остаточных назофарингеальных мазков от госпитализированных больных COVID-19 на платформе MGI с последующим биоинформатическим анализом и применением анализа обогащения генов.

Результат: впервые проведено экзомное секвенирование из назофарингеальных мазков от 163 пациентов, госпитализированных с COVID-19 с использованием платформы MGI. В основной группе выявлены уникальные варианты генетических полиморфизмов, в том числе 6 ранее не описанных.

Заключение: генетические варианты, статистически значимо ассоциированные со случаями однократной или неоднократной госпитализации больных COVID-19 в ходе исследования, выявлены не были. Назофарингеальные мазки могут быть использованы для проведения полноэкзомного секвенирования. Необходимы дальнейшие исследования для идентификации уникальных генетических вариантов, ответственных за восприимчивость к инфекционным заболеваниям.

Ключевые слова: полноэкзомное секвенирование, редкие генетические варианты, назофарингеальные мазки, MUC5B, SERPINE1.

Abstract

The aim of this controlled study was to identify polymorphisms in the genome of COVID-19 patients associated with the frequency of hospitalization.

Materials and methods: Two groups of patients were formed: the main group — 56 patients with COVID-19, hospitalized at least twice during the study period, and the control group — 107 patients for whom only one hospitalization with COVID-19 was confirmed during the study period. Whole-exome sequencing of residual nasopharyngeal swabs from patients hospitalized with COVID-19 was performed on the MGI platform, followed by bioinformatics analysis and gene enrichment analysis.

Result: For the first time, exome sequencing was performed from oropharyngeal swabs from 163 patients hospitalized with COVID-19 using the MGI platform. In the main group, unique variants of genetic polymorphisms were identified, including six previously undescribed ones.

Conclusion: No genetic variants were statistically significantly associated with single or multiple hospitalizations of COVID-19 patients in the study. Nasopharyngeal swabs can be used for whole exome sequencing. Further studies are needed to identify unique genetic variants responsible for susceptibility to infectious diseases. Nasopharyngeal swabs can be used for whole-exome sequencing. Further studies are needed to identify unique genetic variants responsible for susceptibility to infectious diseases.

Key words: whole exome sequencing, rare genetic variants, nasopharyngeal swabs, MUC5B, SERPINE1.

Введение

Развитие методов геномного анализа дало мощный импульс исследованиям в области генетики человека [1] и поиску генетических детерминант, определяющих восприимчивость к тем или иным инфекционным болезням, и прогнозу их течения. Пандемия COVID-19 привела к значительному росту числа исследований, связанных как с изучением генома SARS-CoV-2 (впервые было получено свыше 17 млн полных последовательностей генома для одного единственного возбудителя в рекордно короткие сроки [2]), так и с идентификацией генетических маркеров предрасположенности к возбудителю COVID-19 и характером течения заболевания [3].

Стремительное внедрение методов секвенирования нового поколения позволяет исследователям в разных странах мира анализировать различные по объему выборки здоровых и больных людей и выявлять связи между геном человека и восприимчивостью к инфекциям. Большинство крупных исследований в настоящее время являются полногеномным анализом ассоциаций (GWAS — genome-wide association studies). Именно они сыграли важную роль в выявлении генетических факторов, связанных с восприимчивостью к COVID-19. Так, крупнейшее исследование COVID-19 HG project [4, 5] рассматривало связь генетических вариантов, включающих множество локусов, с симптоматическим течением и тяжестью COVID-19. Другие исследования позволили идентифицировать ранее не выявленную группу генов, определённые аллельные варианты в которых приводят к крайне тяжелому течению заболевания. В исследовании [6] были идентифицированы 23 независимых варианта генов, которые определяли тяжелое течение COVID-19, в том числе 16 новых независимых ассоциаций, включая варианты, участвующие в передаче сигналов интерферона (IL10RB и PLSCR1) и дифференциации лейкоцитов (BCL11A).

Исследования GWAS позволили установить, что многие из выявленных в геноме человека локусов связаны с функцией иммунной системы, включая гены, участвующие в пути интерферона, который играет решающую роль в реакции организма на вирусные инфекции. Варианты в этих генах могут влиять на то, насколько эффективно иммунная система человека может реагировать на SARS-CoV-2 [9].

Необходимо отметить, что большую роль в генетической предрасположенности людей к инфекциям и, в частности, к SARS-CoV-2, играет их этническая принадлежность, поскольку разные этносы обладают разными аллелями в генах, имеющих ключевое значение для развития COVID-19.

Еще в начале пандемии COVID-19 было установлено, что исход заболевания значительно различается в разных этнических группах [7]. GWAS подтвердил, что генетический риск тяжелого течения COVID-19 может различаться среди разных популяций. Эта изменчивость может быть связана с различиями в частотах аллелей в разных этнических группах. В одном из крупных исследований было установлено различное распространение отдельных аллельных вариантов ACE2 у разных пациентов, часть из которых приводили к сверхэкспрессии основного рецептора SARS-CoV-2, тем самым увеличивая риск возникновения инфекции [8]. В настоящее время опубликовано множество работ, посвященных отдельным этническим группам населения, описывающих аллельные варианты в тех или иных генах, ответственных за повышенную восприимчивость к COVID-19.

Исследования GWAS способствовали более глубокому пониманию сложного взаимодействия между генетикой индивидуума и тяжестью течения COVID-19; именно такие исследования должны стать частью персонализированной медицины, поскольку они обеспечивают ценную информацию не только для отдельно взятого человека, но и для стратегий общественного здравоохранения.

Вместе с тем, полногеномный анализ ассоциаций обычно фокусируется на распространенных вариантах, не позволяя описывать более редкие варианты, часто определяющие восприимчивость отдельных пациентов к заболеванию. При этом ожидается, что варианты, оказывающие наибольшее влияние на исходы COVID-19, будут встречаться в популяции редко [10]. Для идентификации таких вариантов лучше подходят исследования экзому человека, которые позволяют идентифицировать редкие варианты аллелей в геноме и выявлять их значимость для иммунного ответа и тяжести заболевания.

В настоящее время опубликовано большое количество работ, посвященных анализу данных экзомного секвенирования и связи с тяжестью течения COVID-19. Так, например, установлено, что у носителей полиморфизмов в гене, кодирующем рецептор интерферона альфа IFNAR2, COVID-19 протекает тяжело [11]. В дальнейшем такие же данные были получены в отдельных странах — Бразилии [12], Мексике [13], Палестине [14].

