

УДК 616.36.-002-036.22-084

*В. В. ДАВЫДОВ, С. В. ЖАВОРОНОК*

## ГЕНЕТИЧЕСКИЕ МАРКЕРЫ АДАПТАЦИИ ВИРУСА ГЕПАТИТА Е К ВНУТРИКЛЕТОЧНОЙ СРЕДЕ ХОЗЯИНА

*Учреждение образования «Белорусский государственный медицинский университет»,  
Минск, Республика Беларусь*

Проведены биоинформатические исследования 309 последовательностей генома вируса гепатита Е (ВГЕ) 1-го и 3-го генотипов, выделенных из организма человека и животных. Определены предпочтения в использовании синонимичных кодонов ВГЕ, по сравнению с широким кругом хозяев. Рассчитаны индексы CAI, RCDI и SiD. Установлены различия в адаптированности ВГЕ к хозяину. Наиболее высокую степень адаптации последовательности ВГЕ, имеют к использованию кодонов в клетках приматов, грызунов, кроликов и птиц. Использование синонимичных кодонов в клетках одомашненных свиней и верблюдов имеет более высокую степень сходства с ВГЕ, чем экспрессионные системы дикого кабана, благородного оленя и лося. Последовательности ВГЕ-3, характеризуются более высокой адаптированностью к экспрессии в клетках разных хозяев, чем последовательности ВГЕ-1, которые характеризуются равно высокой степенью адаптированности к экспрессии генов как в клетках человека, так и в клетках разных видов животных, что является результатом эволюции этого генотипа вируса, в результате которой произошло сужение круга его потенциальных хозяев.

*Ключевые слова:* биоинформатический анализ, вирус гепатита Е, использование синонимичных кодонов, адаптация вируса к организму хозяина.

**Введение.** Аминокислоты являются мономерами белковых молекул. В соответствии с центральной догмой молекулярной биологии передача генетической информации в живых системах происходит от ее носителя, которым является ДНК или РНК, к полипептиду в соответствии с генетическим кодом, который обладает свойством триплетности – одну аминокислоту кодирует три нуклеотида (кодон). В стандартном генетическом коде набор из 61 кодона кодирует 20 аминокислот. Для кодирования одной аминокислоты в теоретической модели все синонимичные кодоны должны использоваться равновероятно, однако это не так. Смещение или предвзятость использования кодонов (от англ. «Codon Usage Bias, CUB») – это явление, при котором в геноме организма или в различных генах одного организма определенные кодоны более предпочтительно используются в кодирующих областях, вместо их синонимичных альтернатив. Определенные кодоны могут эффективнее использоваться рибосомами, это может быть связано с количеством и аффинностью транспортных РНК (тРНК) к определенным кодонам. Установлена связь между уровнем экспрессии генов тРНК и использованием определенных кодонов в наборе высоко экспрессирующихся генов, специфичных для определенных тканей [13], что свидетельствует о тканеспецифичности CUB у многоклеточных организмов. Определенные кодоны могут улучшать точность трансляции и снижать вероятность ошибок, тогда как использование других кодонов, наоборот, может ее ускорять. Существует предположение, что использование редких кодонов замедляет скорость трансляции, чтобы обеспечить оптимальные уровни экспрессии белка и его правильную укладку [10]. Мутации и генетический дрейф являются основными факторами, определяющими конфигурацию паттернов использования кодонов в геномах различных вирусов. Исследование ограничений в использовании кодонов дает информацию о молекулярной эволюции генов вирусов, регуляции их экспрессии, а также полезно для биотехнологии. Экспрессия белков вируса в гетерологичных хозяевах может сопровождаться значительными трудностями. Их гены могут содержать кодоны, которые редко используются у предполагаемого хозяина, трансгены могут происходить от организмов, использующих неканонический код, или содержать регуляторные элементы, ограничивающие экспрессию, в кодирующей последовательности [4]. Таким образом изучение различий CUB геномов вирусов и их хозяев открывает возможности для разработки дизайна рекомбинантных вакцин, для которых необходима активная экспрессия вирусных белков для создания эффективных иммунных ответов во внутриклеточной среде хозяина.

Цель исследования – изучение генетических маркеров ВГЕ, свидетельствующих о степени

адаптации генома вируса к экспрессии его генов в организме разных хозяев.

