

*На правах рукописи*

**ШКУРНИКОВ**

**Максим Юрьевич**

**РОЛЬ ГЕНОТИПА ГЛАВНОГО КОМПЛЕКСА  
ГИСТОСОВМЕСТИМОСТИ КЛАССА 1 И ПРОФИЛЯ МИКРОРНК В  
ПАТОГЕНЕЗЕ ТЯЖЕЛОЙ И КРАЙНЕ ТЯЖЕЛОЙ ФОРМ COVID-19**

3.3.3. Патологическая физиология

**АВТОРЕФЕРАТ**

диссертации на соискание ученой степени

доктора медицинских наук

Иркутск – 2023

Работа выполнена в Федеральном государственном автономном образовательном учреждении высшего образования «Национальный исследовательский университет «Высшая школа экономики» и Федеральном государственном бюджетном научном учреждении «Научный центр проблем здоровья семьи и репродукции человека».

**Научные консультанты:** **Колесников Сергей Иванович**  
доктор медицинских наук, профессор,  
академик РАН, Заслуженный деятель науки РФ

**Тоневицкий Александр Григорьевич**  
доктор биологических наук, профессор,  
член-корреспондент РАН, Заслуженный деятель науки РФ

**Официальные оппоненты:**

**Мальшев Игорь Юрьевич** – доктор медицинских наук, профессор, ФГБОУ ВО «Московский государственный медико-стоматологический университет имени А. И. Евдокимова» Минздрава России, кафедра патологической физиологии, заведующий.

**Попугаев Константин Александрович** – доктор медицинских наук, профессор, ФГАУ «Национальный медицинский исследовательский центр здоровья детей» Минздрава России, заместитель директора по научной работе.

**Яблонский Пётр Казимирович** – доктор медицинских наук, профессор, заслуженный врач РФ, ФГБУ «Санкт-Петербургский научно-исследовательский институт фтизиопульмонологии» Минздрава России, директор.

**Ведущая организация:** Федеральное государственное автономное образовательное учреждение высшего образования «Российский Национальный Исследовательский Медицинский Университет им. Н.И. Пирогова» Министерства здравоохранения Российской Федерации.

Защита состоится «\_\_» \_\_\_\_\_ 202\_\_ года в \_\_\_\_ часов на заседании диссертационного совета 24.1.187.01 на базе Федерального государственного бюджетного научного учреждения «Научный центр проблем здоровья семьи и репродукции человека» по адресу: 664003, г. Иркутск, ул. Тимирязева, 16.

С диссертацией можно ознакомиться в библиотеке Федерального государственного бюджетного научного учреждения «Научный центр проблем здоровья семьи и репродукции человека» по адресу: 664003, г. Иркутск, ул. Тимирязева, 16 и на сайте <http://health-family.ru>.

Автореферат разослан «\_\_» \_\_\_\_\_ 2023 г.

Ученый секретарь  
диссертационного совета,  
доктор биологических наук

**Гребенкина Людмила Анатольевна**

## ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

### Актуальность темы исследования

SARS-CoV-2 представляет собой одноцепочечный РНК-вирус с положительной цепью, который вызывает у людей вирусную пневмонию, сопряженную с тяжелым респираторным синдромом [Kim и др., 2020]. По данным ВОЗ эпидемия коронавирусной инфекции (COVID-19), вызванная этим вирусом, превратилась в пандемию и унесла более 6,9 миллиона жизней во всем мире в период с декабря 2019 года по май 2023 года. И несмотря на то, что ВОЗ объявила об окончании пандемии, вспышки заболевания и смертность от него продолжают наблюдаться по всему миру, что не снижает актуальность исследования этого заболевания.

В ряде исследований была продемонстрирована взаимосвязь тяжести течения COVID-19 с генотипом главного комплекса гистосовместимости класса I (ГКГС-I или HLA-I). Прежде всего была выявлена взаимосвязь частоты встречаемости аллелей HLA-A\*01:01 и HLA-A\*02:01 с числом случаев заражения и уровнем смертности от COVID-19 в различных регионах Италии [Pisanti и др., 2020]. Также на выборке пациентов с различной тяжестью течения COVID-19 [Iturrieta-Zuazo и др., 2020] были подтверждены результаты теоретического моделирования взаимодействия пептидов SARS-CoV-2 с различными аллельными вариантами молекул ГКГС-I. Была продемонстрирована взаимосвязь тяжести COVID-19 и числа вирусных пептидов с высокой аффинностью к индивидуальному набору молекул ГКГС-I: чем больше таких пептидов связываются с ГКГС-I, тем легче течение заболевания. Позднее нами была показана взаимосвязь генотипа ГКГС-I пациента с риском смерти от COVID-19 в возрасте до 60 лет [Shkurnikov и др., 2021].

Молекулы ГКГС-I являются ключевыми медиаторами первых звеньев развития специфического иммунного ответа на заражение SARS-CoV-2. После попадания в клетку SARS-CoV-2 индуцирует трансляцию своих белков, часть из которых попадает в протеасомы инфицированной клетки, расщепляется до пептидов длиной 8-12 аминокислотных остатков и связывается с молекулами ГКГС-I. После связывания комплекс из молекулы ГКГС-I и пептида вируса переносится на поверхность зараженной клетки, где может быть распознан T-клеточным рецептором CD8<sup>+</sup> T-лимфоцитов. В результате зрелый CD8<sup>+</sup> T-лимфоцит активируется и уничтожает зараженную клетку с помощью перфоринов и сериновых протеаз [Wherry, Ahmed, 2004], прерывая размножение вируса.

В организме человека присутствует три основных вида молекул ГКГС-I, кодируемых генами HLA-A, HLA-B, HLA-C. Каждый из этих генов у отдельного человека может быть представлен в двух вариантах – аллелях, наследуемых от родителей. Таким образом генотип ГКГС-I у отдельного человека может включать от трех до шести аллельных вариантов. В человеческой популяции существуют сотни аллельных вариантов генов HLA-A, HLA-B, HLA-C. Каждый из них кодирует молекулу ГКГС-I, обладающую индивидуальной способностью к узнаванию различных чужеродных белков. Распределение аллелей специфично для разных стран и популяций [Wang и др., 2009].

От индивидуальной комбинации молекул ГКГС-I зависит тяжесть течения ряда инфекционных заболеваний (малярия, туберкулез, СПИД, вирусные гепатиты) [Wang и др., 2009]. Продемонстрирована связь между генотипом ГКГС-I и чувствительностью к SARS-CoV, вирусу возбудителю первого тяжелого острого респираторного синдрома. Так аллели HLA-B\*07:03 [Ng и др., 2004], HLA-B\*46:01 [Lin и др., 2003], HLA-C\*08:01 [Chen и др., 2006] являются факторами предрасположенности к развитию заболевания в тяжелой форме. В свою очередь аллель HLA-C\*15:02 ассоциирован с легким течением заболевания [Wang и др., 2011].

За более чем три года пандемии COVID-19 накоплена значительная информация о фактических эпитопах различных вариантов SARS-CoV-2 [Vita и др., 2019], особенностях формирования Т-клеточной памяти [Francis и др., 2022], тенденциях в частоте мутаций в вирусе [Nersisyan и др., 2022a]. Однако взаимосвязь генотипа ГКГС-I и течения COVID-19 преимущественно анализировали по данным первой волны пандемии, вызванной исходным вариантом вируса [Langton и др., 2021; Lorente и др., 2021; Tavasolian и др., 2021; Venet и др., 2022; Novhannisyanyan и др., 2022].

Более того, при исследованиях ассоциаций генотипа и тяжести течения COVID-19 практически не учитывался возраст переболевших, сопутствующие заболевания. Между тем, возраст является значимым фактором в формировании иммунного ответа к COVID-19 [McGroder и др., 2021; Promislow, 2020; Sanchez-Vazquez и др., 2021]. В частности показано, что у людей старше 60 лет в значительной степени падает длина теломер наивных Т-лимфоцитов, что приводит практически к десятикратному падению их способности к делению при активации [Anderson и др., 2022], сокращается репертуар Т-клеточных рецепторов [Britanova и др., 2014].

Всё это обуславливает необходимость тщательного изучения взаимосвязи генотипа ГКГС-I с тяжестью COVID-19, особенностями формирования иммунитета к вирусу, эволюцией SARS-CoV-2, влияющей на его иммуногенные эпитопы. Пробелы в изучении вклада генотипа ГКГС-I в патогенез COVID-19,

отсутствие обобщенного отечественного и зарубежного опыта по использованию данных генотипирования в клинической практике, высокая практическая востребованность новых разработок в представленной области в условиях продолжающегося персистирования вируса SARS-CoV-2 в человеческой популяции, делают крайне актуальным и своевременным проведение настоящего исследования.

### **Степень разработанности темы исследования**

В ходе пандемии COVID-19 беспрецедентные усилия были направлены на изучение вируса SARS-CoV-2 и патогенеза COVID-19, молекулярно-биологических механизмов формирования иммунного ответа при данном заболевании, факторов риска тяжелого течения заболевания. В частности, предпринимались многочисленные попытки выявить генетические факторы риска заражения SARS-CoV-2 и тяжелого течения COVID-19, но таких клинически значимых факторов не найдено.

В то же время анализ отечественной и зарубежной литературы показал, что от индивидуальной комбинации молекул ГКГС-I зависит тяжесть течения ряда инфекционных заболеваний, в том числе продемонстрирована связь между генотипом ГКГС-I и чувствительностью к SARS-CoV.

Аналогичные наблюдения были опубликованы и для SARS-CoV-2, но большинство исследователей анализировали частоты аллелей ГКГС-I и практически не уделяли внимания особенностям варианта SARS-CoV-2, вызвавшего вирусную пневмонию, а также способности индивидуального набора молекул ГКГС-I пациентов представлять пептиды вируса на поверхности зараженной клетки, влиянию мутаций вируса на эффективность врожденного и приобретенного иммунитета, вкладу этих процессов в патогенез COVID-19.

Практически не исследована роль микроРНК как молекул, влияющих на вирус SARS-CoV-2 и участвующих в патогенезе COVID-19 на уровне целостного организма.

С этих позиций рассматриваемая в диссертации научная проблема является недостаточно проработанной и слабо изученной.

### **Цель исследования**

Установить роль генотипа главного комплекса гистосовместимости класса I и профиля микроРНК пациентов в патогенезе тяжелой и крайне тяжелой форм новой коронавирусной инфекции COVID-19 для оптимизации оценки тяжести и прогноза течения заболевания.

### **Задачи исследования**

1) Выявить взаимосвязь частоты генотипов главного комплекса гистосовместимости класса I с особенностями течения заболевания у пациентов с вирусной пневмонией, вызванной SARS-CoV-2 в период 2020 – 2021 гг.

2) Разработать способ оценки риска развития тяжелой и крайне тяжелой форм COVID-19 на основе анализа генотипа главного комплекса гистосовместимости класса 1 у пациентов с вирусной пневмонией, вызванной SARS-CoV-2.

3) Определить возможное влияние микроРНК ткани лёгких на эволюцию генома SARS-CoV-2.

4) Выявить взаимосвязь тяжести и исхода заболевания COVID-19 с профилем циркулирующих в плазме крови пациентов микроРНК.

5) Оценить у пациентов, переболевших COVID-19, зависимость Т-клеточного ответа на пептиды SARS-CoV-2 от генотипа главного комплекса гистосовместимости класса 1.

6) Разработать патогенетически обоснованный алгоритм диагностических мероприятий для прогноза тяжелого и крайне тяжелого течения COVID-19 на основе анализа генотипа комплекса гистосовместимости класса 1.

### **Научная новизна исследования**

Впервые установлено, что у пациентов в возрасте до 60 лет с вирусной пневмонией, вызванной двумя наиболее распространенными вариантами штамма Дельта SARS-CoV-2, в случае инфицирования вариантом AY.122 число высокоаффинных для их индивидуального набора молекул ГКГС-I пептидов вируса SARS-CoV-2 значительно сократилось по сравнению с инфицированием вариантом B.1.617.2. Это прежде всего связано с мутацией G8R в белке NS8.

Впервые продемонстрировано, что среди госпитализированных в третью волну пандемии COVID-19 пациентов с вирусной пневмонией, вызванной SARS-CoV-2, значительно снижено число носителей аллеля HLA-A\*01:01 ГКГС-I. Установлено, что носители аллеля HLA-A\*01:01 формируют Т-клеточный иммунитет к пептидам, кодируемым геном ORF1ab вируса SARS-CoV-2. Гены SARS-CoV-2 мутируют с разной скоростью. Ген ORF1ab наиболее консервативен среди них, а значит носители аллеля HLA-A\*01:01 могут длительное время сохранять Т-клеточный иммунитет.

Продemonстрирована более высокая частота встречаемости иммунодоминантных эпитопов из белков гена ORF1ab вируса SARS-CoV-2 у носителей HLA-A\*01:01 по сравнению с эпитопами из белков этого гена у носителей HLA-A\*02:01 в когорте выздоровевших пациентов первой волны COVID-19.

Впервые на основании исследования взаимосвязи генотипа ГКГС-I с тяжестью течения заболевания у пациентов с вирусной пневмонией, вызванной SARS-CoV-2, продемонстрирована взаимосвязь вероятности летального исхода заболевания и числа пептидов вируса SARS-CoV-2, высокоаффинных для индивидуального набора молекул ГКГС-I пациента.

Впервые изучено возможное влияние микроРНК ткани лёгких на эволюцию генома вируса SARS-CoV-2. Показано, что 5p- и 3p-нетранслируемые области вируса SARS-CoV-2 практически не имеют регионов связывания с характерными для ткани лёгких микроРНК, но вирус обладает значительным числом мест связывания с микроРНК в регионе NSP3-NSP5, ответственном за аутопротеолиз вирусных полипептидов и формирование вирионов. В вариантах штамма Омикрон произошло значимое снижение числа мест связывания с микроРНК клеток хозяина, что могло способствовать снижению патогенности данного штамма.

Впервые выявлен механизм регуляции экспрессии генов ACE2 и TMPRSS2 человека молекулами микроРНК. Показано, что лизин-специфическая деметилаза 5B (JARID1B), кодируемая геном KDM5B, может опосредованно влиять на экспрессию ACE2/TMPRSS2 путем репрессии транскрипции hsa-let-7e/hsa-miR-125a и hsa-miR-141/hsa-miR-200. Впервые продемонстрирована взаимосвязь уровня циркулирующих в плазме крови микроРНК hsa-miR-25-3p и hsa-miR-19b-3p с крайне тяжелым течением COVID-19.

Впервые показано, что мутации вируса SARS-CoV-2 практически не затрагивают иммунодоминантные для российской популяции пептиды вируса SARS-CoV-2.

Впервые разработан и апробирован патогенетически обоснованный алгоритм диагностических мероприятий для прогноза тяжелого и крайне тяжелого течения COVID-19 на основе анализа генотипа ГКГС-I пациента. Если значение «Индекса риска» более 89, пациент относится к группе с высоким риском тяжелого течения COVID-19.

### **Теоретическая и практическая значимость работы**

Полученные в ходе исследования новые научные данные вносят вклад в углубление фундаментальных знаний о патогенезе новой коронавирусной инфекции COVID-19 с позиции обоснования роли ГКГС-I и микроРНК в формировании специфического иммунитета к вирусу SARS-CoV-2, а также развитию и стойкости постинфекционного иммунитета.

