

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
имени М.В. ЛОМОНОСОВА

На правах рукописи

Вакуленко Юлия Александровна

**Закономерности рекомбинации и эволюции у ряда социально
значимых (+)РНК-вирусов млекопитающих**

1.5.10 – вирусология

АВТОРЕФЕРАТ

диссертации на соискание ученой степени

кандидата биологических наук

Москва – 2023

Диссертация подготовлена на кафедре вирусологии биологического факультета МГУ имени М.В. Ломоносова и в Научно-исследовательском отделе института медицинской паразитологии, тропических и трансмиссивных заболеваний им. Е.И. Марциновского ФГАОУ ВО Первый МГМУ им. И.М.Сеченова Минздрава России (Сеченовский Университет).

Научный руководитель – *Лукашев Александр Николаевич, д.м.н., член-корр. РАН*

Официальные оппоненты:

Альховский Сергей Владимирович, д.б.н., член-корр. РАН, заведующий лабораторией биотехнологии ФГБУ "Национальный исследовательский центр эпидемиологии и микробиологии имени почётного академика Н.Ф. Гамалеи" Министерства здравоохранения РФ

Тикунова Нина Викторовна, д.б.н., доцент, заведующая лабораторией молекулярной микробиологии ФГБУ науки Института химической биологии и фундаментальной медицины Сибирского отделения Российской академии наук (ИХБФМ СО РАН)

Литов Александр Геннадьевич, к.б.н., старший научный сотрудник лаборатории биологии арбовирусов ФГАНУ «Федеральный научный центр исследований и разработки иммунобиологических препаратов им. М.П. Чумакова РАН» (Институт полиомиелита)

Защита диссертации состоится «27» апреля 2023 г. в 16 часов 00 минут на заседании диссертационного совета МГУ.015.4 Московского государственного университета имени М.В. Ломоносова по адресу: 119234, Москва, Ленинские горы МГУ, д. 1, стр. 12, биологический факультет, ауд. 389. E-mail: mgu.03.01.dissovet@gmail.com.

С диссертацией можно ознакомиться в отделе диссертаций научной библиотеки МГУ имени М.В. Ломоносова (Ломоносовский просп., д. 27) и на портале: <https://dissovet.msu.ru/dissertation/015.4/2404>.

Автореферат разослан «___» _____ 2023 г.

Ученый секретарь
диссертационного совета,
доктор биологических наук



Т.В. Комарова

ВВЕДЕНИЕ

1. Актуальность

Вирусы млекопитающих с РНК-геномом положительной полярности относительно мРНК, или (+)РНК-вирусы, вызывают широкий спектр заболеваний человека и животных, например, полиомиелит, гастроэнтерит, болезни органов дыхания, ящур и др. Кроме того, (+)РНК-вирусы являются частым источником новых инфекций, таких как менингоэнцефалит, вызываемый энтеровирусом 71 типа, SARS, MERS и COVID-19. Рекомбинация широко распространена у (+)РНК-вирусов и, наряду с высокой скоростью накопления замен, является фактором, обеспечивающим их генетическое разнообразие. Несмотря на то, что большое количество исследований посвящено рекомбинации и эволюции конкретных патогенов человека и животных, знания об эволюции (+)РНК-вирусов в целом остаются фрагментарными. Последний систематический анализ рекомбинации у пяти основных семейств (+)РНК-вирусов был опубликован в 2006 году. За последние пятнадцать лет количество вирусных последовательностей в базе данных GenBank выросло в сотни раз, что позволяет провести комплексный анализ на новом уровне. Рекомбинация может привести к более быстрому по сравнению с накоплением мутаций изменению фенотипа вируса, что позволяет вирусам менять клеточный тропизм, хозяев и осваивать новые экологические ниши, поэтому анализ закономерностей рекомбинации необходим для понимания закономерностей возникновения новых вирусов в частности и принципов существования генетической информации в мире РНК-вирусов в целом.

Одной из наиболее актуальных групп (+)РНК-вирусов как источника возникающих заболеваний являются энтеровирусы человека. Как правило, энтеровирусная инфекция проходит без клинических проявлений, но иногда энтеровирусы могут вызывать тяжелые заболевания, такие как полиомиелит, серозный менингит, энцефалит, сепсис-подобное заболевание новорожденных и др. Регулярно открывают новые серотипы, раз в несколько десятилетий возникают новые клинические формы энтеровирусной инфекции, однако понимания механизмов происхождения и возможности исчезновения серотипов на данный момент нет.

2. Степень разработанности темы

Рекомбинация была описана практически у всех (+)РНК-вирусов. Большинство работ сосредоточено на очень узких группах вирусов, как правило, уровня рода или вида. Аналогии между профилями рекомбинации у разных семейств проводились очень редко и достаточно давно (Simmonds & Welch, 2006), отсутствуют систематические исследования рекомбинации у разных семейств с использованием идентичной методологии.

Мощным инструментом для решения фундаментальных и прикладных задач вирусологии являются Байесовские филогенетические методы, которые вошли в практику около 15 лет назад (Nascimento et al., 2017). Байесовские методы позволяют восстановить филогенетическую историю вирусов во времени и пространстве (Baele et al., 2013; Russel et al., 2019). Вопрос времени возникновения современного

разнообразие вирусов также остается открытым. Для некоторых типов энтеровирусов (из более чем 100) было показано, что общий предок вирусов, принадлежащих одному типу, возник не более 150 лет назад (Savolainen-Kopra et al., 2011; Tee et al., 2010). Систематического анализа динамики и закономерностей эволюции типов энтеровирусов в глобальном масштабе не проводилось. Кроме того, требовалось исключить влияние артефактов анализа на результаты молекулярного датирования событий в эволюции энтеровирусов.

Таким образом, имеется систематический пробел в наших знаниях о генетике циркулирующих вирусов.

3. Цели и задачи

Цель работы – провести анализ генетического разнообразия, изменчивости и рекомбинации у ряда социально значимых (+)РНК-вирусов млекопитающих, принадлежащих семействам *Picornaviridae*, *Caliciviridae*, *Astroviridae* и *Coronaviridae*.

Задачи:

- 1) Изучить закономерности естественной рекомбинации у вирусов, принадлежащих семействам *Caliciviridae*, *Astroviridae* и *Coronaviridae* на разных таксономических уровнях и с учетом динамики во времени.
- 2) Изучить принципы происходящей за счет рекомбинации модульной эволюции геномов вирусов, принадлежащих семействам *Caliciviridae*, *Astroviridae* и *Coronaviridae*.
- 3) Систематически изучить филодинамику наиболее распространенных типов неполиомиелитных энтеровирусов человека видов А-D.
- 4) Определить нуклеотидную последовательность типизирующего фрагмента VP1 и изучить филогенетические взаимоотношения изолятов коксакивируса А2 (CVA2), выделенных от пациентов с острыми вялыми параличами, в сопоставлении с другими последовательностями CVA2, доступными в базе GenBank.

Для достижения целей работы также было необходимо решить следующие вспомогательные задачи:

- 1) Разработать инструментарий для автоматизации работы с нуклеотидными и аминокислотными последовательностями вирусов.
- 2) Разработать инструментарий для детекции и визуализации рекомбинационных событий в условиях частой рекомбинации, разрушающей филогенетический сигнал.
- 3) Оценить влияние артефактов в геномных данных и их метаданных на результаты молекулярного датирования событий в эволюции энтеровирусов.

4. Научная новизна

Проведен систематический анализ рекомбинации в семействах (+)РНК-вирусов *Caliciviridae*, *Astroviridae* и *Coronaviridae*. Показано, что профили рекомбинации внутри участков генома, кодирующих структурные и неструктурные белки, различаются между семействами, но сходны внутри одного семейства РНК-вирусов.

Для калицивирусов характерен обмен целыми блоками генов неструктурных и структурных белков и крайне редкая рекомбинация внутри участков, кодирующих структурные и неструктурные белки. Для астровирусов характерен обмен целыми блоками генов неструктурных и структурных белков и умеренная рекомбинация внутри структурных и неструктурных генов. У коронавирусов наиболее часто рекомбинация происходит между геном структурного белка S и другими участками генома. Рекомбинация внутри гена белка S происходит чаще, чем в неструктурной области, причем с большей вероятностью происходит обмен целыми доменами белка S. Гены остальных структурных белков (N, E, M) по профилю естественной рекомбинации соответствуют генам неструктурных белков.

Впервые выполнено систематическое изучение филодинамики типов (фрагментов генома, кодирующих капсидные белки) энтеровирусов человека в глобальном масштабе. Показано, что глобальная популяция энтеровирусов разных типов регулярно подвергалась «бутылочным горлышкам», что приводило к вымиранию линий. Отдельные типы могут регулярно подвергаться почти полному исчезновению практически в глобальном масштабе даже в течение десятков лет, однако не зафиксировано ни одного примера безвозвратного исчезновения типа энтеровируса. Определены последовательности типизирующих фрагментов VP1 коксакивирусов A2, ассоциированных со случаями острого вялого паралича, выделенных в рамках надзора за неполиомиелитными энтеровирусами с 2001 по 2020 год. Опровергнута гипотеза о возникновении варианта коксакивируса A2 с измененными патогенетическими характеристиками.

5. Теоретическая и практическая значимость

Теоретическая значимость работы заключается в том, что были выявлены особенности естественной рекомбинации трех семейств (+)РНК-вирусов – *Caliciviridae*, *Astroviridae* и *Coronaviridae*, а также закономерности эволюции в глобальном масштабе участков, кодирующих капсидный белок VP1 неполиомиелитных энтеровирусов (семейство *Picornaviridae*).

Разработаны и опубликованы в журнале «Инфекция и иммунитет» НИИ эпидемиологии и микробиологии имени Пастера Роспотребнадзора практические рекомендации по подготовке и аннотации последовательностей геномов вирусов для филогенетического анализа.

6. Методология и методы исследования

При выполнении исследований использовались специализированные программы, для некоторых задач были разработаны программные сценарии на языках R и Python.