**Материалы и методы исследования.** Полные кодирующие последовательности (CDS) открытых рамок считывания (ORC1, ORC2 и ORC3) геномов ВГЕ 1-го (n=61) и 3-го генотипов, выделенных из организма человека (n=148) и животных (n=100) были извлечены из базы данных GenBank в Национальном центре биотехнологической информации (NCBI) (<http://www.ncbi.nlm.nih.gov>). Перед проведением исследований из ORC1 во всех последовательностях была удалена гипервариабельная область. Из ORC2 была удалена область, перекрывающаяся с ORC3. В анализе использовались только такие записи, которые имели одинаковую длину, были точно кратны трем нуклеотидам и имели правильный стартовый и стоп-кодон. Все последовательности были выровнены с помощью программы MEGA 11.0 [9] с использованием алгоритма CLUSTAL W [12]. Всего в этом исследовании было проанализировано 309 последовательностей генома ВГЕ.

Индекс адаптации кодонов (от англ. «Codon adaptation index», CAI) – это показатель, характеризующий степень сходства между использованием синонимичных кодонов исследуемого гена вируса и эталонного набора, представляющего собой таблицу использования кодонов в организме хозяина. Значение CAI находится в диапазоне от 0 до 1. Более высокие значения CAI указывают на более высокую адаптивность вируса к внутриклеточной среде хозяина и более высокий уровень экспрессии его генов. Значения CAI рассчитывали на Сервере CAIcal (<http://genomes.urv.cat/CAIcal/>) [7] с использованием эталонных таблиц использования кодонов для видов, являющихся хозяевами для ВГЕ.

Для определения возможной скорости трансляции вирусного гена в клетке хозяина определяли индекс относительной деоптимизации кодонов (от англ. «Relative codon deoptimization index», RCDI) [6]. Значения RCDI рассчитывали на сервере RCDI/eRCDI (<http://genomes.urv.cat/CAIcal/>) [8] с использованием эталонных таблиц использования кодонов для видов, являющихся хозяевами для ВГЕ. Эталонные таблицы использования кодонов в организме хозяина ВГЕ и гепатоцитах человека загружали из специализированной базы в сети Интернет «CoCoPUTs (Codon Pair and Dinucleotide Usage Data)», доступной по адресу [https://dnahive.fda.gov/dna.cgi?cmd=codon\\_usage&id=537&mode=cocoputs](https://dnahive.fda.gov/dna.cgi?cmd=codon_usage&id=537&mode=cocoputs). Значение RCDI варьируется от 1 (использование кодонов полностью оптимизировано для использования кодонов в эталонном геноме) до более высоких значений при деоптимизации тестируемой последовательности [2].

Индекс подобия (от англ. «Similarity Index», SiD) [14] рассчитывался по следующей формуле:

$$R(A, B) = \frac{\sum_{i=1}^{59} a_i \times b_i}{\sqrt{\sum_{i=1}^{59} a_i^2 \times \sum_{i=1}^{59} b_i^2}}$$

где  $a_i$  — значение RSCU для конкретного кодона для кодирующей последовательности ВГЕ,  $b_i$  — значение RSCU для аналогичного кодона хозяина.

В соответствии с данным методом, сходство использования кодонов хозяином и вирусом  $R(A, B)$  определяется как значение косинуса угла, включенного между пространственными векторами  $A$  и  $B$  путем рассмотрения 59 кодонов как 59 различных пространственных векторов.  $D(A, B)$  представляет собой величину силы потенциального влияния общего использования кодонов хозяином на ВГЕ.