Полученные результаты по зависимости Т-клеточного ответа на вирус SARS-CoV-2 от генотипа ГКГС-I могут стать основой для создания высокоэффективных вакцин на основе пептидов ORF1ab.

Практическая значимость заключается в том, что разработанные в процессе исследования и научно обоснованные патофизиологические подходы к анализу индивидуальных особенностей генотипа ГКГС-I пациентов позволят совершенствовать прогноз развития тяжелых и крайне тяжелых форм COVID-19, а значит более эффективно оказывать медицинскую помощь в условиях продолжающейся эпидемии и в периоды ее повторного возникновения.

Разработанные в исследовании предложения и практические рекомендации могут быть внедрены в деятельность медицинских организаций, оказывающих медицинскую помощь в стационарных условиях при новой коронавирусной инфекции COVID-19.

### **Методология и методы исследования**

В исследовании были использованы данные о генотипе добровольцев из Федерального регистра доноров костного мозга (РНИМУ им. Н.И. Пирогова). Группа пациентов первой волны была сформирована в период с мая по август 2020 года. Группа пациентов третьей волны была сформирована в период с июня по июль 2021 года. Все пациенты, участвовавшие в исследовании, были госпитализированы с COVID-19 в ГБУЗ «ГКБ № 15 ДЗМ». ДНК из образцов крови выделялась набором реагентов QIAamp DNA Mini Kit (QIAGEN, Германия). Подготовка библиотек для секвенирования генов ГКГС осуществлялась с помощью набора реагентов для подготовки библиотек фрагментов ДНК генов ГКГС I и II классов для генотипирования высокопроизводительным секвенированием (NGS) «HLA-Эксперт» (ДНК-Технология, Россия). Секвенирование библиотек фрагментов ДНК генов ГКГС I и II классов осуществлялось с помощью секвенатора Illumina MiSeq (Illumina, США) на стандартной проточной ячейке набора реагентов MiSeq Reagent Kit v2 500-cycles (Illumina, США) в режиме парных прочтений 2x250. Для исследования образцов на наличие РНК SARS-CoV-2 применяли различные наборы реагентов для ОТ-ПЦР и комплекты к ним наборы для выделения РНК: ПОЛИВИР SARS-CoV-2 Express (ЛИТЕХ, Россия), АмплиПрайм SARS-CoV-2 DUO / МагноПрайм ФАСТ-Р (НекстБио, Россия) и SARS-CoV-2 «CoV-2-Тест» (ТестГен, Россия). После подготовки библиотек фрагментов вируса SARS-CoV-2 и их секвенирования на секвенаторе Illumina MiSeq с использованием набора реактивов MiSeq Reagent Kit v3 600-cycles, запускаемого в режиме чтения 2x250, последовательности аннотировались с применением баз данных Rango-lineage и NextStain. Выделение мононуклеаров периферической крови осуществляли в градиенте плотности фикола (ПанЭко, Россия). Пептиды с чистотой не менее 95% были синтезированы Peptide 2.0 Inc. или в ФГБУ науки Институт биоорганической химии им. академиков М.М. Шемякина и Ю.А. Овчинникова Российской академии наук. Протеомы вариантов SARS-CoV-2 были получены из базы данных GISAID. В качестве эталонного вируса был использован вариант Wuhan-Hu-1 (EPI\_ISL\_402125). Прогноз аффинности взаимодействия молекула ГКГС-I – пептид осуществляли с помощью netMHCpan версии 4.1. Для статистического анализа использовались следующие функции из библиотеки stats в R: fisher.test для точного критерия Фишера, wilcox.test для U-критерия Манна-Уитни. Кроме того, использовалась поправка на множественность



сравнений Бенджамини-Хохберга. Графики были построены с помощью библиотек ComplexHeatmap, pheatmap и ROCit.

### **Положения, выносимые на защиту**

1) Тяжесть течения заболевания пациентов с вирусной пневмонией, вызванной SARS-CoV-2, связана с генотипом ГКГС-I: чем больше вирусных пептидов с высокой аффинностью связываются с ГКГС-I, тем легче течение заболевания.

2) Среди госпитализированных в третью волну пандемии COVID-19 пациентов с вирусной пневмонией, вызванной SARS-CoV-2, значимо снижено число носителей аллеля HLA-A\*01:01 ГКГС-I. Носители аллеля HLA-A\*01:01 формируют Т-клеточный иммунитет к пептидам, кодируемым геном ORF1ab вируса SARS-CoV-2. Гены SARS-CoV-2 мутируют с разной скоростью. Ген ORF1ab наиболее консервативен среди них, а значит носители аллеля HLA-A\*01:01 могут длительное время сохранять Т-клеточный иммунитет.

3) МикроРНК ткани лёгких могут влиять на эволюцию SARS-CoV-2, поскольку вирус обладает значительным числом мест связывания с микроРНК в регионе NSP3-NSP5, ответственном за аутопротеолиз вирусных полипептидов и формирование вирионов. Наблюдается значимая тенденция к уменьшению числа мест связывания с микроРНК в геноме вариантов вируса разных волн.

4) Экспрессия генов ACE2 и TMPRSS2 человека регулируется молекулами микроРНК через репрессию транскрипции hsa-let-7e/hsa-miR-125a и hsa-miR-141/hsa-miR-200 лизин-специфической деметилазой 5B. Высокий уровень циркулирующих микроРНК hsa-miR-25-3p и hsa-miR-19b-3p в плазме крови ассоциирован с благоприятным исходом крайне тяжелой формы COVID-19.

### **Степень достоверности и апробация результатов**

Диссертация выполнена в лаборатории исследований молекулярных механизмов долголетия Национального исследовательского университета «Высшая школа экономики», а также в отделе персонализированной и профилактической медицины Федерального государственного бюджетного научного учреждения «Научный центр проблем здоровья семьи и репродукции человека» и является частью тем государственных заданий № 11816п-П8, № 75-00672-22-ПР, а также гранта Минобрнауки России № 075-15-2021-1049.

Достоверность и объективность полученных результатов подтверждается репрезентативностью выборочных совокупностей объектов исследования, достаточным объемом наблюдений, а также использованием генеральных совокупностей и адекватных методов исследования. За основу статистических материалов взят обоснованный массив фактических данных с применением современных подходов к обработке исходной информации. Статистическая

обработка материалов осуществлялась в среде для математического анализа R версий 3.6.3 – 4.2.2.

### **Личный вклад автора**

Автором работы проведен анализ отечественных, зарубежных научных источников литературы, официальных источников информации. Диссертантом самостоятельно разработана методика настоящего исследования. Организован и осуществлен сбор первичного материала. Автор принимал непосредственное участие в проведении экспериментов, анализе полученных данных и подготовке публикаций. Автором диссертационного исследования лично проводилась математико-статистическая обработка результатов работы. Автором проведена аналитическая работа с последующей интерпретацией полученных результатов исследования, сформулированы выводы и практические рекомендации.

### **Публикации**

По материалам диссертационного исследования были опубликованы 23 печатные работы, в том числе 18 статей в научных журналах и изданиях, входящих в международные базы данных и системы цитирования Scopus и Web of Science, три патента на изобретения, одни клинические рекомендации, одно свидетельство о регистрации программы для ЭВМ.

### **Структура и объём работы**

Диссертационная работа состоит из введения, обзора литературы, описания материалов и методов исследования, четырех глав собственных результатов и их обсуждения, выводов, заключения, списка используемых сокращений и списка использованной литературы, приложения. Работа изложена на 216 страницах и содержит 32 рисунка, 20 таблиц и 419 литературных ссылок.

### **Апробация работы**

Основные положения и результаты работы были доложены и обсуждены на научно-практической конференции «Инфекционные заболевания в XXI веке. Современные подходы к диагностике и лечению» (онлайн, 23 апреля 2021 года), Российском конгрессе по клинической фармакогеномике (г. Москва, 05 февраля 2022 года), научном семинаре с участием сотрудников Института молекулярной биологии НАН РА, Армянского института биоинформатики (АБИ) на базе Русско-Армянского университета (г. Ереван, Республика Армения, 20 апреля 2022 года), конгрессе с международным участием «Инновационная кардиология» (г. Минск, Республика Белоруссия, 21 октября 2022 года), II интернет-конференции «Покровские чтения» (онлайн, 01 ноября 2022 года), XI конгрессе «Национальной Ассоциации Фтизиатров» (г. Санкт-Петербург, 25 ноября 2022 года).

## СОДЕРЖАНИЕ РАБОТЫ

### Объекты и методы исследования

В настоящем диссертационном исследовании применен комплексный подход к решению проблемы стратификации пациентов по группам риска тяжести течения острой коронавирусной инфекции COVID-19 в зависимости от генетических особенностей пациента и вызвавшего заболевание варианта вируса SARS-CoV-2. Схема дизайна исследования представлена на Рисунке 1. Произведено подробное описание методических основ оценки тяжести течения COVID-19 в зависимости от генотипа ГКГС-I и варианта вируса SARS-CoV-2. Также в исследовании представлены результаты оценки диагностической значимости предложенного подхода, оценен вклад микроРНК в патогенез COVID-19.

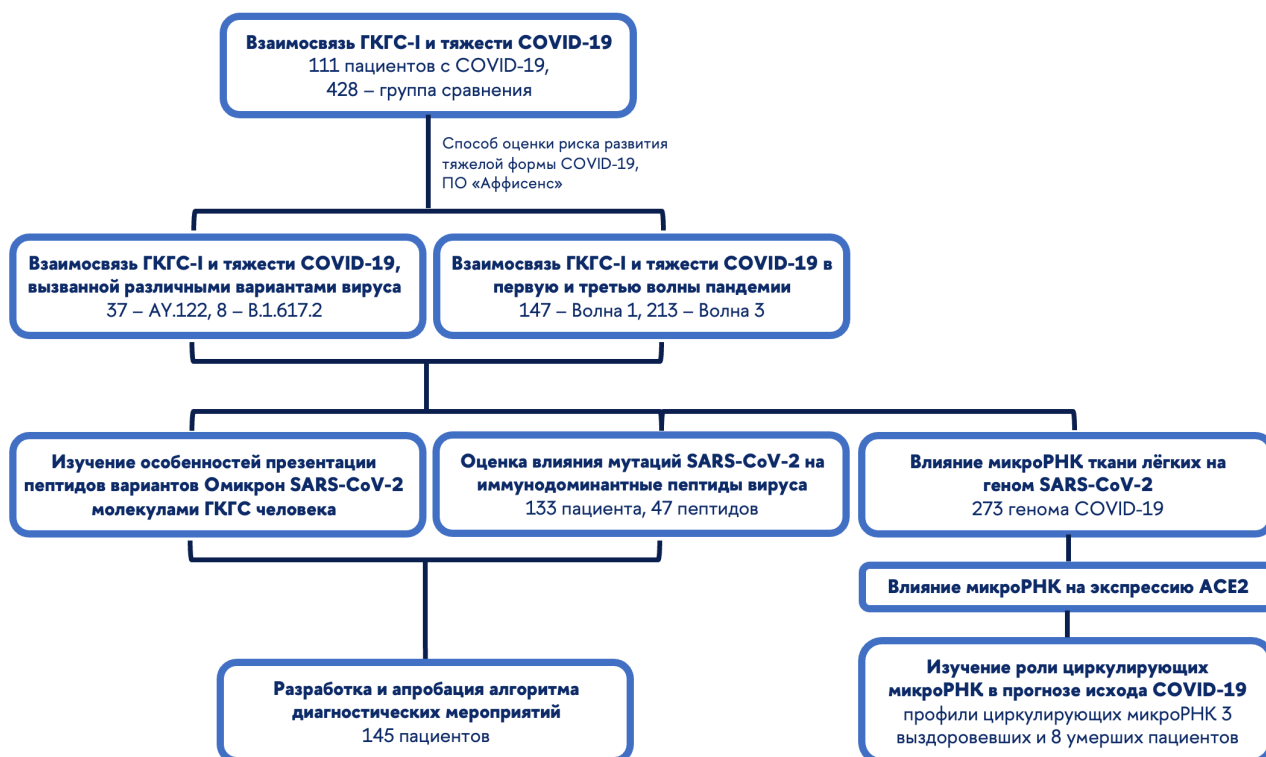


Рисунок 1 – Схема дизайна исследования

Период проведения исследования затронул первую, третью и четвертую волны пандемии, пришедшиеся на период с марта 2020 года по декабрь 2022 года. Для осуществления исследования потребовалось применение междисциплинарного подхода на базе патофизиологии с привлечением молекулярной биологии и математического анализа. В этой связи в рамках настоящей диссертации была разработана комплексная методика, позволившая изучить особенности течения COVID-19 в зависимости от генотипа ГКГС-I, выявить иммунодоминантные эпитопы вируса, оценить их стабильность,

продемонстрировать влияние микроРНК альвеоцитов человека на геном SARS-CoV-2, разработать программное обеспечение для прогноза тяжести течения COVID-19 и клинические рекомендации по его применению, достичь поставленной цели и решить задачи исследования. Программа исследования предусматривала выбор его объектов в зависимости от задач и этапов его проведения, обоснование единиц и объема необходимых наблюдений, разработку программ наблюдения и соответствующих им методов.

Объектом исследования на разных этапах явились: пациенты с новой коронавирусной инфекцией COVID-19 и её основным осложнением – вирусной пневмонией, вызванной SARS-CoV-2; их генотипы ГКГС-I; законченные случаи их лечения в условиях стационара; последовательности генома SARS-CoV-2; профиль микроРНК ткани лёгких и плазмы крови.

Предметом исследования явились: генотипы ГКГС-I и клиническая характеристика пациентов с вирусной пневмонией, вызванной SARS-CoV-2, и наличие у них конкурирующей патологии; исходы госпитализации при вирусной пневмонии, вызванной SARS-CoV-2; результаты компьютерного моделирования взаимодействия пептидов вируса SARS-CoV-2 с различными аллельными вариантами молекул ГКГС-I; интенсивность Т-клеточного ответа в ответ на стимуляцию различными пептидами SARS-CoV-2; геном различных вариантов вируса SARS-CoV-2; экспрессия микроРНК в ткани лёгких, представленность микроРНК в плазме крови.

В исследовании были использованы сплошной и выборочный способы проведения наблюдения. Предусматривалось использование методов статистического анализа, математического моделирования, клинического обследования, секвенирования следующего поколения, генотипирования ГКГС-I, оценки Т-клеточного ответа.

Обработка собранного первичного материала, а также математико-статистический анализ полученных результатов осуществлялись в среде для математического анализа R версий 3.6.3 – 4.2.2 с использованием следующих пакетов: readr, dplyr, reshape2, ggpubr, stringr, ComplexHeatmap, pheatmap, ROCit и иных связанных пакетов.