В России также проводятся исследования, направленные на выявление генов, обогащенных редкими генетическими вариантами, связанными с тяжестью течения COVID-19, на данных российской популяции при помощи репликационного анализа [15]. Более того, развитие исследований в области экзомного секвенирования позволило создать отечественную базу данных RUSeq — открытый крупномасштабный референтный на-

бор генетических вариантов на основе 7452 образцов экзомов, собранных в Москве и Санкт-Петербурге, что позволяет анализировать частоты встречаемости отдельных генетических вариантов непосредственно в российской популяции [16].

Большинство проведенных исследований по анализу экзомов были направлены на поиск связи между отдельными аллельными вариантами и тяжестью течения COVID-19.

Вместе с тем, в популяции реконвалесцентов COVID-19 следует выделить несколько групп, например, лиц, у которых зарегистрирована однократная госпитализация, и лиц, которые были госпитализированы дважды или многократно.

Цель исследования — выявление полиморфизмов в геноме больных COVID-19, ассоциированных с разной частотой госпитализации.

Материалы и методы исследования

В исследование включены назофарингеальные мазки 163 больных COVID-19 (U07.1; U07.2), госпитализированных в Клиническую инфекционную больницу им. С.П. Боткина (Санкт-Петербург) в марте 2020 г. — декабре 2023 г. (период исследования). Мазок из ротоглотки брали при поступлении больного в стационар.

Выборка исследования сформирована из 2 групп: 1) основная — 56 больных COVID-19, госпитализированных не менее двух раз за период исследования; 2) контрольная — 107 пациентов, для которых был подтверждена однократная госпитализация с COVID-19 за период исследования.

Секвенирование. Выделение геномной ДНК проводили из орофарингеальных мазков набором DU-250 (Биолабмикс, Россия). Подготовку ДНК-библиотек выполняли с использованием набора NadPrep EZ DNA Library Preparation Kit v2 (Nanodigmbio, КНР). Далее готовые библиотеки обогащали экзонами с помощью набора реагентов NadPrep ES Hybrid Capture (Nanodigmbio, КНР) и набора зондов NEXome Plus Panel (Nanodigmbio, КНР). Размер экзомной панели составляет 35,2 млн нуклеотидов, включая участки HLA, наиболее распространённые в популяции однонуклеотидные варианты и полный экзом человека. Секвенирование выполнено в парноконцевом режиме с длиной прочтения 150 п.о. (суммарно 300 п.о.) на секвенаторе DNBSEQ-G400 (MGI, КНР) с целевым покрытием 100x.

Биоинформатический анализ. Полученные «сырые» данные секвенирования прошли оценку качества, выравнивание на референсный геном и аннотацию. Отбор генов-кандидатов для исследования потенциально значимых взаимосвязей был проведен на основе анализа оригинальной литературы и, в частности, наиболее цитируемых

обзоров по сочетанию ключевых слов «GWAS», «SARS-CoV-2», «COVID-19», «Exome sequencing» в базах данных публикаций PubMed и Google Scholar. Избранные гены также проверялись на наличие в экзомной панели.

Обработка и интерпретация данных секвенирования проведена с использованием ПО «Mutracker» (РУ № РЗН 2024/23293 от 07.08.2024) производства ФГАОУ ВО РНИМУ им. Н.И. Пирогова Минздрава России.

Для оценки качества «сырых» данных секвенирования использовали программу FastQC [17].

Определение вариантов осуществляли с помощью программ bcftools [Li H., 2011] и DeepVariant [18].

Для дальнейшего анализа путём двухэтапного процесса отбора был составлен список из 134 генов, ассоциированных с инфекцией SARS-CoV-2. Первый этап включал в себя отбор части генов из базы данных MalaCards [19], которая представляет собой интегрированный ресурс для изучения взаимосвязей между различными заболеваниями, молекулярными путями и генами. Из этой базы данных было извлечено 89 генов, которые имеют теоретические или документированные ассоциации с патогенезом COVID-19 и его осложнениями. Второй этап включал систематический литературный поиск для выявления дополнительных генов, связанных с заболеванием. Поиск проводился в базах данных PubMed, Google Scholar и ScienceDirect с использованием ключевых слов: «SARS-CoV-2», «COVID-19», «Генетические ассоциации», «Варианты генов», «Иммунный ответ». Особое внимание уделяли исследованиям, в которых рассматривались генетические варианты (например, однонуклеотидные полиморфизмы) и их связь с восприимчивостью к инфекции или тяжестью течения заболевания. В процессе поиска были отобраны 45 генов.

Анализ обогащения генов. Анализ обогащения генов, ассоциированных с SARS-CoV-2, проводился комплексно, путём поиска взаимодействий как самих белковых продуктов, так и биологических функциональных путей. Для выявления потенциальных взаимодействий между белками использовали базу данных STRING [Szklarczyk D. et al., 2023]. Результаты анализа включали построение сети взаимодействий белок — белок. Были исключены слабые взаимодействия и белки с низким уровнем доказательности ассоциаций. Для дальнейшего анализа использовали только высокоуверенные взаимодействия (порог схожести >0,8). Для анализа молекулярных путей использовалась база данных WikiPathways [22], которая интегрирована в STRING. В результате для каждого гена был проведён анализ обогащения на основе его участия в известных молекулярных путях. Порог

значимости для обогащения был установлен на уровне $p\text{-value} < 0,05$ с поправкой на множественные тестирования методом Бенджамини — Хохберга. Особое внимание уделяли путям, связанным с иммунным ответом, воспалительными процессами и вирусными инфекциями. Полученная сеть белок — белок взаимодействий визуализировалась с помощью Cytoscape [23].

Результаты и обсуждение

Описание выборки

Медиана возраста пациентов составила 69 лет (19, 97), мужчин — 37% (60 чел.). Зарегистрировано в контрольной группе 2 случая смерти (летальность в общей выборке — 1,9%), средняя дли-

тельность госпитализации — $5,2 \pm 4,2$ дня. Против COVID-19 были вакцинированы 23% (13 чел.) пациентов основной и 40% (43 чел.) контрольной группы. На рисунке 1 представлена хронология госпитализаций больных.