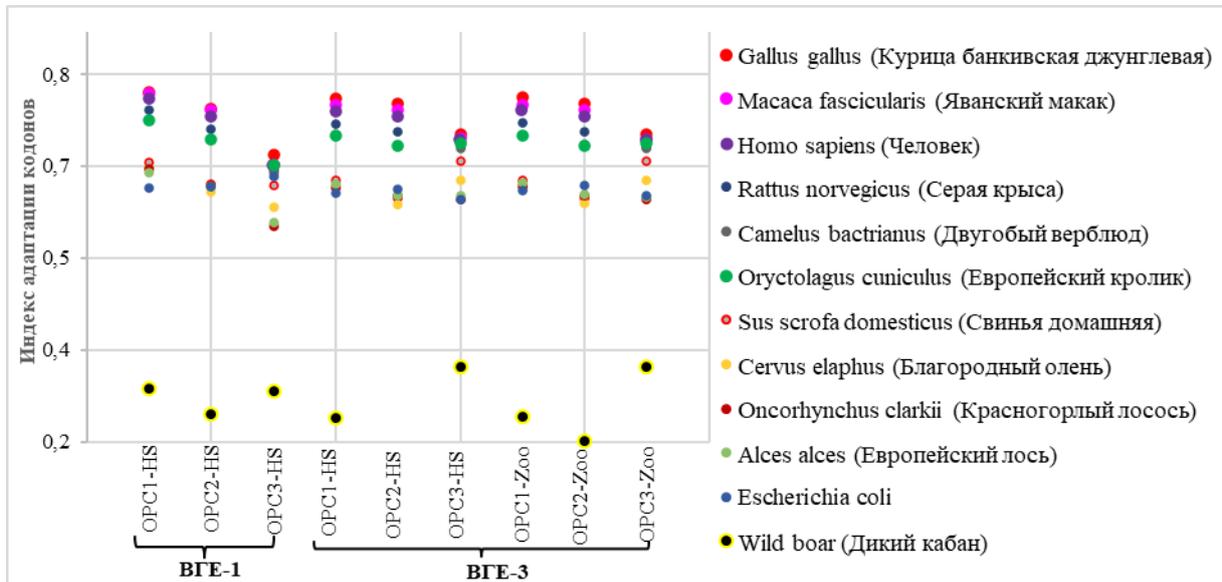
$$D(A, B) = \frac{1 - R(A, B)}{2}$$

Значение SiD колеблется от 0 до 1,0. Чем больше значение индекса, тем больше влияние хозяина на использование кодонов ВГЕ.

**Результаты и их обсуждение.** В результате проведенных исследований установлено, что наиболее высокие значения CAI были характерны для эталонных наборов использования кодонов в клетках Курицы банкивской джунглевой, который по результатам эпидемиологических исследований на является хозяином ВГЕ (рисунок 1). Кроме этого вида также высокие значения CAI были характерны для Приматов, в том числе и Человека, а также для Кролика и Грызунов. Другим неожиданным результатом были очень низкие значения индекса для клеток дикого кабана, которые оказались даже ниже, чем в клетках кишечной палочки, которая не рассматривалась как потенциальный хозяин и была включена в исследование только в качестве группы сравнения.

Значение индекса CAI существенно зависело от исследуемой ОРС и генотипа вируса.

Наибольшие значения показателя были характерны для OPC1. Индекс адаптации последовательностей OPC2 находился на более низком уровне. Значения САИ для OPC3, в большинстве случаев, был еще ниже. Исключение составляли дикие кабаны. Показатель адаптированности последовательностей OPC3 к этим хозяевам был более высоким, чем для последовательностей OPC2.

