УДК: 577.2:519.23

Структурно-статистические свойства геномов флавивирусов

Тюлько Ж.С.^{*1}, Кутыркин В.А.^{**2}, Чалей М.Б.^{***3}

¹ОмГМУ Минздрава России, Омск, 644099, Россия

²МГТУ им. Н.Э. Баумана, Москва, 107005, Россия

³ИМПБ РАН – филиал ИПМ им. М.В. Келдыша РАН, Пущино, Московская область,
142290, Россия

Аннотация. На основе спектрально-статистического подхода исследуются основные структурно-статистические свойства кодирующих районов из геномов флавивирусов. Анализируются как полноразмерные кодирующие районы полипротеинов, так и их отдельные структурные сегменты. Показано, что структурно-статистические свойства геномных последовательностей флавивирусов, в целом, аналогичны выявленным ранее свойствам 3-регулярности и скрытой триплетной профильности в кодирующих районах геномов прокариот и эукариот. С другой стороны, в отдельных, кодирующих белки, структурных сегментах из геномов флавивирусов полностью отсутствует двухуровневая организация кодирования, и в их значительной части проявляется свойство однородности последовательности. Такие особенности кодирующих районов флавивирусов объясняются простой структурой и высокой скоростью мутаций их геномов.

Ключевые слова: геном флавивирусов, геном прокариот, геном эукариот, спектрально-статистический подход, скрытая профильная периодичность, свойство 3-регулярности.

ВВЕДЕНИЕ

Ранее исследовались структурно-статистические свойства кодирующих районов из геномов прокариотических [1] и эукариотических организмов [2]. Было показано, что практически во всех кодирующих районах распознаётся новый тип скрытой периодичности, названный скрытой профильной периодичностью. Для описания этого типа скрытой периодичности в работе [3] были введены стохастические модели профильной периодичности, в которых паттерном периодичности является случайная строка, состоящая из независимых случайных букв, с заданными вероятностями принимающих значения из алфавита ДНК. В общем случае, последовательности, в которых распознаётся скрытая профильная периодичность, не являются размытыми тандемными повторами. Поэтому для распознавания скрытой профильной периодичности был предложен спектрально-статистический подход, опирающийся на статистические методы и критерии [3]. На основе этого подхода в кодирующих районах геномов различных организмов были выявлены следующие структурно-статистические свойства. Оказалось, что порядка 75 % кодирующих районов (CDSs – Coding DNA Sequences) в геномах бактерий и высших организмов обладают скрытой триплетной

* tjs@omsk-osma.ru

** vkutyarkin@yandex.ru

*** maramaria@yandex.ru

профильной периодичностью (скрытой триплетной профильностью) [2]. Кроме того, порядка 13 % CDSs в этих геномах обладают скрытой профильной периодичностью с периодом, кратным, но не равным трём. В таких CDSs наблюдалось свойство 3-регулярности, которое может быть названо ослабленной триплетной профильностью. Таким образом, в 13 % CDSs из рассмотренных геномов наблюдается двухуровневая организация кодирования. Первый уровень, обусловленный генетическим кодом, проявляется в свойстве 3-регулярности. Профильный период, кратный, но не равный трём, выявляет второй уровень организации кодирования, который часто коррелирует с размером повторяющихся доменов во вторичной структуре белков [2, 3]. Практически во всех остальных CDSs (~12 %) наблюдается только свойство 3-регулярности. Таким образом, выявленная в рассмотренных кодирующих районах скрытая триплетная профильная периодичность и регулярность являются аргументом против гипотезы об эволюции кодирующих последовательностей в геномах организмов путём какого-либо сдвига рамки считывания. Нарушение триплетной профильной периодичности в кодирующих районах, вероятнее всего, обусловлено точечными мутациями нуклеотидов в виде вставок и делеций.