Для получения информации о вирусных последовательностях из записей GenBank, выделения необходимых участков полногеномных вирусных последовательностей, полуавтоматического формирования выравниваний последовательностей полных геномов вирусов автором были разработаны скрипты на

языке Python. Для множественного выравнивания последовательностей использовалась программа mafft v.7.304.

Оценка временного сигнала в наборах вирусных последовательностей осуществлялась в программе TempEst v.1.5.3. Филогенетический анализ нуклеотидных последовательностей вирусов был выполнен с помощью программ IQ-TREE v.1.6.12, BEAST v.1.8.3 и v.1.10.4. Для визуализации филогенетических деревьев использовались возможности программы FigTree v.1.4.4 и пакета ggtree (язык R). Визуализация результатов Байесовского филогенетического анализа для разных выборок энтеровирусных последовательностей осуществлялась средствами пакета matplotlib языка Python.

Для анализа рекомбинации в вирусных последовательностях использовались общепринятые методы, реализованные в программе RDP4. Визуализация результатов программы RDP4 была осуществлена с помощью пакета ggplot2 языка R. Также автором был реализован метод для детекции и визуализации рекомбинации с помощью матриц отклонений попарных расстояний и графиков соответствия попарных генетических расстояний. Метод был имплементирован на языке R в качестве пакета RecDplot и интерактивного приложения ShinyApp.

Статистические тесты выполнялись в программе GraphPad Prism v.8.

Последовательности типизирующего фрагмента энтеровирусов типа коксакивирус A2, выделенных от пяти пациентов с острыми вялыми параличами, были определены с помощью секвенирования по Сэнгеру.

7. Положения, выносимые на защиту

1. Общим свойством эволюции вирусов с несегментированным (+)РНК-геномом семейств *Caliciviridae*, *Astroviridae*, *Coronaviridae* является высокая частота рекомбинации между участками, кодирующими структурные и неструктурные белки. Профили рекомбинации внутри участков генома, кодирующих структурные и неструктурные белки, различаются между семействами, но сходны внутри одного семейства вирусов.
2. У коронавирусов наиболее часто рекомбинация происходит между геном структурного белка S и другими участками генома. Рекомбинация внутри гена белка S происходит чаще, чем в неструктурной области, причем с большей частотой происходит обмен целыми доменами белка S.
3. Участки генома, кодирующие капсидные белки энтеровирусов, могут регулярно проходить «бутылочное горлышко», т.е. подвергаться почти полному исчезновению и замещаться новыми вариантами практически в глобальном масштабе даже в течение десятков лет.
4. Выявлено отсутствие связи вирусов типа коксакивирус A2, выделенных от больных с острыми вялыми параличами, с конкретной филогенетической группой, что опровергает гипотезу о возникновении варианта вируса с измененными патогенетическими характеристиками.

8. Степень достоверности

Достоверность результатов определяется применением общепринятых методов обработки и анализа геномных данных, автоматизированных методов подготовки данных, исключая «человеческий фактор», статистических критериев при анализе данных, проведением симуляционных исследований для оценки влияния возможных ошибок в исходных данных на результат. Достоверность результатов и выводов также подтверждена процедурой независимого рецензирования публикаций по теме диссертации, согласованностью с независимыми исследованиями в данной области, значительным уровнем цитирования публикаций.

9. Объект и предмет исследования

Объект исследования – нуклеотидные последовательности геномов вирусов, принадлежащих семействам *Picornaviridae*, *Caliciviridae*, *Astroviridae*, *Coronaviridae*; энтеровирусы типа коксакивирус А2, выделенные от больных с острыми вялыми параличами.

Предмет исследования – закономерности естественной рекомбинации представителей семейств *Caliciviridae*, *Astroviridae* и *Coronaviridae*, эволюция участков, кодирующих капсидные белки энтеровирусов (семейство *Picornaviridae*).

10. Связь с государственными программами

Диссертационная работа выполнена при поддержке грантов РФФИ 19-15-00055 (руководитель А.Н. Лукашев), РФФИ 19-115-50403 (руководитель А.А. Девяткин), РФФИ 22-15-00230 (руководитель А.Н. Лукашев).

11. Личный вклад

Соискателем были получены основные результаты представленной работы. Соискателем были выполнены экспериментальные исследования по определению и анализу частичной нуклеотидной последовательности геномов коксакивирусов А2. Соискателем были разработаны методы детекции и визуализации рекомбинации, используемые в данной работе. Интерпретация результатов, представление результатов на научных конференциях и подготовка их к публикации в рецензируемых научных журналах также составляют личный вклад автора.

12. Публикации

По материалам диссертации опубликовано 5 экспериментальных и 2 обзорные статьи в журналах, индексируемых в RSCI, WoS или Scopus.

13. Апробация работы

Апробация основных результатов диссертации проведена в 2022 году в рамках защиты научно-квалификационной работы, подготовленной соискателем при обучении в аспирантуре биологического факультета МГУ. Результаты работы были представлены соискателем на российских и международных конференциях «Europic», 2018, Эгмонд-ан-Зе (устный доклад); «14th International Conference on Molecular Epidemiology and Evolutionary Genetics of Infectious Diseases», 2018, Ситжес (устный доклад); «II Student Conference Life Sciences in the 21st Century: Looking into the Future

(SCLS 2019)», 2019, Москва (устный доклад); «Студенческая конференция», 2021, Москва (стендовый доклад). Основные результаты диссертации были опубликованы в 7 статьях в научных рецензируемых журналах, входящих в Q1, Q2 и Q4. Публикации по теме диссертации проходили процедуру независимого рецензирования.

14. Структура и объем диссертации

Работа состоит из введения, обзора литературы, описания материалов и методов, результатов и их обсуждения, выводов и списка цитируемой литературы, содержащего 348 ссылок. Работа изложена на 200 страницах текста, содержит 9 таблиц и 48 рисунков.

СОДЕРЖАНИЕ РАБОТЫ

МАТЕРИАЛЫ И МЕТОДЫ ИССЛЕДОВАНИЯ

Данные

Подготовка выравненных полных геномов для анализа рекомбинации

Для анализа рекомбинации все доступные полные геномные последовательности вирусов, принадлежащих семействам *Caliciviridae* (рода *Norovirus* (N=3439), *Lagovirus* (N=507), *Sapovirus* (N=216)), *Astroviridae* (род *Mamastrovirus* (N=382)) и *Coronaviridae* (рода *Alphacoronavirus* (N=1260), *Betacoronavirus* (N=17083), *Gammacoronavirus* (N=452), *Deltacoronavirus* (N=162)) были выгружены из базы данных нуклеотидных последовательностей GenBank (апрель 2019-август 2020). Затем с помощью разработанного автором набора скриптов на языке Python были подготовлены выравнения нуклеотидных последовательностей исследуемых вирусов. Для астровирусов и калицивирусов были подготовлены выравнения слитых нуклеотидных открытых рамок считывания (ОРС), а для коронавирусов – выравнения слитых нуклеотидных последовательностей пяти основных генов (полная ORF1ab, S, E, M и N), которые присутствуют в геномах всех родов коронавирусов. После фильтрации и оптимизации данных в итоговых выравнениях содержалось 1084 последовательности для рода *Norovirus*, 338 для рода *Lagovirus*, 148 для рода *Sapovirus*, 319 для рода *Mamastrovirus*, 164 для рода *Alphacoronavirus*, 122 для рода *Alphacoronavirus*, 260 для рода *Gammacoronavirus* и 56 для рода *Deltacoronavirus*.

Изоляты CVA2, выделенные от больных с острыми вялыми параличами

Впервые были исследованы 5 изолятов энтеровирусов типа коксакивирус А2 (CVA2), выделенных на территории Российской Федерации в 2001-2020 годах от больных с острыми вялыми параличами (ОВП). Материалы для исследования были любезно предоставлены Ивановой О.Е. (ФГАУ «Федеральный научный центр исследований и разработки иммунобиологических препаратов им. М.П. Чумакова РАН» (Институт полиомиелита))

Подготовка выравниваний для филогенетического анализа типов энтеровирусов

Для анализа были выбраны 44 типа энтеровирусов человека, для каждого из которых было доступно более 50 последовательностей, кодирующих капсидный белок VP1, в базе данных GenBank. Автоматическая подготовка выравниваний полного и типизирующего фрагментов VP1 энтеровирусов была выполнена с использованием программы, разработанной автором (доступна по ссылке <https://github.com/v-julia/GenAlignment>). Для типа CVA2 в выборку последовательностей типизирующего фрагмента VP1 были добавлены три последовательности вирусов, выделенных от пациентов с ОВП, клиническая картина которых соответствовала паралитическому полиомиелиту. Эти последовательности были определены в рамках диссертационной работы.

Определение нуклеотидной последовательности типизирующего фрагмента VP1 изолятов CVA2, выделенных от больных с ОВП

Типизирующие фрагменты VP1 изолятов, выделенных от пяти пациентов с ОВП, были отсеквенированы согласно (Oberste et al., 1999). Последовательность типизирующего фрагмента VP1 была амплифицирована с помощью полимеразной цепной реакции (ПЦР) с обратной транскрипцией со специфическими праймерами к типизирующему фрагменту. Продукты амплификации были визуализированы с помощью электрофореза в 1% агарозном геле на трис-ацетатном буфере и окрашены бромистым этидием. Полоски, соответствующие ПЦР-продуктам, были вырезаны из геля, очищены набором «QIAquick Gel Extraction kit» (фирма QIAGEN). Определение нуклеотидной последовательности ПЦР-продуктов осуществлялось методом Сэнгера на секвенаторе Applied Biosystems 3100.

Анализ рекомбинации

Для анализа общих закономерностей рекомбинации были рассчитаны матрицы филогенетической совместимости, реализованные в программе RDP4, а также матрицы отклонения попарных расстояний (матрицы ОНР), реализованные автором работы в виде пакета и веб-приложения на языке R (Рис. 1).