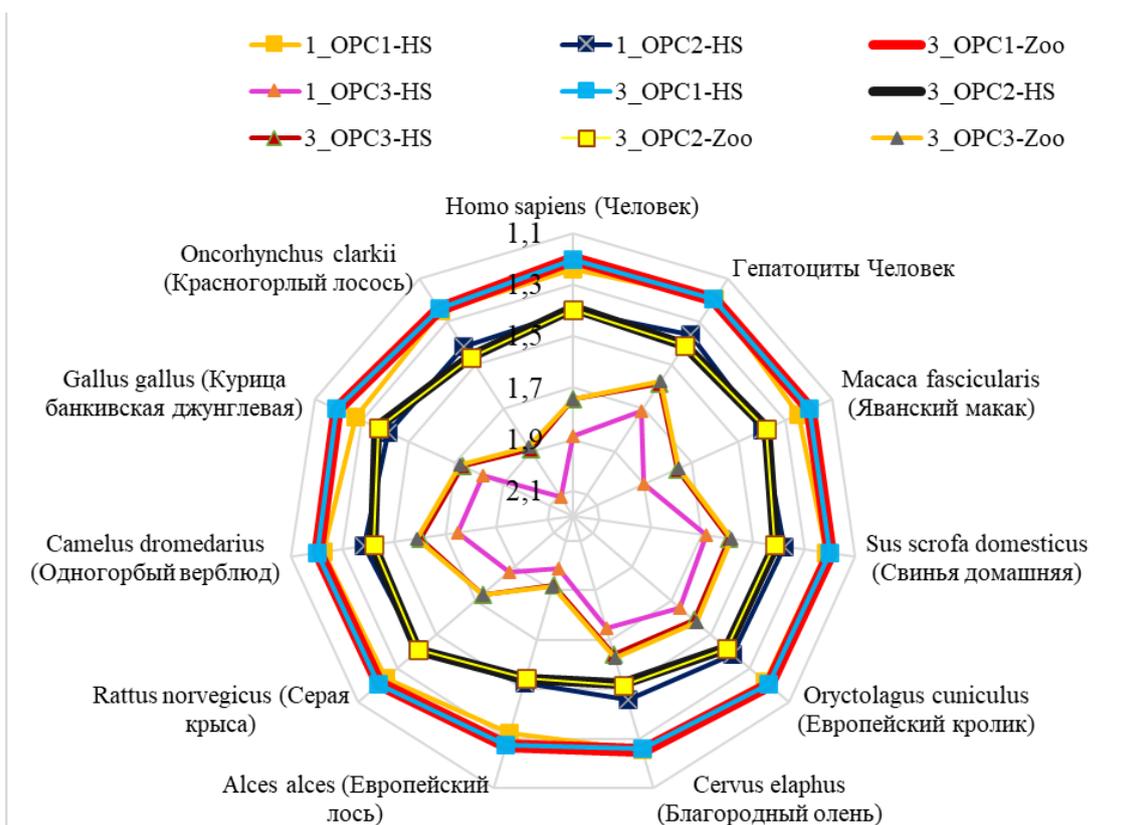


**Рис. 1.** Значения индекса адаптации кодонов в последовательностях OPC1, OPC2 и OPC3 ВГЕ-1 и ВГЕ-3 к использованию кодонов в различных хозяевах ВГЕ

Наиболее высокие значения RCDI для всех OPC всех генотипов (хозяев) были установлены для Дикого кабана. Значения индекса для этого хозяина в большинстве случаев превышали показатель всех других изученных видов, являющихся потенциальными хозяевами ВГЕ, в два раза, что свидетельствует об очень низкой адаптации диких кабанов к обеспечению экспрессии генов ВГЕ в их клетках (рисунок 2). Выявлены незначительные различия показателя RCDI последовательностей ВГЕ-1 и ВГЕ-3, заключающиеся в более низком значении индекса для ВГЕ-3 по отношению к большинству хозяев.

Индекс RCDI имел наименьшее значение для гена, отвечающего за репликацию РНК ВГЕ (OPC1) и не имел существенных отличий у разных хозяев вируса, что свидетельствует о высоком и потенциально одинаковом уровне экспрессии этого гена в клетках их организма.

Индекс RCDI имел наименьшее значение для гена, отвечающего за репликацию РНК ВГЕ (OPC1) и не имел существенных отличий у разных хозяев вируса, что свидетельствует о высоком и потенциально одинаковом уровне экспрессии этого гена в клетках их организма.