Настоящая работа посвящена исследованию нуклеотидных последовательностей одноцепочечных РНК-геномов арбовирусов рода *Flavivirus* из семейства *Flaviviridae*. В состав рода *Flavivirus* входят более 70 видов. Большинство флавивирусов является арбовирусами и передается теплокровным хозяевам путем биологической трансмиссии, посредством членистоногих переносчиков [4, 5]. Некоторые флавивирусы открытые в последние годы специфичны только для насекомых и не передаются теплокровным [6, 7]. Многие флавивирусы, являются опаснейшими патогенами для человека и при укусе насекомых (клещи, комары) могут вызывать парезы, параличи, энцефалиты и геморрагические лихорадки с высокой летальностью [4]. Наибольшее беспокойство в настоящее время вызывает распространение по всему миру опасных болезней тропического и субтропического пояса, вызываемых такими флавивирусами, как вирусы желтой лихорадки, Зика, Дэнге, Западного Нила, японского энцефалита [8]. Для России особенно актуальна проблема эпидемического распространения вирусов клещевого энцефалита, Западного Нила, Повассан [9] и омской геморрагической лихорадки [10], заболеваемость которыми приносит ущерб не только здоровью населения, но и экономике страны [11]. Особый интерес вызывает изучение флавивирусов в связи с имеющимися данными о возможности их интродукции в новые экосистемы и способности приспосабливаться к новым переносчикам, сохраняя структуру генома [12, 13].

В быстро мутирующих геномах РНК-вирусов [14] можно ожидать отклонение от закономерностей структурно-статистических свойств, выявленных в геномах прокариот и эукариот. В первую очередь, возможна потеря свойства регулярности в отдельных кодирующих сегментах вирусного генома.

Исходные последовательности геномов флавивирусов были получены из базы GenBank [15] в виде отдельных файлов. В настоящей работе мы ограничились рассмотрением геномов, в которых CDSs специфичных белков вируса последовательно, без перекрытия располагались в CDS общего полипептида. В соответствии с данными файлов GenBank, аннотированные кодирующие последовательности вирусных белков также как и полная кодирующая последовательность полипептида выделялись для анализа.

Две выборки кодирующих геномных последовательностей флавивирусов анализировались в работе. Одна выборка содержала CDSs из геномов вирусов, переносимых комарами и клещами, и передаваемых человеку при укусе насекомого. В неё, в частности, вошли кодирующие последовательности из геномов вирусов жёлтой лихорадки, лихорадки денге, японского энцефалита и энцефалита долины Муррея,

энцефалита Сент-Луис и вируса Зика. Общая численность этой выборки составила 384 последовательности. Другая выборка содержала CDSs из геномов вирусов, специфичных только для комаров и не вызывающих заболеваний у теплокровных. В эту выборку вошли кодирующие последовательности геномов вирусов *Aedes flavivirus*, *Culex flavivirus* и *Cell fusing agent* – вируса, первоначально выделенного из линии клеток комаров *Aedes aegypti*. В этой выборке находилось 95 последовательностей.

Результаты анализа структурно-статистических свойств CDSs, полученные для геномов флавивирусов сравнивались с аналогичными результатами, полученными ранее для CDSs из геномов бактерий и многоклеточных организмов [1, 2].

МАТЕРИАЛЫ И МЕТОДЫ

Спектрально-статистический подход

Для исследования структурно-статистических свойств, как отдельных кодирующих сегментов, так и полных геномов вирусов, исследуемых в настоящей работе, применялся спектрально-статистический подход (2С-подход) [3].

При распознавании скрытой профильной периодичности в нуклеотидной последовательности этот подход опирается на анализ статистических спектров, аналитический вид которых подробно описан ранее [3]. Поэтому в настоящей работе будет использоваться только анализ графических представлений статистических спектров 2С-подхода.

Согласно 2С-подходу, оценка длины скрытого профильного периода в анализируемой нуклеотидной последовательности осуществляется на основе анализа её характеристического спектра, обозначаемого символом **C**. Однако такой анализ проводится только в том случае, когда на основе соответствующего статистического критерия последовательность признаётся неоднородной. Этот критерий использует анализ спектра отклонения этой последовательности от однородности, обозначаемого как **D₁**. Пример такого спектра для кодирующей последовательности (CDS) тетрапирольной метилазы (tetrapyrrole methylas family protein) генома бактерии *Mycoplasma mycoides* из базы данных KEGG [16] показан на рисунке 1,а. Последовательность признаётся неоднородной, если значения спектра **D₁** превышают единицу не менее чем на 5 % тест-периодов в анализируемом диапазоне. Так анализируемая последовательность, спектр **D₁** которой показан на рисунке 1,а, признаётся неоднородной.