Генетические варианты

Полноэкзомное секвенирование и таргетный анализ 134 генов, отобранных по данным MalaCards и литературным данным, позволили идентифицировать 51 ген, несущий вариации у пациентов основной группы (табл. 1). Всего в данных генах было выявлено 89 вариантов, большинство из которых составляли однонуклеотидные полиморфизмы (82 SNV, single nucleotide variants).

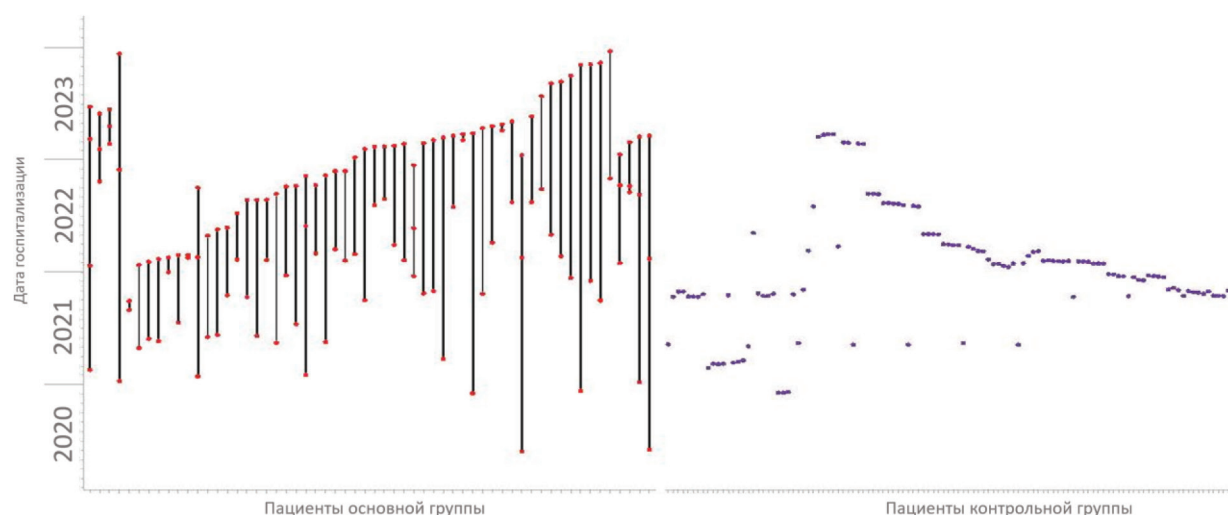


Рис. 1. Хронология госпитализаций больных COVID-19 основной (слева) и контрольной (справа) группы. Точки показывают дату госпитализации для каждого включенного в исследование пациента. В основной группе длина линии указывает на длительность промежутка между эпизодами госпитализаций (каждая точка — отдельный эпизод)

Таблица 1

Генетические варианты, выявленные в основной группе больных COVID-19

Ген	Хромосома	rsID	Тип варианта	Количество вариантов
MUC5B	11	rs199749553, rs199704716, rs546308873, rs200554635, rs201750183, rs566783011, rs201530286, rs190819085, rs201532622, rs200620306, rs750562955, rs200138247, rs200770447, rs56159668	Миссенс	14
HLA-C	6	rs1143551, rs41560617, rs141142418	Миссенс	3
ACE	17	rs199591851, rs201804955	Миссенс	2
ADAR	1	rs958924655, rs761657434	Миссенс	2
CPN1	10	rs1589481488, rs748121885	Миссенс	2
IRF7	11	rs201379782, rs201036875	Миссенс	2
NRP1	10	rs201570278, rs2228637	Миссенс	2
SLC6A19	5	rs765798920, rs1746128246	Миссенс	2
TLR4	9	rs200195902, rs5030724	Миссенс	2
UNC93B1	11	rs144399212, rs146593182	Миссенс	2
VWF	12	rs61750580, rs781111573	Миссенс	2

Окончание таблицы 1

Ген	Хромосома	rsID	Тип варианта	Количество вариантов
ABO	9	rs555009598	Миссенс	1
AGTR2	X	rs138706948	Миссенс	1
AIRE	21	rs121434254	Нонсенс	1
ALB	4	rs79228041	Миссенс	1
BDKRB1	14	rs2229459	Миссенс	1
BDKRB1	14	rs145322761	Нонсенс	1
BIRC2	11	rs139349997	Миссенс	1
BMAL1	11	novel	Миссенс	1
CRP	1	rs77832441	Миссенс	1
CYP24A1	20	rs777676129	Делеция без сдвига рамки считывания	1
EGFR	7	rs773596817	Миссенс	1
F2	11	rs756544259	Миссенс	1
F3	1	rs5901	Миссенс	1
FCGR2A	1	rs148465413	Миссенс	1
FOXP3	X	rs782188307	Миссенс	1
FYCO1	3	rs4356851	Миссенс	1
HLA-A	6	rs757622018	Инсерция со сдвигом рамки считывания	1
HLA-C	6	rs148706212	Инсерция без сдвига рамки считывания	1
IFIH1	2	rs35744605	Нонсенс	1
IFNA1	9	rs28383794	Миссенс	1
IFNAR1	21	rs1257850924	Миссенс	1
IL17A	6	rs1444067564	Миссенс	1
MAVS	20	rs201823260	Миссенс	1
MAVS	20	rs751608344	Делеция без сдвига рамки считывания	1
MUC5B	11	rs754593750	Делеция без сдвига рамки считывания	1
NLRP1	17	rs1173780449	Миссенс	1
NRP2	2	novel, rs200406312	Миссенс	1
OAS1	12	rs774712843	Сайт сплайсинга	1
PLG	6	rs4252070	Миссенс	1
REL	2	rs199772834	Миссенс	1
REN	1	rs776126542	Миссенс	1
SCARB1	12	rs74830677	Миссенс	1
SERPINE1	7	rs1247505082	Миссенс	1
SLC6A20	3	rs199831562	Миссенс	1
STAT2	12	novel	Миссенс	1
STING1	5	novel	Миссенс	1
TICAM1	19	rs141394423	Миссенс	1
TLR2	4	rs121917864	Миссенс	1
TLR7	X	rs750434070	Миссенс	1
TYK2	19	rs35018800	Миссенс	1

У 37 больных основной группы (66%; 95% ДИ 53,7 – 78,5%) было выявлено как минимум 2 редких генетических варианта, и 24 пациента имели три и более варианта (42,8%; 95% ДИ 29,9 – 55,8%). Особо следует отметить пациента, в экзOME которого было выявлено 22 редких генетических варианта (в том числе 8 в гене главного комплекса гистосовместимости HLA-C и 6 в гене муцина MUC5B), из них 3 варианта в настоящее время отсутствуют в российской и международных базах данных.