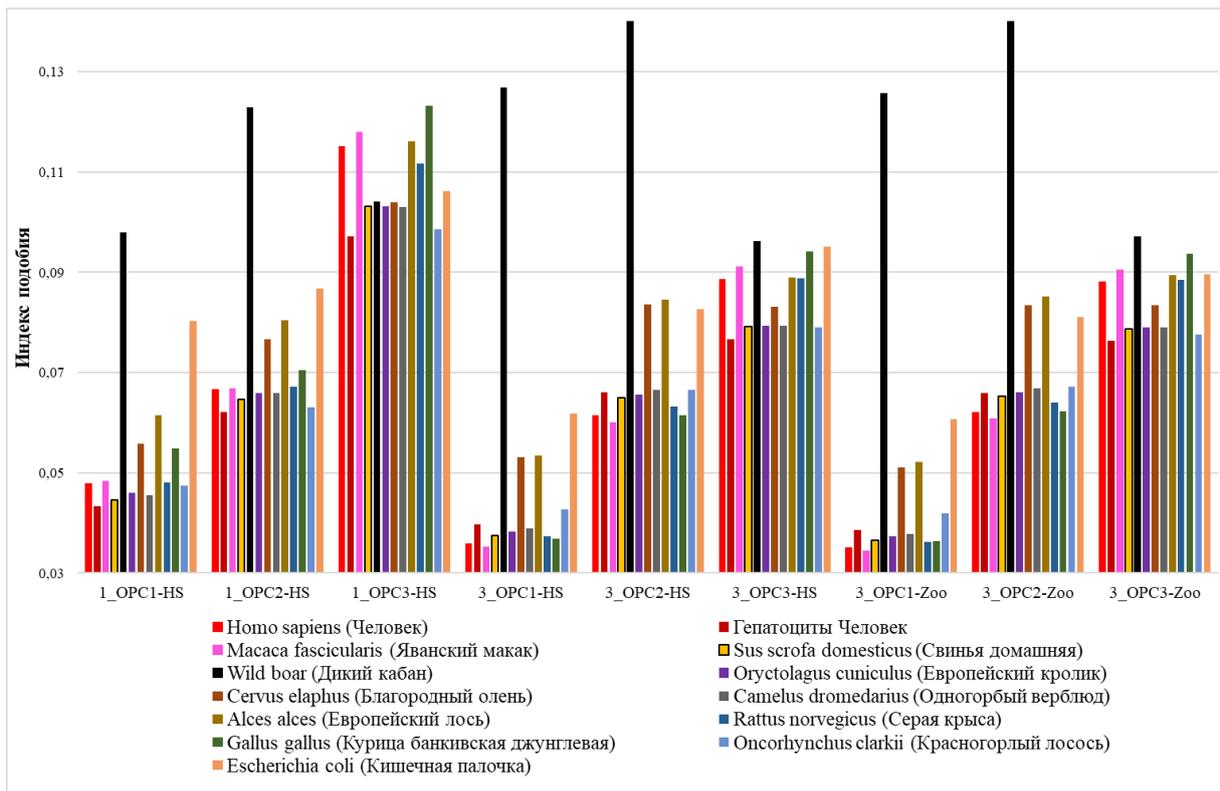


**Рис. 2.** Значения индекса относительной деоптимизации кодонов в последовательностях ОРС1, ОРС2 и ОРС3 ВГЕ-1 и ВГЕ-3 к использованию кодонов в различных хозяевах ВГЕ

Значение RCDI гена, кодирующего белок капсида вируса (ОРС2), находилось на более высоком уровне, чем ОРС1, что свидетельствует о менее выраженной адаптации к экспрессии этого гена в организме изученных хозяев. Для всех изученных хозяев (за исключением дикого кабана) индекс деоптимизации кодонов, как для последовательностей ОРС1, так и ОРС2, приобретал близкие значения, что свидетельствует о сходных уровнях возможной экспрессии этих генов в клетках организм изученного круга хозяев. Значение RCDI для ОРС3 было наибольшим для всех изученных хозяев, особенно для последовательности ВГЕ-1, что является свидетельством меньшей адаптированности гена вируса, отвечающего за выход вируса из клетки, по сравнению с геном, отвечающим за синтез белка его капсида. Значения индекса последовательностей ОРС3 для разных хозяев не были близкими по значению, как в случае ОРС1 и ОРС2. Наибольшее значение этого показателя (кроме дикого кабана) было установлено для Красногорлого лосося, Европейского лося и, что было неожиданным, для Яванского макака. Внутриклеточная среда гепатоцитов человека оказалась наиболее благоприятной для синтеза белка ОРС3 ВГЕ, по сравнению с клетками человека в целом.

Изменения индекса подобия (SiD) имели выраженную тенденцию роста от ОРС1 к ОРС2 и ОРС3 (рисунок 3). Данная особенность была характерна для большинства изученных видов-хозяев ВГЕ, в том числе и для кишечной палочки, которая была включена в исследование в качестве внешней группы. Эта особенность не была характерна для диких кабанов, для которых показатели SiD имели иную картину. Значения индекса для этого хозяина, рассчитанные для ОРС2 имели более высокие значения, чем для ОРС3.

Не смотря на сходный паттерн значений индекса подобия для разных ОРС, выявлены отличия в адаптированности ВГЕ к использованию синонимичных кодонов его разными хозяевами. Наибольшим сходством в использовании кодонов по сравнению с ВГЕ характеризовались такие виды, как Человек, Японский макак, одомашненные животные (свиньи, кролики, куры и верблюды), а также синантропные серые крысы. SiD красногорлого лосося также находился на низком уровне. Дикие кабаны, благородный олень и лось имели высокие значения индекса во всех генотипах и рамках считывания.



**Рис. 3.** Значения индекса подобия использования синонимичных кодонов в последовательностях ОРС1, ОРС2 и ОРС3 ВГЕ-1 и ВГЕ-3 к использованию кодонов в различных хозяевах ВГЕ

Значение индекса подобия ВГЕ-1 для печеночной ткани ожидаемо было самым низким по отношению ко всем изученным хозяевам, в том числе ниже, значения SiD, рассчитанного для организма человека в целом. Однако, для ВГЕ-3 значение SiD для гепатоцитов было более высокими, чем для организма человека в целом и даже выше, чем в организме домашней свиньи. Это может свидетельствовать о более широких адаптивных возможностях ВГЕ-3 к репликации в различных тканях хозяина, что может объяснить существование внепеченочных проявлений ВГЕ, сопровождающихся паразитированием вируса в клетках нервной ткани и клетках матки.