Если анализируемая последовательность признана неоднородной, то в качестве оценки длины скрытого профильного периода в этой последовательности выбирается тот тест-период, на котором впервые (с учётом статистической погрешности) достигается максимальное значение её характеристического спектра **C**. На рисунке 1,б показан характеристический спектр для CDS тетрапирольной метилазы *M. mycoides*. Согласно этому спектру в качестве оценки длины скрытого профильного периода следует выбрать тест-период в 12 нукл.

Для подтверждения выбранной оценки **L** длины скрытого профильного периода в 2С-подходе используется спектр **D_L** отклонения анализируемой последовательности от **L**-профильности (скрытой **L**-профильной периодичности). В последовательности признаётся наличие скрытой **L**-профильности, если значения спектра **D_L** не превышают единицу не менее чем на 95 % тест-периодов в анализируемом диапазоне. На рисунке 1,в показан спектр **D₁₂** отклонения от 12-профильности анализируемой кодирующей последовательности из генома *M. mycoides*. Согласно этому спектру в ней распознаётся 12-профильность. Вместе с тем, согласно спектру **D₃** скрытая 3-профильность в этой последовательности отсутствует (см. рис. 1,г).

Рис. 1. Анализ CDS тетрапирольной метилазы (KEGG ENTRY MSC_0564, 888 нукл.) из генома бактерии *M. mycoides* с помощью 2C-подхода. (а) Спектр отклонения от однородности. (б) Характеристический спектр. (в) Спектр отклонения от скрытой 12-профильной периодичности. (г) Спектр отклонения от скрытой триплетной профильной периодичности.

Характеристический спектр CDS тетрапирольной метилазы *M. mycoides* обладает ещё одним свойством: Практически все его максимумы (пики) наблюдаются на тест-периодах, кратных трём. Такая особенность была названо в работах [1–3] свойством 3-регулярности анализируемой последовательности. В работах [1, 2] было показано, что свойство 3-регулярности в кодирующих районах обусловлено размером кодона генетического кода в 3 нукл. Этот вывод был подтверждён на основе численных экспериментов с бинарно перекодированными абзацами литературных текстов. Кроме того, в кодирующих районах прокариотических и эукариотических организмов практически все длины скрытых профильных периодов кратны трём [1, 2]. Ранее мы отмечали наличие двухуровневой организации кодирования в CDS, если в последовательности наблюдается свойство 3-регулярности и длина скрытого профильного периода кратна, но не равна трём [1–3]. Отметим, что это явление наблюдалось в значительной части CDSs геномов прокариот и эукариот [1, 2]. Согласно такому подходу, в последовательности CDS тетрапирольной метилазы *M. mycoides* обнаружена двухуровневая организация кодирования (см. рис. 1).

Согласно спектрам 2C-подхода, показанным на рисунке 2, в последовательности CDS для агматиновой деиминазы (agmatine deiminase) *M. mycoides* выявляется свойство 3-регулярности и распознаётся скрытая триплетная профильность, т.е. двухуровневая организация кодирования отсутствует.

Для сравнения структурно-статистических свойств геномов вирусов с аналогичными свойствами в кодирующих районах прокариот и эукариот использовались результаты, полученные ранее в работах [1, 2]. В этих работах были проанализированы 10 геномов различных бактерий и пять геномов эукариот, включая геном человека.

Рис. 2. Анализ спектрально-статистических свойств CDS агматинной деаминазы *M. mycoides* (KEGG ENTRY MSC_0701, 1095 нукл.) с помощью 2С-подхода. (а) Спектр отклонения от однородности. (б) Характеристический спектр. (в) Спектр отклонения от скрытой 3-профильной периодичности.

Геномные последовательности флавивирусов

Одноцепочечная РНК позитивной полярности ($\approx 11\ 000$ нукл.), представляющая собой геном флавивирусов, является инфекционной. Флавивирусный вирион имеет сферическую форму (диаметр $\approx 40\text{--}50$ нм) и состоит из нуклеокапсида и покрывающей его липопротеиновой оболочки. Нуклеокапсид включает в себя структурный элемент капсида – капсидный белок С и геномную одноцепочечную РНК. Нуклеокапсид окружен липидной мембраной, в которую включены мембранный белок М и оболочечный белок Е, взаимодействующие при сборке вириона. Вирусные белки С, М, Е и семь неструктурных белков (NS1, NS2A, NS2B, NS3, NS4A, NS4B, NS5), необходимые для размножения вируса в клетках хозяина [4], последовательно считываются в единой рамке считывания при синтезе полипротеина, процессинг которого происходит ко- и пост-трансляционно.

