

РАСПРОСТРАНЕННОСТЬ В ГЕНОМЕ ВИРУСОВ ЧЕЛОВЕКА МАЛЫХ ГОМОЛОГИЧНЫХ И КОМПЛЕМЕНТАРНЫХ ФРАГМЕНТОВ И ВОЗМОЖНАЯ ИХ РОЛЬ

Е.П. Харченко

ФГБУН Институт эволюционной физиологии и биохимии им. И.М. Сеченова РАН, Санкт-Петербург, Россия

Резюме. С помощью компьютерного анализа исследована распространенность аутокомплементарных, комплементарных и гомологичных последовательностей (КП и ГП) (длиной в 21 нуклеотид) в геномах 14 вирусов, вызывающих наиболее распространенные инфекции у человека. Выборка вирусов включала вирусы с (+) и (–) односпиральной РНК и ДНК-содержащий вирус гепатита В. Выполненный анализ по распространенности ГП свидетельствует о существовании двух ее экстремальных показателей: с одной стороны, имеет место наличие у одного и того же вируса ГП практически ко всем другим вирусам (например, вирусы Эбола, тяжелого острого респираторного синдрома и паротита) и многочисленности у него же ГП к одному и тому же другому вирусу (у вируса тяжелого острого респираторного синдрома особенно к вирусу Денге, вирусу Эбола и полиовирусу), и, с другой стороны, редкая встречаемость и немногочисленность ГП для некоторых вирусов (вирус краснухи, вирус гепатита А и гепатита В). Сходная картина отмечается и в распространенности КП. Вирус краснухи, имеющий резко отличный от других РНК-содержащих вирусов нуклеотидный состав его генома, также отличался по максимальному числу вирусов, к которым он не имеет КП. Большинство же исследованных вирусов по обоим показателям ГП и КП имеют промежуточные величины между экстремальными значениями. АутоКП длиной в ≤ 19 нуклеотидов были многочисленны у всех исследованных вирусов, что предполагает наличие у вирусов с онРНК разветвленной вторичной структуры. Помимо возможной роли в рекомбинации внутритиповых штаммов, аутоКП могли бы через фолдинг геномов и мРНК быть регуляторами скорости трансляции вирусных белков, обеспечивая оптимальное количественное соотношение их для сборки вирионов. Обнаружение распространенности коротких ГП и КП среди РНК- и ДНК-содержащих вирусов можно рассматривать как результат многократной рекомбинации между ними, свершавшейся в прошлом и возможной в настоящем и определяющей их изменчивость и адаптацию. Рекомбинация могла происходить при коинфицировании ими человека или других общих для них хозяев. Включение в геном вирусов ГП и КП не только обновляло их, но и могло бы служить и памятью о существовании конкурента (противника) за овладение хозяином, и средством противодействия конкуренту при коинфицировании, являясь аналогией системы CRISPR/Cas бактерий и архей.

Ключевые слова: рекомбинация вирусов, геном вирусов, гомологичные последовательности, комплементарные последовательности, аутокомплементарные последовательности, интерференция вирусов.

Адрес для переписки:

Харченко Евгений Петрович
194223, Россия, Санкт-Петербург, пр. Мориса Тореза, 44,
ИЭФБ им. И.М. Сеченова РАН.
Тел./факс: 8 (812) 552-70-31 (служебн.); 8 904 338-22-80 (моб.).
E-mail: neuro.children@mail.ru

Contacts:

Eugene P. Kharchenko
194223, Russian Federation, St. Petersburg, Toreza pr., 44,
Sechenov Institute of Evolutionary Physiology and Biochemistry.
Phone/Fax: +7 (812) 552-70-31 (office); +7 904 338-22-80 (mobile).
E-mail: neuro.children@mail.ru

Библиографическое описание:

Харченко Е.П. Распространенность в геноме вирусов человека малых гомологичных и комплементарных фрагментов и возможная их роль // Инфекция и иммунитет. 2017. Т. 7, № 4. С. 393–404. doi: 10.15789/2220-7619-2017-4-393-404

Citation:

Kharchenko E.P. Occurrence of small homologous and complementary fragments in human virus genomes and their possible role // Russian Journal of Infection and Immunity = Infektsiya i immunitet, 2017, vol. 7, no. 4, pp. 393–404. doi: 10.15789/2220-7619-2017-4-393-404

OCCURRENCE OF SMALL HOMOLOGOUS AND COMPLEMENTARY FRAGMENTS IN HUMAN VIRUS GENOMES AND THEIR POSSIBLE ROLE

Kharchenko E.P.

I. Sechenov Institute of Evolutionary Physiology and Biochemistry, Russian Academy of Sciences, St. Petersburg, Russian Federation

Abstract. With computer analysis occurrence of small homologous and complementary fragments (21 nucleotides in length) has been studied in genomes of 14 human viruses causing most dangerous infections. The sample includes viruses with (+) and (–) single stranded RNA and DNA-containing hepatitis A virus. Analysis of occurrence of homologous sequences has shown the existence two extreme situations. On the one hand, the same virus contains homologous sequences to almost all other viruses (for example, Ebola virus, severe acute respiratory syndrome-related coronavirus, and mumps virus), and numerous homologous sequences to the same other virus (especially in severe acute respiratory syndrome-related coronavirus to Dengue virus and in Ebola virus to poliovirus). On the other hand, there are rare occurrence and not numerous homologous sequences in genomes of other viruses (rubella virus, hepatitis A virus, and hepatitis B virus). Similar situation exists for occurrence of complementary sequences. Rubella virus, the genome of which has the high content of guanine and cytosine, has no complementary sequences to almost all other viruses. Most viruses have moderate level of occurrence for homologous and complementary sequences. Autocomplementary sequences are numerous in most viruses and one may suggest that the genome of single stranded RNA viruses has branched secondary structure. In addition to possible role in recombination among strains autocomplementary sequences could be regulators of translation rate of virus proteins and determine its optimal proportion in virion assembly with genome and mRNA folding. Occurrence of small homologous and complementary sequences in RNA- and DNA-containing viruses may be the result of multiple recombinations in the past and the present and determine their adaptation and variability. Recombination may take place in coinfection of human and/or common hosts. Inclusion of homologous and complementary sequences into genome could not only renew viruses but also serve as memory of existence of a competitor for host and means of counteraction against a competitor in coinfection being an analogy of the bacterial CRISPR/Cas system.