В рамках работы был также реализован подход для обнаружения и визуализации рекомбинационных событий – график соответствия попарных генетических расстояний (график СППР) (Рис. 1б). При отсутствии рекомбинации попарные генетические расстояния между парами последовательностей в двух разных регионах следуют линейной зависимости (Рис. 1б, ген 1 против гена 2). В случае рекомбинации, попарные расстояния, рассчитанные по рекомбинантной и нерекомбинантной областям, будут отклоняться от линии регрессии (Рис. 1б, ген 1 против гена 3, ген 2 против гена 3). Среднеквадратическое отклонение (СКО) пар попарных расстояний от линии регрессии характеризует рекомбинацию между двумя участками генома в целом, что позволяет построить матрицу отклонений попарных расстояний для оценки уровня рекомбинации по геному.

Детекция единичных рекомбинационных событий также проводилась с помощью 9 методов, реализованных в программе RDP4: RDP, GENECONV, Bootscan, Maxchi, Chimaera, SiScan, PhylPro, LARD, 3Seq.

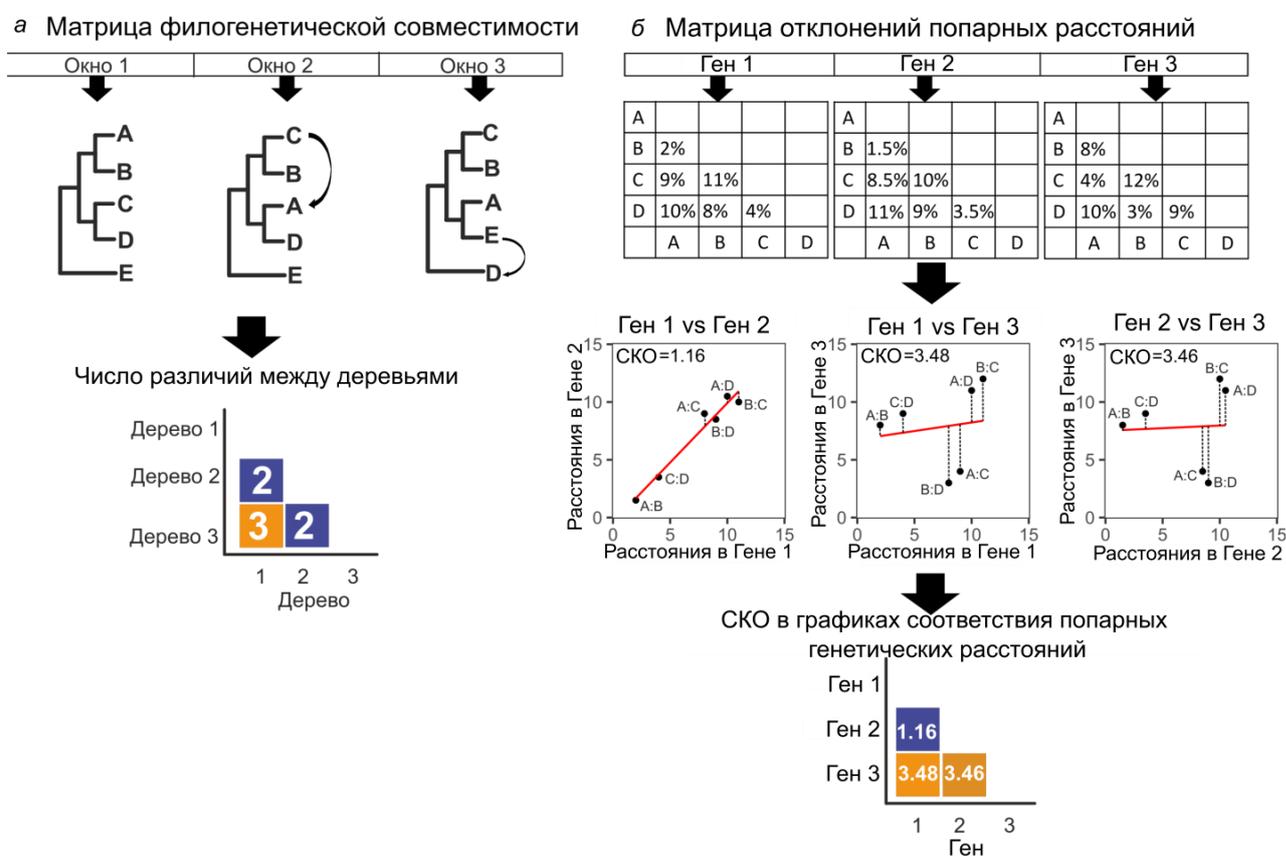


Рисунок 1. Принципы методов обнаружения рекомбинации. (а) Иллюстрация алгоритма построения матрицы филогенетической совместимости. Выравнивание последовательностей разделяется на перекрывающиеся окна определенной длины. Для каждого окна в выравнивании строится филогенетическое дерево, затем количество различий между полученными деревьями рассчитывается и сохраняется в матрице. Количество различий нормализуют и визуализируют в виде тепловой карты. (б) Алгоритм построения графиков соответствия попарных генетических расстояний (СПГР) и матриц отклонений попарных расстояний (ОПР). Для построения графика СПГР генетические расстояния между парой последовательностей в области генома 1 (Ген 1) и области генома (Ген 2) отмечаются на осях x и y соответственно. Общее отклонение всех попарных расстояний от линии регрессии оценивается как среднеквадратичное отклонение (СКО). СКО на графиках для всех возможных пар областей генома визуализируются в виде тепловой карты для построения матрицы ОПР.

Филогенетический анализ

Построение филогенетических деревьев максимального правдоподобия осуществлялось с помощью программы IQ-TREE. Поддержки узлов деревьев рассчитывались с помощью сверхбыстрой аппроксимации бутстрепа UFBoot. Выбор модели нуклеотидных замен осуществлялся с помощью модуля ModelFinder программы IQ-TREE.

Наличие временного сигнала было проверено с помощью программы TempEst v.1.5.3. Построение Байесовских филогенетических деревьев с разверткой по

времени осуществлялось в программе BEAST v.1.8.3, v.1.10.4. Выбор лучшей комбинации модели молекулярных часов и ветвления деревьев был выполнен на основании значения коэффициента Байеса.

Визуализация филогенетических деревьев проводилась с помощью программы FigTree v1.4.4 и пакета ggtree языка программирования R.

Для статистического анализа была использована программа GraphPad Prism.

Вычисление времени полужизни рекомбинантных форм норовирусов

Для оценки времен полужизни рекомбинантных форм норовирусов было использовано два подхода. Первый подход заключался в построении деревьев по нуклеотидным последовательностям участков генома VP1 и РНК-зависимой РНК-полимеразы (RdRp) с разверткой по времени и вычислении медианы возрастов совпадающих поддеревьев. В рамках второго подхода для всех пар последовательностей с одинаковым генотипом VP1 были вычислены расстояния между VP1 методом Maximum Composite Likelihood (MCL- расстояния) в программе MEGA v.7, а затем рассчитаны доли сравнений, в которых изоляты имели одинаковые генотипы по VP1 и разные Р-типы (генотипы по RdRp) для диапазонов MCL-дистанций. С помощью регрессионного анализа, выполненного средствами модуля sklearn языка Python, было вычислено MCL расстояние, которое соответствует 50% рекомбинантных форм среди всех сравнений пар последовательностей. Период полужизни рекомбинантных форм был вычислен путем умножения этого значения расстояния MCL на скорость накопления замен в гене VP1, полученной с помощью BEAST (4.34×10^{-3} замен/сайт/год).

РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

Метод для детекции и визуализации рекомбинации

Общепринятым методом анализа общей распространенности рекомбинации в геноме является визуализация пропорции консервативных филогенетических групп, реализованная в матрицах филогенетической совместимости (Simmonds & Welch, 2006) и сканировании TreeOrder (Simmonds & Midgley, 2005). Недостатком этого метода является то, что он не показывает «глубину» событий рекомбинации, то есть насколько близкородственны вирусы, между которыми проходила рекомбинация. Также в случае многочисленных событий рекомбинации между сравниваемыми фрагментами филогенетический сигнал может теряться, что затрудняет детекцию рекомбинации. Одной из задач данной работы была разработка метода детекции и визуализации рекомбинации, который бы позволил выявлять многочисленные события, а также показывал бы «глубину» этих событий. Для этого подход, использованный ранее (Lukashev et al., 2014), был развит и реализован в виде пакета на языке R и веб-приложения. Метод заключается в вычислении попарных расстояний между последовательностями вирусов в разных регионах генома и построении графиков СПГР. Координаты (генетические расстояния) отклоняющихся точек могут указывать на время рекомбинационных событий в прошлом. Например, грубая средняя оценка скорости накопления замен у коронавируса 5×10^{-4}

замен/сайт/год (Su et al., 2016) предполагает, что самый недавний общий предок коронавирусов, отличающийся по нуклеотидной последовательности на 10^{-2} (1%), существовал около десяти лет назад. Для того чтобы глобально оценить несогласованность попарных генетических расстояний, можно визуализировать множество графиков СПГР для всех возможных пар областей генома в виде матрицы ОНР. Для этого матрицы генетических расстояний строятся для всех регионов генома с использованием скользящего окна, затем для каждой возможной пары регионов (один график СПГР) вычисляется СКО точек от линии регрессии, и СКО из всех возможных пар регионов генома визуализируются в виде тепловой карты (Рис. 1б). Необходимость разработки этого метода была обусловлена очень высокой частотой рекомбинации у многих РНК-вирусов, ограничивавшей применение классических подходов для анализа. Затем новый метод был использован в комбинации с классическими подходами для систематического изучения роли рекомбинации в эволюции вирусов, принадлежащих нескольким широко распространенным и актуальным семействам (+)РНК-вирусов – *Caliciviridae*, *Astroviridae* и *Coronaviridae*.

Общие закономерности рекомбинации семейств *Caliciviridae*, *Astroviridae* и *Coronaviridae*

Закономерности рекомбинации были проанализированы для отдельных родов вирусов млекопитающих, принадлежащих семействам *Caliciviridae* (рода *Norovirus*, *Lagovirus*, *Sapovirus*), *Astroviridae* (род *Mamastrovirus*) и *Coronaviridae* (рода *Alpha-*, *Beta-*, *Gamma-*, *Deltacoronavirus* – α -CoV, β -CoV, γ -CoV и δ -CoV соответственно). Выбор вирусов был обусловлен социально-экономическим значением, количеством доступных полногеномных последовательностей. Исследование закономерностей рекомбинации является продолжением работ, выполненных ранее (Lukashev, 2010).