## **РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ**

### **Взаимосвязь ГКГС-I с тяжестью COVID-19**

Разработан способ оценки риска развития тяжелого течения COVID-19 для российской популяции. Данный способ включает забор биологического материала пациента, выделение геномной ДНК с последующим генотипированием аллелей генов HLA-A, HLA-C, обработку результатов генотипирования и прогнозирование по полученным данным риска развития

тяжелой формы COVID-19 с высокой вероятностью летального исхода. Обработку результатов генотипирования осуществляют с использованием метода главных компонент посредством определения для каждого аллеля значений 2-ой и 3-ей главных компонент гена HLA-A и значения 4-ой главной компоненты гена HLA-C из базы данных аллелей генов HLA-A, HLA-C, предварительно определенных для популяции, в которую входит исследуемый индивидуум с последующим суммированием всех значений главных компонент, нормированием значения полученной суммы и сравнением с пороговым значением. Прогнозируют риск развития тяжелой формы COVID-19 при получении нормированного значения суммы выше порогового значения.

База данных аллелей генов HLA-A, HLA-C со значениями главных компонент была сформирована по итогам соответствующего генотипирования популяции, включая пациентов, умерших от COVID-19. С помощью программы netMHCpan [Reynisson и др., 2020] был осуществлен расчет коэффициентов связывания молекул ГКГС-I, кодируемых аллелями генов HLA-A и HLA-C, с пептидами вируса SARS-CoV-2. Результаты расчета были обработаны методом главных компонент. При этом обработку осуществляли отдельно для каждого гена HLA-A, HLA-C, а количество главных компонент определяли таким образом, чтобы доля дисперсии, обусловленной каждой компонентой, составляла не менее 5% от исходной дисперсии.

Нормирование суммы полученных значений главных компонент ( $S$ ) аллелей генов HLA-A, HLA-C индивидуума реализовано по формуле  $N = (S - S_{min}) / (S_{max} - S_{min})$ , где  $S_{min}$  – сумма всех минимальных значений главных компонент по каждому аллелю из базы данных,  $S_{max}$  – сумма всех максимальных значений главных компонент по каждому аллелю из базы данных,  $N$  – нормированный индекс риска (или просто индекс риска).

Пороговое значение было определено на основании использования значений главных компонент умерших пациентов в возрасте до 60 лет из базы данных аллелей генов HLA-A, HLA-C, при этом для каждого умершего пациента определяли значения 2-ой и 3-ей главных компонент гена HLA-A и значение 4-ой главной компоненты гена HLA-C из базы данных аллелей генов HLA-A, HLA-C, с последующим суммированием всех значений главных компонент, нормированием полученных сумм (по формуле выше), после чего из полученных значений  $N_1, N_2, \dots, N_k$  ( $N_i$  – нормированный индекс риска для  $i$ -го пациента,  $k$  – общее число пациентов) в качестве порогового значения выбрали значение, характеризующее максимальную долю пациентов, умерших до 60 лет, в группе пациентов, с  $N_i$  выше порогового значения. При этом определение максимальной доли пациентов, умерших до 60 лет, было реализовано с помощью точного теста Фишера.

Данный способ был разработан по итогам проведенных исследований генетического материала 539 пациентов, включая генетический материал умерших пациентов в возрасте до 60 лет и после 60 (111 пациентов). В результате были выявлены возможные связи между генотипами умерших пациентов, их возрастом на момент смерти с критическим течением COVID-19.

Использование в данном способе данных HLA-генотипа пациентов умерших от COVID-19 в возрасте до 60 лет и старше 60 лет, значимо повлияло на повышение точности получаемых оценок. Смерть от COVID-19 является следствием сочетания сопутствующих заболеваний и неадекватного ответа иммунной системы, одной из причин которого является генетическая предрасположенность. Поэтому сравнение умерших от COVID-19 в возрасте до 60 лет с умершими в возрасте старше 60 лет позволило выявить преимущественный вклад наследственного фактора.

Кроме того, в разработанном способе предложены иные, нежели в доступных источниках, критерии оценки аффинности пептидов вируса к молекулам ГКГС-I, кодируемым аллелями генов HLA. В работе Iturrieta-Zuazo и соавторов [Iturrieta-Zuazo и др., 2020] упоминается два способа интерпретации результатов математического моделирования аффинности пептида и молекул ГКГС-I: при аффинности менее 50 нМоль взаимодействие считалось возможным, при более и равно 50 нМоль – нет; при аффинности менее 500 нМоль взаимодействие считалось возможным, при более и равно 500 нМоль – нет. Разработанный в рамках исследования способ использует непрерывную шкалу значений аффинности от 1 до 5000 в отличие от способа Iturrieta-Zuazo и соавторов, который бинаризует (0 или 1) значение аффинности.

Разработанный способ оценки риска развития тяжелой формы COVID-19 на основе анализа генотипа ГКГС-I демонстрирует более высокую точность прогнозных оценок течения COVID-19 по сравнению со способом-прототипом, что является важным для выбора тактики лечения пациента. Тем не менее, необходимо учитывать, что в популяции постепенно формируется иммунитет и улучшаются протоколы лечения пациентов с COVID-19, но вирус SARS-CoV-2 постоянно мутирует, что может значительно сказываться на точности прогноза.

Разработанный способ был реализован в форме программного обеспечения «Аффисенс» (ПО) для клинической оценки риска развития тяжелой формы течения коронавирусной инфекции COVID-19. Свидетельство о регистрации программы для ЭВМ № 2023619424, 11.05.2023.

### Влияние на тяжесть течения COVID-19 различных вариантов вируса

Ранее в данной работе на выборке пациентов первой волны нами было показано, что генотип ГКГС-I является значимым фактором риска тяжелого течения COVID-19 только у пациентов в возрасте не старше 60 лет. Нами была изучена взаимосвязь генотипа ГКГС-I этих пациентов с тяжестью течения COVID-19, вызванной двумя наиболее распространенными летом 2021 года вариантами штамма Дельта SARS-CoV-2: AY.122 и B.1.617.2.

В августе 2021 года была сформирована выборка из 45 пациентов, госпитализированных со среднетяжелым и тяжелым течением COVID-19, проведено HLA-генотипирование и определена штаммовая принадлежность вируса SARS-CoV-2. Характеристика пациентов представлена в Таблице 1.

Таблица 1 – Характеристика выборки пациентов с COVID-19

Показатель	Вариант AY.122	Вариант B.1.617.2	<i>p</i>
Число пациентов	37	8	
Возраст, медиана [25%-75%]	35 [32 – 46]	45 [35 – 51,8]	0,26
Пол (Муж./Жен.)	10/27	1/7	0,65
Тяжесть COVID-19			0,04
Средняя	14	8	
Тяжелая	23	0	
Максимальная степень поражения лёгких по КТ			0,37
КТ-0	3	0	
КТ-1	10	4	
КТ-2	5	1	
КТ-3	9	0	
КТ-4	3	0	
н.д.	7	3	
Исход			1
Летальный	1	0	
Выписан	36	8	
Сопутствующие заболевания			
Диабет	8	2	1
Ожирение	7	1	1
Гипертония	8	2	1

Секвенирование РНК, выделенной из мазков из носоглотки, позволило установить, что пациенты были инфицированы вариантами AY.122 и B.1.617.2 штамма Дельта SARS-CoV-2. По данным портала GISAID [Elbe, Buckland-Merrett, 2017] выявленное нами соотношение вариантов SARS-CoV-2 характерно

для московского региона в августе 2021 года (ОШ = 2,2,  $p = 0,1$ , 95% ДИ: 0,8 – 5,4).

Нам удалось установить, что тяжесть течения COVID-19 значительно различалась для больных с вариантами AY.122 и B.1.617.2 (ОШ неопр.,  $p = 0,4$ , 95% ДИ 0,9 – неопр.). Мы проанализировали различия в профиле мутаций этих вариантов SARS-CoV-2 (Рисунок 2). В варианте AY.122 выявлено 12 мутаций: 6 мутаций в ORF1ab, две мутации во вспомогательных белках и четыре мутации в структурных белках. В свою очередь вариант B.1.617.2 несет 24 мутации: пять мутаций в ORF1ab, пять мутаций во вспомогательных белках и 14 мутаций в структурных белках. Необходимо отметить, что два варианта SARS-CoV-2 объединяют только две мутации: D614G в спайк-белке и P323L в РНК-зависимой РНК полимеразе из ORF1ab. Обе эти мутации закрепились в геноме SARS-CoV-2 еще в начале 2020 года и ассоциированы с тяжестью течения COVID-19 [Biswas, Mudi, 2020].

Учитывая, что подавляющее большинство мутаций в вариантах вируса различаются, мы проанализировали их влияние на способность молекул ГКГС-I взаимодействовать с мутантными пептидами каждого варианта.

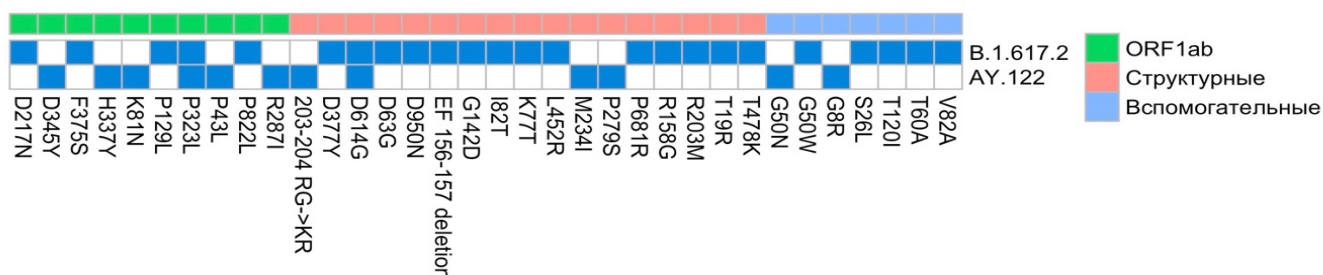


Рисунок 2 – Профиль мутаций в вариантах AY.122 и B.1.617.2 SARS-CoV-2

Генотипирование ГКГС-I пациентов показало, что наиболее частыми аллелями в выборке были HLA-A\*02:01 (частота 0,26) и HLA-A\*01:01 (частота 0,25), что характерно для популяции московского региона [Shkurnikov и др., 2021]. Малый объем выборки не позволил выявить взаимосвязи гаплотипов и отдельных аллелей генов ГКГС-I с тяжестью течения COVID-19 для анализируемых вариантов SARS-CoV-2. Тем не менее, был оценен индекс риска тяжелого течения COVID-19 [Shkurnikov и др., 2021], отражающий интегральную способность индивидуального набора молекул ГКГС-I презентировать пептиды SARS-CoV-2 (Рисунок 3, панель А). Различия в «Индексе риска» между пациентами, переболевшими вариантом AY.122 в среднетяжелой и тяжелой форме, оказались недостоверными ( $p = 0,47$ ). Но у пациентов с тяжелым течением есть тенденция к более высокому индексу риска.



Анализ влияния мутаций на число вирусных пептидов, взаимодействующих с индивидуальным набором молекул ГКГС-I с аффинностью менее 50 нМоль, показал, что различия достоверны для вспомогательных белков вируса (Рисунок 3, панель Б). Так, у инфицированных вариантом AY.122 число высокоаффинных пептидов значительно сократилось по сравнению с инфицированными вариантом B.1.617.2 ( $p < 0,01$ ).

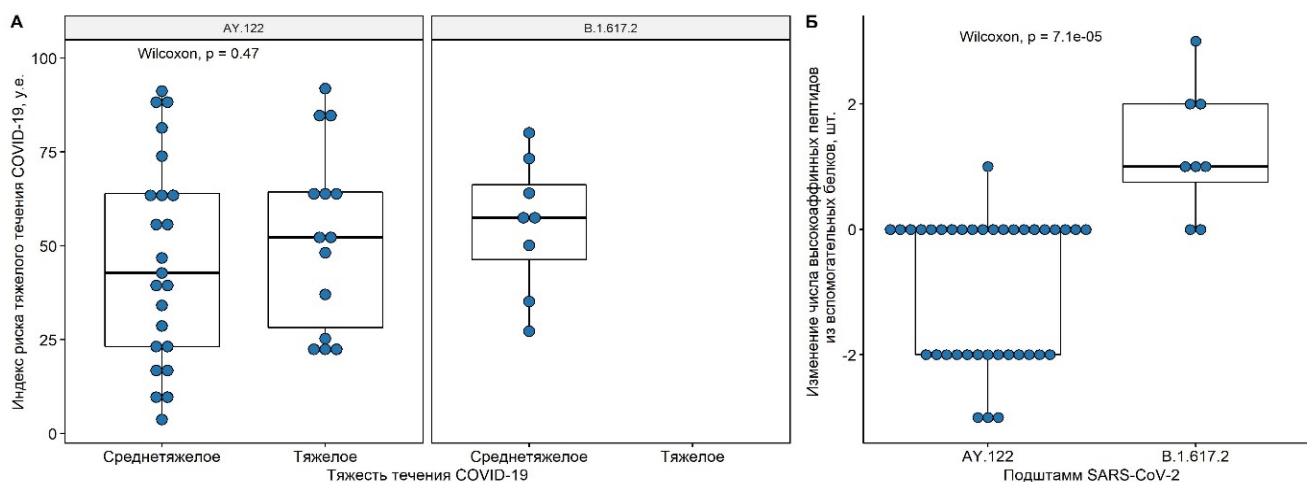


Рисунок 3 – Панель А – Взаимосвязь индекса риска тяжелого течения COVID-19 с фактическим течением заболевания. Панель Б – Влияние мутаций во вспомогательных белках вариантов AY.122 и B.1.617.2 на аффинность их взаимодействия с индивидуальным набором молекул ГКГС

Значимое снижение числа высокоаффинных пептидов в варианте AY.122 прежде всего связано с мутацией G8R в белке NS8. Данная мутация вызывает падение аффинности взаимодействия пептидов FLGIITTV (с 42 нМоль до 121 нМоль) и FLVFLGIITTV (с 25 нМоль до 59 нМоль) с молекулой ГКГС-I, кодируемой наиболее частым для нашей популяции аллелем HLA-A\*02:01. Ранее была показана иммуногенность пептида FLGIITTV у пациентов, переболевших COVID-19 [Saini и др., 2021]. Особого внимания заслуживает тот факт, что белок NS8 (ORF8) способен подавлять созревание молекул ГКГС-I и их транслокацию на поверхность зараженной клетки [Matsuoka и др., 2022; Zhang и др., 2021]. Данные обстоятельства могли внести весомый вклад в превалирование варианта AY.122 в российской популяции и более тяжелое, по сравнению с другими вариантами штамма Дельта, течение вызванной им COVID-19. Предлагаемый подход к анализу мутаций на уровне вариантов SARS-CoV-2 представляется очень важным, так как SARS-CoV-2 продолжает активно мутировать.

Кроме того, были изучены особенности генотипа ГКГС-I пациентов в первую и третью волну COVID-19. Как было показано ранее, штамм Дельта

коронавируса SARS-CoV-2, вызвавший третью волну пандемии COVID-19 в России, в значительной степени отличался от циркулировавших ранее штаммов вируса [Klink и др., 2022]. Всплеск заболеваемости был связан с высокой трансмиссивностью этого штамма по сравнению с Альфа-вариантом [Callaway, 2021]. Возрастание трансмиссивности во многом было обусловлено увеличением до 6 раз по сравнению со штаммом Альфа количества выдыхаемых на пике инфекции зараженным человеком вирусных частиц [Earnest и др., 2022]. Дельта-вариант увеличил не только риск госпитализации по поводу COVID-19, но и риск смерти у непривитых пациентов [Bast и др., 2021].