Key words: *recombination of viruses, virus genome, homologous sequences, complementary sequences, autocomplementary sequences, virus interference.*

Молекулярная биология вирусов/бактериофагов часто свидетельствует о наличии у них прообразов структур и процессов, свойственных их более высокоорганизованным хозяевам, подтверждая тем самым универсальность принципов молекулярной организации живых существ. Наше сознание обычно оказывается скованным в признании существования у вирусов систем, свойственных их хозяевам, из-за малых размеров геномов и простоты организации вирусов. Лишь в последнее десятилетие стала вырисовываться сложная сеть молекулярных взаимодействий компонентов вируса в поражаемой клетке и, наконец, получила признание вовлеченность в инфекционный процесс вирусных некодирующих РНК (внРНК). Круг вирусов, у которых выявлены внРНК разных размеров, все больше расширяется, и в их числе представители *Adenoviridae*, *Herpesviridae*, *Flaviviridae*, *Rhabdoviridae* и *Orthomyxoviridae*. В жизненном цикле этих вирусов внРНК играют роль в регуляции репликации, персистенции, клеточной трансформации и ускользании от иммунной системы хозяина [17].

Применение методов секвенирования нового поколения позволило выявить при инфекциях разными вирусами у различных организмов (от клеща до мыши) образование и накопление в клетке малых вирусных РНК (мвРНК) в результате деградации различными путями синтезируемых вирусных РНК. Хотя длина, полярность и нуклеотидный состав мвРНК ва-

рировали в зависимости от вирусного субстрата и механизма деградации их хозяином, среди них четко выступало преобладание мвРНК размером в 20–23 нуклеотидов. Характерно, что появление мвРНК у разных животных часто обнаруживает профиль, соответствующий активации образования малых интерферирующих РНК [4]. В этой связи возник вопрос, не могут ли содержаться у вирусов среди их мвРНК такие, которые гомологичны или комплементарны последовательностям генома другого вируса или аутокомплемментарны. Их наличие позволило бы расширить понимание роли мвРНК в инфекционном процессе и обосновать возможное соучастие их в других феноменах вирусов, таких как рекомбинация, изменчивость и интерференция. Поэтому цель настоящего исследования состояла в изучении распространенности среди патогенных для человека вирусов с однонитевыми РНК (онРНК) и ДНК малых гомологичных (ГП) и комплементарных (КП) нуклеотидных последовательностей, а также в пределах генома каждого вируса аутокомплемментарных последовательностей (аутоКП).

Методы

Для компьютерного анализа были использованы последовательности геномов 13 вирусов с онРНК и 1 вирус с неполной двуспиральной ДНК [вирус гепатита В (штамм ауw)], патогенные для человека. В их числе вирусы с (–)

онРНК [вирус кори (штамм ICHINOSE-BA), паротита (штамм MIYAHARA VACCINE), Эбола (Zaire ebolavirus strain Ebola virus/H sapiens-rec/LBR/2014/Makona-L2014_ZsG), Борна (штамм V)], вирусы с (+) онРНК [вирусы гепатитов А (генотип IB, изолят HM175) и С (генотип Ia, изолят Н), краснухи (штамм Therien), полиовирус (штамм Sabin), вирус лихорадки Денге (штамм Nauru/West Pac/1974), вирус Зика (штамм Mr 766) и вирус тяжелого острого респираторного синдрома (ВТОРС) (изолят Tor2)], вирус с фрагментированным (–)онРНК геномом [вирус гриппа (A/California/08/2009(H1N1)] и вирус с обратной транскриптазой [вирус иммунодефицита человека 1 (ВИЧ 1, группа М подтип В (изолят HXB2)], Источником первичных структур вирусных геномов и их общих параметров служили доступные в Интернете базы данных (www.ncbi.nlm.nih.gov, <http://viralzone.expasy.org>).

Нуклеотидные последовательности геномов вирусов подвергались компьютерному анализу на присутствие аутоКП, а между парами вирусов выявляли взаимно КП и ГП длиной в 21 нуклеотид, допуская наличие в них соответственно 1–3 неспаренных нуклеотидов и 1–3 неидентичных позиций.

Выбор длины фрагментов в 21 нуклеотид был мотивирован с распространенностью ее у мвРНК [4], длиной малых интерферирующих РНК эукариот [12] и последовательностей в *tracrRNA* (адаптивная система CRISPR/Cas бактерий и архей), которые осуществляют комплементарное связывание с геномом фагов [5], и внРНК вируса гриппа А [13, 14, 20].

Результаты

Выполненный анализ по распространенности ГП свидетельствует о существовании двух ее экстремальных показателей: с одной стороны, имеет место наличие у одного и того же вируса ГП практически ко всем другим вирусам и многочисленности у него же ГП к одному и тому же другому вирусу (далее кратко оба показателя будем именовать соответственно «множественностью ГП» и «многократностью ГП»), и, с другой стороны, редкая встречаемость и немногочисленность ГП для некоторых вирусов. То же самое отмечается и в распространенности КП. Большинство же исследованных вирусов по обоим показателям ГП имеют промежуточные величины между экстремальными значениями.

Даже при известной высокой скорости репликации и изменчивости спонтанное возникновение в пределах генома одного и того же вируса более 40 ГП и не меньшего числа КП (например, у ВТОРС) длиной в 21 нуклеотид и уровнем идентичности $\geq 85\%$ (при минимальности множественности и многократности ГП у другого вируса, например у вируса краснухи) представляется маловероятным, и единственной

альтернативой объяснению существования ГП (как и КП) у вирусов остается возникновение их за счет рекомбинации между ними.

Анализ геномов различных вирусов другими авторами также показал, что между разными по своей природе вирусами происходила рекомбинация геномов [8, 10, 17]. Рекомбинировали между собой также РНК- и ДНК-содержащие вирусы. В частности, у вирусов с двуспиральной ДНК выявлены последовательности генов вирусов с онДНК (например в геноме вируса герпеса 7 выявлены последовательности генома парвовируса), а у вирусов с онДНК выявлены последовательности генов вирусов с онРНК [17]. У более чем 3000 вирусных геномов, охватывающих 120 000 генов, 10% вирусных генов являются химерами. В целом (+) онРНК- и онДНК-содержащие вирусы были более склонны обладать химерными генами, чем (–) онРНК-содержащие вирусы [8]. Поэтому далее по тексту изложение результатов и обсуждение их ведется именно в аспекте распространенности рекомбинации среди вирусов, происходившей многократно в прошлом.

Для удобства и упрощения визуализации сети связей между вирусами по наличию в них ГП и КП в рисунках 1 и 3 были использованы две параллельные идентичные колонки, в строках которых приведены названия вирусов. Строки между колонками соединялись попарно в соответствии с наличием у вирусов ГП (или КП). Для каждого вируса общее число других вирусов, с которыми он разделяет ГП (или КП), равно сумме лучей, исходящих от него в правой и левой колонках.