Белок-предшественник ргМ участвует в начальных этапах фолдинга белка Е [17]. Белок Е отвечает за сборку вириона, слияние мембран и рецепторное связывание [18]. В случае нейтрального значения рН гликопротеин Е представляет собой димер, чьи мономеры состоят из трех отдельных доменов. Считается, что мутации, влияющие на патогенность вируса, связаны с изменениями кодирующей последовательности этих трех доменов белка Е [17, 19, 20]. NS2A, NS2B, NS4A и NS4B – это низкомолекулярные гидрофобные белки, функции которых в развитии вируса в клетках пока до конца не ясны, но предполагается, что они могут взаимодействовать с другими вирусными белками и РНК. NS5 – это РНК-зависимая РНК-полимераза [4].

В настоящей работе с помощью спектрально-статистического подхода анализировались 62 последовательности полных геномов флавивирусов, представленных одноцепочечной РНК позитивной полярности. Одновременно анализировались 415 отдельных, неперекрывающихся кодирующих последовательностей из этих геномов. Более подробные данные о количественном составе проанализированных CDSs представлены в таблице 1.

Таблица 1. Основной количественный состав анализируемых CDSs из геномов флавивирусов. Вирусы, указанные под номерами 10–13 специфичны только для комаров

Protein CDS																
№	вирус	Polyprotein	anchC	C	preM	prM	M	E	NS1	NS2A	NS2B	NS3	NS4A	NS4B	NS5	2K
1.	Dengue	9	4	9	1	7	3	9	9	9	9	9	9	9	9	4
2.	Japanese encephalitis	10	0	10	0	10	1	10	10	10	10	10	10	10	10	0
3.	Kunjin	1	0	1	0	1	1	1	1	1	1	1	1	1	1	0
4.	St. Louis encephalitis	3	0	1	0	1	1	1	1	1	1	1	1	1	1	1
5.	Murray Valley encephalitis	5	1	5	1	1	1	5	5	5	5	5	5	5	5	4
6.	Yellow fever	6	0	2	0	1	2	2	2	2	2	2	2	2	2	1
7.	Zika	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8.	West Nile	2	0	2	0	2	0	2	2	2	2	2	2	2	2	0
9.	Usutu	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1
10.	Aedes flavivirus	5	1	2	1	2	2	2	2	2	2	2	2	2	2	1
11.	Cell fusing agent	4	1	3	1	3	3	3	3	3	3	3	3	3	3	3
12.	Culex flavivirus	9	0	1	0	1	1	1	2	0	1	1	1	1	1	1
13.	Palm Creek	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Обозначения: anchC – anchored capsid protein C; C – capsid protein C; preM – premembrane protein; prM – membrane glycoprotein precursor; M – membrane glycoprotein M; E – envelope protein E; NS1 – nonstructural protein NS1; NS2A – nonstructural protein NS2A; NS2B – serine protease subunit NS2B; NS3 – serine protease NS3; NS4A – nonstructural protein NS4A; NS4B – nonstructural protein NS4B; NS5 – RNA-dependent RNA polymerase NS5; 2K – peptide 2k.

Таблица 2. Коды доступа (Accession Numbers) файлов GenBank, содержащих анализируемые геномные последовательности вирусов