Более высокую адаптированность к репликации в клетках большинства хозяев продемонстрировали ВГЕ 3-го генотипа. По сравнению с ВГЕ-1 во всех рамках считывания они характеризовались более низкими значениями SiD для всех изученных хозяев (не только для печеночной ткани), кроме диких кабанов, для которых показатели индекса для ВГЕ-3 были более высокими, чем для ВГЕ-1. Самой высокой адаптированностью к хозяину характеризовались последовательности ВГЕ-3 ОРС1. Различий в степени адаптированности последовательностей ВГЕ-3, выделенных из разных хозяев (человек или животные) выявлено не было.

Зависимость вирусов от ключевых внутриклеточных процессов хозяина, таких как репликация и синтез белка, обуславливают прямую зависимость приспособленности, выживаемости и эволюции вируса от использования кодонов вирусом и его хозяином. Информация о том, как вирус использует кодоны может дать представление о факторах, обуславливающих предвзятость использования кодонов, и регуляции экспрессии генов. В этом исследовании мы использовали широкий спектр методов для изучения ключевых факторов, ответственных за предвзятость использования кодонов ВГЕ, вклад использования синонимичных кодонов в эволюционные процессы ВГЕ и приспособленность ВГЕ к различным хозяевам.

В целях определения возможного уровня экспрессии белков вируса у хозяина и адаптации вируса к хозяину был использован индекс адаптации кодонов. Значения индекса CAI в ОРС разных генотипов (хозяев) ВГЕ формировали характерный паттерн. Значения CAI ОРС1 были более высокими, по сравнению с ОРС2, которые в свою очередь были выше показателя ОРС3. Данная картина была характерна для всех изученных последовательностей ВГЕ. Результаты исследований свидетельствуют о широких адаптивных возможностях ВГЕ, особенно выраженных в последовательностях генома вируса, отвечающих за его репликацию и образование капсида (ОРС1 и ОРС2). ОРС3 контролирует выход вируса из клетки организма хозяина и в определенной степени

ограничивает спектр его возможных хозяев. В нашем исследовании мы не обнаружили существенно более высокой адаптированности последовательностей ОРС1 ВГЕ-1 к человеку по сравнению с другими хозяевами, о которой сообщалось S. Vaha и др. Мы не можем подтвердить вывод этих исследователей о том, что установленная ими «адаптация ОРС1 ВГЕ-1 к людям может объяснить эпидемический характер инфекций ВГЕ генотипа 1 (в переводе с англ.)» [1]. Полученный в наших исследованиях результат о высокой адаптированности последовательностей ВГЕ к экспрессии его генов в клетках курицы требует дальнейшего изучения.

Индекс RCDI является интегрирующим показателем, позволяющим оценить кумулятивные эффекты смещения использования синонимичных кодонов в геноме вируса, путем сравнения с аналогичным показателем у того или иного хозяина. Полученные нами результаты изучения RCDI в ОРС ВГЕ, в целом, дополняют данные об индексе адаптации использования кодонов (CAI) в последовательностях ВГЕ и его хозяев. Подтвержден очень низкий уровень адаптированности всех последовательностей ВГЕ не зависимо от ОРС и генотипа к синтезу белка в организме диких кабанов. Также полученные данные о значениях RCDI для различных ОРС ВГЕ показывают более высокий уровень адаптированности последовательностей ОРС1 и ОРС2 и низкий – для ОРС3. Выявленные генотип специфические особенности индекса деоптимизации, заключавшиеся в более низком значении RCDI для последовательностей ВГЕ-3 по сравнению с ВГЕ-1, которые противоречат результатам изучения индекса адаптации кодонов CAI, обладали статистической достоверностью только для ОРС3. Не смотря на большое количество факторов, влияющих на значение индекса, величины RCDI для ВГЕ-3, выделенных из разных хозяев, рассчитанные для ОРС1 (ОРС2) имели отличия только в тысячных долях его значения, причем для всех изученных хозяев. Этот невероятный факт, вероятно, свидетельствует об одинаковой степени адаптации ВГЕ-3, выделенных из разных хозяев, и нуждается в подтверждении. Похожие исследования были проведены группой ученых из Китая. В своей работе В. Li и др. [5] рассчитывали RCDI только для тех генотипов и хозяев, которые соответствовали результатам эпидемиологических исследований. Так для ВГЕ-1 показатель был рассчитан только для человека, а для ВГЕ-3 для человека и свиньи и т.д. Результаты исследований, охарактеризованные в нашем исследовании, имеют более обширный характер. Используя более широкий круг хозяев для расчета RCDI, мы демонстрируем практически одинаковую степень адаптированности последовательностей ВГЕ-1 и особенно ВГЕ-3, выделенных из разных хозяев, к трансляции в их клетках, которые ранее не относились к хозяевам ВГЕ.