Для анализа глобальных закономерностей рекомбинации исследуемых вирусов были построены матрицы филогенетической совместимости и матрицы ОНР (Рис. 2). Общим свойством всех семейств было то, что как на матрицах филогенетической совместимости, так и на матрицах ОНР можно было выделить два блока генома, между которыми рекомбинация происходит чаще всего. У калицивирусов и астровирусов эти блоки соответствовали участкам, кодирующим неструктурные белки (ORF1 у норовирусов, часть ORF1 у саповирусов и лаговирусов, ORF1ab у астровирусов) и структурные белки (ORF2-ORF3 у норовирусов, часть ORF1 и ORF2 у саповирусов и лаговирусов, ORF3 у астровирусов) (Рис. 2, стрелки).

У коронавирусов рекомбинация наиболее часто происходила между геном шиповидного белка S и остальными участками генома. На матрицах филогенетической совместимости четко прослеживалась граница между S и другими участками для всех родов коронавирусов, кроме γ -CoV, у которых наблюдалось большое количество филогенетических конфликтов между всеми областями генома, то есть классический метод не позволял эффективно анализировать рекомбинацию у этого рода.

Матрицы ОНР, основанные на несогласованности генетических расстояний, позволили четко выявить более частую рекомбинацию у γ -CoV между геном S и

остальным геномом, а также показали, что в остальных трех родах (особенно у β -CoV) в 5'-части S наблюдалась большая, чем в 3'-части, степень несогласованности попарных дистанций по отношению к другим участкам генома. Таким образом, матрицы ОНР предоставили дополнительную информацию о закономерностях рекомбинации у γ -CoV.

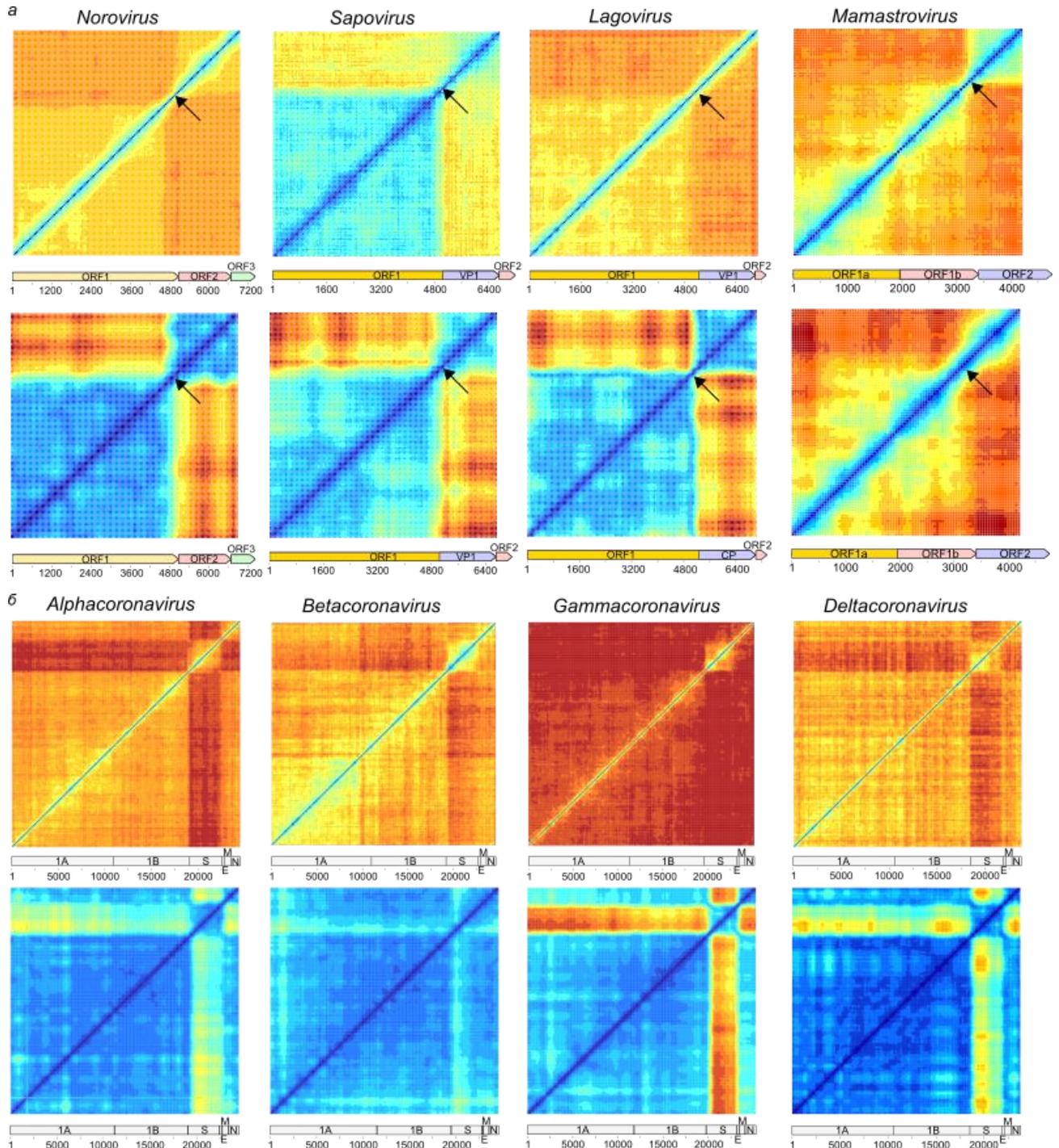


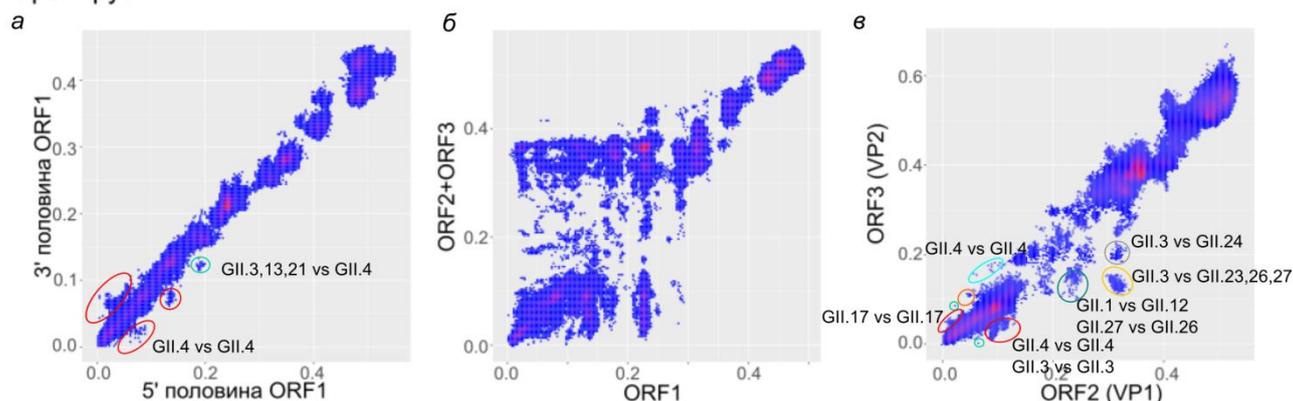
Рис. 2 Матрицы филогенетической совместимости на основе нормализованных расстояний Робинсона-Фулдса (верхняя панель) и матрицы ОНР (нижняя панель) для родов безоболочечных (+)РНК-вирусов (а) и родов семейства *Coronaviridae* (б). Цвета отражают значения нормированных расстояний Робинсона-Фулдса (красный – высокие) и СКО в графиках СПГР, соответственно. Под матрицами размещены схемы сконкатенированных ОРС, цифры под ними соответствуют координатам позиций в выравниваниях.

Рекомбинация у калицивирусов

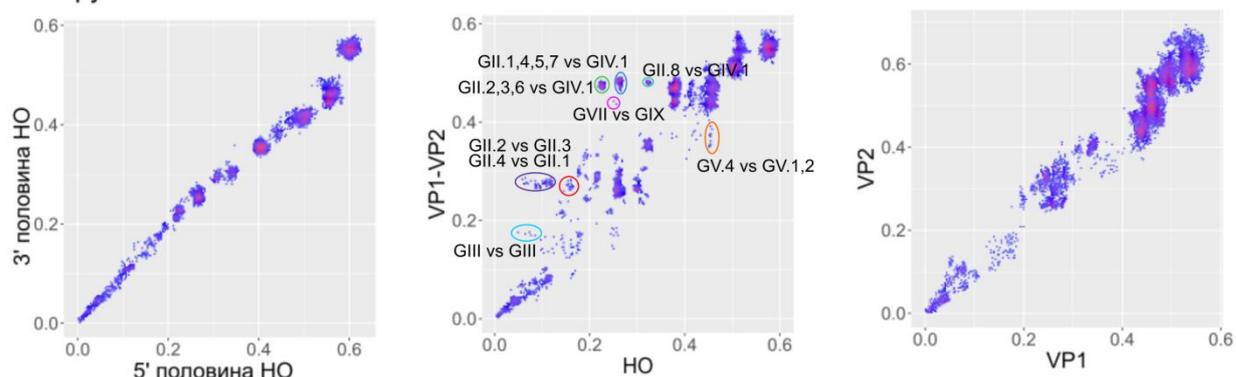
Для подробного исследования всех событий рекомбинации внутри родов были использованы все доступные методы анализа рекомбинации в программе RDP4. У всех трех родов калицивирусов большинство рекомбинантных фрагментов составляли почти полные области, кодирующие белки VP1 и VP2, что подтверждает частую рекомбинацию на стыке ORF1/ORF2, которая является известной чертой норовирусов (Begall et al., 2018; Bull et al., 2007), а также была описана ранее в нескольких сообщениях у лаговиров (Le Pendu et al., 2017) и у саповирусов (dos Anjos et al., 2011; Kuroda et al., 2017). Единичные события рекомбинации также были выявлены внутри неструктурной (кодирующей белки репликативного комплекса) и структурной (кодирующей белки капсида) областей генома (НСО и СО, соответственно) у норовирусов и лаговиров, в то время как у саповирусов рекомбинационные события внутри этих областей детектированы не были. Для того чтобы получить общее представление о динамике рекомбинации были построены графики СПГР, которые визуализируют отношения генетических расстояний в двух областях генома. Графики СПГР согласовывались с результатами, полученными классическими методами: наблюдалась линейная зависимость между попарными нуклеотидными дистанциями в двух половинах НСО и двух половинах СО генома с отклонением небольших групп точек у норовирусов и лаговиров. Графики СПГР между НСО и СО генома характеризовались выраженным несоответствием попарных расстояний во всех родах калицивирусов. У норовирусов и лаговиров попарные генетические расстояния образовывали симметричное облако в пределах 30-40% нуклеотидной дистанции, что может свидетельствовать о многочисленных (фактически, неисчислимых) событиях рекомбинации (Рис. 3). Также для этих родов было характерно наличие недавних событий рекомбинации между вирусами, отличающимися на несколько процентов в одной области генома и до 30-40% в другой. У саповирусов, напротив, недавних событий рекомбинации не наблюдалось. В отличие от норовирусов, у саповирусов и лаговиров VP1 кодируется не отдельной ОРС, а является частью полипротеина ORF1. Таким образом, место «горячей» точки рекомбинации определялось не границами ОРС, а функциональностью белков.