У 6 пациентов были выявлены не депонированные ранее в базу данных gnomAD генетические варианты (табл. 2). Все идентифицированные варианты были однонуклеотидными полиморфизмами, приводящими к аминокислотным заменам. Была проведена проверка выявленных вариантов по российской базе данных RUSeq. 2 варианта (в генах BMAL1 и NRP2) были ранее идентифицированы и депонированы в RUSeq, причем варианты обнаруживались только в Сибири и на Дальнем Востоке и частоты аллелей были очень низкими.

Контрольная группа пациентов также была генетически гетерогенной. У 57 человек (58%; 95% ДИ 48,6%–67,3%) установлены 2 и более однонуклеотидных вариантов, и 45 больных (42%, 95% ДИ, 32,7 – 51,4%) являлись носителями 3 и более вариантов в исследованных генах. Таким образом, по количеству выявленных генетических вариантов пациенты основной и контрольной групп отличались.

Необходимо отметить, что нами выявлено 13 генов, полиморфизмы в которых обнаруживались только у пациентов основной группы, однако представленность конкретных генетических вариантов в группе пациентов была низкой.

По всем выявленным вариантам, для которых имелись данные, было проведено сравнение частот в базе данных gnomAD (Non-Finnish European population, NFE, европейцы, не финны) с базой

данных RUSeq. Были выявлены различия в частоте ряда вариантов, не превышающие 0,2 – 0,6%. Так, частота варианта rs2228637 в гене NRP1, кодирующем ко-рецептор SARS-CoV-2 белок NRP1, по данным RUSeq, составляет 1,7%, в то время как в базе gnomAD по группе NFE она составляет около 1,1%. Таким образом, для большинства исследованных генов различия в частотах полиморфизмов между российской популяцией (по данным RUSeq) и европейской не финской популяцией (по данным gnomAD) являются незначительными. Наиболее часто встречающимися в популяции генетическими вариантами из исследованного набора генов оказались rs199749553 (MUC5B) – 16% и rs28383794 (IFNA1) – 7%.

Функциональная аннотация и анализ молекулярных путей в STRING

База данных STRING содержит данные о физических и функциональных взаимодействиях белков. Анализ даёт возможность определить сверхпредставленные биологические процессы среди исследуемых белков.

Для каждого гена, для которого в группе пациентов с неоднократной госпитализацией были выявлены генетические варианты, был проведён анализ обогащения на основе его участия в известных молекулярных путях согласно базе данных WikiPathways (рис. 2). Визуализация отображает 10 наиболее обогащённых наборов генов, выявленных в результате анализа. Наиболее обогащёнными оказались наборы генов, связанные с вирусной инфекцией (разными патогенами), системой врожденного иммунитета, индукцией интерферонов I типа, сигнальными путями TLR, а также ренин-ангиотензиновой системой и сигнальными путями брадикинина, что представляет особый интерес в связи с важной ролью брадикинин-ренин-ангиотензиновой системы в патогенезе COVID-19 (теория брадикининового шторма) [24].

Таблица 2

Выявленные генетические варианты, отсутствующие в базе данных gnomAD

Ген	Хромосома	Положение	Тип	Референс	SNV	Замена	Патогенный потенциал*	RUSeq
BMAL1	chr11	13366675	Миссенс	T	C	I188T	Вероятно патогенный	AF = 0,0019 (Восточная Россия)
MUC5B	chr11	1242750	Миссенс	C	G	S1957C	Неизвестно	Нет в RUSeq
NRP2	chr2	205716374	Миссенс	G	A	G145S	Вероятно патогенный	AF = 0,0019 (Восточная Россия)
SERPINE1	chr7	101130594	Миссенс	T	C	F149L	Вероятно патогенный	Нет в RUSeq
STAT2	chr12	56351432	Миссенс	C	G	K267N	Вероятно патогенный	Нет в RUSeq
STING1	chr5	139480889	Миссенс	G	C	P141A	Доброкачественный	Нет в RUSeq
TLR2	chr4	153704717	Миссенс	C	T	L604F	Вероятно патогенный	Нет в RUSeq
UNC93B1	chr11	67998449	Миссенс	T	C	S231G	Вероятно патогенный	Нет в RUSeq

* Согласно алгоритму PolyPhen2.

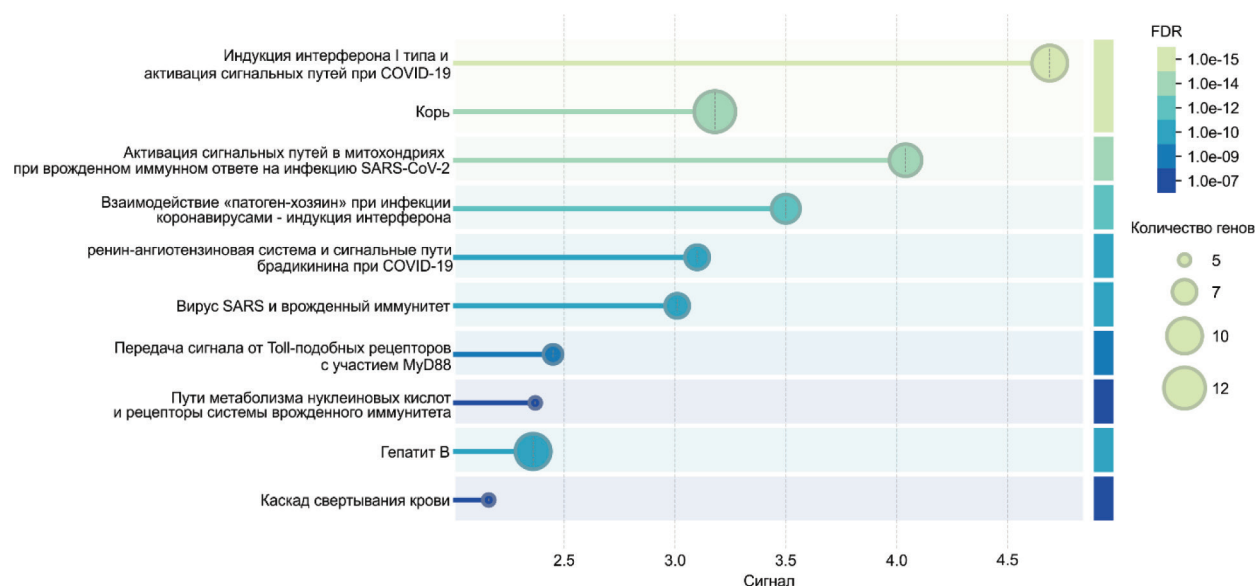


Рис. 2. Результаты анализа обогащения генов, ассоциированных с SARS-CoV-2, выполненного через интерфейс WikiPathways в базе данных STRING: FDR — скорректированное р-значение, отражающее значимость обогащения для каждого пути