Использование метода расчета индекса подобия (SiD) позволяет провести всестороннюю оценку степени сходства общего шаблона использования кодонов между вирусом и хозяином, включая в оценку все 59 синонимичных кодонов одновременно как разные 59 пространственные вектора. Преимущество этого подхода заключается в том, что вместо прямой оценки использования каждого синонимичного кодона используется сравнительное общее использование всех кодонов одновременно. Таким образом, этот метод позволяет избежать ситуации, когда вариации значений использования 59 синонимичных кодонов искажают целостную оценку влияния хозяина на вирус.

На основании сходства показателя SiD к хозяевам антропонозного ВГЕ-1 могут быть отнесены и все животные, в которых могут паразитировать ВГЕ-3. Это домашние свиньи, кролики, верблюды, являющиеся хозяином для ВГЕ-7 и даже серые крысы, которые являются хозяином для ВГЕ-С1. С другой стороны, такой хозяин ВГЕ-3, как дикий кабан, продемонстрировал очень высокие уровни SiD, что свидетельствует о низкой адаптации вируса к паразитированию в его клетках. Однако данные, полученные в нашей работе, о более низких значениях индекса для ОРС3 по отношению к ОРС2 для диких кабанов, являющиеся принципиально иным паттерном, по сравнению с остальными изученными хозяевами, свидетельствует о том, что гены ВГЕ, отвечающие за разные стадии жизненного цикла вируса, в организме этого хозяина имеют другое направление адаптации. Недостаток адаптированности генов ВГЕ, отвечающих за синтез белков капсида (ОРС2), может быть в данном случае «компенсирован» более высокой адаптацией последовательности ОРС3, обладающей широким спектром функций. Как известно, ОРС3 может участвовать в инкапсидации геномной РНК вируса и сборке его капсида, выходе вирионов из клетки хозяина и, возможно, в проникновении вируса в новые клетки, что в данном случае могло обеспечить расширение круга хозяев ВГЕ и включение в него диких кабанов. Вывод, сделанный в работе В. Li и др., о более высокой адаптированности ОРС1 по сравнению с ОРС-2 полностью соответствует нашим результатам.

Таким образом, на основании анализа индекса деоптимизации кодонов и индекса подобия наше исследование показало, что предвзятость использования кодонов у многих видов животных была практически идентична таковой у зоонозных и антропонозных генотипов ВГЕ, что указывает на

возможность того, что эти животные являются хозяевами для ВГЕ. Предвзятость использования кодонов может частично объяснить предпочтения хозяина зоонозного ВГЕ и дать ориентиры для поиска новых животных-хозяев ВГЕ, но для этого требуются дополнительные биоинформатические, экспериментальные и эпидемиологические данные.

Использование синонимичных кодонов для построения ОРС не является случайным, существуют механистические или эволюционные ограничения, которые определяют степень свободы для построения кодирующей последовательности [11]. Другими словами, каждый организм, в том числе и вирус, использует определенный набор правил для построения последовательностей своих ОРС, которые ограничивают общее количество вариантов, предоставляемых вырожденностью генетического кода. Использование кодонов и предвзятость контекста пар нуклеотидов отражают действие двух основных эволюционных сил: отбора по эффективности трансляции мРНК и мутационного дрейфа, действующего без разбора на кодирующую и некодирующую последовательность [3].

**Заключение.** Биоинформатический анализ, основанный на установлении сходства в использовании синонимичных кодонов и определении уровня возможной белок синтетической активности генов ВГЕ в разных системах экспрессии (индексы CAI, RCDI и SiD), позволил установить различия в адаптированности ВГЕ к широкому кругу хозяев. Наиболее высокую степень адаптации последовательности ВГЕ, имеют к использованию кодонов в клетках приматов, грызунов, кроликов и птиц. Возможная экспрессия генов ВГЕ в клетках парнокопытных не имеет однозначной оценки. Использование синонимичных кодонов в клетках одомашненных свиней и верблюдов характеризуются высокой степенью сходства с ВГЕ, в то время как адаптация ВГЕ к экспрессионным системам дикого кабана, благородного оленя и лося находятся на низком уровне, что свидетельствует об относительно недавней истории включения в круг хозяев ВГЕ этих видов, вызванного распространением вируса на новые территории и недостаточной сформированности взаимоотношений в этих системах «паразит-хозяин». Последовательности ВГЕ-3, являющегося полигостальным патогеном, характеризуются более высокой адаптированностью к экспрессии в клетках разных хозяев, чем последовательности ВГЕ-1. Последовательности ВГЕ-1, которые по данным эпидемиологических исследований поражают только человека, характеризуются равно высокой степенью адаптированности к экспрессии генов как в клетках человека, так и в клетках разных видов животных, что является результатом эволюции этого генотипа вируса, в результате которой произошло сужение круга его потенциальных хозяев.