В штамме Дельта выявлено 18 мутаций, повлиявших на аминокислотную последовательность его белков, в значительной степени поменявших течение заболевания [Nersisyan и др., 2022b]. Пять из них пришлись на спайк-белок, что ослабило эффективность гуморального иммунитета, сформированного как естественным путем (переболевшие пациенты), так и после вакцинации [Bian и др., 2021]. Кроме того, одна из этих мутаций (D614G) повысила сродство рецептор-связывающего домена (RBD) спайк-белка к рецептору ACE2 [Ozono и др., 2021]. Необходимо отметить, что изменилась и скорость репликации вируса. Так, штамм Дельта реплицируется в два раза медленнее по сравнению со штаммом Альфа в первые 8 часов после заражения [Shuai и др., 2022]. Данное обстоятельство крайне важно, так как продуцируемый SARS-CoV-2 вспомогательный белок ORF8 способен напрямую взаимодействовать с молекулами ГКГС-I и подавлять их созревание. Так, уже через 18 часов после начала экспрессии в клетке ORF8 практически полностью прекращается экспорт на её поверхность молекул ГКГС-I [Zhang и др., 2021].

Накоплен значительный объем информации об актуальных эпитопах различных вариантов SARS-CoV-2 [Vita и др., 2019], особенностях формирования Т-клеточной памяти [Francis и др., 2022], тенденциях в частоте мутаций различных регионов вируса [Nersisyan и др., 2022b]. Особого внимания заслуживают эпитопы, произошедшие из белков, кодируемых в области ORF1ab SARS-CoV-2. Геном SARS-CoV-2 включает большую область ORF1ab, четыре гена структурных белков (S, E, M, N) и несколько ORF, кодирующих вспомогательные белки [Arya и др., 2021]. Ген ORF1ab кодирует несколько неструктурных белков, участвующих в различных молекулярных процессах, которые необходимы для размножения SARS-CoV-2, включая экспрессию и репликацию его генома [Badua, Baldo, Medina, 2021]. Сразу после трансляции полученная полипептидная цепь расщепляется вирусными протеазами и образуются 16 неструктурных белков (NSP1-NSP16). Известно, что ген ORF1ab очень консервативен по сравнению с другими участками генома SARS-CoV-2 [Jaroszewski и др., 2021; Lubin и др., 2022].

В данном исследовании мы сравнили особенности генотипов ГКГС-I пациентов в возрасте не старше 60 лет, переболевших COVID-19 в первую и третью волну, оценили влияние мутаций в вирусе SARS-CoV-2 на иммуногенные эпитопы CD8<sup>+</sup> Т-лимфоцитов.

Было проведено генотипирование ГКГС-I у 147 пациентов, переболевших в первую волну COVID-19 в г. Москва в период с мая по август 2020 года, – группа Волна 1. Также были генотипированы 219 пациентов, переболевших COVID-19 в период с июня по июль 2021, – группа Волна 3 (Таблица 2).

Таблица 2 – Клиническая характеристика групп сравнения

Параметр	Группа		<i>p</i>
	Волна 1	Волна 3	
Число пациентов	147	219	
Возраст, лет	43,4 ± 11	42,9 ± 5,2	0,4
Пол			
Мужской	73 (49,7%)	123 (56,2%)	0,24, ОШ = 0,77
Женский	74 (50,3%)	96 (43,8%)	
Вакцинация	0	19 (8,7%)	4,9e-05
Тяжесть COVID-19			
Лёгкая/средняя	120 (81,6%)	192 (87,7%)	0,13, ОШ = 0,63
Тяжёлая	27 (18,4%)	27 (12,3%)	
Исход заболевания			
Выздоровление	127 (86,4%)	197 (90%)	0,32, ОШ = 0,71
Летальный	20 (13,6%)	22 (10%)	
Реанимация	28 (19%)	51 (23,3%)	0,36, ОШ = 1,3
Кислородная поддержка			
Без	98 (66,7%)	140 (63,9%)	8,8e-05
Низкопоточная	24 (16,3%)	38 (17,4%)	
Прон позиция	0	11 (5%)	
Высокопоточная	16 (10,9%)	4 (1,8%)	
ИВЛ	9 (6,1%)	26 (11,9%)	
Сопутствующие заболевания			
ХОБЛ	0	9 (4,1%)	0,01
Ожирение	4 (2,7%)	21 (9,6%)	0,01, ОШ = 3,8
Заболевания сердца	20 (9,1%)	14 (9,5%)	1, ОШ = 0,95
Гипертоническая болезнь	16 (10,9%)	50 (22,8%)	0,003, ОШ = 2,4
Злокачественные новообразования	3 (2%)	9 (4,1%)	0,4, ОШ = 2,1

Мы не выявили значимых различий в возрасте пациентов, включенных в группы сравнения, а также в соотношении полов в группах. Доля вакцинированных пациентов в группе Волна 3 была незначительна и составляла

8,7%, что несколько ниже общегородского уровня вакцинации на июнь 2021 года – 15% [Boguslavsky, Sharova, Sharov, 2022].

Необходимо отметить значительное увеличение доли пациентов с хронической обструктивной болезнью лёгких ( $p = 0,01$ ), ожирением ( $p = 0,01$ , ОШ = 3,8), гипертонической болезнью ( $p = 0,003$ , ОШ = 2,4) в третьей волне COVID-19. При оценке вклада коморбидностей в риск смерти от COVID-19 показано, что если в первую волну COVID-19 поражения миокарда ( $p = 5,2e-5$ , ОШ = 13) и гипертоническая болезнь ( $p = 0,045$ , ОШ = 3,5) были значимыми факторами риска смерти, то в третью волну подобного влияния не отмечено.

Была проверена гипотеза о различии в частотах единичных аллелей у пациентов групп сравнения: пациентов с COVID-19 первой волны, пациентов с COVID-19 третьей волны и популяционной группы. Распределение частот основных аллелей HLA-A, HLA-B и HLA-C в этих трех группах обобщенно представлено на Рисунке 4.

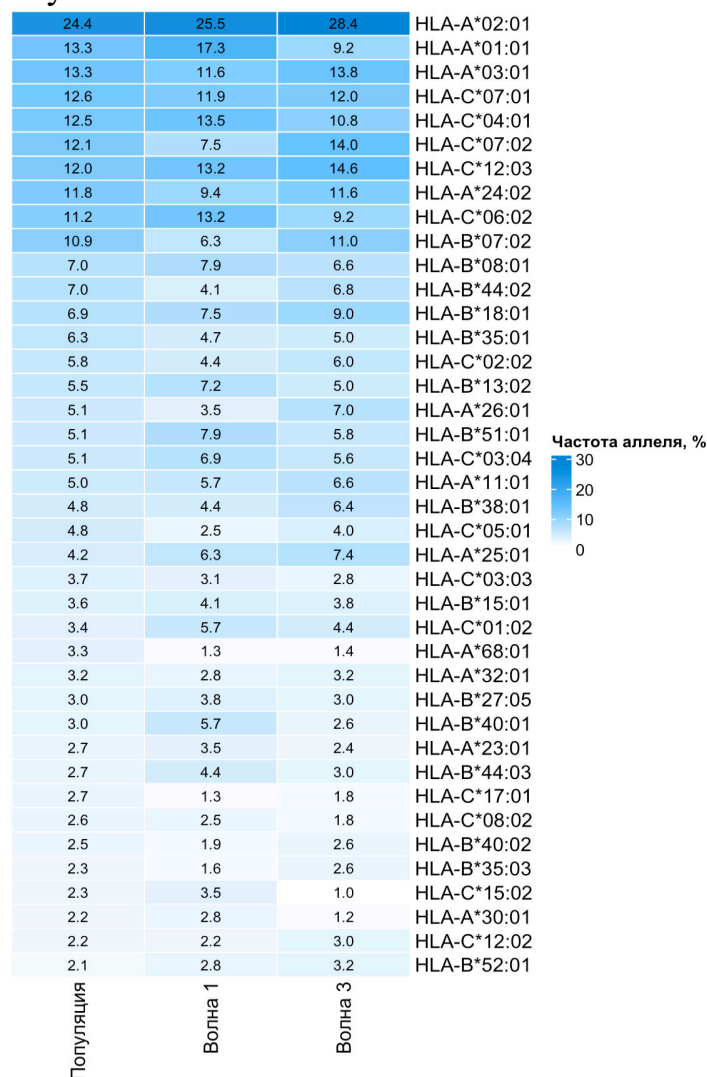


Рисунок 4 – Распределение частот аллелей HLA-A, HLA-B и HLA-C в группах сравнения. Представлены аллели с частотой не менее 5% хотя бы в одной из групп сравнения

Выявлено, что для всех возможных групповых сравнений только один аллель из двенадцати наиболее частых аллелей имел высокое отношение шансов, которое можно считать статистически значимым после поправки на множественность сравнений. В частности, частота аллеля HLA-A\*01:01 снизилась с 17,3% в группе Волна 1 до 9,8% в группе Волна 3 ( $p = 0,025$  с поправкой на множественность сравнений, ОШ = 0,5).

Мы предположили, что снижение частоты встречаемости носителей HLA-A\*01:01 среди пациентов третьей волны могло быть связано с особенностями Т-клеточного ответа, формируемого у ее носителей. Для проверки этой гипотезы были использованы компьютерные методы анализа взаимодействия молекула ГКГС-I – пептид вируса, влияния мутаций в основных вариантах SARS-CoV-2 на эти взаимодействия, а также результаты экспериментальной проверки иммуногенности пептидов у пациентов, переболевших в первую волну.

Анализ числа пептидов вируса SARS-CoV-2, взаимодействующих с 12 наиболее частыми аллелями ГКГС-I в европейской популяции [Gonzalez-Galarza и др., 2019], показал, что аллель HLA-A\*01:01 относится к числу аллелей со средней способностью к взаимодействию с пептидами как структурных, так и неструктурных белков вируса. Он способен провзаимодействовать с аффинностью менее 50 нМоль с 28 пептидами из белков, кодируемых ORF1ab, и 8 пептидами из структурных и вспомогательных белков. В свою очередь аллель HLA-A\*02:01 имеет 207 высокоаффинных пептидов из ORF1ab и 56 высокоаффинных пептидов из других белков вируса.

Также были проанализированы различия в способности ГКГС-I презентировать вирусные пептиды между генотипами, которые включают HLA-A\*01:01, и генотипами, которые не включают этот аллель. Носители аллеля HLA-A\*01:01, имели в среднем меньшее количество высокоаффинных к их набору молекул ГКГС-I пептидов вируса по сравнению с не-носителями независимо от волны ( $p < 0,02$ ). Следует отметить, что носители аллеля HLA-A\*02:01, в отличие от носителей HLA-A\*01:01, имели большее количество высокоаффинных к их набору молекул ГКГС-I вирусных пептидов ( $p < 3,8e-08$ ).

Был проведен анализ иммуногенных эпитопов вируса SARS-CoV-2 по данным Immune Epitope Database (IEDB) для аллеля HLA-A\*01:01 и аллеля HLA-A\*02:01. На момент анализа (май 2022 года) в базе содержалась информация об ответе на 365 пептидов. Два эпитопа были проверены для обоих аллелей: S<sub>136–144</sub>(CNDPFLGVY), M<sub>170–178</sub>(VATSRTLSTY). Для аллеля HLA-A\*01:01 была проверена иммуногенность 50 пептидов из ORF1ab и 33 пептидов не из ORF1ab. Для аллеля HLA-A\*02:01 было проверено 139 пептидов из ORF1ab и 145 пептидов не из ORF1ab. Среди проверенных пептидов были определены иммунодоминантные эпитопы (вызывали ответ Т-лимфоцитов не менее чем в

50% проверенных образцов). Для аллеля HLA-A\*01:01 было выявлено 10 иммунодоминантных эпитопов, кодируемых ORF1ab, и только 3 эпитопа не из ORF1ab. Для аллеля HLA-A\*02:01 был выявлен 51 иммунодоминантный эпитоп из ORF1ab и 59 эпитопов не из ORF1ab. Соотношение иммунодоминантных эпитопов из ORF1ab для аллеля HLA-A\*01:01 было в 3,8 раз больше по сравнению с аллелем HLA-A\*02:01 ( $p = 0,04$ ). Этот результат свидетельствует о том, что среди иммунодоминантных пептидов для аллеля HLA-A\*01:01 преобладают пептиды, кодируемые в ORF1ab и характеризующиеся низкой частотой мутаций.

Для проверки гипотезы о том, что у выздоровевших носителей аллеля HLA-A\*01:01 первой волны было большее количество иммуногенных эпитопов из белков ORF1ab, проанализировали Т-клеточные ответы 28 пациентов носителей хотя бы одного из двух наиболее распространенных аллелей в европейской популяции с подтвержденным COVID-19 во время первой волны пандемии: HLA-A\*01:01 (13 пациентов) и HLA-A\*02:01 (15 пациентов). Для каждого образца Т-клеток была протестирована индивидуальная иммуногенность 15 подтвержденных эпитопов ORF1ab SARS-CoV-2. Пептиды из этой панели не индуцировали Т-клеточный ответ у пациентов, ранее не болевших COVID-19 [Titov и др., 2022]. Количество эпитопов для HLA-A\*01:01 равнялось семи, а для HLA-A\*02:01 – восьми.

Поскольку время от появления симптомов до момента забора крови для анализа иммуногенности пептидов могло повлиять на силу Т-клеточного ответа, мы сравнили это время между группами носителей HLA-A\*01:01 и HLA-A\*02:01. Для носителей HLA-A\*01:01 среднее время после появления симптомов составило 34 дня, а для носителей HLA-A\*02:01 – 41 день; различия были незначительными.

Для оценки силы Т-клеточного иммунитета, сформированного пептидами из ORF1ab у носителей аллелей HLA-A\*01:01 и HLA-A\*02:01, оценивали частоту истинно положительных ответов (чувствительность) для каждого из эпитопов. Этот показатель отражает отношение истинно положительных ответов к ожидаемому количеству положительных ответов. Частота истинно положительных ответов значительно различались для эпитопов из ORF1ab ( $p = 0,017$ ). Три из семи протестированных пептидов для носителей аллеля HLA-A\*01:01 были иммунодоминантными (ORF1ab<sub>1637-1646</sub> TTDPSFLGRY, ORF1ab<sub>1636-1646</sub> HTTDPSFLGRY, ORF1ab<sub>1321-1329</sub> PTDNYITTY), но не было ни одного иммунодоминантного пептида из восьми протестированных для носителей аллеля HLA-A\*02:01.

Помимо данных IEDB и собственных данных проверки эпитопов, были проанализированы данные scRNA-seq CD8<sup>+</sup> Т-клеток переболевших COVID-19,

активированных отдельными пептидами [Francis и др., 2022]. Рассматриваемый набор данных включал набор фенотипированных Т-клеточных клонов, отвечающих на полный набор эпитопов, происходящих от SARS-CoV-2, ассоциированных с четырьмя основными аллелями ГКГС-I (HLA-A\*01:01, HLA-A\*02:01, HLA-A\*24:02 и HLA-B\*07:02). Было проанализировано распределение эпитопов, происходящих из SARS-CoV-2, которые вызывали ответ Т-клеток, в зависимости от области генома, из которой они произошли. В соответствии с результатами, упомянутыми выше, большинство иммуногенных эпитопов для носителей HLA-A\*01:01 происходило из ORF1ab. Для остальных аллелей такой тенденции не наблюдалось (Рисунок 5, панель А).