№	вирус	Accession Numbers
1.	Dengue	AB122020.1, AF180817.1, AF208496.1, AF326573.1, AF326825.1, AF350498.1, AF375822.1, AY099336.1, AY648961.1
2.	Japanese encephalitis	AF080251.1, AF098735.1, AF098736.1, AF098737.1, AF221499.1, AF221500.1, AF315119.1, EF571853.1, HM596272.1, JN864064.1
3.	Kunjin	D00246.1
4.	St. Louis encephalitis	DQ359217.1, FJ753286.2, KM267635.1
5.	Murray Valley encephalitis	JX123032.1, KF751869.1, KF751870.1, KF751871.1, NC_000943.1
6.	Yellow fever	DQ100292.1, DQ118157.1, DQ235229.1, FJ654700.1, JF912182.1, JN620362.1
7.	Zika	AY632535.2, DQ859059.1, EU545988.1, KF268949.1, KJ776791.2, KU497555.1
8.	West Nile	JX123030.1, JX123031.1
9.	Usutu	AY453411.1
10.	Aedes flavivirus	AB488408.1, KC181923.1, KJ741266.1, KU201526.1, NC_012932.1, NC_034017.1
11.	Cell fusing agent	AF411835.1, KJ741267.1, KU936054.1, NC_001564.2
12.	Culex flavivirus	EU879060.1, FJ663034.1, KT726939.1, AB262759.2, AB377213.1, AB701766.1, EU879060.1, FJ502995.1, FJ663034.1
13.	Palm Creek	KC505248.1

Исходные файлы с данными были получены из базы GenBank (см. таблицу 2). В основном, в работе рассматривались вирусы, представляющие опасность для теплокровных и, в частности, для человека, переносчиками которых являются комары и клещи. В отдельную группу были выделены вирусы, найденные только у комаров.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Согласно результатам работ [1, 2] практически все кодирующие последовательности (CDSs) из геномов эукариот и прокариот обладают свойством 3-регулярности. Кроме того, в более 90 % этих последовательностей распознаётся скрытая профильная периодичность с длиной скрытого профильного периода кратной трём. При этом около 77 % кодирующих районов обладают скрытой триплетной профильной периодичностью и, приблизительно, в 13% кодирующих последовательностей распознаётся двухуровневая организация кодирования с длиной периода, кратной, но не равной трём. Ранее отмечалось, что в кодирующих последовательностях с двухуровневой организацией кодирования может наблюдаться корреляция длины скрытого периода с длиной повторяющегося домена в кодируемом белке.

По результатам анализа, выполненного в настоящей работе, на рисунке 3 представлена дендрограмма структурно-статистических свойств кодирующих последовательностей из геномов флавивирусов, опасных для теплокровных животных и, в частности, для человека, которые передаются при укусах комаров и клещей. Из этой дендрограммы следует, что во всех CDSs, кодирующих полипротеин распознаётся скрытая триплетная профильная периодичность. Однако, среди CDSs отдельных структурных единиц полипротеина скрытая триплетная профильная периодичность наблюдается только в 86 % последовательностей. Остальные 14 % CDSs отдельных структурных единиц не обладают свойством 3-регулярности. Около 9 % CDSs отдельных структурных белков вируса оказываются однородными, 4 % CDSs не обладают скрытой профильностью и 0.3 % CDSs обладают скрытой профильностью с длиной периода, некратной трём.

Рис. 3. Дендрограмма структурно-статистических свойств CDSs из геномов флавивирусов, переносимых комарами и клещами и поражающих теплокровных.

На рисунке 4 приведена аналогичная дендрограмма для CDSs вирусов, не поражающих теплокровных, и специфичных для комаров. Здесь также во всех CDSs полипротеинов распознаётся скрытая триплетная профильная периодичность. Однако, в CDSs отдельных структурных единиц на 20 % снижен уровень последовательностей, в которых распознаётся скрытая триплетная профильная периодичность; кроме того, среди этих CDSs процент однородных последовательностей в три раза выше в сравнении с аналогичными CDSs вирусов, поражающих теплокровных.

Рис. 4. Дендрограмма структурно-статистических свойств CDSs из геномов флавивирусов, специфичных для комаров.

Сравнивая структурно-статистические свойства CDSs геномов прокариот и эукариот [1, 2] с аналогичными свойствами CDSs белков вирусного генома, можно отметить следующие особенности. В CDSs белков вирусного генома отсутствует двухуровневая организация кодирования, что, вероятнее всего, связано с отсутствием повторяющихся структур в CDSs вирусов. Кроме того, в CDSs генома флавивирусов наблюдается значительная доля однородных последовательностей. Возможно, это связано с высокой мутабельностью вируса [21], что может объяснять и другие особенности в отклонении от 3-профильности в CDSs вирусов (см. рис. 3 и рис. 4). С другой стороны, полученные результаты анализа структурно-статистических свойств в геномах вирусов, прокариот и эукариот не подтверждают гипотезу эволюции белков путем сдвига рамки считывания. Поскольку практически все CDSs прокариот и эукариот обладают свойством 3-регулярности, скрытой 3-профильностью или скрытой профильностью с длиной периода, кратной трём.