Практически все рекомбинационные события, детектированные у калицивирусов, происходили между вирусами, принадлежащими одной геногруппе. Межгеногрупповые события рекомбинации у норовирусов и саповирусов были детектированы между вирусами, заражающими одинаковых хозяев и отличающимися друг от друга на 30-45% нуклеотидной последовательности в исследуемых областях, поэтому, возможно, они произошли до разделения вирусов на геногруппы. Таким образом, можно говорить об отсутствии рекомбинации между вирусами разных геногрупп калицивирусов.

Норовирус



Саповирус



Лаговирус

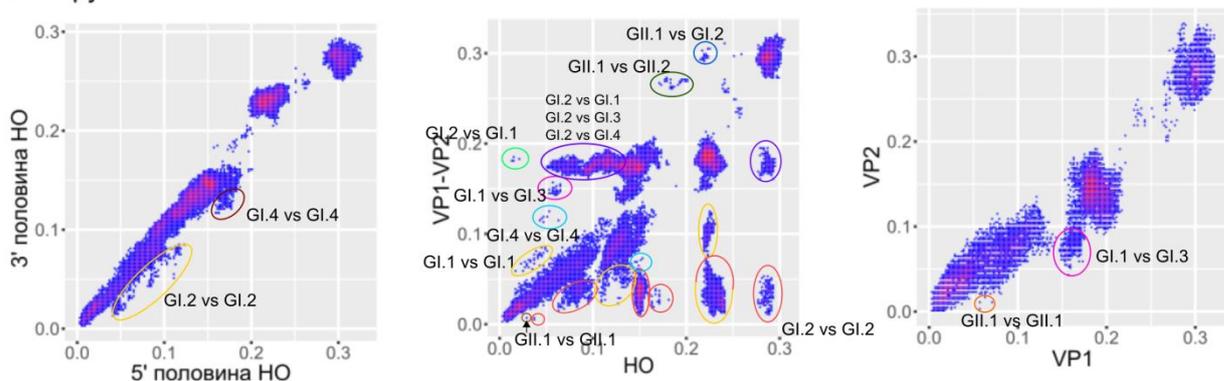


Рис. 3. Соответствие попарных нуклеотидных расстояний в 5' и 3' половинах неструктурных областей генома (НО) (а), НО и структурной области генома (СО), кодирующей белки VP1 и VP2 (б), а также между областями, кодирующими структурные белки VP1 и VP2 (в) соответственно, у норовирусов, саповирусов и лаговирусов. Отдельные события рекомбинации, обсуждаемые в диссертации, выделены овалами.

Для более точной оценки временной динамики рекомбинации для двух самых распространенных геногрупп норовирусов – GI и GII – было вычислено время полужизни рекомбинантных форм, которое составило 10.4 и 8.4–11.3 лет соответственно. Время полужизни рекомбинантных форм норовирусов, заражающих человека, несколько больше, чем у энтеровирусов (1.3, 9.8, 3.1 лет у типов E9, E11, E30, соответственно (Lukashev et al., 2014; McWilliam Leitch et al., 2009, 2010, 2012)), но находятся в близком диапазоне. Таким образом, норовирусы рекомбинируют примерно так же часто, как и энтеровирусы человека, которые принадлежат к тому же таксономическому отряду (*Picornavirales*) и сходны с норовирусами с точки зрения

глобальной распространенности, пути передачи, сайтов репликации и генетического разнообразия.

Рекомбинация у астровирусов

Как на матрицах филогенетической совместимости, так и на матрицах ОНР для рода *Mamastrovirus* видно, что события рекомбинации были более равномерно распространены по всему геному по сравнению с калицивирусами. Можно выделить области с несколько ограниченной рекомбинацией внутри них, соответствующие половинам ORF1a, ORF1b и ORF2 (Рис. 2). Наибольшая несогласованность филогенетических деревьев и попарных генетических дистанций наблюдалась при сравнении НСО (ORF1ab) со СО (ORF2), причем, в отличие от калицивирусов, четкой границы между этими участками генома не наблюдалось. Графики СПГР подтвердили умеренную рекомбинацию внутри ORF1a, ORF2, а также между ORF1a и ORF1b (Рис. 4а–в). График СПГР ORF1ab и ORF2 отличался от других отклонением большого количества попарных генетических расстояний между вирусами одного вида, образующих «облако», причем среди астровирусов встречались вирусы с очень близкими неструктурными белками и дивергировавшими белками капсида, но не наоборот (Рис. 4г). Это наблюдение соответствует большей вариабельности рекомбинационно-мобильных структурных генов у астровирусов и гена S у коронавирусов, как будет показано далее. Таким образом, для астровирусов млекопитающих характерна модульная эволюция генома, но в отличие от калицивирусов нет признаков строгой невозможности рекомбинации внутри СО или НСО.

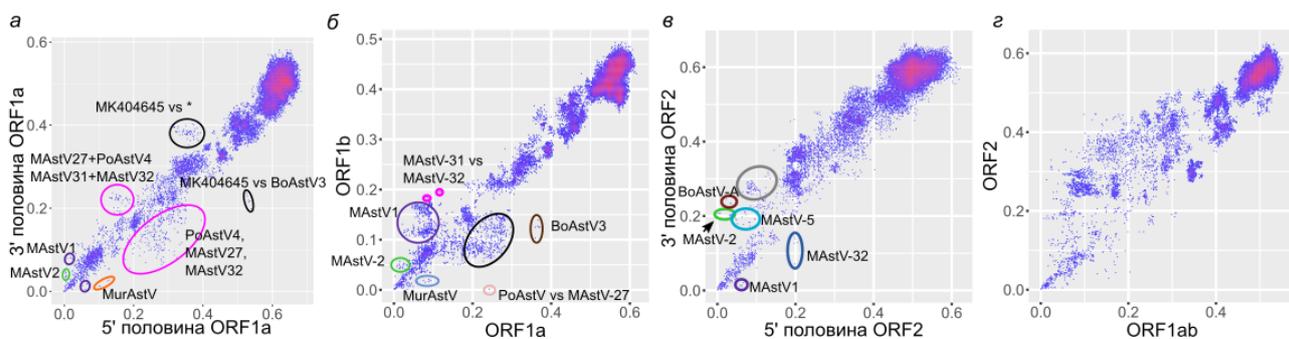


Рис. 4. Соответствие попарных нуклеотидных расстояний между 5' и 3' половинами ORF1a (а), ORF1a и ORF1b (б), между 5' и 3' половинами ORF2 (в), а также ORF1ab и ORF2 (г) у вирусов, принадлежащих роду *Mamastrovirus*.

Рекомбинация у коронавирусов

Матрицы ОНР указывали на умеренную распространенность рекомбинации внутри ORF1 во всех четырех родах коронавирусов, что подтвердили графики СПГР ORF1a и ORF1b (Рис. 5а). У α -CoV и β -CoV в целом наблюдалась линейная зависимость генетических расстояний в ORF1a и ORF1b с отклонениями некоторых групп точек, соответствующих сравнениям вирусов, принадлежащих одному виду (Рис. 5а). В то же время, у γ -CoV наблюдалось облако пар генетических расстояний для нуклеотидных дистанций менее 10%, что свидетельствует о частой рекомбинации среди близкородственных вирусов.

Общей чертой всех коронавирусов была несогласованность генетических расстояний в ORF1ab и S между вирусами, принадлежащими одинаковым под родам (до 40% нуклеотидных различий) (Рис. 5б). Также у всех родов коронавирусов наблюдалось полное отсутствие корреляции между попарными генетическими расстояниями для нуклеотидных дистанций менее 10%, то есть существовало «облако генов», подразумевающее большое количество рекомбинационных событий. В родах α -CoV и γ -CoV наблюдалась относительно недавняя рекомбинация между вирусами со сравнительно сильно дивергировавшими S (расстояния более 25%) и схожими ORF1ab (менее 15%), но не наоборот. Одним из ярких примеров таких событий является рекомбинация между представителями *Avian coronavirus*, выделенными из курицы и индейки, рода γ -CoV (Рис. 5б, красный круг) (Jackwood et al., 2010).

Несоответствие попарных генетических расстояний между двумя доменами белка S, S1 и S2, было более выражено, чем внутри ORF1, но менее выражено, чем между ORF1 и S (Рис. 5в). У γ -CoV на графике СПГР доменов S1 и S2 наблюдалось симметричное «облако» генетических дистанций до 20-25%, что может свидетельствовать о частой рекомбинации между доменами (Рис. 5в), включая недавние события между достаточно дивергентными вирусами с генетическими расстояниями <2% в одной половине S и >15% в другой. В α -CoV и β -CoV «облако» генетических дистанций, образующихся в результате рекомбинации между S1 и S2, было менее выраженным, но затрагивало большие генетические расстояния (до 40%) и было асимметричным. Последовательности с расстояниями до 40% в α -CoV и 30% в β -CoV в доменах S1 могли сочетаться с близкородственными (до 10% различий) S2, но не наоборот. Внутри субдоменов S1-CTD, S1-NTD и домена S2 было заметно меньше различных событий рекомбинации, чем между ними, причем такие события затрагивали только вирусы, принадлежащие одному виду.