Также были проанализированы фенотипы реагирующих на стимуляцию пептидами клонов Т-клеток (Рисунок 5, панель Б). Эпитопы, ассоциированные с HLA-A\*01:01, вызывали ответы в основном от субпопуляции Т-клеток центральной памяти (T<sub>cm</sub>). В то же время доля отвечающих T<sub>cm</sub> клеток, ассоциированных с другими аллелями, была значительно меньше (попарные сравнения HLA-A\*01:01 с HLA-A\*02:01, HLA-B\*07:02 и HLA-A\*24:02,  $p < 0,02$ ). Вместе с наблюдением, что большинство известных эпитопов, связанных с HLA-A\*01:01, происходят из консервативной области ORF1ab, эти результаты свидетельствуют о том, что люди, являющимися носителями HLA-A\*01:01, могут иметь более высокие шансы развить сильный иммунный ответ при вторичном инфицировании SARS-CoV-2 из-за уже существующей у них популяции T<sub>cm</sub> клеток иммунной памяти.

Затем были оценены изменения аффинности взаимодействий высокоаффинных эпитопов аллелей HLA-A\*01:01 и HLA-A\*02:01, ассоциированные с мутациями в основных вариантах SARS-CoV-2. Ни один из высокоаффинных эпитопов HLA-A\*01:01 не был затронут мутациями в существующих актуальных на момент исследования вариантах вируса: Дельта G/478K.V1 и Омикрон (BA.1-BA.4). В то же время семь высокоаффинных эпитопов HLA-A\*02:01 были затронуты мутациями. Шесть из них были расположены в Спайк-белке и один в белке NS3. Все мутации, кроме одной, приводили к снижению аффинности взаимодействия с HLA-A\*02:01. Так, в варианте Омикрон BA.1 пептид S<sub>495-503</sub> YGFQPTNGV имел три замены (G496S, Q498R, N501Y), что привело к увеличению аффинности его взаимодействия с HLA-A\*02:01 (с 2160 нМоль до 87 нМоль). В вариантах Омикрон BA.1 и BA.3 делеции A67V и HV 69-70 привели к исчезновению иммунопреобладающего эпитопа S<sub>62-70</sub> VTWFHAIHV.

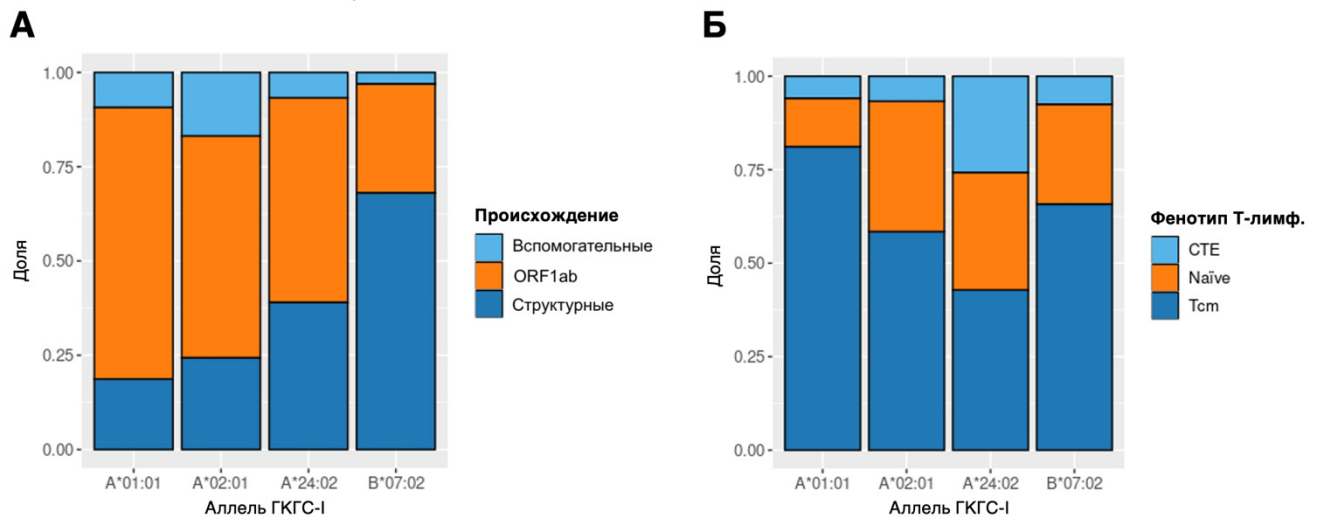


Рисунок 5 – (А) Распределение иммуногенных эпитопов, полученных из SARS-CoV-2, в зависимости от их происхождения. Эпитопы, связанные с аллелем HLA-A\*01:01, как правило, происходят в основном из консервативной области ORF1ab SARS-CoV-2. (Б) Распределение фенотипов Т-клеток, активированных полным набором эпитопов SARS-CoV-2. Tcm – субпопуляция Т-лимфоцитов центральной памяти; CTE – цитотоксические Т-эффекторные клетки. Эпитопы, связанные с аллелем HLA-A\*01:01, вызывают более сильный ответ Tcm по сравнению с другими аллелями (парные точные значения критерия Фишера:  $5,423e-05$  для HLA-A\*01:01 по сравнению с HLA-A\*02:02),  $0,0178$  для HLA-A\*01:01 по сравнению с HLA-B\*07:02,  $6,196e-05$  для HLA-A\*01:01 по сравнению с HLA-A\*24:02)

Кроме того, было проанализировано влияние мутаций в основных вариантах SARS-CoV-2 на аффинность связывания всех возможных вирусных пептидов с 12 указанными выше аллелями ГКГС-I. HLA-A\*01:01 имел 51 высокоаффинный пептид из ORF1ab и 13 высокоаффинных пептидов из других белков (структурных и вспомогательных) в варианте Wuhan-Hu-1 SARS-CoV-2. Отличительной особенностью этого аллеля было наличие относительно большого количества высокоаффинных для него пептидов, не подверженных мутациям в основных вариантах SARS-CoV-2. Из 64 высокоаффинных пептидов HLA-A\*01:01 только один пептид значительно изменил аффинность в 16 проанализированных вариантах вируса. В то же время аналогичная по количеству высокоаффинных пептидов молекула ГКГС-I, кодируемая аллелем HLA-A\*24:02, имела 27 измененных пептидов из ORF1ab ( $p = 3e-09$ , ОШ = 0,02). Следует отметить, что высокоаффинные пептиды из ORF1ab для всех проанализированных аллелей были менее подвержены мутациям по сравнению с остальными пептидами ( $p = 4,3e-05$ ).

Было продемонстрировано значительное снижение частоты встречаемости носителей аллеля HLA-A\*01:01 среди госпитализированных пациентов во время



третьей волны пандемии COVID-19. Расчет аффинности взаимодействия между молекулами ГКГС-I и пептидами SARS-CoV-2 выявил возможную причину уменьшения частоты встречаемости носителей аллеля HLA-A\*01:01. Разные гены и, следовательно, белки SARS-CoV-2 мутируют с разной скоростью. Ген ORF1ab крайне консервативен по сравнению с другими генами SARS-CoV-2. Носители аллеля HLA-A\*01:01 имеют значительное количество высокоаффинных пептидов из этого гена.

В когорте выздоровевших пациентов первой волны COVID-19 были подтверждены результаты компьютерного моделирования и продемонстрирована более высокая частота встречаемости иммунодоминантных эпитопов из белков гена ORF1ab вируса SARS-CoV-2 у носителей HLA-A\*01:01 по сравнению с эпитопами этого гена у носителей HLA-A\*02:01.

Более того, анализ результатов одноклеточного фенотипирования Т-клеток у выздоровевших пациентов показал, что преобладающим фенотипом у носителей HLA-A\*01:01 являются Т-клетки центральной памяти. Преобладание Т-лимфоцитов данного фенотипа может способствовать формированию длительного Т-клеточного иммунитета у носителей данного аллеля. Наши результаты могут стать основой для создания высокоэффективных вакцин на основе пептидов ORF1ab.

### **Роль микроРНК в регуляции взаимодействия вируса и макроорганизма**

Показано, что (+)РНК вирусы способны эволюционировать под влиянием микроРНК клеток-хозяина [Scheel и др., 2016]. Однако в данных исследованиях не учитывалась способность вируса SARS-CoV-2 менять последовательность регионов связывания с seed-регионами микроРНК человека в результате мутаций.

Необходимо отметить, что SARS-CoV-2 прежде всего размножается в альвеоцитах второго типа ткани лёгких, что сужает разнообразие взаимодействующих с ним микроРНК с 2000-3000 до 200-300 видов, характерных именно для этого типа ткани [Nersisyan и др., 2022c].

За более чем два года пандемии вирус SARS-CoV-2 значительно мутировал, причем далеко не все его мутации привели к изменению последовательности вирусных белков в силу вырожденности генетического кода [Nersisyan и др., 2022a]. Не исключено, что данные мутации могли закрепляться в геноме вируса под влиянием микроРНК альвеоцитов. Мы проверили гипотезу об эволюционном давлении микроРНК, характерных для ткани лёгких человека, на геном вируса SARS-CoV-2.

Среди более чем 129 миллионов последовательностей SARS-CoV-2 были отобраны 273, отражающие ключевые штаммы и их варианты, циркулировавшие в Москве и Московском регионе с начала пандемии и до сентября 2022 года. Отобранные варианты SARS-CoV-2 были разделены на две группы хронологически:

– группа «Ранние штаммы» включала в себя штаммы Альфа (5 вариантов) и Дельта (52 варианта), циркулировавшие в московском регионе с начала эпидемии до января 2022 года;

– группа «Штаммы Омикрон» включала 216 вариантов штамма Омикрон.

Анализ результатов секвенирования образцов ткани лёгких позволил выявить более чем 300 видов микроРНК. В то же время, всего на 32 из них пришлось более 95% всех молекул микроРНК в этой ткани. Список высокопредставленных микроРНК и их доля от всех молекул микроРНК в ткани лёгких приведён на Рисунке 6. Для высокопредставленных микроРНК были определены регионы вирусной РНК, обратно-комплиментарные их seed-региону.

Для проверки гипотезы об эволюционном давлении микроРНК, характерных для ткани лёгких человека, на геном вируса SARS-CoV-2, было оценено изменение числа регионов связывания между группами сравнения. В группе «Ранние штаммы» взвешенное среднее статистически значимо меньше по сравнению с группой «Штаммы Омикрон» (критерий Манна-Уитни,  $p = 0,0002$ ). Таким образом, более поздние мутации в РНК SARS-CoV-2 приводят к возможной потере регуляторной активности микроРНК.

Анализ вклада отдельных микроРНК в данное снижение показал, что оно прежде всего связано с hsa-miR-21-5p, hsa-miR-30a/e-3p, hsa-miR-451a (Рисунок 6). Необходимо отметить, что в группе «Штаммы Омикрон» возросло влияние hsa-miR-24-3p. Тем не менее, вопрос о том, какие именно из кодирующих участков SARS-CoV-2 теряют позиции связывания, остается открытым.

Выравнивание анализируемых штаммов вируса относительно штамма «Wuhan-Hu-1» позволило сравнить распределение позиций регионов связывания с учетом экспрессии микроРНК (Рисунок 7). Более строго, каждый регион связывания, обусловленный некоторой отдельной микроРНК, вносил равный экспрессии этой микроРНК вклад в распределение. Для сравнения распределений регионов связывания между группами «Штаммы Омикрон» и «Ранние штаммы», распределения по всем штаммам, принадлежащим к соответствующей группе, были усреднены. Данный анализ позволил установить, что регионы связывания распределены неравномерно по кодирующим участкам РНК.

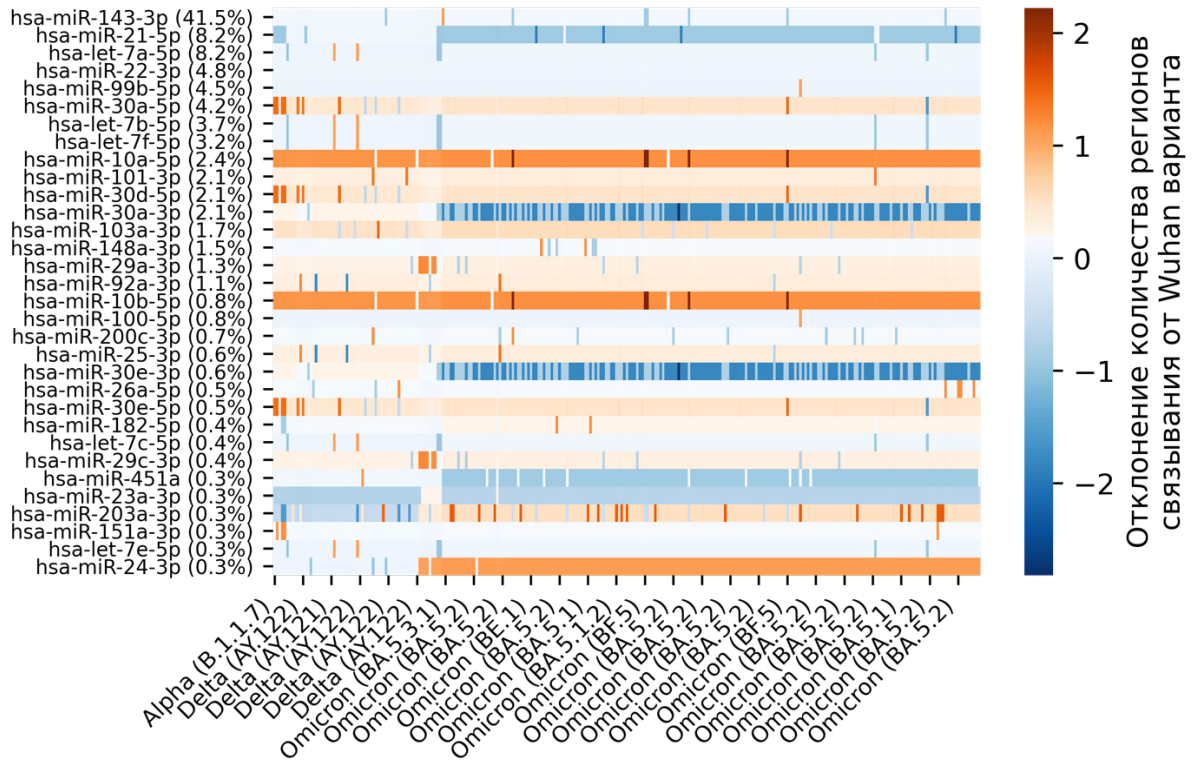


Рисунок 6 – Отклонение количества регионов связывания каждой высоко-экспрессированной микроРНК в каждом из анализируемых вариантов вируса от штамма «Wuhan-Hu-1», приходящихся на 30331 нуклеотид (длина штамма «Wuhan-Hu-1»). МикроРНК упорядочены по их экспрессии в ткани лёгких. Рядом с названием микроРНК указана их доля от общего числа молекул микроРНК в ткани

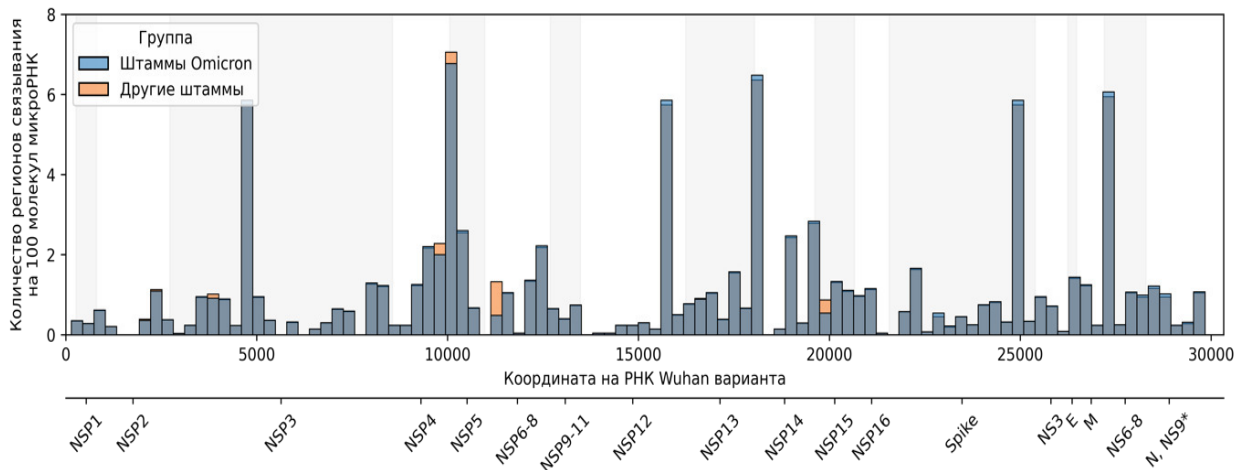


Рисунок 7 – Взвешенное среднее (по микроРНК с учетом их экспрессии) количества регионов связывания по группам «Ранние штаммы» и «Штаммы Омикрон»

В 5'- и 3'-нетранслируемых областях вируса практически полностью отсутствуют места связывания с микроРНК ткани лёгких. При этом можно выделить четыре белок-кодирующие области вируса, на которые приходится подавляющее число мест связывания с микроРНК: NSP3, NSP4, NSP12 и NSP14.