Возможно, что нарушение основных свойств кодирующих районов в геноме вируса вызвано высокой скоростью накопления повреждающих мутаций. Приведем пример того, как кодирующая последовательность неструктурного белка NS2B со скрытой триплетной профильностью теряет это свойство в результате двух вставок или двух делеций (см. рис. 5), превращаясь в однородную последовательность. Исходная последовательность длиной 375 нуклеотидов получена из GenBank (Accession Number J741266). На рисунке 5,а спектр отклонения от однородности D_1 исходной последовательности фиксирует её неоднородность. Спектр D_3 отклонения от триплетной профильности в этой последовательности (см. рис. 5,б) показывает наличие в ней скрытой триплетной профильности. Вставки двух нуклеотидов в исходную последовательность, показанные на рисунке 5,в, приводят к тому, что последовательность становится однородной, что подтверждает спектр D_1 на рисунке 5,г. Аналогичный эффект наблюдается при удалении двух нуклеотидов (см. рис. 5,д), что подтверждает спектр D_1 на рисунке 5,е.

Рис. 5. Иллюстрация искажения скрытой 3-профильности (а, б) до однородности в исходной последовательности CDS для неструктурного белка NS2B специфичного вируса комаров *Aedes* в результате двух вставок (в, г) и двух делеций (д, е). Вставленные и делетированные нуклеотиды выделены красным шрифтом на желтом фоне.

ЗАКЛЮЧЕНИЕ

В настоящей работе были проанализированы структурно-статистические свойства кодирующих последовательностей геномов флавивирусов и проведён сравнительный анализ с выявленными ранее свойствами кодирующих районов геномов прокариот и эукариот.

В большинстве кодирующих районов геномов флавивирусов, как и в кодирующих последовательностях геномов прокариот и эукариот, выявляются свойства 3-регулярности и скрытой триплетной профильной периодичности. Однако в кодирующих последовательностях геномов флавивирусов полностью отсутствует двухуровневая организация кодирования, которая свойственна части кодирующих районов геномов прокариотических и эукариотических организмов и обусловлена повторяющимися структурными доменами в кодируемых белках. Кроме того, для весьма значительной доли кодирующих последовательностей геномов флавивирусов характерна однородность нуклеотидной последовательности. Показано, что эффект однородности в кодирующих последовательностях геномов флавивирусов может быть вызван, как минимум, двумя точечными мутациями в виде вставок и делеций. Выявленные особенности в геномах флавивирусов могут быть обусловлены простой структурой и высокой скоростью мутаций вирусных геномов.

Работа была выполнена при частичной поддержке гранта РФФИ № 15-07-05783.