В целом исследование показало, что ГП охватывают разные пары вирусов и каждый вирус отличается по множественности и многократности ГП. В анализируемой нами выборке большинство вирусов с (+)онРНК, что отражает существенное их преобладание среди вирусов не только человека, но и эукариот. Из многочисленного списка выявленных пар ГП фрагментов на рисунке 2 представлены некоторые из них. Представителей *Mononegavirales* отличает близкое генетическое родство, что четко прослеживается в порядке расположения генов в их геномах и наличием протяженных гомологичных последовательностей особенно в гене РНК-полимеразы. Поэтому на рисунке 2 не представлены те пары ГП, которые образованы двумя представителями *Mononegavirales* (вирусы кори, паротита, Эбола и Борна).

На рисунке 1, отображающем попарно связи между всеми исследованными нами вирусами, видно, что вирусы Эбола, паротита и ТОРС имеют ГП с наибольшим числом вирусов, в то время как вирусы краснухи, гепатита А и гепатита В не имеют их с большинством вирусов. Поскольку первые принадлежат к вирусам с онРНК разной полярности, то из этого можно было бы предположить, что полярность РНК не является барьер-

ром для рекомбинации между вирусами. Однако, в противоположность вирусу Эбола, другие представители семейства *Mononegavirales* — вирусы кори, паротита, Борна — чаще всего не разделяли ГП с вирусами, имеющими (+) онРНК. Общая особенность вирусов ТОРС, Эбола и паротита — самые большие геномы среди исследованных нами вирусов (соответственно 29,75; 18,96 и 15 kb), а первые два из них имеют наибольшие среди РНК-содержащих вирусов геномы, что, по-видимому, определяет их повышенный потенциал к рекомбинации и «ассимилированию» внедряющихся в их геном новых фрагментов. Для ВТОРС характерна и многократность ГП к большинству вирусов и особенно к вирусам денге, Эбола и полиовирусу (до 10 разных ГП). Меньшие показатели по многократности ГП имеют пары ВТОРС—вирус паротита, ВТОРС—ВИЧ 1, ВИЧ 1—вирус денге, вирус гепатита С—вирус Эбола, вирус денге—вирус Эбола. Следует подчеркнуть, что связность в парах вирусов по обладанию ими ГП (рис. 1) в значительной степени обусловлена внедрением в вирусы с крупным геномом фрагментов генома разных вирусов (и с крупным, и с малым геномом) и, в меньшей степени, приобретением вирусами с малыми геномами фрагментов генома других вирусов.

Наиболее выраженные ограничения с рекомбинацией у вирусов краснухи, гепатита А и гепатита В (у всех троих геном (+)онРНК) по-видимому связаны с разными причинами, в числе которых их уникальность. К примеру, вирус краснухи выделяется среди всех РНК-содержащих вирусов наиболее высоким процентом ГЦ (69,6%), и вторжение в его геном фрагментов с более низким процентом ГЦ, как и вторжение в геном других вирусов фрагментов РНК с высоким процентом ГЦ, привело бы к структурно-функциональным коллизиям. РНК вируса гепатита А отличается наиболее низким процентом ГЦ (37,9%) и малой длиной. Для вирус гепатита В свойственны очень маленький размер генома, уникальная среди вирусов его организация (частично двуспиральная ДНК) и наиболее сложный среди других вирусов жизненный цикл, который оказался в эволюции малоперспективным, судя по крайней малочисленности группы VII по классификации Балтимора, в которой семейство *Hepadnaviridae* представлено лишь вирусом гепатита В.

Картина распространенности среди вирусов КП частично перекрывается с таковой для ГП, но в целом, как и в случае ГП, КП охватывают разные пары вирусов и каждый вирус отличает-

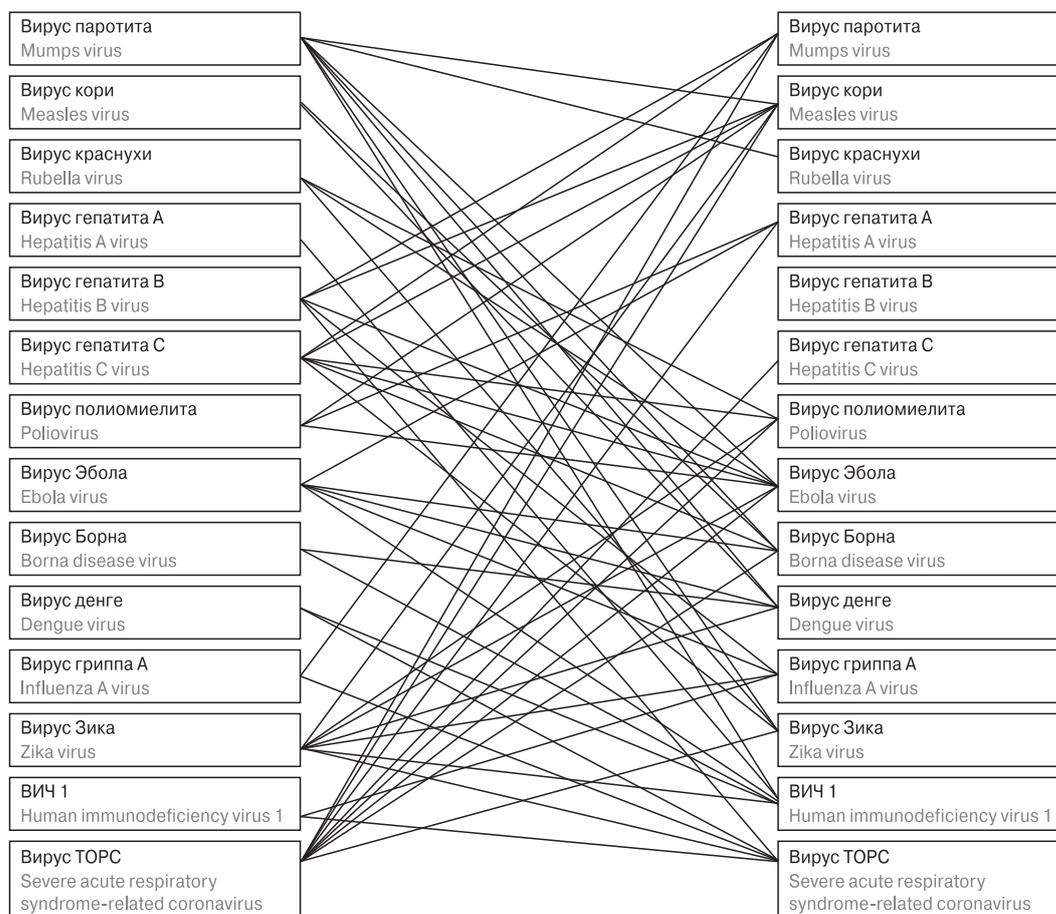


Рисунок 1. Схема связности вирусов по наличию в их геномах гомологичных последовательностей
 Figure 1. Connection scheme of viruses according to the presence of homologous sequences in their genomes