#### Литература:

- [1] *Baha S.* Comprehensive analysis of genetic and evolutionary features of the hepatitis E virus // *BMC Genomics*. 2019. Vol. 20. P. 790–812.
- [2] *Butt A.* Evolution of codon usage in Zika virus genomes is host and vector specific // *Emerg. Microbes. Infect.* 2016. Vol. 5, P. e107.
- [3] *Fedorov A.* Regularities of context-dependent codon bias in eukaryotic genes // *Nucleic Acids Research*. 2002. Vol. 30. P. 1192–1197.
- [4] *Gustafsson C.* Codon bias and heterologous protein expression // *Trends in Biotechnology*. 2004. Vol. 22. P. 346–353.
- [5] *Li B, Wu H.* Codon Usage of Hepatitis E Viruses: A Comprehensive Analysis // *Front. Microbiol.* 2022. Vol. 13.
- [6] *Mueller S.* Reduction of the Rate of Poliovirus Protein Synthesis through Large-Scale Codon Deoptimization Causes Attenuation of Viral Virulence by Lowering Specific Infectivity // *Journal of Virology*. 2006. Vol. 80. P. 9687–9696.
- [7] *Puigbò P.* CAIcal: A combined set of tools to assess codon usage adaptation // *Biology Direct*. 2008. Vol. 3, № 1. P. 38.
- [8] *Puigbò P.* RCDI/eRCDI: a web-server to estimate codon usage deoptimization // *BMC Res Notes*. 2010. Vol. 3, № 1. P. 87.
- [9] *Tamura K.* MEGA11: Molecular Evolutionary Genetics Analysis Version 11 // *Molecular Biology and Evolution*. 2021. Vol. 38. P. 3022.
- [10] *Thanaraj T. A.* Ribosome-mediated translational pause and protein domain organization // *Protein Science*. 1996. Vol. 5. P. 1594–1612.
- [11] *Thomas L.* Codon choice and gene expression: synonymous codons differ in their ability to direct aminoacylated-transfer RNA binding to ribosomes in vitro // *Proceedings of the National Academy of Sciences*. 1988. Vol. 85. P. 4242–4246.
- [12] *Thompson J.* CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice // *Nucleic Acids Res.* 1994. Vol. 22. P. 4673–4680.
- [13] *Wright F.* The ‘effective number of codons’ used in a gene // *Gene*. 1990. Vol. 87, № 1. P. 23–29.

[14]Zhou J. The Distribution of Synonymous Codon Choice in the Translation Initiation Region of Dengue Virus // PLoS ONE. 2013. Vol. 8. P. e77239.

*V. V. DAVYDOV, S. V. ZHAVORONOK*

**GENETIC MARKERS OF HEPATITIS E VIRUS ADAPTATION TO THE HOST'S INTRACELLULAR ENVIRONMENT**

*Educational institution "Belarusian State Medical University", Minsk, Republic of Belarus*

**Summary**

Bioinformatic studies of 309 genome sequences of hepatitis E virus (HEV) genotypes 1 and 3 isolated from humans and animals were conducted. Preferences in the use of synonymous HEV codons were determined, compared with a wide range of hosts. CAI, RCDI and SiD indices were calculated. Differences in HEV adaptability to the host were established. The highest degree of adaptation of the HEV sequence to codon use was observed in the cells of primates, rodents, rabbits and birds. Synonymous codon usage in domesticated pig and camel cells has a higher degree of similarity to HEV than the expression systems of wild boar, red deer and elk. HEV-3 sequences are characterized by a higher adaptability to expression in cells of different hosts than HEV-1 sequences, which are characterized by an equally high degree of adaptability to gene expression in both human cells and cells of different animal species, which is a result of the evolution of this genotype of the virus, as a result of which the range of its potential hosts has narrowed.

*Keywords:* bioinformatics analysis, hepatitis E virus, synonymous codon usage bias, virus adaptation to the host organism.