СПИСОК ЛИТЕРАТУРЫ

1. Кутыркин В.А., Чалей М.Б. Модель организации кодирования в прокариотических организмах. *Математическая биология и биоинформатика*. 2016. Т. 11. № 1. С. 24–45. doi: [10.17537/2016.11.24](https://doi.org/10.17537/2016.11.24)
2. Chaley M., Kutyrkin V. Stochastic model of homogeneous coding and latent periodicity in DNasequences. *Journal of Theoretical Biology*. 2016. V. 390. P. 106–116.
3. Кутыркин В.А., Чалей М.Б. Спектрально-статистический подход к распознаванию скрытой профильной периодичности в последовательностях ДНК. *Математическая биология и биоинформатика*. 2014. Т. 9. № 1. С. 33–62. doi: [10.17537/2014.9.33](https://doi.org/10.17537/2014.9.33)
4. *Руководство по вирусологии: Вирусы и вирусные инфекции человека и животных*. Под. ред. Львова Д.К. М.: Медицинское информационное агентство, 2013.
5. Grard G., Moureau G., Charrel R.N., Lemasson J.J., Gonzalez J.P., Gallian P., Gritsun T.S., Holmes E.C., Gould E.A., de Lamballerie X. Genetic characterization of tick-borne flaviviruses: New insights into evolution, pathogenetic determinants and taxonomy. *Virology*. 2007. V. 361. P. 80–92.
6. Blitvich B.J., Firth A.E. Insect-specific flaviviruses: a systematic review of their discovery, host range, mode of transmission, superinfection exclusion potential and genomic organization. *Viruses*. 2015. V. 7. P. 1927–1959. doi: [10.3390/v7041927](https://doi.org/10.3390/v7041927)
7. Calzolari M., Zé-Zé L., Vázquez A., Seco M.P.S., Amaro F., Dottori M. Insect-specific flaviviruses, a worldwide widespread group of viruses only detected in insects. *Infection, Genetics and Evolution*. 2016. V. 40. P. 381–388. doi: [10.1016/j.meegid.2015.07.032](https://doi.org/10.1016/j.meegid.2015.07.032)
8. Holbrook M.R. Historical Perspectives on Flavivirus Research. *Viruses*. 2017. V. 9. No. 5. P. 97. doi: [10.3390/v9050097](https://doi.org/10.3390/v9050097)
9. Subbotina E.L., Loktev V.B. Molecular evolution of the tick-borne encephalitis and Powassan viruses. *Molecular Biology*. 2012. V. 46. P. 75–84.
10. Ястребов В.К., Якименко В.В. Омская геморрагическая лихорадка: итоги исследований (1946-2013). *Вопросы вирусологии*. 2014. Т. 59. № 6. С. 5–11.
11. Платонов А.Е., Авксентьев Н.А., Авксентьева М.В., Деркач Е.В., Платонова О.В., Титков А.В., Колясникова Н.М. Социально-экономическое бремя пяти природно-очаговых инфекций в Российской Федерации. *Фармакоэкономика. Современная фармакоэкономика и фармакоэпидемиология*. 2015. Т. 8. № 1. С. 47–56.
12. Беликов С.И., Кондратов И.Г., Леонова Г.Н. Молекулярно-генетическая характеристика и эволюция вирусов группы клещевого энцефалита в дальневосточном природном очаге. *Национальные приоритеты России*. 2014. Т. 3. № 13. С. 96–99.
13. Huang Y.-J. S., Higgs S., Horne K.M., Vanlandingham D.L. Flavivirus-Mosquito Interactions. *Viruses*. 2014. V. 6. P. 4703–4730. doi: [10.3390/v6114703](https://doi.org/10.3390/v6114703)
14. Duffy S., Shackelton L.A., Holmes E.C. Rates of evolutionary change in viruses: patterns and determinants. *Nature Reviews Genetics*. 2008. V. 9. P. 267–276.
15. *GenBank*. URL: <https://www.ncbi.nlm.nih.gov/genbank/> (дата обращения 12.06.2017).
16. *KEGG: Kyoto Encyclopedia of Genes and Genomes*. URL: <http://www.kegg.jp> (дата обращения: 10.05.2017).
17. Lorenz I.C., Allison S.L., Heinz F.X., Helenius A. Folding and Dimerization of Tick-Borne Encephalitis Virus Envelope Proteins prM and E in the Endoplasmic Reticulum. *J. of Virology*. 2002. V. 76. No. 11. P. 5480–5491.
18. Chambers T.J., Hahn C.S., Galler R., Rice C.M. Flavivirus genome organization, expression and replication. *Ann. Rev. Microbiol.* 1990. V. 44. P. 649–688.
19. Dunster L.M., Gibson C.A., Stephenson J.R., Minor P.D., Barrett A.D. Attenuation of virulence of flaviviruses following passage on HeLa cells. *J. Gen. Virol.* 1990. V. 71.

- P. 601–607.
20. McMinn P.S., Marshall I.D., Dalgarno L. Neurovirulence and neuroinvasiveness of Murrey Valley encephalitis virus mutants selected by passage in a monkey kidney cell line. *J. Gen. Virol.* 1995. V. 76. P. 865–872.
 21. Combe M, Sanjuán R. Variation in RNA virus mutation rates across host cells. *PLoS Pathog.* 2014. V. 10. P. e1003855. doi: [10.1371/journal.ppat.1003855](https://doi.org/10.1371/journal.ppat.1003855)

Рукопись поступила в редакцию 08.09.2017.

Дата опубликования 12.10.2017.