AATCAGACAAATTGTATCATAA . . . AATGAGACAAATTGTTTCAATCA	(722–743) Полиовирус / Poliovirus (11693–11714) Вирус кори / Measles virus
AGGCATTCTTTCATCTCCTTTT . . AGGCATTTTCTTCTCCTTTT	(10706–10727) Вирус Эбола / Ebola virus (4678–4699) Вирус гепатита А / Hepatitis A virus
TCAGCTCACACCCCTTGAGAG . . . TCAGCTACCACCGCTTGAGAG	(7379–7400) Вирус Эбола / Ebola virus (8066–8087) ВИЧ 1 / Human immunodeficiency virus 1
AAAGAAAAGCTGGCCTAACAA . . . AAGGAAAAGCTGGCCTCTCAA	(3105–3126) Вирус Эбола / Ebola virus (4122–4143) Вирус денге / Dengue virus
AAAAGTGATACAGCAAATTTT . . . AAAAC TGATACACCAAGATTTT	(11914–11935) Вирус кори / Measles virus (2232–2253) Вирус денге / Dengue virus
CAGAGGAGAGCAAGAAATGGA . . CAGAGGAGAGGAAGAAATGGA	(5360–5381) ВИЧ 1 / Human immunodeficiency virus 1 (5887–5908) Вирус денге / Dengue virus
CTTTTGGAAGAGAAACAGTTA . . . CTTTTGAAAGAGCAACCGTTA	(2227–2248) Вирус гепатита В / Hepatitis B virus (1255–1276) Вирус гриппа А, белок NP / Influenza A virus, NP protein
AGAATCAAAACTATGCCTTG C . . . AGAATCAAAACTAGGCTTTT C	(5573–5594) Вирус паротита / Mumps virus (507–528) Вирус гриппа А, ген PA / Influenza A virus, PA gene
ACCSTTTAGAGACTATGTAGA . . . AAACTTTAGAGCCTATGTAGA	(1207–1228) ВИЧ 1 / Human immunodeficiency virus 1 (680–701) Вирус гриппа А, ген PA / Influenza A virus, PA gene
TATATGAACTCCAGGGGCTT . . . TACATGAACACCCCGGGGCTT	(11658–11679) Вирус ТОРС / Severe acute respiratory syndrome-related coronavirus (4964–4985) Вирус гепатита С / Hepatitis C virus
TGTCATAAАСТАСТААААТСТС . . . TGTCATAACAАСТGACATCTC	(8450–8471) Вирус ТОРС / Severe acute respiratory syndrome-related coronavirus (5814–5835) Вирус Зика / Zika virus
TTCTTGCTGCATTGGTTTGT . . . TTCTTGCTGCATTGGTTAAAT	(8536–8557) Вирус ТОРС / Severe acute respiratory syndrome-related coronavirus (3586–3607) Вирус гепатита А / Hepatitis A virus
AGGAACTGTATCCTGGAATTT . . . AGGAACTGTATCCTTTAACTT	(4295–4316) Вирус ТОРС / Severe acute respiratory syndrome-related coronavirus (1776–1797) ВИЧ 1 / Human immunodeficiency virus 1

Рисунок 2 (окончание, 1-й фрагмент на с. 397). Гомологичные фрагменты геномов вирусов

Figure 2 (last part, 1st part see on page 397). The homologous fragments of virus genomes

ся по множественности и многократности КП. На рисунке 4 представлены примеры пар КП РНК-содержащих вирусов и вирус гепатита В, а на рисунке 3 схематично представлена связность каждого вируса с теми вирусами, с которыми он разделяет КП.

Аналогично распространенности ГП среди вирусов, один и тот же вирус может иметь в его геноме несколько КП к геному другого вируса или вовсе их не иметь. Заведомо было предсказуемо, что вирус краснухи, имеющий резко отличный от других РНК-содержащих вирусов

нуклеотидный состав генома, будет отличаться по максимальному числу вирусов, к которым он не имеет КП. По этому показателю вирус краснухи минимально уступают лишь вирусы гепатита В, Зика, Борна и ВИЧ 1. ВТОРС, как и в случае распространенности ГП, имеет КП к геномам всех других исследованных вирусов (за исключением вируса краснухи), и ему уступают лишь вирусы паротита, кори, гриппа А и денге. Многократность КП у ВТОРС особенно высока к вирусам гепатита С и Эбола, поменьше она к вирусам денге, вирус кори, вирус гепатита А и вирус парото-

тата. По несколько КП выявлены в таких парах вирусов как вирус гепатита С—полиовирус, вирус гепатита С—вирус денге, вирус паротита—вирус денге, вирус паротита — ВИЧ 1, вирус кори—вирус гепатита А, вирус Эбола—вирус гепатита А.

Что касается распространенности аутоКП длиной в 21 нуклеотид с уровнем комплементарности ≥ 18 нуклеотидов, то лишь в геноме ВТОРС, вирусов краснухи и гепатита С их несколько, а у остальных по 1 либо 0. Однако аутоКП меньшей длины (< 18 нуклеотидов) были многочисленны у всех исследованных вирусов, что предполагает, с одной стороны, наличие у вирусов разветвленной вторичной структуры онРНК, и, с другой стороны, возможной способности РНК дуплексироваться при репликации.