Графики СПГР между N-концевым субдоменом (S1-NTD), C-концевым субдоменом (S1-CTD) домена S1 и доменом S2 (Рис. 5г, д) показали, что большинство событий рекомбинации, обнаруженных при сравнении S1 и S2 (Рис. 5в) в β -CoV (но не в других родах коронавирусов) соответствовали переносу только S1-NTD, а не всего S1. Большинство таких событий рекомбинации в β -CoV произошло между представителями *SARS-related coronavirus*, выделенными из летучих мышей, людей и панголинов (Рис. 5в, г, черный круг). Это довольно неожиданно, поскольку, по крайней мере, у SARS-подобных коронавирусов рецептор-связывающий домен картирован в S1-CTD (Li, 2015; Wang et al., 2020). Также примечательно то, что в α -CoV и γ -CoV рекомбинация между далекими вирусами затрагивала полноразмерный S, а не его отдельные домены. Модульная структура S обсуждалась ранее на основании экспериментов по реконструкции рекомбинантных коронавирусов и заменой отдельных фрагментов шиповидного белка (Graham & Baric, 2010). Данное исследование показывает, что обмен доменами S происходит систематически в масштабе рода во всех родах коронавирусов в естественных условиях, и модульная эволюция генома и гена S является общей закономерностью изменчивости коронавирусов.

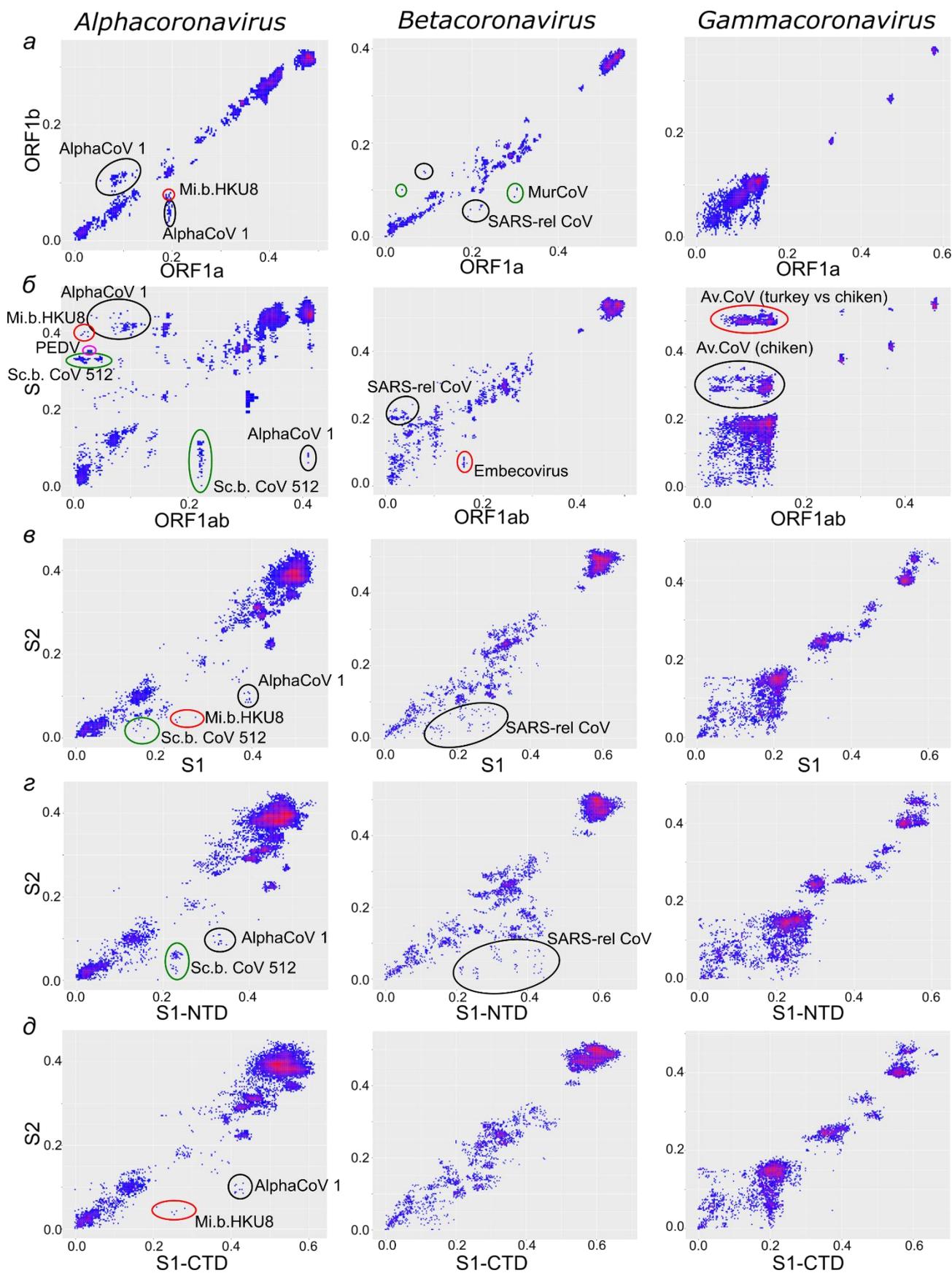


Рис. 5. Рекомбинация чаще встречается между ORF1 и Spike, чем внутри ORF1. Соответствие попарных нуклеотидных расстояний между ORF1a и ORF1b (а), ORF1ab и Spike (б), участками S1 и S2 белка Spike (в), субдоменом S1-NTD и участком S2 (г), доменом S1-CTD и участком S2 (д). Обозначения: AlphaCoV 1, *Alphacoronavirus 1*; Mi.b.HKU8, *Miniopterus bat coronavirus HKU 8*; Sc.b.CoV 512, *Scotophilus bat coronavirus 512*; SARS-rel.,

SARS-related coronavirus; MurCoV, *Murine coronavirus*; Av.CoV, *Avian coronavirus*. Точки, которые соответствуют парам вирусов, подвергшимся рекомбинации, и обсуждаются в тексте диссертации, отмечены кружками.

Согласно матрицам филогенетической совместимости и ОПП, рекомбинация между генами структурных белков E, M, N и ORF1ab не была заметно более интенсивной, чем внутри ORF1ab (Рис. 2), что также подтверждается графиками СПГР (приведены в тексте диссертации). Таким образом, с точки зрения общих для изученных РНК-вирусов закономерностей рекомбинации структурные гены E, M и N фактически эквивалентны НСО коронавирусов, и только ген S соответствует структурным генам других семейств с точки зрения иной рекомбинационной мобильности.

Молекулярная эволюция типов энтеровирусов человека

Для изучения истории происхождения энтеровирусов, которые включают более 100 типов, был проведен Байесовский филогенетический анализ типов (N=32), для которых было доступно более 35 последовательностей для «типизирующего» фрагмента капсидного белка VP1. Результаты, опубликованные ранее, указывали на независимую эволюционную историю структурных и неструктурных областей генома даже на короткой временной шкале. Также ранее было показано, что рекомбинация у энтеровирусов распространена в области генома, кодирующей неструктурные белки, и практически отсутствует в структурной области. Это стало основанием выбора участка генома, кодирующего капсидный белок VP1, для изучения молекулярной эволюции типов отдельно от других частей генома. Виды энтеровирусов имеют сложную генетическую структуру, и входящие в них типы могут иметь разные закономерности эволюции и необязательно составляют одну популяцию, поэтому сначала каждый тип был проанализирован по-отдельности.

Количество последовательностей в выравнивании типизирующих фрагментов VP1 отдельных типов составило от 35 до 629. Прототипные последовательности, или самые первые выделенные последовательности каждого типа, датируются 1947–1970 годами. Большая часть последовательностей участка генома, кодирующего VP1, из базы данных GenBank была выделена после 2000 года. Время между самым «старым» и самым «молодым» изолятом, или так называемая глубина выборки, отличалось у типов и составило в среднем 64 года.

Согласно полученным в работе расчетам, время существования последнего общего предка отдельных типов находилось в пределах 55–200 лет назад, а скорость накопления замен у разных типов составила $0.4\text{--}1.21 \times 10^{-2}$ замен/сайт/год. При этом выявлены следующие закономерности: скорость накопления замен не коррелирует с количеством последовательностей в выравниваниях (размером выборки). Для тех типов, которые представлены более чем 100 последовательностями (11 из 32 типов), статистически значимая корреляция между высотой корня и объемом выборки также отсутствует. Интервал между временем выделения прототипной последовательности и самой старой непрототипной отличается у разных типов от -2 до 53 лет и не

коррелирует ни с высотой корня, ни со скоростью замен. После удаления прототипных штаммов из выборок разброс времени самого раннего известного генома в выборках ставится значительно больше. При этом расчетная скорость замен незначительно увеличилась (в среднем на 0.2×10^{-3} замен/сайт/год) и коррелирует с уменьшением глубины выборки (Spearman $r=0.41$, $p=0.02$), а время последнего общего предка типов уменьшилось на 2–163 года (в среднем на 24.6 лет). При отсутствии в выравнивании прототипного штамма последний общий предок типа оказывается моложе примерно на 25 лет. При удалении из выборки прототипных штаммов наблюдалась корреляция между глубиной выборки и высотой корня (Pearson $r=-0.52$, $p\text{-value}<0.05$) для типов без учета прототипных штаммов. Следовательно, оценка влияния неравномерности некоторых параметров выборок последовательностей генов VP1 для исследуемых типов энтеровирусов показала, что неравномерная представленность последовательностей по годам практически не влияет на оценку возраста корня, а определяющим фактором для оценки корня является возраст самого старого изолята.