Все они принадлежат к гену ORF1ab. Ген ORF1ab кодирует ряд неструктурных белков, ответственных за дальнейшую экспрессию структурных и вспомогательных белков, а также репликацию вируса [Badua, Baldo, Medina, 2021]. Его трансляция осуществляется вирусом в первую очередь. В результате синтезируется полипептид, который после аутопротеолиза разрезается на 16 отдельных белков (NSP1 – NSP16) [Hartenian и др., 2020]. Аутопротеолическая активность обусловлена наличием папаин-подобных протеолитических доменов в мультидоменном белке NSP3 [Hartenian и др., 2020], на который по нашим данным приходится наибольшее число регионов связывания с микроРНК клетки-хозяина.

Возможно, это обусловлено эволюционным приспособлением вируса к микроРНК-окружению и использованию его для эпигенетической регуляции своего генома. Более того, в области, кодирующей NSP15, NS8, и NSP6, значительно снижается количество регионов связывания с микроРНК в «Штаммах Омикрон» ( $p < 0,01$ ). Ранее было показано, что снижение числа таких регионов приводит к снижению патогенности вируса восточного конского энцефалита (EEEV) [Trobaugh и др., 2014]. Более того, белок NS8 (ORF8) способен подавлять созревание молекул ГКГС-I и их транслокацию на поверхность зараженной клетки [Matsuoka и др., 2022; Zhang и др., 2021].

Можно заключить, что вирус SARS-CoV-2 практически не имеет регионов связывания в 5'- и 3'-нетранслируемых областях с микроРНК, характерными для ткани лёгких. Тем не менее вирус обладает значительным числом мест связывания с микроРНК в регионе NSP3-NSP5, ответственном за аутопротеолиз вирусных полипептидов и формирование вирионов. В вариантах штамма Омикрон произошло значимое снижение мест связывания с микроРНК клеток-хозяина, что могло способствовать снижению патогенности данного штамма.

Также была изучена роль циркулирующих микроРНК в прогнозе исхода COVID-19. Для анализа были использованы первичные данные секвенирования микроРНК (GSE195898), выделенных из плазмы крови из трех выздоровевших и восьми умерших пациентов (группа сравнения) с крайне тяжелой формой COVID-19, сопоставимых по полу и возрасту, проходивших лечение в отделении интенсивной терапии IRCCS Policlinico San Donato (Милан, Италия). Для каждого пациента был оценен профиль микроРНК плазмы крови на момент поступления в больницу (T0) и перед выпиской или смертью (T1).

Результаты секвенирования показали, что в момент поступления в лечебное учреждение (T0) в плазме крови пациентов в количестве более 150 прочтений на миллион (rpm) присутствовало 932 вида микроРНК и их 5'-изоформ. К наиболее высоко представленным в T0 и T1 можно отнести следующие микроРНК: hsa-miR-22-3p|0, hsa-miR-339-3p|0, hsa-miR-451a|0.

Анализ различий в профиле микроРНК у выздоровевших и не выздоровевших пациентов показал, что значимо изменилась экспрессия 46 микроРНК. В T0 у выздоровевших впоследствии пациентов была выше экспрессия hsa-miR-19b-3p|0 (в 4,5 раза,  $p = 0,017$ ), hsa-miR-25-3p|+1 (в 4,8 раза,  $p = 0,047$ ). В T1 значимо изменялась представленность 291 микроРНК. Наибольшие различия представленности в плазме между группой выздоровевших пациентов и группой сравнения были отмечены для следующих микроРНК: hsa-miR-451a|0 (в 13 раз,  $p = 7,65e-07$ ), hsa-miR-22-3p|0 (в 4,3 раза,  $p = 7,67e-05$ ), hsa-miR-19b-3p|0 (в 14 раз,  $p = 1,23e-06$ ).

Сравнение множеств микроРНК, различающихся между выздоровевшими пациентами и группой сравнения, показал, что только две микроРНК сонаправленно изменялись в T0 и T1 (Таблица 3).

Таблица 3 – Представленность сонаправленно изменившихся микроРНК в группах сравнения в T0 и T1

МикроРНК	Представленность микроРНК в плазме в T0, log2RPM			Представленность микроРНК в плазме в T1, log2RPM			
	Выздоровевшие	Скончавшиеся	$p$	Выздоровевшие	Скончавшиеся	$p$	$p$ с поправкой на множественность сравнений
hsa-miR-19b-3p 0	11,7 ± 2,9	10,6 ± 1,5	0,017	14,3 ± 1,8	10,6 ± 1,4	1,23e-06	0,003
hsa-miR-25-3p +1	3,4 ± 1,2	1,9 ± 0,8	0,047	4,7 ± 1,2	2 ± 0,7	2,25e-05	0,046

Были оценены возможные места связывания с геномом SARS-CoV-2 множества микроРНК, различающегося между группой выздоровевших пациентов и группой сравнения в T0 и T1. Ряд микроРНК не имел мест связывания с геномом вируса: hsa-miR-1225-3p|+3, hsa-miR-4498|+1, hsa-miR-6787-5p|+2, hsa-miR-1538|+1, hsa-miR-1307-5p|+1, hsa-miR-7111-5p|+2. Число

мест связывания микроРНК hsa-miR-19b-3p|0 и hsa-miR-25-3p|+1 составляло 12 и 9 соответственно. При этом медиана числа мест связывания остальных микроРНК находилась на уровне 3. Можно сделать вывод о том, что у выздоровевших пациентов был значимо повышен уровень микроРНК, имеющих значительное число мест связывания с геномом SARS-CoV-2 ( $p = 0,048$ ).

Также было проанализировано расположение возможных мест связывания микроРНК с геномом вируса. Наибольшее число мест связывания было расположено в регионе ORF1ab, кодирующем неструктурные белки. У микроРНК hsa-miR-19b-3p|0 наибольшее число мест связывания пришлось на короткий регион (338 нуклеотидов), кодирующий белок NSP9. При этом обе анализируемые микроРНК имели по два возможных места связывания с протяженным регионом (5834 нуклеотидов), кодирующим белок NSP3. Ни одна из микроРНК не имела мест связывания в 5' и 3'-нетранслируемых областях генома SARS-CoV-2.

Таким образом, представленность hsa-miR-19b-3p|0 и hsa-miR-25-3p|+1 значимо различалась в плазме крови пациентов с различным исходом COVID-19, как в момент поступления в стационар, так и в момент выписки или смерти. МикроРНК могут выступать в качестве компонента внутриклеточного иммунитета, регулируя трансляцию и репликацию (+) РНК-вирусов и изменяя патогенез вирусных инфекций [Huang и др., 2007; Ingle и др., 2015]. К числу наиболее высоко представленных в плазме крови относится hsa-miR-19b-3p|0, значимо повышенная в группе выздоровевших пациентов. На неё приходится более 1,5% всех циркулирующих молекул микроРНК. В рамках данной работы показано, что эта микроРНК обладает значительным числом мест связывания с геномом SARS-CoV-2, причем в наиболее его стабильной части – ORF1ab.

Наибольшее число мест связывания hsa-miR-19b-3p|0 приходится на регион длиной 338 нуклеотидов, кодирующий белок NSP9. Белок NSP9 способен связываться с 7SL РНК, входящей в частицы узнавания сигнала, тем самым нарушая транспорт белков в эндоплазматический ретикулум и на мембрану клетки [Banerjee и др., 2020]. Одним из семейств белков, созревающих в эндоплазматическом ретикулуме, является семейство молекул ГКГС-I. Нарушение их созревания может способствовать нарушению противовирусной активности цитотоксических Т-лимфоцитов. Кроме того, hsa-miR-19b-3p|0 способна связываться с регионом, кодирующим белок NSP3. Белок NSP3 совместно с NSP4 отвечает за формирование в зараженной клетке двухмембранных везикул, защищающих вирус от механизмов внутриклеточного врожденного иммунитета [Klatte, Shields, Agoni, 2022]. Таким образом, высоко представленная в плазме крови выздоровевших пациентов с COVID-19 микроРНК hsa-miR-19b-3p|0 способна связываться с регионами вируса,

кодирующими белки, ответственные за подавление механизмов внутриклеточного иммунитета.

Кроме того, hsa-miR-19b-3p|0 способна усиливать активность цитотоксических Т-лимфоцитов. Было обнаружено, что уровни hsa-miR-19b-3p|0 значительно повышены в мононуклеарных клетках периферической крови пациентов с длительной ремиссией ВИЧ. Сверхэкспрессия hsa-miR-19b-3p|0 способствует пролиферации CD8+Т-клеток, а также экспрессии интерферона- $\gamma$  и гранзима В, ингибируя апоптоз CD8+Т-клеток, индуцированный стимуляцией анти-CD3/CD28. Было обнаружено, что мишенью miR-19b является ген PTEN [Yin и др., 2019].

Можно заключить, что у пациентов, выздоровевших после крайне тяжелой формы COVID-19, в плазме крови значимо повышен уровень hsa-miR-19b-3p|0. Данная микроРНК представлена в плазме крови в значимых количествах и способна связываться с регионами SARS-CoV-2, кодирующими белки, подавляющие внутриклеточные механизмы иммунитета. Кроме того, данная микроРНК способна стимулировать функциональную активность и пролиферацию цитотоксических Т-лимфоцитов – одного из ключевых компонентов адаптивного клеточного иммунитета против SARS-CoV-2. Результаты исследования могут быть использованы при разработке противовирусных препаратов на основе РНК-интерференции, а также при разработке прогностических тест-систем для оптимизации тактики лечения пациентов с COVID-19.

### **Разработка и проверка клинической значимости патогенетически обоснованного алгоритма диагностических мероприятий для прогноза тяжелого и крайне тяжелого течения COVID-19 на основе анализа генотипа ГКГС-I**

В сформированной выборке подавляющее большинство пациентов выздоровели после COVID-19 (Таблица 4). Летальный исход был отмечен у 10,3% пациентов. Необходимо отметить, что в период формирования выборки в московском регионе преобладал вариант AY.122, значительно отличающийся от исходного штамма «Wuhan-Hu-1» вируса SARS-CoV-2 [Klink и др., 2022].

Группы сравнения (Таблица 4) были сопоставимы по возрасту и половому составу. Ожидаемо в группе пациентов с тяжелым течением COVID-19 преобладали пациенты с объемным поражением лёгких ( $p = 5,8e-05$ ) и ожирением ( $p = 0,0002$ ). Также в группе с тяжелым течением COVID-19 зафиксирована высокая частота летальных исходов ( $p = 6,0e-06$ ).

Таблица 4 – Характеристика выборки пациентов с COVID-19

Показатель	Тяжелое течение	Среднетяжелое течение	<i>p</i>
Количество пациентов	69	76	
Возраст, медиана [25%-75%]	52 [41 – 66]	49 [41 – 57]	0,064
Пол (Муж./Жен.)	37/32	34/42	0,32
Максимальная степень поражения лёгких по КТ			
КТ-0	0	2 (2,6%)	5,8e-05
КТ-1	12 (17,4%)	18 (23,7%)	
КТ-2	12 (17,4%)	32 (42,1%)	
КТ-3	17 (24,6%)	14 (18,4%)	
КТ-4	14 (20,3%)	0	
н.д.	14 (20,3%)	10 (13,2%)	
Летальный исход	15 (21,7%)	0	6,0e-06
Ожирение	11 (16,0%)	0	0,0002
Гипертония	16 (23,2%)	22 (29,0%)	0,45

Были оценены чувствительность и специфичность «Индекса риска» при прогнозе летального исхода COVID-19 и тяжелого течения COVID-19 (Рисунок 8).