Обсуждение

Роль вРНК рассматривалась неоднократно в обширных обзорах в разных аспектах [5, 11, 19], и в обсуждении полученных нами результатов будут затронуты другие аспекты их возможных функций. Выявленные в настоящем исследовании в геномах вирусов 3 категории последовательностей — аутоКП, ГП и КП — потенциально

могут играть роль в различных процессах жизненного цикла вирусов и соответственно влиять на течение инфекционного процесса, который они вызывают у человека. Так аутоКП и КП, в дополнение к точечным мутациям, коротким делециям и вставкам в геноме вируса, способны обеспечить внутригеномную рекомбинацию между разными его штаммами и обеспечивать тем самым более быстрый (скачкообразный) механизм изменчивости вируса, резко изменяя его патогенность и антигенность. Обновленная антигенность будет проявляться в случае сдвига рамки транслирования нового фрагмента генома, обретенного в результате рекомбинации, на 1–2 нуклеотида. Многочисленность коротких аутоКП и КП в геномах вирусов, в частности у вирусов с онРНК, позволяет формировать множество вариантов дуплексов между геномами при инфекции клетки, при этом результатом рекомбинации может быть как приобретение его генами новых фрагментов либо утраты в них отдельных фрагментов. Хотя редким, но возможным событием могло бы быть участие аутоКП и КП в РНК/РНК-рекомбинации, приводящей к реактивации дефектных персистирующих вирусов, в частности вирусных штаммов, используемых в живых

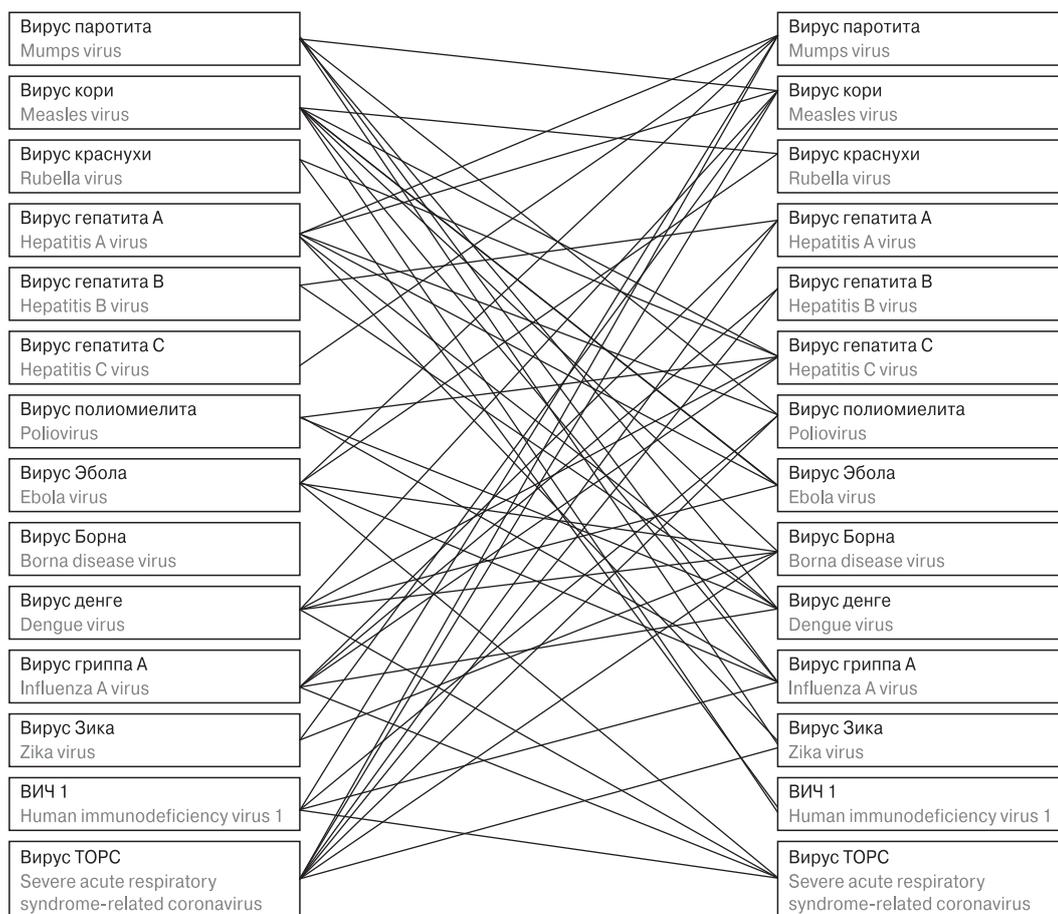


Рисунок 3. Схема связности вирусов по наличию в их геномах комплементарных последовательностей

Figure 3. Connection scheme of viruses according to the presence of complementary sequences in their genomes

A C A T T T T T T T G T C T A T T C A A A . . . T G T T A A A A A C A G G T T A G T T T	(11713–11731) Вирус кори / Measles virus
A G G T G G A T C T C C T C T A C G C A T . . . T C C A T C T A G A G G A G A A G A G T A	(7049–7031) Вирус гепатита А / Hepatitis A virus
A G G T G G A T C T C C T C T A C G C A T . . . T C C A T C T A G A G G A G A A G A G T A	(1743–1761) Вирус Борна / Borna disease virus
T G G G C A T C C A T G T T G C T G G A G . . . A C C C G G A G G T A C A A C C A C T T C	(6150–6132) Вирус денге / Dengue virus
T G G G C A T C C A T G T T G C T G G A G . . . A C C C G G A G G T A C A A C C A C T T C	(5852–5870) Вирус гепатита А / Hepatitis A virus
A A T G C T T A T T A T A T T T T G G T T . . T T A C C A A T A A T A G A A A C C A A	(9422–9404) Вирус денге / Dengue virus
A A T G C T T A T T A T A T T T T G G T T . . T T A C C A A T A A T A G A A A C C A A	(12706–12724) Вирус ТОРС / Severe acute respiratory syndrome-related coronavirus
T T C C T T C T G G C C C T G C T C T C T . . . A A C T A A G A A C G G G A C G A G A G A	(29–11) Вирус гриппа А, ген NA / Influenza A virus, NA gene
T T C C T T C T G G C C C T G C T C T C T . . . A A C T A A G A A C G G G A C G A G A G A	(869–887) Вирус гепатита С / Hepatitis C virus
G T T T T T G C C T T C T G A C T T C T T . . . T A A A A A C G G C A G A C T C A A G A A	(514–496) Вирус гриппа А, ген PA / Influenza A virus, PA gene
G T T T T T G C C T T C T G A C T T C T T . . . T A A A A A C G G C A G A C T C A A G A A	(1952–1970) Вирус гепатита В / Hepatitis B virus
G T T T T T G C C T T C T G A C T T C T T . . . T A A A A A C G G C A G A C T C A A G A A	(2271–2253) Вирус гриппа А, ген PB1 / Influenza A virus, PB1 gene

Рисунок 4 (окончание). Комплементарные фрагменты геномов вирусов

Figure 4 (last part). The complementary fragments of virus genomes

вакцинах. Напоминанием о возможности такого сценария служит известный печальный пример с живой полиомиелитной вакциной. Использование живых гриппозных вакцин также таит опасность реверсирования вакцинальных штаммов в дикие. Оно может реализоваться по давно обсуждаемому механизму, согласно которому РНК-зависимая РНК-полимераза, скользящая в процессе репликации по одной нити РНК, может переместиться (неоднократно) на другую комплексированную с ней нить РНК, обеспечивая обмен генетическими фрагментами между реплицируемыми нитями РНК [15]. Очевидно, что успех и частота рекомбинации между штаммами будет определяться уровнем консервативности аутоКП и КП в их геномах.