Объяснить влияние величины возраста самого старого изолята только на высоту корня, но не на скорость замен, можно прохождением популяций энтеровирусов через «бутылочные горлышки» и вымиранием линий, что приводит к сбросу возраста корня наблюдаемого дерева. Это не могло не отразиться на топологии филогенетических деревьев типов: вирусы после гипотетического «бутылочного горлышка» группировались бы вместе на филогенетическом дереве. Действительно, в 16 из 32 типов (CVA2, CVA5, CVA10, CVA16, E1, E2, E7, E14, E18, E19, E20, E24, E25, E29, E33, EV-D68) прототипный штамм занимал базальное положение, причем в эховирусах чаще, чем в коксакивирусах. Такая топология филогенетических деревьев косвенно подтверждает гипотезу о прохождении глобальной популяции энтеровирусов через «бутылочные горлышки» и вымирании большинства вариантов вирусов, которые гипотетически существовали в 1950-х годах (пример топологии на Рис. 6а). В 8 типах (CVA24, CVB5, E3, E11, E13, E21, E30, EV-A71) современные вирусы представляют собой несколько превалирующих линий, которые разошлись после 1950-х годов (Рис. 6б). Эта группа включает наиболее изученные типы энтеровирусов. Параллельная циркуляция нескольких линий, возникших после 1950-х годов, может быть связана как с большей репрезентативностью выборок последовательностей, так и с большей относительной приспособленностью этих типов. Таким образом, характер группирования внутри разных типов показывает возможные эволюционные сценарии, которые варьируются от балансирования на грани вымирания единственной линии до разделения на несколько глобально распространенных линий. Эффект «бутылочного горлышка» и вымирания линий вирусов может происходить в глобальной популяции энтеровирусов в любое время и с любой частотой, что ведет к преобладанию одной или нескольких линий современных вирусов и гипотетическому исчезновению «старых» линий.

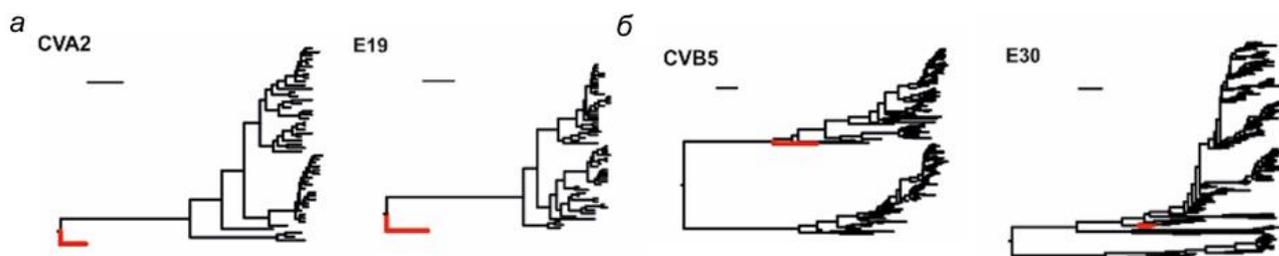


Рис. 6. Байесовские филогенетические деревья, построенные по типизирующему фрагменту VP1 (294–339 нт) в программе BEAST v.1.8.3 для четырех из 32 типов энтеровирусов – CVA2, E6, CVB5, CVA4, – характеризующие два типа топологий филогенетических деревьев типов: а – базальное положение прототипного штамма, б – существование нескольких превалирующих линий энтеровирусов, разошедшихся после 1950-х годов. Прототипные штаммы показаны красным цветом. Длины ветвей деревьев соответствуют времени. Длина отрезка-индикатора масштаба соответствует 10 годам.

При прохождении глобальной популяции энтеровирусов через «бутылочные горлышки» молекулярные часы могут постоянно сбрасываться, поэтому для определения времени возникновения типов был проведен Байесовский филогенетический анализ на уровне видов. Результаты такого анализа необходимо интерпретировать с осторожностью, потому что полный участок генома VP1 разных типов необязательно представляет целостную популяцию, а также потому, что механизмы происхождения типов могут включать интенсивное накопление мутаций за короткий промежуток времени (Smuga et al., 2011). Филогенетические деревья имели хорошее разрешение, апостериорные вероятности большинства узлов на уровне выше индивидуальных типов были равны 1, параметры расчетов соответствовали другим исследованиям (Рис. 7).

Времена последних общих предков отдельных типов, выведенные при анализе полных последовательностей VP1, попадали в 95% доверительный интервал времен, выведенных с помощью анализа выравнивания коротких последовательностей. Ветви, ведущие к отдельным типам, были значительно длиннее (80–427 лет), чем ветви внутри типов. Такая топология деревьев совместима с гипотетическими «бутылочными горлышками», которые могли приводить к вымиранию большинства линий, предшествовавших современным вирусам. С другой стороны, самое недавнее разделение двух отдельных типов (общий предок E21 и E30) датируется всего лишь 214 [167–264] годами назад, а время последних общих предков ряда других пар типов было менее 300 лет назад. Таким образом, регулярное прохождение через «бутылочные горлышки» могло существенно влиять на эволюцию типов, а также есть свидетельства того, что известное разнообразие типов энтеровирусов могло возникнуть в течение нескольких последних столетий.

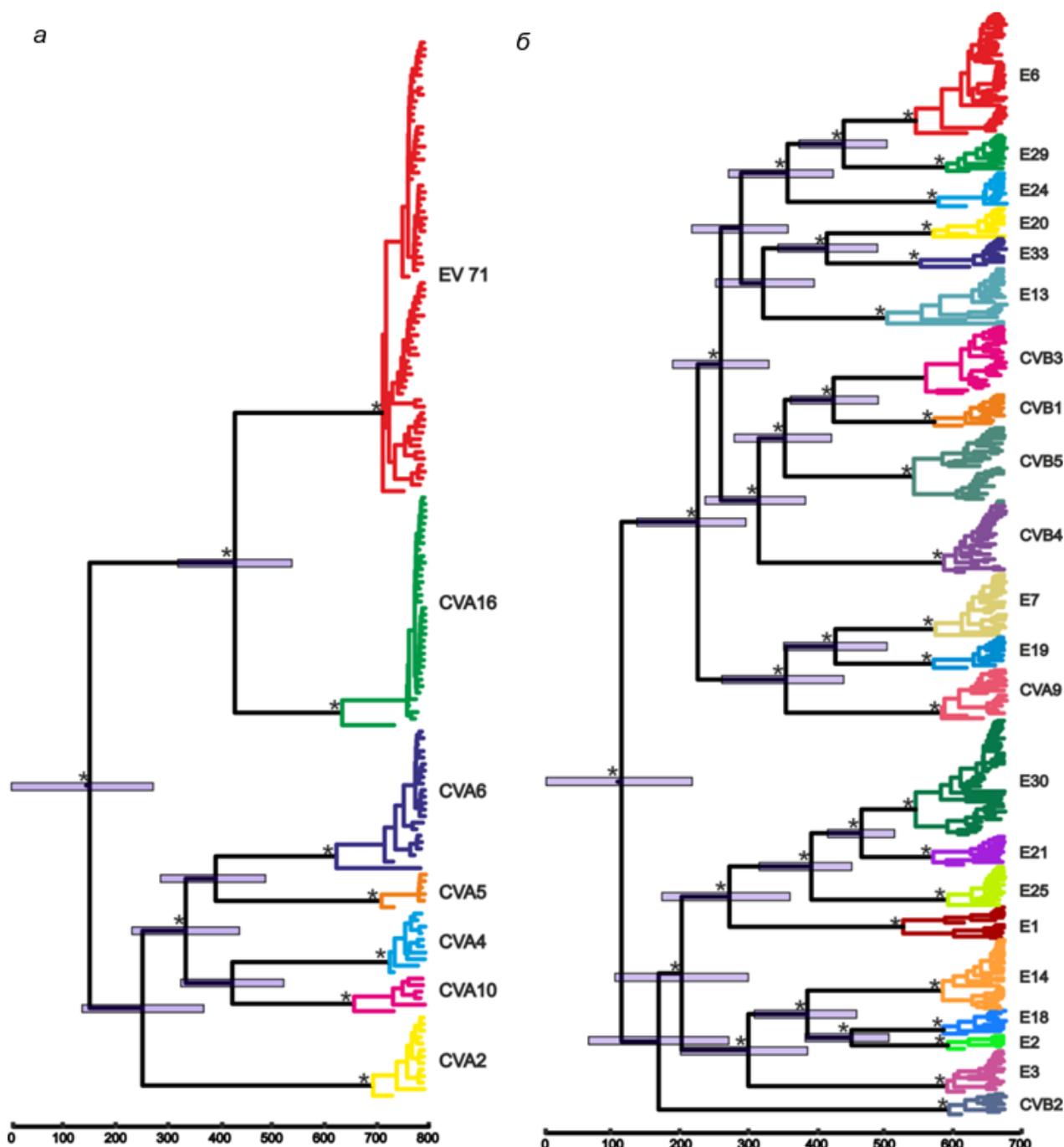


Рис. 7. Байесовские филогенетические деревья видов *Enterovirus A* (а) и *Enterovirus B* (б), построенные по полному гену VP1 в программе BEAST v.1.8.3. Шкала показывает время в годах. Планки погрешностей на узлах выше уровня типов показывают 95% интервал наибольшей плотности апостериорной вероятности. Типы покрашены разными цветами. Звездочками отмечены узлы с апостериорной вероятностью более 0.9.

Молекулярно-генетический анализ клинических изолятов коксакивируса A2, выделенных от детей с острым вялым параличом

В рамках надзора за неполиомиелитными энтеровирусами с 2001 по 2020 год в 5 из 7280 случаев острого вялого паралича (ОВП) у детей от 1 до 4 лет были выделены вирусы типа коксакивирус A2 (CVA2) в 2008 (два случая), 2015 (один случай) и 2019 (два случая) годах. Способность неполиомиелитных вирусов вызывать ОВП до сегодняшнего дня является предметом обсуждений и не доказана

окончательно. Для проверки гипотезы о возможном возникновении варианта CVA2 с измененными патогенетическими свойствами в этой работе были определены нуклеотидные последовательности типизирующего фрагмента VP1 выделенных изолятов. В соответствии с практическими рекомендациями ВОЗ, три из пяти случаев можно было отнести к паралитическому полиомиелиту (больные Н.Т., К.О. и Д.А.). Для того чтобы понять филогенетические взаимоотношения этих изолятов в сопоставлении с другими последовательностями CVA2, выделенными в России и других странах мира, доступными в базе GenBank, было построено Байесовское филогенетическое дерево с разверткой по времени (Рис. 8).