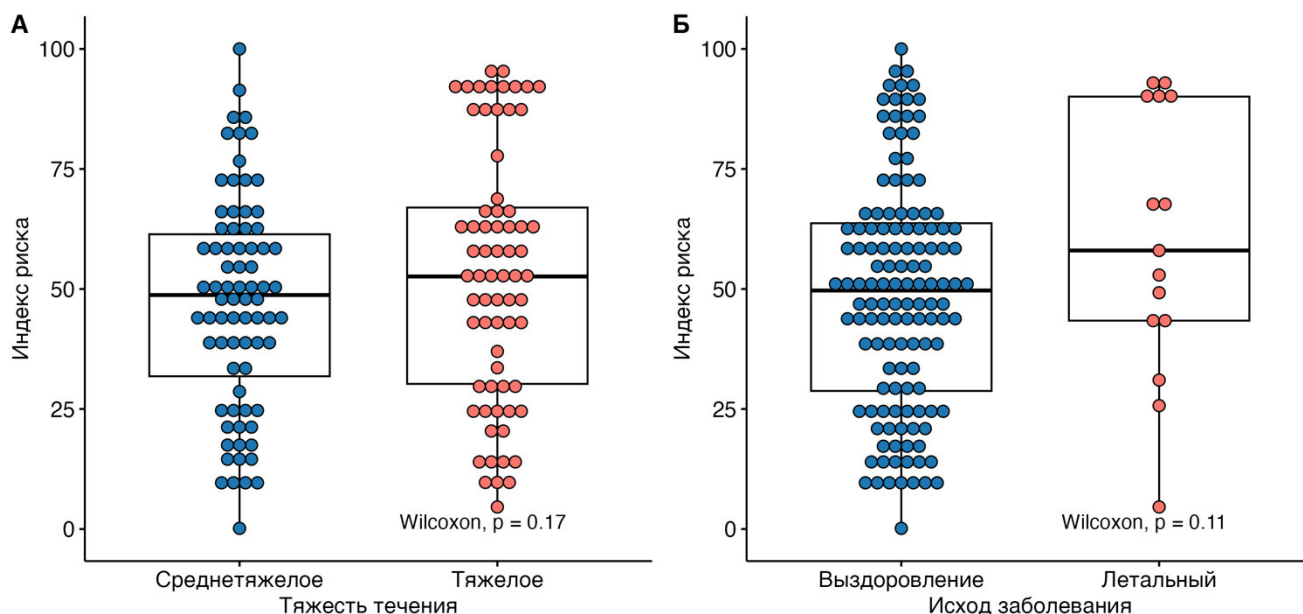


Рисунок 8 – Результаты оценки «Индекса риска»: А – пациенты с COVID-19 среднетяжелого и тяжелого течения; Б – выздоровевшие пациенты и пациенты с летальным исходом



Были оценены характеристики прогноза тяжести течения COVID-19 в зависимости от порогового уровня при интерпретации «Индекса риска». Так при пороговом уровне не менее 41 чувствительность составила 0,7, специфичность 0,36, ожидаемая доля ложных отклонений (FDR) – 0,51. Отношение шансов составило 1,25 ( $p = 0,6$ , 95% ДИ 0,59 – 2,69), что не позволяет рассматривать данный пороговый уровень в качестве значимого. При пороговом уровне не менее 89 чувствительность составила 0,14, специфичность 0,97, FDR – 0,17. ОШ составило 6,2 ( $p = 0,014$ , 95% ДИ 1,25 – 60,4).

Аналогичные характеристики были оценены для прогноза летального исхода COVID-19. При пороговом уровне не менее 41 чувствительность составила 0,8, специфичность 0,35, FDR – 0,88. ОШ составило 2,11 ( $p = 0,39$ , 95% ДИ 0,53 – 12,23), что не позволяет рассматривать данный пороговый уровень в качестве значимого. При пороговом уровне не менее 89 чувствительность составила 0,27, специфичность 0,94, FDR – 0,67. ОШ составило 5,44 ( $p = 0,023$ , 95% ДИ 1,03 – 24,7).

Таким образом была доказана эффективность разработанного на основе полученных в исследовании данных подхода к диагностике и прогнозу тяжести COVID-19.

Исходя из полученных результатов был предложен следующий алгоритм диагностических мероприятий для прогноза тяжелого и крайне тяжелого течения COVID-19 на основе анализа генотипа ГКГС-I:

1) в приемном отделении необходимо провести взятие образца периферической крови или букального соскоба для проведения исследования генов ГКГС-I пациента методом секвенирования следующего поколения;

2) для пациентов, поступивших в приёмные отделения, с подтвержденным COVID-19:

а) выполнить генетическое исследование генов ГКГС-I с целью определения варианта генотипа пациента;

б) в результате обработки ПО результатов генотипирования генов HLA-A, HLA-B, HLA-C пациенту присваивается так называемый «Индекс риска» – число в диапазоне от 0 (нуля) до 100, что пропорционально связано с повышением риска неблагоприятного течения заболевания;

в) если значение «Индекса риска» более 89, пациент относится к группе с высоким риском тяжелого течения COVID-19.

Можно заключить, что был разработан патогенетически обоснованный алгоритм диагностических мероприятий для прогноза тяжелого и крайне тяжелого течения COVID-19 на основе анализа генотипа ГКГС-I.

## ЗАКЛЮЧЕНИЕ

Значительные усилия мирового научного сообщества были направлены на решение вопроса о генетических детерминантах тяжести COVID-19. Полногеномное исследование ассоциаций было наиболее очевидным решением данной проблемы. Однако интерпретация выявленных генетических детерминант, связанных с COVID-19, остается сложной задачей, поскольку большинство генетических вариантов расположены в некодирующих регуляторных областях с высоким неравновесным сцеплением [Smemo и др., 2014; Kumar, Wijmenga, Withoff, 2012; Claussnitzer и др., 2015] и функциональны во временном и пространственном контекстах [The FANTOM Consortium и др., 2014].

Альтернативные подходы к поиску генетических детерминант восприимчивости к COVID-19 базировались на понимании механизмов формирования специфического иммунного ответа на SARS-CoV-2. Интенсивность и характер клеточного иммунного ответа на инфекцию SARS-CoV-2 обусловлены разнообразием как репертуара T-клеточных рецепторов, так и генотипом ГКГС-I, аллели которого определяют презентацию антигенов при заболевании, а вместе с тем и восприимчивость к заболеванию и исход вирусных инфекций [Zunec, 2020; MacDonald и др., 2000; Ochoa и др., 2020].

В рамках данной работы была разработана и оценена клиническая информативность способа оценки риска развития тяжелого течения COVID-19. Данный способ оценивает способность молекул ГКГС-I, кодируемых генами HLA-A и HLA-C, презентировать пептиды SARS-CoV-2.

Были оценены характеристики прогноза тяжести течения COVID-19 в зависимости от порогового уровня при интерпретации «Индекса риска». При пороговом уровне не менее 41 чувствительность составила 0,7, специфичность 0,36, FDR – 0,51. При пороговом уровне не менее 89 чувствительность составила 0,14, специфичность 0,97, FDR – 0,17. Отношение шансов составило 6,2 ( $p = 0,014$ , 95% ДИ 1,25 – 60,4).

Аналогичные характеристики были оценены для прогноза летального исхода COVID-19. При пороговом уровне не менее 41 чувствительность составила 0,8, специфичность 0,35, FDR – 0,88. Отношение шансов составило 2,11 ( $p = 0,39$ , 95% ДИ 0,53 – 12,23). При пороговом уровне не менее 89 чувствительность составила 0,27, специфичность 0,94, FDR – 0,67. Отношение шансов составило 5,44 ( $p = 0,023$ , 95% ДИ 1,03 – 24,7). Можно заключить, что разработанный на первом этапе работы способ оценки риска развития тяжелого течения COVID-19 на основе анализа генотипа ГКГС-I позволяет

прогнозировать тяжелое течение заболевания или летальный исход при пороговом уровне не менее 89.

В рамках исследования была определена штаммовая принадлежность вируса SARS-CoV-2 и изучена взаимосвязь генотипа ГКГС-I с тяжестью течения COVID-19, вызванной распространенными летом 2021 года вариантами штамма Дельта SARS-CoV-2: AY.122 и B.1.617.2. У пациентов в возрасте не старше 60 лет со среднетяжелым и тяжелым течением COVID-19 генотипирование ГКГС-I показало, что наиболее частыми аллелями в выборке были HLA-A\*02:01 (частота 0,26) и HLA-A\*01:01 (частота 0,25), что характерно для популяции Московского региона. Анализ влияния мутаций на число вирусных пептидов, взаимодействующих с индивидуальным набором молекул ГКГС-I с аффинностью менее 50 нМоль, показал, что различия достоверны для вспомогательных белков вируса. Так, у инфицированных вариантом AY.122 число высокоаффинных пептидов значительно сократилось по сравнению с инфицированными вариантом B.1.617.2 ( $p < 0,01$ ).

Значимое снижение числа высокоаффинных пептидов в варианте AY.122 прежде всего связано с мутацией G8R в белке NS8. Данная мутация вызывает падение аффинности взаимодействия пептидов FLGIITTV и FLVFLGIITTV с молекулой ГКГС-I, кодируемой наиболее частым для российской популяции аллелем HLA-A\*02:01. Особого внимания заслуживает тот факт, что белок NS8 (ORF8) способен подавлять созревание молекул ГКГС-I и их транслокацию на поверхность зараженной клетки [Matsuoka и др., 2022; Zhang и др., 2021]. Данные обстоятельства могли внести весомый вклад в превалирование варианта AY.122 в российской популяции и более тяжелое, по сравнению с другими вариантами штамма Дельта, течение вызванной им COVID-19.

Как уже упоминалось ранее, молекулы ГКГС-I являются одним из ключевых медиаторов первых шагов в развитии специфического иммунного ответа на COVID-19. В данном исследовании были проанализированы различия в генотипе ГКГС-I пациентов в возрасте не старше 60 лет, переболевших COVID-19 в первую и третью волну, оценено влияние мутаций в вирусе SARS-CoV-2 на иммуногенные эпитопы CD8<sup>+</sup> Т-лимфоцитов.

Установлено, что среди госпитализированных в третью волну пандемии COVID-19 пациентов с вирусной пневмонией, вызванной SARS-CoV-2, значительно снижено число носителей аллеля HLA-A\*01:01 ГКГС-I. Установлено, что носители аллеля HLA-A\*01:01 формируют Т-клеточный иммунитет к пептидам, кодируемым геном ORF1ab вируса SARS-CoV-2. Гены SARS-CoV-2 мутируют с разной скоростью. Ген ORF1ab наиболее консервативен среди них, а значит носители аллеля HLA-A\*01:01 могут длительное время сохранять Т-клеточный иммунитет.

Продемонстрирована более высокая частота встречаемости иммунодоминантных эпитопов из белков гена ORF1ab вируса SARS-CoV-2 у носителей HLA-A\*01:01 по сравнению с эпитопами этого гена у носителей HLA-A\*02:01 в когорте выздоровевших пациентов первой волны COVID-19.

Более того, анализ результатов одноклеточного фенотипирования Т-клеток у выздоровевших пациентов показал, что преобладающим фенотипом у носителей HLA-A\*01:01 являются Т-клетки центральной памяти. Преобладание Т-лимфоцитов данного фенотипа может способствовать формированию длительного Т-клеточного иммунитета у носителей данного аллеля и соответственно снижению заражаемости и тяжести заболевания.

Также в рамках данного исследования было оценено влияние мутаций на 25 иммунодоминантных эпитопов вируса SARS-CoV-2. Оказалось, что мутации, произошедшие с вирусом за более чем два года пандемии, затрагивают только четыре иммунодоминантных эпитопа (16 %).

В рамках данного исследования проверена гипотеза об эволюционном давлении на геном вируса SARS-CoV-2 микроРНК, характерных для ткани лёгких. В результате анализа было установлено, что вирус SARS-CoV-2 практически не имеет регионов связывания в 5'- и 3'-нетранслируемых областях с характерными для ткани лёгких микроРНК. Тем не менее вирус обладает значительным числом мест связывания с микроРНК в регионе NSP3-NSP5, ответственном за аутопротеолиз вирусных полипептидов и формирование вирионов. В вариантах штамма Омикрон произошло значимое снижение мест связывания с микроРНК клеток хозяина, что могло способствовать снижению патогенности данного штамма.

Нами продемонстрировано, что у пациентов, выздоровевших после крайне тяжелой формы COVID-19, в плазме крови значимо повышен уровень hsa-miR-19b-3p и hsa-miR-25-3p. Данные микроРНК представлены в плазме крови в значимых количествах, способны связываться с регионами SARS-CoV-2, кодирующими белки подавляющие внутриклеточные механизмы иммунитета. Кроме того, hsa-miR-19b-3p способна стимулировать функциональную активность и пролиферацию цитотоксических Т-лимфоцитов – одного из ключевых компонентов адаптивного клеточного иммунитета против SARS-CoV-2.

Можно заключить, что индивидуальные особенности генотипа ГКГС-I, а также микроРНК вносят значимый вклад в патогенез COVID-19 и формирование Т-клеточного иммунитета к нему (Рисунок 9).

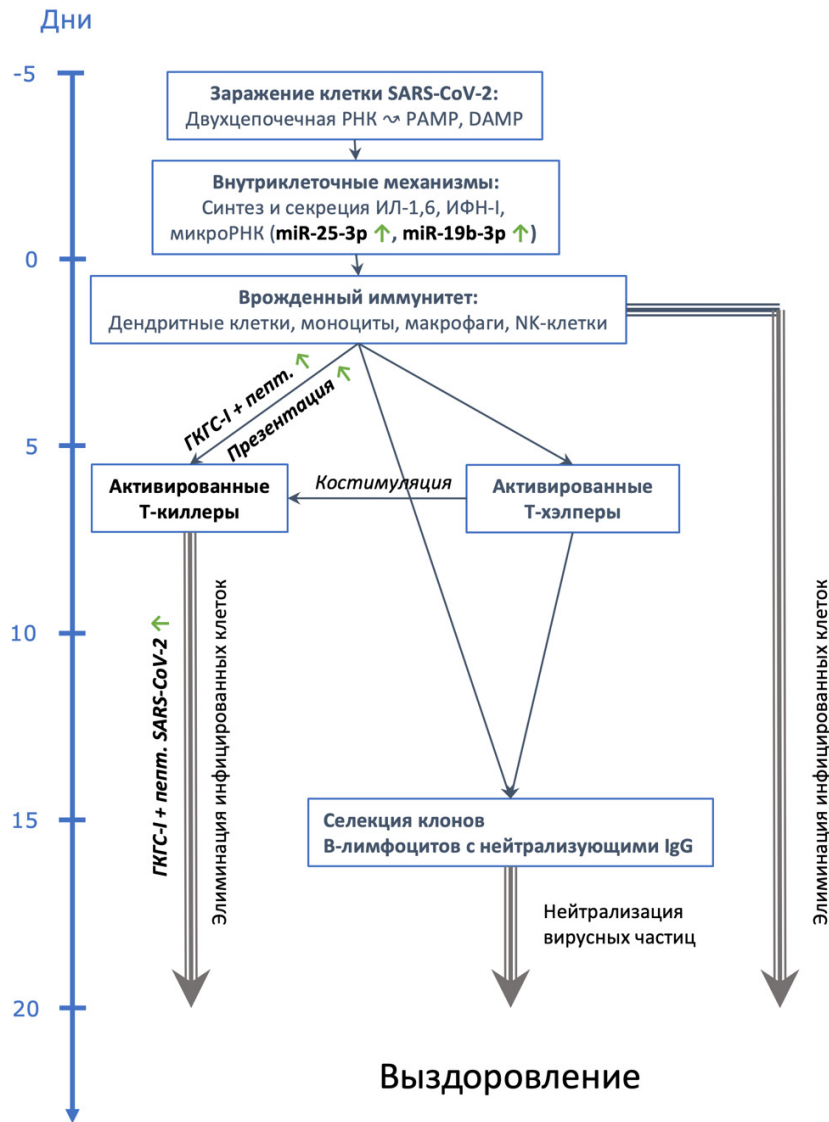


Рисунок 9 – Патогенетический вклад генотипа ГКГС-I в COVID-19 средней степени тяжести

Вклад в тяжелое и крайне тяжелое течение COVID-19 вносят следующие патогенетические факторы: мутации SARS-CoV-2, в том числе и под влиянием микроРНК клеток хозяина, приводящие к исчезновению иммунодоминантных для генотипа ГКГС-I пациента эпитопов, особенности экспрессии микроРНК и генотип ГКГС-I пациента, приводящие к слабому цитотоксическому Т-клеточному ответу на начальных этапах заболевания.

Совокупность этих факторов приводит к тому, что в промежутке с 3 – 4 до 14 – 16 дня заболевания возникает «иммунологический провал» в который с SARS-CoV-2 борются только компоненты врожденного иммунитета (Рисунок 10).