Поскольку онРНК-содержащие вирусы имеют множество аутоКП, что позволяет формировать их геномам различные варианты фолдинга и асимметричные из них дуплексы, то при заражении организма даже одним вирионом может развиваться хронический инфекционный процесс с репликацией множества вариантов генома заразившего организм вириона, в которых делеции и вставки будут порождены по упомянутому механизму смены матрицы РНК-зависимой РНК-полимеразой. Такой сценарий не исключен при ВИЧ-инфекции, поскольку в 70–80% случаев достаточно одного вириона либо инфицированной им клетки, для того чтобы установилась продуктивная клиническая инфекция, и вторгшийся в организм вирион вскоре эволюционирует в разные подтипы [9].

Если аутоКП хорошо объяснима изменчивостью подтипов вируса, то при коинфекции КП между разными вирусами (с онРНК) позволяют реализовать между ними РНК/РНК рекомбина-

цию по механизму смены матрицы репликации, даже если геномы вирусов разной полярности, что является наикратчайшим путем создания многообразия вирусов в эволюции. Заметим, что эволюция вирусов служит убедительным примером того, что природа не роскошествует своими ресурсами и принцип экономии действовал при создании их многообразия. Характерно, что на уровне бактерий преобладают бактериофаги с двуспиральной ДНК, а на уровне эукариот и особенно млекопитающих преобладает распространенность вирусов с онРНК, причем численность вирусов с (+)онРНК намного выше таковой у вирусов с (–)онРНК. Почему такая разница? Во-первых, «материальные» и временные затраты на воспроизведение ДНК-содержащих вирусов выше, чем на вирусы с онРНК. Во-вторых, создание многообразия РНК-содержащих вирусов и проще, и быстрее у вирусов с онРНК обеих полярностей. В-третьих, преобладание численности вирусов с (+) онРНК над вирусами с другим типом нуклеиновой кислоты можно было бы объяснить временным преимуществом — их геном при попадании в клетку может прямо функционировать как мРНК.

Существование в вирусных геномах обеих полярностей как аутоКП, так и КП, позволяющих реализоваться механизму смены матрицы репликации, сыграло, по-видимому, решающую роль в возникновении многообразия их представителей, например в минус-геномных семействах порядка *Mononegavirales* или плюс-геномном семействе *Picornaviridae*, отличающихся однотипностью организации порядка генов в их геномах и варьированием размеров генов и выраженной гомологией их РНК-зависимых РНК-полимераз. Помимо участия КП в форми-

ровании многообразия вирусов в процессе эволюции, в случае вирусов, имеющих фрагментарный геном и множество КП между фрагментами их геномов, КП могли бы служить как факторы, обеспечивающие при сборке вирионов комплектацию полным набором их генов, например в случае вируса гриппа.

Обнаружение ГП среди РНК-содержащих вирусов подтверждает существование в природе множественной глубокой рекомбинации среди вирусов [8, 10, 17], которая могла бы происходить при коинфицировании ими человека или в других общих для них хозяевах. Не следует полагать, что наличие в геномах вирусов ГП обязательно отображается в гомологичные пептидные последовательности тех белков вирусов, на гены которых приходится соответствующие ГП, так как рамка считывания рекомбинированных фрагментов может быть сдвинута на 1 или 2 нуклеотида и ГП в геноме могут соответствовать совершенно несходные последовательности пептидных фрагментов. Поэтому судить о происхождении белков только по сравнению их первичных структур было бы недостаточно. В случае же отсутствия сдвига рамки считывания рекомбинированных фрагментов генов у разных вирусов, в соответствующих белках, функции которых совершенно различны, могут быть выявлены гомологичные или даже идентичные гептапептиды, что может послужит дополнительным источником полиреактивности антител и ошибок иммунодиагностики [2, 3]. По существу распространенность ГП в геномах вирусов предполагает наличие пептидного континуума родства белков вирусов [2]. Кодированные ГП гептапептиды могут составлять «ядро» эпитопов, узнаваемых обоими классами главного комплекса гистосовместимости, и быть причиной гетерологичного иммунитета, под которым понимается феномен реактивации вторым неродственным вирусом Т-клеток памяти, генерированных в ответ на ранее перенесенную инфекцию другим (первым) вирусом, содержащим в его белках гомологичные ко второму вирусу эпитопы [16]. Проявление гетерологичного иммунитета может быть двояким. Если индуцированные вторым вирусом антитела либо эффекторные CD8 Т-лимфоциты к первому вирусу обладают протективным эффектом против самого второго вируса, то они ослабят либо вовсе заблокируют развитие инфекционного процесса у пациента, а при слабой их аффинности индуцируемая ими реактивация иммунной системы лишь усилит ее повреждающее действие на организм, обуславливая более тяжелое течение инфекции, вызванной вторым вирусом.

В практическом аспекте заслуживают внимательного изучения проявления взаимного гетерологичного иммунитета при инфицировании детей вирусами кори и паротита, поскольку белки этих вирусов содержат даже длинные идентичные последовательности. Известно, что аттенуирован-

ная коревая вакцина обладает протективным эффектом, уменьшая заболеваемость и смертность от неродственных патогенов [6]. Наличие множества ГП между вирусом кори и к вирусу паротита может послужить перспективой для создания пептидной, рекомбинантной или генной вакцины, активной против них обоих. Поскольку наиболее крупные РНК-содержащие вирусы (ВТОРС, вирусы Эбола и паротита) отличаются наиболее высоким содержанием ГП к другим вирусам, и, следовательно, имеют наиболее представительный набор иммунных эпитопов, родственных таковым у разных вирусов, то, имея в виду способность вакцин вызывать гетерологичные эффекты, было бы интересно проследить, как вакцинация против вирусов Эбола и паротита ассоциируется у иммунизированных лиц с противостоянием к другим неродственным инфекциям и появлением в их адаптивной иммунной системе клеток, реактивных к их возбудителям.