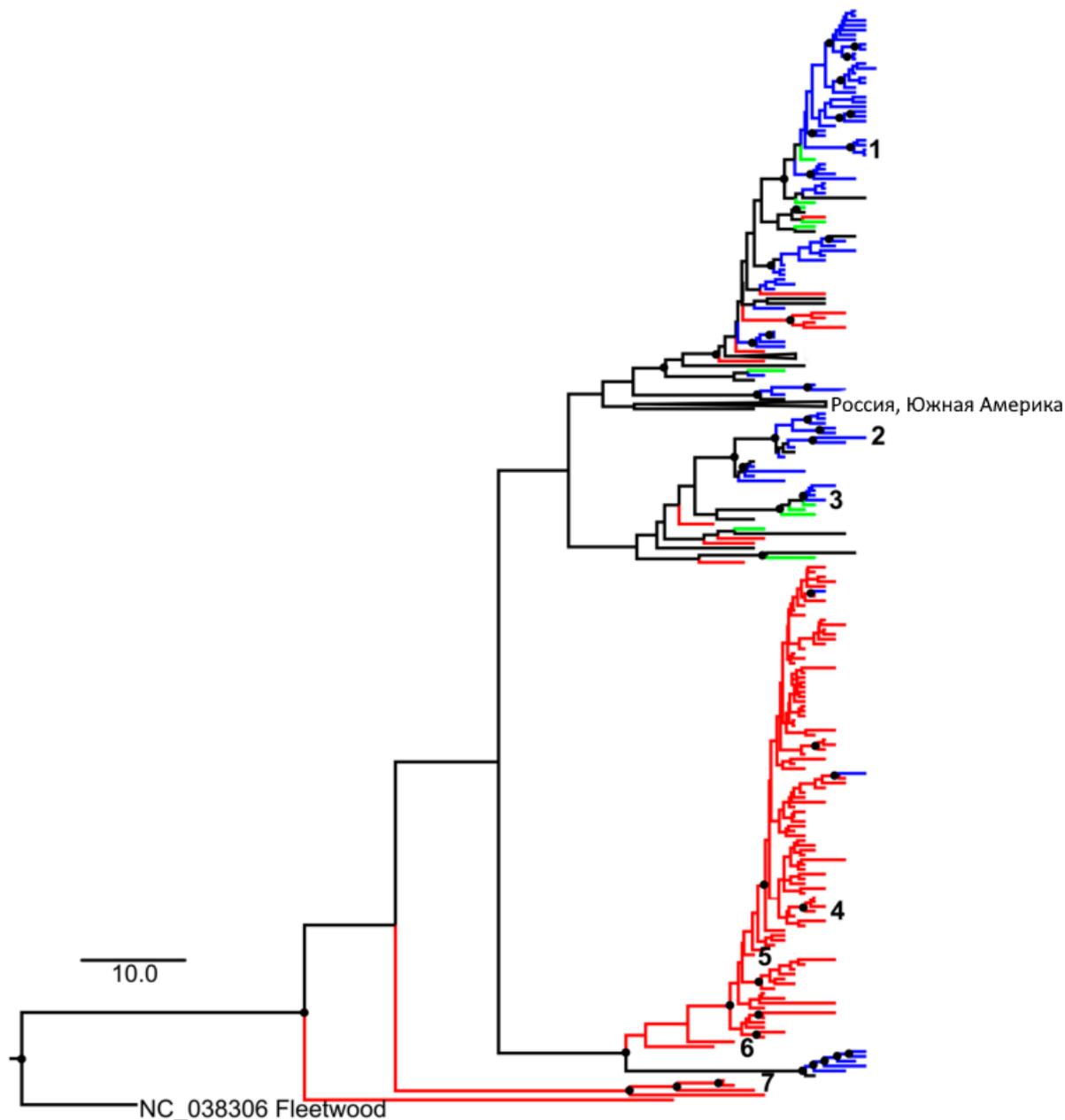


Рис. 8. Байесовское филогенетическое дерево, построенное по последовательностям типизирующего фрагмента VP1 типа CVA2, доступным в базе Genbank. Ветви, ведущие к последовательностям, собранным в России, Восточной Азии (Китай, Япония) и Европейском союзе, окрашены синим, красным и зеленым цветами соответственно. Последовательности изолятов, ассоциированных с ОВП, пронумерованы следующим образом: (1) – больной Н.Т.;

2 — больной К.О.; (3) – больной Д.А.; (4) – 2014 г., изолят из Тайваня (Yen и др., 2017); (5) – 2008 г., изолят из Китая; (6), (7) – 2006 г, изолят из Китая. Узлы с апостериорной вероятностью выше 0.95 отмечены кругами. Длина отрезка-индикатора масштаба соответствует 10 годам. Клады дерева, не содержащие изолятов, ассоциированных с ОВП, сжаты.

В целом, последовательности CVA2 группировались на филогенетическом дереве по географическим регионам, в которых они были выделены. На филогенетическом дереве можно выделить две большие клады. Одна кллада включала преимущественно последовательности из Китая, а вторая – из России, Европы, стран Северной и Южной Америк и Азии. Среди всех последовательностей CVA2, доступных в Genbank, только 7 были ассоциированы со случаями ОВП. Эти изоляты были равномерно распределены на дереве. Вирусы, выделенные в данном исследовании от больных с ОВП, были филогенетически близки с вирусами, выделенными в разных регионах России и ассоциированными с острыми респираторными заболеваниями, герпетической ангиной, ящуроподобным заболеванием и редко – с асептическим вирусным менингитом. Таким образом, филогенетический анализ вирусов CVA2, выделенных от больных с ОВП, не выявил их связи с конкретной филогенетической группой.

ВЫВОДЫ

1. Общим свойством эволюции вирусов с несегментированным РНК-геномом положительной полярности относительно мРНК является высокая частота рекомбинации между участками, кодирующими структурные и неструктурные белки. Профили рекомбинации внутри участков генома, кодирующих структурные и неструктурные белки, различаются между семействами, но сходны внутри одного семейства РНК-вирусов (*Caliciviridae*, *Astroviridae*, *Coronaviridae*).
2. Для калицивирусов характерен обмен целыми блоками неструктурных и структурных генов и крайне редкая рекомбинация внутри участков, кодирующих структурные и неструктурные белки. Для астровирусов характерен обмен целыми блоками неструктурных и структурных генов и умеренная рекомбинация внутри структурных и неструктурных генов.
3. У коронавирусов наиболее часто рекомбинация происходит между геном структурного белка S и другими участками генома. Рекомбинация внутри гена белка S происходит чаще, чем в неструктурной области, причем с большей частотой происходит обмен целыми доменами белка S. Гены остальных структурных белков (N, E, M) по профилю естественной рекомбинации соответствуют неструктурным генам.
4. Участки генома, кодирующие капсидные белки энтеровирусов, благодаря рекомбинации имеют эволюционную историю, практически оторванную от генов неструктурных белков, и существуют в биосфере как высоко динамичные пулы, в пределах которых отдельные типы могут регулярно подвергаться почти полному исчезновению практически в глобальном масштабе даже в течение десятков лет.
5. Филогенетический анализ вирусов CVA2, выделенных от больных с острым вялым параличом, не выявил их связи с конкретной филогенетической группой, что опровергает гипотезу о возникновении варианта вируса с измененными патогенетическими характеристиками.

НАУЧНЫЕ СТАТЬИ ПО ТЕМЕ ДИССЕРТАЦИИ, ОПУБЛИКОВАННЫЕ В ЖУРНАЛАХ SCOPUS, WOS или RSCI

1. Lukashev A., **Vakulenko Y.** Molecular evolution of types in non-polio enteroviruses // J Gen Virol. Microbiology Society, 2017. Vol. 98, № 12. P. 2968–2981. IF Scopus = 2.78
2. Lukashev, A.N.; **Vakulenko, Y.A.**; Turbabina, N.A.; Deviatkin, A.A.; Drexler, J.F. Molecular epidemiology and phylogenetics of human enteroviruses: Is there a forest behind the trees? // Rev. Med. Virol. 2018. Vol. 28, № 6. P. e2002. IF = 3.943
3. **Vakulenko Y.**, Deviatkin A., Lukashev A. The effect of sample bias and experimental artefacts on the statistical phylogenetic analysis of picornaviruses // Viruses. 2019. Vol. 11, № 11. IF Scopus = 4.013
4. **Вакуленко Ю.А.**, Лукашев А.Н., Девяткин А.А. Использование методов статистической филогенетики в вирусологии // ИНФЕКЦИЯ И ИММУНИТЕТ. 2021. Vol. 11, № 1. P. 42–56. IF Scopus = 0.598
5. **Vakulenko Y.**; Deviatkin A.; Drexler J.F.; Lukashev A. Modular Evolution of Coronavirus Genomes // Viruses. 2021. Vol. 13, № 7. P. 1270. IF Scopus = 5.708
6. Ivanova O.E., Shakaryan A.K., Morozova N.S., **Vakulenko Y.A.**, Ereemeeva T.P., Kozlovskaya L.I., Baykova O.Y., Shustova E.Y., Mikhailova Y.M., Romanenkova N.I., Rozaeva N.R., Dzhaparidze N.I., Novikova N.A., Zverev V.V., Golitsyna L.N., Lukashev A.N. Cases of Acute Flaccid Paralysis Associated with Coxsackievirus A2: Findings of a 20-Year Surveillance in the Russian Federation // Microorganisms. 2022. Vol. 10, № 1. P. 112. IF Scopus = 4.782
7. **Vakulenko Y.**, Orlov A., Lukashev A. Patterns and temporal dynamics of natural recombination in noroviruses // Viruses 2023. Vol. 15, № 2. P. 372. IF Scopus = 5.712