Обсуждение

В мировой практике наиболее часто используемым типом биоматериала для проведения экзомного секвенирования является кровь, реже для неинвазивного тестирования — буккальные мазки. В онкогинекологической практике описан успешный опыт использования цервикальных мазков [25]. За время пандемии COVID-19 в мире были собраны большие коллекции назо- и орофарингеальных мазков от больных COVID-19, однако примеров использования респираторных мазков для экзомного секвенирования нами в литературе не найдено. В результате проведенного исследования впервые было показано, что назофарингеальные мазки пригодны для проведения полноэкзомного секвенирования ДНК человека и поиска редких генетических вариантов.

Нами не установлены генетические варианты, статистически значимо ассоциированные со случаями однократной или неоднократной госпитализации больных COVID-19, что может объясняться как отсутствием прямой связи, так и недостаточностью выборки. В то же время важным результатом исследования является идентификация ранее не описанных однонуклеотидных вариантов в 6 генах, которые могут быть тканеспецифичными соматическими мутациями [29], что нуждается в дополнительной экспериментальной проверке. Так, впервые описан полиморфизм гена *SERPINE1*, кодирующего ингибитор активатора плазминогена PAI-1 — chr7 101130594: T>C (F149L). Ранее в исследованиях по направленному мутагенезу было показано, что замена F126A (нумерация с сигнальным пептидом — F149A) может приводить к снижению чувстви-

тельности к ингибитору XR5118 [26]. Кроме того, белок с данной заменой выступает в качестве субстрата, в результате чего серпин расщепляется, и интактный активатор плазминогена высвобождается. Таким образом, данный полиморфизм может способствовать снижению тромбообразования.

Секретируемые муцины MUC5AC и MUC5B — это крупные гликопротеины, которые играют важнейшую защитную роль, участвуя в захвате патогенов и мукоцилиарном клиренсе. Гены муцинов, в частности MUC5B, крайне полиморфны. В настоящее время описано 1135 полиморфизмов в гене MUC5B [27]. В исследовании [28] показано, что вариант MUC5B rs35705950-T ассоциирован со снижением риска госпитализации при COVID-19, но не дает преимуществ против тяжелого течения заболевания. Однако учитывая колоссальное разнообразие однонуклеотидных полиморфизмов, описанных для MUC5B, требуются дополнительные популяционные исследования для установления однозначной связи конкретных мутаций с восприимчивостью к COVID-19 и течением заболевания.

Полиморфизм в генах TLR и, в частности, TLR2, описан у пациентов с пневмониями, вызванными как SARS-CoV-2, так и бактериальными патогенами [31]. Гены *STAT2* и *STING1* кодируют белки, вовлеченные в обнаружение вирусной РНК, и являются интерферон-стимулируемыми генами. Белок, кодируемый *STING1*, участвует в сигнальном пути, который контролирует иммунитет к цитозольной ДНК и является важнейшим фактором aberrантных реакций интерферона I типа, в том числе при COVID-19. Также известно, что неструктурный белок NSP6 SARS-CoV-2 может инду-

цировать аутофагию, в результате которой происходит деградация STING1 и снижение клеточного противовирусного ответа [30].

Таким образом, исследования экзоста человека являются мощным инструментом для раскрытия генетических основ предрасположенности к возбудителям инфекционных болезней, в частности к SARS-CoV-2, и вариантов течения инфекционного процесса. Получаемые в результате секвенирования ДНК человека данные могут быть использованы для популяционного моделирования и разработки профилактических и терапевтических стратегий персонализированной медицины.

Исследование выполнено при поддержке гранта НЦМУ «Передовые цифровые технологии» (в рамках Соглашения о предоставлении гранта от 20.04.2022 №075-15-2022-313).