ГП и КП, образующиеся в процессе деградации вирусной геномной или мРНК, либо транскрипты с них от одного вируса, связываясь с белками хозяина, могли бы блокировать репликацию, транскрипцию и трансляцию других вирусов при коинфекции, что можно рассматривать как проявление интерференции. Интерференция вирусов — широко распространенный феномен среди вирусов животных (первоначально был описан у бактериофагов) и определяется как взаимодействие вирусов, при котором происходит подавление одним вирусом репродукции другого вируса и изменение им инфекционного процесса посредством соучастия различных молекулярных механизмов [1]. При распространенности везикулярного транспорта молекул и даже вирионов из клетки в клетку коинфицирование одной и той клетки двумя разными вирусами могло бы быть нередким событием, однако встречается оно редко, что могло бы быть обусловлено существованием интерференции между вирусами, реализуемый через механизм, подобный системе CRISPR/Cas бактерий и архей. Не ясно, который из пары вирусов с ГП и КП будет доминирующим при коинфицировании, но очевидно то, что исход интерференции будет определяться множеством факторов и в их числе жизненный цикл вирусов, фолдинг и доступность их РНК и транскриптов с ГП и КП, не исключаяющими вовлеченность и гетерологичного иммунитета. К примеру, весь жизненный цикл вирус Борна протекает в ядре, а вирус гепатита С — в цитоплазме, то есть их жизненные циклы разобщены в клетке и наличие в их геномах ГП и КП может не проявиться в интерференции.

Как известно, существующая у бактерий и архей адаптивная защитная система CRISPR/Cas построена на механизме включения в их геном коротких фрагментов генома поражающих их вирусов и использования транскриптов с этих фрагментов для специфической посадки на геном вторгнувшегося вируса нуклеазы, расщепля-

ющей его [5, 7]. Открытие CRISPR/Cas свидетельствует о том, что способность распознавания «не-своего» и защита от него возникла уже на самых ранних этапах эволюции и представлена множеством механизмов. Этот эффективный и экономичный принцип защиты от вирусной инфекции не был утрачен в эволюции эукариот. Он активно используется для регуляции транскрипции и трансляции и при том не только с участием малых РНК. Регуляторная роль РНК представлена сложной сетью взаимодействия различных типов РНК: кодирующими белок матричными РНК и некодирующими РНК (длинными некодирующими РНК, псевдогенами и циркуляторными РНК). Эти РНК-транскрипты выступают как конкурирующие эндогенные «губки» микроРНК: они сообщаются и регулируют друг друга, конкурируя за связывание с комплементарными к ним микроРНК. Открытые первоначально в клетках как регуляторы экспрессии генома и трансляции белков микроРНК циркулируют, как оказалось, в стабильной форме в разных жидкостях тела, включая кровь, и могут служить в качестве нового поколения патогенетических и прогностических биомаркеров [18].

Для аргументации возможной роли ГП и КП в интерференции обратимся к вирусам герпеса и краснухи, отличающимся среди всех известных организмов самым высоким процентом ГЦ в их геноме и влекущим два разных исхода при инфекции. Особенность вирусов группы герпеса — заселение ими почти каждого из нас после рождения на всю жизнь, и они не блокируемы инфицированием организма другими вирусами, что можно было бы объяснить отсутствием у других вирусов (из-за резкого отличия нуклеотидного состава их геномов) общих с вирусами герпеса ГП и КП. Вирус краснухи отличается от других исследованных нами вирусов отсутствием к большинству из них ГП и КП или наличием лишь минимума их к остальным вирусам и поражает почти каждого из нас в детстве или, реже, во взрослом состоянии, вызывая пожизненный иммунитет. Напрашивается предположение — вирусы с высоким процентом ГЦ в геноме вне конкуренции (интерферируют) с другими вирусами за инфицирование ими хозяина.

Завершая обсуждение распространенности ГП и КП геноме вирусов, нельзя не упомянуть в этой связи о коллизиях, которые они могли бы создать для методов диагностики природы вирусной инфекции, основанных на использо-

вании ПЦР. В случае совпадения праймера с ГП или КП диагностика с помощью ПЦР может привести к ложным результатам. Риск возникновения ошибок особенно велик для вирусов, характеризующихся множественностью и многократностью ГП и КП.

В заключение хотелось бы подчеркнуть, что биоинформатика позволяет анализировать одновременно множество вирусов на уровне и нуклеиновых кислот, и белков, выявляя новые данные, позволяющие сделать обобщающие выводы о происхождении и эволюции вирусов, которые невозможно извлечь экспериментально [8, 10, 17, 19], и по-новому взглянуть на существующие представления о вирусных инфекциях и последствиях вакцинации против них. Выявленная компьютерным анализом распространенность малых ГП и КП среди геномов разных вирусов человека можно рассматривать как результат многократной рекомбинации между ними, свершавшейся в прошлом и возможной в настоящем и определяющей их изменчивость и адаптацию. Чтобы ускользнуть от иммунной системы хозяина, природа наделила вирусы также способностью к молекулярной мимикрии. Она заключается в подгонке структуры белков вируса к структуре белков хозяина. Такая маскировка обеспечивает вирусу мир не только с надзирающей иммунной системой хозяина, но и большую совместимость его белков с клеточной средой, в которой он обитает, и многосторонняя генетическая рекомбинация наиболее быстро обеспечивает такую адаптацию вируса. Включение в геном вирусов ГП и КП не только обновляло их, но и могло бы служить и памятью о существовании конкурента (противника) за овладение хозяином, и средством противодействия конкуренту при коинфицировании, являясь аналогией системы CRISPR/Cas бактерий и архей. Не исключено, что ГП и КП, высвобождающиеся в результате деградации части реплицированных геномов или их мРНК, вовлечены и в другие функции в жизненном цикле вируса, тем самым давая основание предполагать, что и метаболизм вирусных компонентов является важным звеном для реализации инфекции. В аспекте полифункциональности можно также рассматривать и существование аутоКП. Помимо их возможной роли в рекомбинации внутритиповых штаммов, они могли бы через фолдинг геномов и мРНК быть регуляторами скорости трансляции вирусных белков, обеспечивая оптимальное количественное соотношение их для сборки вирионов.

Список литературы/References

1. Стент Г. Молекулярная биология вирусов и бактерий. М.: Мир, 1965. 467 с. [Stent G. Molekulyarnaya biologiya virusov i bakterii [Molecular biology of viruses and bacteria]. Moscow: Mir, 1965. 467 p.]
2. Харченко Е.П. Возможные коллизии в иммунодиагностике вирусных инфекций и вакцинации // Инфекция и иммунитет. 2016. Т. 6, № 2. С. 157–164. [Kharchenko E.P. Immune epitope continuum of the protein relationships, poly- and autoreactivity of antibodies. *Infektsiya i immunitet = Russian Journal of Infection and Immunity*, 2016, vol. 6, no. 2, pp. 157–164. doi: 10.15789/2220-7619-20162-157-164 (In Russ.)]