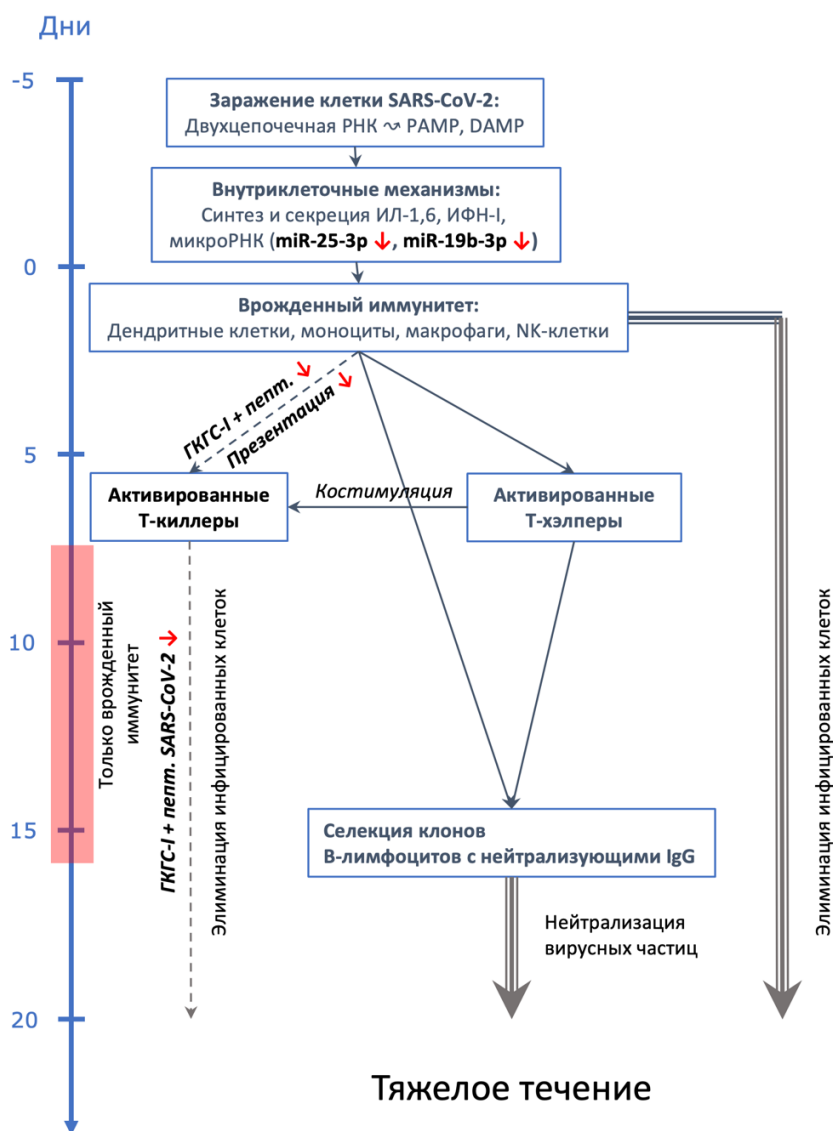


Рисунок 10 – Патогенетический вклад генотипа ГКГС-I в COVID-19 тяжелой и крайне тяжелой степени тяжести

## ВЫВОДЫ

1) У пациентов в возрасте до 60 лет генотип ГКГС-I связан с тяжестью течения COVID-19, вызванной вариантом AY.122 штамма Дельта SARS-CoV-2. Значимое снижение числа высокоаффинных пептидов в варианте AY.122 прежде всего связано с мутацией G8R в белке NS8. Данная мутация вызывает падение аффинности взаимодействия пептидов FLGIITTV (с 42 нМоль до 121 нМоль) и FLVFLGIITTV (с 25 нМоль до 59 нМоль) с молекулой ГКГС-I, кодируемой наиболее частым для российской популяции аллелем HLA-A\*02:01.

2) Среди госпитализированных в третью волну пандемии COVID-19 пациентов с вирусной пневмонией, вызванной SARS-CoV-2, значимо снижено число носителей аллеля HLA-A\*01:01 ГКГС-I ( $p = 0,025$  с поправкой на множественность сравнений, ОШ = 0,5). Установлено, что носители аллеля

HLA-A\*01:01 формируют Т-клеточный иммунитет преимущественно к пептидам, кодируемым геном ORF1ab вируса SARS-CoV-2.

3) Мутации практически не затрагивают пептиды вируса SARS-CoV-2, вызывающие CD8<sup>+</sup> Т-клеточный ответ у пациентов российской популяции. За два года пандемии из 25 иммунодоминантных пептидов вируса мутировали только четыре, что может способствовать длительному сохранению эффективного Т-клеточного иммунитета против SARS-CoV-2.

4) МикроРНК ткани лёгких влияют на эволюцию вируса SARS-CoV-2, вызывая снижение мест связывания вируса с микроРНК клеток хозяина, что значимо проявилось в случае со штаммом Омикрон.

5) Высокая экспрессия циркулирующих в плазме крови микроРНК hsa-miR-25-3p hsa-miR-19b-3p ассоциирована с благоприятным исходом тяжелого течения COVID-19.

6) Число высокоаффинных пептидов вируса SARS-CoV-2 для молекул ГКГС-I обратно пропорционально вероятности летального исхода заболевания у пациента. Специфичность прогноза тяжелого течения COVID-19 составляет 0,97, отношение шансов – 6,2 ( $p = 0,014$ , 95% ДИ 1,25 – 60,4). Специфичность прогноза летального исхода COVID-19 равна 0,94, отношение шансов – 5,44 ( $p = 0,023$ , 95% ДИ 1,03 – 24,7).

7) В результате проведенных исследований разработано и зарегистрировано программное обеспечение для прогноза тяжести и исхода течения вирусной пневмонии, вызванной SARS-CoV-2.

## ПРАКТИЧЕСКИЕ РЕКОМЕНДАЦИИ

1) При выборе тактики лечения пациента с вирусной пневмонией, вызванной SARS-CoV-2, следует руководствоваться результатами исследования генов ГКГС-I, позволяющими оценить возможную перспективу и тенденцию развития коронавирусной инфекции.

2) Поток пациентов, поступающих в ЛПУ с диагнозом COVID-19, может быть разделен на ранних этапах на группы пациентов с потенциально легким или тяжелым течением вирусной инфекции на основании значения «Индекса риска».

Для этого в приемном отделении необходимо провести взятие образца периферической крови или букального соскоба для проведения исследования генов ГКГС-I пациента методом секвенирования следующего поколения.

3) Можно рекомендовать проводить поступающим в приёмные отделения пациентам с подтвержденным COVID-19:

а) выполнить генетическое исследование генов ГКГС-I с целью определения варианта генотипа пациента;

б) в результате обработки ПО результатов генотипирования генов HLA-A, HLA-B, HLA-C пациенту присваивается так называемый «Индекс риска» – число в диапазоне от 0 (нуля) до 100, что пропорционально связано с повышением риска неблагоприятного течения заболевания;

в) если значение «Индекса риска» более 89, пациент относится к группе с высоким риском тяжелого течения COVID-19.

## СПИСОК РАБОТ, ОПУБЛИКОВАННЫХ ПО ТЕМЕ ДИССЕРТАЦИИ

### Работы, опубликованные в научных рецензируемых изданиях

- 1) **Шкурников, М.Ю.** Роль циркулирующей микроРНК miR-19b в прогнозе исхода COVID-19 / М.Ю. Шкурников, С.И. Колесников // *Acta Biomedica Scientifica (East Siberian Biomedical Journal)*. – 2023. – Т. 8, № 2. – С. 26-32. (Scopus)
- 2) HLA-A\*01:01 allele diminishing in COVID-19 patients population associated with non-structural epitope abundance in CD8+ T-cell repertoire / **M. Shkurnikov**, S. Nersisyan, D. Averinskaya [и др.] // *PeerJ*. – 2023. – Т. 11. – С. e14707. (Scopus)
- 3) Жиянов, А.П. Мутации SARS-CoV-2 приводят к снижению числа регионов связывания с микроРНК ткани лёгких / А.П. Жиянов, **М.Ю. Шкурников** // *Клеточные технологии в биологии и медицине*. – 2022. – № 4. – С. 216-221. (Scopus, Web of Science)
- 4) Различия в презентации пептидов вариантов Омикрон-штамма BA.1–BA.5 вируса SARS-CoV-2 молекулами ГКГС / С.А. Нерсисян, **М.Ю. Шкурников**, А.П. Жиянов [и др.] // *Доклады Российской академии наук. Науки о жизни*. – 2022. – Т. 507, № 1. – С. 525-529. (Scopus, ядро РИНЦ)
- 5) Взаимосвязь генотипа главного комплекса гистосовместимости класса I с летальностью COVID-19 у пациентов с сахарным диабетом / **М.Ю. Шкурников**, Д.А. Аверинская, А.Г. Комаров [и др.] // *Доклады Российской академии наук. Науки о жизни*. – 2022. – Т. 507, № 1. – С. 515-519. (Scopus, ядро РИНЦ)
- 6) Взаимосвязь прогноза течения COVID-19 с мутациями белка NS8 SARS-CoV-2 в зависимости от штаммовой принадлежности вируса / **М.Ю. Шкурников**, Д.А. Аверинская, А.Г. Комаров [и др.] // *Доклады Российской академии наук. Науки о жизни*. – 2022. – Т. 507, № 1. – С. 460-464. (Scopus, ядро РИНЦ)
- 7) T-CoV: a comprehensive portal of HLA-peptide interactions affected by SARS-CoV-2 mutations / S. Nersisyan, A. Zhiyanov, **M. Shkurnikov** [и др.] // *Nucleic Acids Research*. – 2022. – Т. 50. – С. D883–D887. (Scopus, Web of Science)



- 8) Alterations in SARS-CoV-2 Omicron and Delta peptides presentation by HLA molecules / S. Nersisyan, A. Zhiyanov, **M. Shkurnikov** [и др.] // PeerJ. – 2022. – Т. 10. – С. e13354. (Scopus)
- 9) Immunogenic epitope panel for accurate detection of non-cross-reactive T cell response to SARS-CoV-2 / A. Titov, R. Shaykhutdinova, **M. Shkurnikov** [и др.] // JCI Insight. – 2022. – Т. 7, № 9. – С. e157699. (Scopus, Web of Science)
- 10) Extracellular miRNAs and Cell–Cell Communication: Problems and Prospects / J. Makarova, A. Turchinovich, **M. Shkurnikov**, A. Tonevitsky // Trends in Biochemical Sciences. – 2021. – Т. 46, №. 8. – С. 640-651. (Scopus, Web of Science)
- 11) Association of HLA Class I Genotypes With Severity of Coronavirus Disease-19 / **M. Shkurnikov**, S. Nersisyan, A. Galatenko [и др.] // Frontiers in Immunology. – 2021. – Т. 12. – С. 641900. (Scopus, Web of Science)
- 12) Integrative analysis of miRNA and mRNA sequencing data reveals potential regulatory mechanisms of ACE2 and TMPRSS2 / S. Nersisyan, E. Knyazev, **M. Shkurnikov** [и др.] // PLoS ONE. – 2020. – Т. 15, №. 7. – С. e0235987. (Scopus, Web of Science)
- 13) Роль регуляции генов АПФ2/TMPRSS2 изоформами микроРНК кишечника в патогенезе COVID-19 / С.А. Нерсисян, **М.Ю. Шкурников**, А.И. Осипьянц, В.И. Вечорко // Вестник РГМУ. – 2020. – №2. – С. 17–20. (Scopus, Web of Science)
- 14) Intracellular and extracellular microRNA: An update on localization and biological role / Yu.A. Makarova, **M.Yu. Shkurnikov**, D. Wicklein [и др.] // Progress in Histochemistry and Cytochemistry. – 2016. – Т. 51, №. 3-4. – С. 33-49. (Scopus, Web of Science)
- 15) Изменение уровня циркулирующих микроРНК hsa-miR-297 и hsa-miR-19b-3p связано с генерализацией рака предстательной железы / А.И. Осипьянц, Е.Н. Князев, **М.Ю. Шкурников** [и др.] // Бюллетень экспериментальной биологии и медицины. – 2016. – Т. 162, № 9. – С. 366-369. (Scopus, Web of Science)
- 16) Внеклеточные микроРНК (обзор) / Ю.А. Макарова, **М.Ю. Шкурников**, А.А. Турчинович [и др.] // Биохимия. – 2015. – Т. 80, № 9. – С. 1344-1355. (Scopus, Web of Science)
- 17) Профиль микроРНК плазмы крови здоровых людей / **М.Ю. Шкурников**, Ю.А. Макарова, Е.Н. Князев [и др.] // Бюллетень экспериментальной биологии и медицины. – 2015. – Т. 160, № 11. – С. 577-579. (Scopus, Web of Science)
- 18) Определение микроРНК плазмы крови, ассоциированных с гемолизом / **М.Ю. Шкурников**, Е.Н. Князев, К.А. Фомичева [и др.] // Бюллетень

экспериментальной биологии и медицины. – 2015. – Т. 160, № 12. – С. 709-711. (Scopus, Web of Science)

### **Методические рекомендации**

- 19) Тактика поведения врача на основании результатов молекулярно-генетических исследований при обследовании пациентов, инфицированных коронавирусной инфекцией: методические рекомендации: Методические рекомендации. – Москва: Государственное бюджетное учреждение города Москвы «Научно-исследовательский институт организации здравоохранения и медицинского менеджмента Департамента здравоохранения города Москвы», 2022. – 32 с.

### **Патенты на изобретения**

- 20) **Шкурников М.Ю.**, А.В. Галатенко Патент № 2777487 С1 Российская Федерация, МПК G01N 33/50, C12Q 1/6804, A61P 11/00. Способ оценки риска необходимости применения кислородной терапии при COVID-19: № 2021136589: заявл. 11.12.2021: опубл. 04.08.2022.
- 21) **Шкурников М.Ю.**, С.А. Нерсисян, А. Г. Тоневицкий. Патент № 2751410 С1 Российская Федерация, МПК А61В 5/00, G01N 33/569, C12Q 1/6858. Способ оценки риска развития тяжелой формы COVID-19: № 2021104170: заявл. 18.02.2021: опубл. 13.07.2021.
- 22) **Шкурников М.Ю.**, С.А. Нерсисян, А.Г. Тоневицкий. Патент № 2747097 С1 Российская Федерация, МПК G01N 33/569, C12Q 1/6858, C12Q 1/6874. Способ оценки риска развития тяжелой формы COVID-19: № 2021103551: заявл. 14.02.2021: опубл. 26.04.2021.

### **Свидетельства о государственной регистрации программы для ЭВМ**

- 23) **Шкурников М.Ю.** Программное обеспечение «Аффисенс» для клинической оценки риска развития тяжелой формы течения коронавирусной инфекции COVID-19. Свидетельство о регистрации программы для ЭВМ № 2023619424, 11.05.2023.

Подписано в печать «10» ноября 2023 г. Формат 60x84/16  
Бумага офсетная. Печать офсетная.  
Усл.авт.л.1,0.

Тираж 150 экз. Заказ №4. Типография НИУ ВШЭ.  
105187, г.Москва, Измайловское шоссе, д.44, стр.2.