Литература

1. The Human Genome Project. Nature milestones. 2021. <https://www.nature.com/articles/d42859-020-00101-9>
2. Brito AF, Semenova E, Dudas G, Hassler GW, Kalinich CC, et al. Global disparities in SARS-CoV-2 genomic surveillance. *Nat Commun*. 2022 Nov 16;13(1):7003. doi: 10.1038/s41467-022-33713-y.
3. The COVID-19 Host Genetics Initiative. The COVID-19 Host Genetics Initiative, a global initiative to elucidate the role of host genetic factors in susceptibility and severity of the SARS-CoV-2 virus pandemic. *Eu. J. Hum. Genet*. 28, 715–718 (2020).
4. COVID-19 Host Genetics Initiative. Mapping the human genetic architecture of COVID-19. *Nature* 600, 472–477 (2021). <https://doi.org/10.1038/s41586-021-03767-x>
5. COVID-19 Host Genetics Initiative. A first update on mapping the human genetic architecture of COVID-19. *Nature* 608, E1–E10 (2022). <https://doi.org/10.1038/s41586-022-04826-7>
6. Kousathanas, A., Pairo-Castineira, E., Rawlik, K. et al. Whole-genome sequencing reveals host factors underlying critical COVID-19. *Nature* 607, 97–103 (2022). <https://doi.org/10.1038/s41586-022-04576-6>
7. Pan D, Sze S, Minhas JS, Bangash MN, Pareek N, Divall P, Williams CM, Oggioni MR, Squire IB, Nellums LB, Hanif W, Khunti K, Pareek M. The impact of ethnicity on clinical outcomes in COVID-19: A systematic review. *EclinicalMedicine*. 2020 Jun 3;23:100404. doi: 10.1016/j.eclinm.2020.100404
8. Vadgama N, Kreymerman A, Campbell J, Shamardina O, Brugger C, Research Consortium GE, Deaconescu AM, Lee RT, Penkett CJ, Gifford CA, Mercola M, Nasir J, Karakikes I. SARS-CoV-2 Susceptibility and ACE2 Gene Variations Within Diverse Ethnic Backgrounds. *Front Genet*. 2022 Apr 27;13:888025. doi: 10.3389/fgene.2022.888025
9. Niemi, M.E.K., Daly, M.J. & Ganna, A. The human genetic epidemiology of COVID-19. *Nat Rev Genet* 23, 533–546 (2022). <https://doi.org/10.1038/s41576-022-00478-5>
10. Butler-Laporte G, Povysil G, Kosmicki JA, Cirulli ET, Drivas T, et al. Exome-wide association study to identify rare variants influencing COVID-19 outcomes: Results from the Host Genetics Initiative. *PLoS Genet*. 2022 Nov 3;18(11):e1010367. doi: 10.1371/journal.pgen.1010367
11. Mieszek SP, Polymeropoulos VM, Xiao C, Polymeropoulos CM, Polymeropoulos MH. Loss-of-function mutations in IFNAR2 in COVID-19 severe infection susceptibility. *J Glob Antimicrob Resist*. 2021 Sep;26:239-240. doi: 10.1016/j.jgar.2021.06.005.
12. Dieter C, de Almeida Brondani L, Lemos NE, Schaeffer AF, Zanotto C, et al. Polymorphisms in *ACE1*, *TM-PRSS2*, *IFIH1*, *IFNAR2*, and *TYK2* Genes Are Associated with Worse Clinical Outcomes in COVID-19. *Genes (Basel)*. 2022 Dec 22;14(1):29. doi: 10.3390/genes14010029.
13. Fricke-Galindo I, Martínez-Morales A, Chávez-Galán L, Ocaña-Guzmán R, Buendía-Roldán I. et al. IFNAR2 relevance in the clinical outcome of individuals with severe COVID-19. *Front Immunol*. 2022 Jul 29;13:949413. doi: 10.3389/fimmu.2022.949413.
14. Abdelhafez M, Nasereddin A, Shamma OA, Abed R, Sinnokrot R. et al. Association of *IFNAR2* rs2236757 and *OAS3* rs10735079 Polymorphisms with Susceptibility to COVID-19 Infection and Severity in Palestine. *Interdiscip Perspect Infect Dis*. 2023 Sep 16;2023:9551163. doi: 10.1155/2023/9551163
15. Апалько, С.В. Поиск генных вариантов, влияющих на тяжесть течения COVID-19, на основе результатов секвенирования клинического экзоста / С.В. Апалько [и др.] // Гены и клетки. — 2024. — Т. 19, № 2. — С. 245–254. — DOI: <https://doi.org/10.17816/gc624810>
16. Barbitoff YA, Khmelkova DN, Pomerantseva EA, Slepchenkov AV, Zubashenko NA et al. Expanding the Russian allele frequency reference via cross-laboratory data integration: insights from 7452 exome samples. *Natl Sci Rev*. 2024 Sep 14;11(10):nwae326. doi: 10.1093/nsr/nwae326
17. Andrews S. (2010). FastQC: a quality control tool for high throughput sequence data. Электронный ресурс. URL: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>. Дата доступа: 10.11.2024
18. Poplin R., Chang P.-C., Alexander D., Schwartz S., Colthurst T. et al. A universal SNP and small-indel variant caller using deep neural networks // *Nature Biotechnology*. — 2018. — V. 36. — P. 983–987. doi: <https://doi.org/10.1038/nbt.4235>
19. Rappaport N., Nativ N., Stelzer G., et al. MalaCards: an integrated compendium for diseases and their annotation // *Database (Oxford)* — 2013. — V.:bat018. doi:10.1093/database/bat018
20. Azzarà A., Cassano I., Paccagnella E., et al. Genetic variants determine intrafamilial variability of SARS-CoV-2 clinical outcomes in 19 Italian families // *PLoS One*. -2022. — V.17. — P.e0275988. doi:10.1371/journal.pone.0275988.
21. Severe Covid-19 GWAS Group; Ellinghaus D, Degenhardt F, Bujanda L, Buti M, et al. Genomewide Association Study of Severe Covid-19 with Respiratory Failure // *N. Engl. J. Med*. — 2020. — V. 383. — P.1522-1534. doi: 10.1056/NEJMoa2020283.
22. Agrawal A., Balci H., Hanspers K., Coort S.L., Martens M. et al. WikiPathways 2024: next generation pathway database // *Nucleic Acids Research*. — 2024. — V.52. — P. D679–D689. <https://doi.org/10.1093/nar/gkad960>
23. Shannon P., Markiel A., Ozier O., Baliga N. S., Wang J. T., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Research*, 13(11), 2498–2504.
24. Garvin M.R., Alvarez C., Miller J., Prates E.T., Walker A.M., et al. A mechanistic model and therapeutic interventions for COVID-19 involving a RAS-mediated bradykinin storm // *Elife*. — 2020. — V.9. — P. e59177. doi: 10.7554/eLife.59177.
25. Bronk J.K., Kapadia C., Wu X., Chapman B.V., Wang R., et al. Feasibility of a novel non-invasive swab technique for serial whole-exome sequencing of cervical tumors during chemoradiation therapy // *PLoS One*. — 2022. — V. 17. — P.: e0274457. doi: 10.1371/journal.pone.0274457.
26. De Taeye B., Gills A., Declerck P.J. The story of the serpin plasminogen activator inhibitor 1: is there any need for another

mutant? // *Thromb. Haemost.* — 2004. — V. 92. — P. 898-924. doi: 10.1160/TH04-05-0269

27. Landrum M.J., Lee J.M., Riley G.R., Jang W., Rubinstein W.S et al. ClinVar: public archive of relationships among sequence variation and human phenotype // *Nucleic Acids Res.* — 2014. — V. 42. — P. D980-5. doi: 10.1093/nar/gkt111.