3. Харченко Е.П. Иммуноэпитопный континуум родства белков и полиреактивность и аутореактивность антител // Медицинская иммунология. 2015. Т. 17, № 4. С. 335–346. [Kharchenko E.P. Immune epitope continuum of the protein relationships, poly- and autoreactivity of antibodies. *Meditsinskaya Immunologiya = Medical Immunology (Russia)*, 2015, vol. 17, no. 4, pp. 335–346. doi: 10.15789/1563-0625-2015-4-335-346 (In Russ.)]
4. Aguiar E.R., Olmo R.P., Marques J.T. Virus-derived small RNAs: molecular footprints of host–pathogen interactions. *Wiley Interdiscip. Rev. RNA*, 2016, vol. 7, iss. 6, pp. 824–837. doi: 10.1002/wrna.1361
5. Barrangou R. CRISPR–Cas systems and RNA-guided interference. *Wiley Interdiscip. Rev. RNA*, 2013, vol. 4, iss. 3, pp. 267–278. doi: 10.1002/wrna.1159
6. Gil A., Kenney L.L., Mishra R., Watkin L.B., Aslan N., Selin L.K. Vaccination and heterologous immunity: educating the immune system. *Trans. R. Soc. Trop. Med. Hyg.*, 2015, vol. 109, no. 1, pp. 62–69. doi: 10.1093/trstmh/tru198
7. Heler R., Marraffini L.A., Bikard D. Adapting to new threats: the generation of memory by CRISPR–Cas immune systems. *Mol. Microbiol.*, 2014, vol. 93, iss. 1, pp. 1–9. doi: 10.1111/mmi.12640
8. Jachiet P.A., Colson P., Lopez P., Baptiste E. Extensive gene remodeling in the viral world: new evidence for nongradual evolution in the mobilome network. *Genome Biol. Evol.*, 2014, vol. 6, iss. 9, pp. 2195–2205. doi: 10.1093/gbe/evu168
9. Keele B.F., Giorgi E.E., Salazar-Gonzalez J.F., Decker J.M., Pham K.T., Salazar M.G., Sun C., Grayson T., Wang S., Li H., Wei X., Jiang C., Kirchherr J., Gao F., Anderson J., Ping L., Swanstrom R., Tomaras G., Blattner W., Goepfert P., Kilby J., Saag M., Delwart E., Busch M., Cohen M., Montefiori D., Haynes B., Gaschen B., Athreya G., Lee H., Wood N., Seoighe C., Perelson A., Bhattacharya T., Korber B.T., Hahn B., Shaw G. Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. *Proc. Natl. Acad. Sci. USA*, 2008, vol. 105, no. 21, pp. 7552–7557. doi: 10.1073/pnas.0802203105
10. Koonin E., Dolja V., Krupovic M. Origins and evolution of viruses of eukaryotes: The ultimate modularity. *Virology*, 2015, vol. 479–480, pp. 2–25. doi: 10.1016/j.virol.2015.02.039
11. Li M.L., Weng K.F., Shih S.R., Brewer G. The evolving world of small RNAs from RNA viruses. *Wiley Interdiscip. Rev. RNA*, 2016, vol. 7, iss. 5, pp. 575–588. doi: 10.1002/wrna.1351
12. Nakanishi K. Anatomy of RISC: how do small RNAs and chaperones activate Argonaute proteins? *Wiley Interdiscip. Rev. RNA*, 2016, vol. 7, iss. 5, pp. 637–660. doi: 10.1002/wrna.1356
13. Perez J.T., Zlatev I., Aggarwal S., Subramanian S., Sachidanandam R., Kim B., Manoharan M., ten Oever B.R. A small-RNA enhancer of viral polymerase activity. *J. Virol.*, 2012, vol. 86, no. 24, pp. 13475–13485. doi: 10.1128/JVI.02295-12
14. Perez J.T., Varble A., Sachidanandam R., Zlatev I., Manoharan M., Garcia-Sastre A., ten Oever B.R. Influenza A virus-generated small RNAs regulate the switch from transcription to replication. *Proc. Natl. Acad. Sci. USA*, 2010, vol. 107, no. 25, pp. 11525–11530. doi: 10.1073/pnas.1001984107
15. Romanova L.I., Blinov V.M., Tolskaya E.A., Viktorova E.G., Kolesnikova M.S., Guseva E.A., Agol V.I. The primary structure of crossover regions of intertypic poliovirus recombinants: a model of recombination between RNA genomes. *Virology*, 1986, vol. 155, no. 1, pp. 202–213.
16. Selin L.K., Wlodarczyk M.F., Kraft A.R., Nie S., Kenney L.L., Puzone R., Celada F. Heterologous immunity: immunopathology, autoimmunity and protection during viral infections. *Autoimmunity*, 2011, vol. 44, pp. 328–347.
17. Stedman K.M. Deep recombination: RNA and ssDNA virus genes in DNA virus and host genomes. *Annu. Rev. Virol.*, 2015, vol. 2, pp. 203–217. doi: 10.1146/annurev-virology-100114-055127
18. Tay Y., Rinn J., Pandolfi P.P. The multilayered complexity of ceRNA crosstalk and competition. *Nature*, 2014, vol. 505, no. 7483, pp. 344–352. doi: 10.1038/nature12986
19. Tycowski K.T., Guo Y., Lee N., Moss W.N., Vallery T.K., Xie M., Steitz J.A. Viral noncoding RNAs: more surprises. *Genes. Dev.*, 2015, vol. 29, pp. 567–584.
20. Umbach J.L., Yen H.L., Poon L.L., Cullen B.R. Influenza A virus expresses high levels of an unusual class of small viral leader RNAs in infected cells. *MBio*, 2010, vol. 1, no. 4: e00204-10. doi: 10.1128/mBio.00204-10

Автор:

Харченко Е.П., д.б.н., ведущий научный сотрудник Института эволюционной физиологии и биохимии им. И.М. Сеченова РАН, Санкт-Петербург, Россия.

Author:

Kharchenko E.P., PhD, MD (Biology), Senior Researcher, Sechenov Institute of Evolutionary Physiology and Biochemistry of the Russian Academy of Sciences, St. Petersburg, Russian Federation.

Поступила в редакцию 05.08.2016
Принята к печати 16.10.2017

Received 05.08.2016
Accepted 16.10.2017