28. Verma A., Minnier J., Wan E.S., Huffman J.E., Gao L., Joseph J., et al. A MUC5B Gene Polymorphism, rs35705950-T, Confers Protective Effects Against COVID-19 Hospitalization but Not Severe Disease or Mortality // *Am J Respir Crit Care Med.* — 2022. — V. 206. — P.1220-1229. doi: 10.1164/rccm.202109-2166OC

29. García-Nieto P.E., Morrison A.J., Fraser, H.B. The somatic mutation landscape of the human body // *Genome Biol.* — 2019. — V.20. — P. 298. <https://doi.org/10.1186/s13059-019-1919-5>.

30. Jiao P., Fan W., Ma X., Lin R., Zhao Y., et al. SARS-CoV-2 nonstructural protein 6 triggers endoplasmic reticulum stress-induced autophagy to degrade STING1. *Autophagy.* 2023 Dec;19(12):3113-3131. doi: 10.1080/15548627

31. Salamaikina S., Karnaushkina M., Korchagin V., Litvinova M., Mironov K., Akimkin V. TLRs Gene Polymorphisms Associated with Pneumonia before and during COVID-19 Pandemic // *Diagnostics (Basel).* — 2022. — V. 13. — P. 121. doi: 10.3390/diagnostics13010121

References

1. The Human Genome Project. Nature milestones. 2021. <https://www.nature.com/articles/d42859-020-00101-9>

2. Brito AF, Semenova E, Dudas G, Hassler GW, Kalinich CC, et al. Global disparities in SARS-CoV-2 genomic surveillance. *Nat Commun.* 2022 Nov 16;13(1):7003. doi: 10.1038/s41467-022-33713-y.

3. The COVID-19 Host Genetics Initiative. The COVID-19 Host Genetics Initiative, a global initiative to elucidate the role of host genetic factors in susceptibility and severity of the SARS-CoV-2 virus pandemic. *Eu. J. Hum. Genet.* 28, 715–718 (2020).

4. COVID-19 Host Genetics Initiative. Mapping the human genetic architecture of COVID-19. *Nature* 600, 472–477 (2021). <https://doi.org/10.1038/s41586-021-03767-x>

5. COVID-19 Host Genetics Initiative. A first update on mapping the human genetic architecture of COVID-19. *Nature* 608, E1–E10 (2022). <https://doi.org/10.1038/s41586-022-04826-7>

6. Kousathanas, A., Pairo-Castineira, E., Rawlik, K. et al. Whole-genome sequencing reveals host factors underlying critical COVID-19. *Nature* 607, 97–103 (2022). <https://doi.org/10.1038/s41586-022-04576-6>

7. Pan D, Sze S, Minhas JS, Bangash MN, Pareek N, Divall P, Williams CM, Oggioni MR, Squire IB, Nellums LB, Hanif W, Khunti K, Pareek M. The impact of ethnicity on clinical outcomes in COVID-19: A systematic review. *EClinicalMedicine.* 2020 Jun 3;23:100404. doi: 10.1016/j.eclinm.2020.100404

8. Vadgama N, Kreymerman A, Campbell J, Shamardina O, Brugger C, Research Consortium GE, Deaconescu AM, Lee RT, Penkett CJ, Gifford CA, Mercola M, Nasir J, Karakikes I. SARS-CoV-2 Susceptibility and ACE2 Gene Variations Within Diverse Ethnic Backgrounds. *Front Genet.* 2022 Apr 27;13:888025. doi: 10.3389/fgene.2022.888025

9. Niemi, M.E.K., Daly, M.J. & Ganna, A. The human genetic epidemiology of COVID-19. *Nat Rev Genet* 23, 533–546 (2022). <https://doi.org/10.1038/s41576-022-00478-5>

10. Butler-Laporte G, Povysil G, Kosmicki JA, Cirulli ET, Drivas T, et al. Exome-wide association study to identify rare variants influencing COVID-19 outcomes: Results from the Host Genetics Initiative. *PLoS Genet.* 2022 Nov 3;18(11):e1010367. doi: 10.1371/journal.pgen.1010367

11. Mieszek SP, Polymeropoulos VM, Xiao C, Polymeropoulos CM, Polymeropoulos MH. Loss-of-function mutations in IFNAR2 in COVID-19 severe infection susceptibility. *J Glob Antimicrob Resist.* 2021 Sep;26:239-240. doi: 10.1016/j.jgar.2021.06.005.

12. Dieter C, de Almeida Brondani L, Lemos NE, Schaeffer AF, Zanotto C, et al. Polymorphisms in ACE1, TM-PRSS2, IFIH1, IFNAR2, and TYK2 Genes Are Associated with Worse Clinical Outcomes in COVID-19. *Genes (Basel).* 2022 Dec 22;14(1):29. doi: 10.3390/genes14010029.

13. Fricke-Galindo I, Martínez-Morales A, Chávez-Galán L, Ocaña-Guzmán R, Buendía-Roldán I. et al. IFNAR2 relevance in the clinical outcome of individuals with severe COVID-19. *Front Immunol.* 2022 Jul 29;13:949413. doi: 10.3389/fimmu.2022.949413.

14. Abdelhafez M, Nasereddin A, Shamma OA, Abed R, Sinnokrot R. et al. Association of IFNAR2 rs2236757 and OAS3 rs10735079 Polymorphisms with Susceptibility to COVID-19 Infection and Severity in Palestine. *Interdiscip Perspect Infect Dis.* 2023 Sep 16;2023:9551163. doi: 10.1155/2023/9551163

15. Apal'ko S.V., Nostaeva A.V., Shimanskij V.S., Sushenceva N.N., Popov O.S. i dr. Poisk gennyh variantov, vliyayushchih na tyazhest' techeniya COVID-19, na osnove rezul'tatov sekvenirovaniya klinicheskogo ekzoma // *Geny i kletki.* 2024. T. 19, № 2. S. 245–254. DOI: <https://doi.org/10.17816/gc624810>

16. Barbitoff YA, Khmelkova DN, Pomerantseva EA, Slepchenkov AV, Zubashenko NA et al. Expanding the Russian allele frequency reference via cross-laboratory data integration: insights from 7452 exome samples. *Natl Sci Rev.* 2024 Sep 14;11(10):nwae326. doi: 10.1093/nsr/nwae326

17. Andrews S. (2010). FastQC: a quality control tool for high throughput sequence data. Электронный ресурс. URL: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>. Дата доступа: 10.11.2024

18. Poplin R., ChangP.-C., Alexander D., Schwartz S., Colthurst T. et al. A universal SNP and small-indel variant caller using deep neural networks // *Nature Biotechnology.* — 2018. — V. 36. — P. 983–987. doi: <https://doi.org/10.1038/nbt.4235>

19. Rappaport N., Nativ N., Stelzer G., et al. MalaCards: an integrated compendium for diseases and their annotation // *Database (Oxford)* — 2013. — V.:bat018. doi:10.1093/database/bat018

20. Azzarà A., Cassano I., Paccagnella E., et al. Genetic variants determine intrafamilial variability of SARS-CoV-2 clinical outcomes in 19 Italian families // *PLoS One.* -2022. — V.17. — P.e0275988. doi:10.1371/journal.pone.0275988.

21. Severe Covid-19 GWAS Group; Ellinghaus D, Degenhardt F, Bujanda L, Buti M, et al. Genomewide Association Study of Severe Covid-19 with Respiratory Failure // *N. Engl. J. Med.* — 2020. — V. 383. — P.1522-1534. doi: 10.1056/NEJMoa2020283.

22. Agrawal A., Balci H., Hanspers K., Coort S.L., Martens M. et al. WikiPathways 2024: next generation pathway database // *Nucleic Acids Research.* — 2024. — V.52. — P. D679–D689. <https://doi.org/10.1093/nar/gkad960>

23. Shannon P., Markiel A., Ozier O., Baliga N. S., Wang J. T., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Research*, 13(11), 2498–2504.

24. Garvin M.R., Alvarez C., Miller J., Prates E.T., Walker A.M., et al. A mechanistic model and therapeutic interventions for COVID-19 involving a RAS-mediated bradykinin storm // *Elife.* — 2020. — V.9. — P. e59177. doi: 10.7554/eLife.59177.

25. Bronk J.K., Kapadia C., Wu X., Chapman B.V., Wang R., et al. Feasibility of a novel non-invasive swab technique for seri-

al whole-exome sequencing of cervical tumors during chemoradiation therapy // PLoS One. — 2022. — V. 17. — P.: e0274457. doi: 10.1371/journal.pone.0274457.

26. De Taeye B., Gils A., Declerck P.J. The story of the serpin plasminogen activator inhibitor 1: is there any need for another mutant? // Thromb. Haemost. — 2004. — V. 92. — P. 898-924. doi: 10.1160/TH04-05-0269

27. Landrum M.J., Lee J.M., Riley G.R., Jang W., Rubinstein W.S et al. ClinVar: public archive of relationships among sequence variation and human phenotype // Nucleic Acids Res. — 2014. — V. 42. — P. D980-5. doi: 10.1093/nar/gkt111.

28. Verma A., Minnier J., Wan E.S., Huffman J.E., Gao L., Joseph J., et al. A MUC5B Gene Polymorphism, rs35705950-T, Confers Protective Effects Against COVID-19 Hospitalization but Not Severe Disease or Mortality // Am J Respir Crit

Care Med. — 2022. — V. 206. — P.1220-1229. doi: 10.1164/rccm.202109-2166OC

29. García-Nieto P.E., Morrison A.J., Fraser, H.B. The somatic mutation landscape of the human body // Genome Biol. — 2019. — V.20. — P. 298. <https://doi.org/10.1186/s13059-019-1919-5>.

30. Jiao P., Fan W., Ma X., Lin R., Zhao Y., et al. SARS-CoV-2 nonstructural protein 6 triggers endoplasmic reticulum stress-induced autophagy to degrade STING1. Autophagy. 2023 Dec;19(12):3113-3131. doi: 10.1080/15548627

31. Salamaikina S., Karnaushkina M., Korchagin V., Litvinova M., Mironov K, Akimkin V. TLRs Gene Polymorphisms Associated with Pneumonia before and during COVID-19 Pandemic // Diagnostics (Basel). — 2022. — V. 13. — P. 121. doi: 10.3390/diagnostics13010121

Авторский коллектив:

Переверий Александр Александрович — лаборант-исследователь лаборатории молекулярной вирусологии Научно-исследовательского института гриппа им. А.А. Смородинцева; тел.: 8(812)499-15-20, e-mail: gilagalex@gmail.com

Комиссаров Андрей Борисович — заведующий лабораторией молекулярной вирусологии Научно-исследовательского института гриппа им. А.А. Смородинцева; тел.: 8(812)499-15-20, e-mail: andrey.komissarov@influenza.spb.ru

Даниленко Дарья Михайловна — заместитель директора по научной работе, руководитель отдела этиологии и эпидемиологии Научно-исследовательского института гриппа им. А.А. Смородинцева, к.б.н.; тел.: 8(812)499-15-25, e-mail: daria.danilenko@influenza.spb.ru

Венев Евгений Валентинович — старший преподаватель Научно-исследовательского института гриппа им. А.А. Смородинцева; врач-инфекционист Клинической инфекционной больницы им. С.П. Боткина; преподаватель кафедры инфекционных болезней и эпидемиологии Первого Санкт-Петербургского государственного медицинского университета им. академика И.П. Павлова; тел.: 8(812)499-15-13

Коноплева Софья Андреевна — студент биологического факультета Санкт-Петербургского государственного университета; тел 8(812) 363-60-00, e-mail: s.konopleva.s@mail.ru

Коржанова Маргарита — лаборант-исследователь лаборатории молекулярной вирусологии Научно-исследовательского института гриппа им. А.А. Смородинцева; тел.: 8(812)499-15-20, e-mail: marory.kmm@gmail.com

Писарева Мария Михайловна — ведущий научный сотрудник лаборатории молекулярной вирусологии Научно-исследовательского института гриппа им. А.А. Смородинцева, к.б.н.; тел.: 8(812)499-15-20, e-mail: maria.pisareva@influenza.spb.ru

Гусев Денис Александрович — главный врач Клинической инфекционной больницы им. С.П. Боткина, д.м.н., профессор; тел. 7(812) 670-78-78, e-mail: gusevden-70@mail.ru

Люзнов Дмитрий Анатольевич — директор Научно-исследовательского института гриппа им. А.А. Смородинцева; заведующий кафедрой инфекционных болезней и эпидемиологии Первого Санкт-Петербургского государственного медицинского университета им. академика И.П. Павлова, д.м.н., профессор; тел.: 8(812)499-15-00, e-mail: office@influenza.spb.ru