

ХАРАКТЕРИЗАЦІЯ ВІДКРИТИХ РАМОК ЗЧИТУВАННЯ ВІРУСУ ГЕПАТИТУ Е СВИНІ ГЕНОТИПУ 3

Лиманська О. Ю.

Національний науковий центр «Інститут експериментальної і клінічної ветеринарної медицини», Харків, Україна, e-mail: olgaliman@ukr.net

Метою роботи було визначення особливостей структурної організації трьох відкритих рамок зчитування ORF1–ORF3 вірусу гепатиту Е свині генотипу 3. Тиск відбору на гени оцінено через проведення тесту Таджими на нейтральність, який вбудовано в програму MEGA 6. Ентропію Шеннона використано для визначення різноманіття в кожній позиції відкритих рамок зчитування. Графіки ентропії для ORF1–ORF3, які представляють кількісну варіабельність для кожної амінокислотної (нуклеотидної) позиції вирівняних послідовностей 11 ізолятів ВГЕ-3 свині, розраховано за допомогою програми BioEdit. Для філогенетичного аналізу нуклеотидні відстані проаналізовано з використанням методу приєднання сусідів. Аналіз нуклеотидних, несинонімічних та синонімічних дистанцій виконано за допомогою MEGA 6. За допомогою тесту Таджими на нейтральність визначено параметри мутаційного аналізу відкритих рамок зчитування ORF1–ORF3 ВГЕ генотипу 3 свині. Аналіз ентропії Шеннона показав, що ORF3 характеризується найбільшим числом варіабельних амінокислотних залишків (21,7 %) порівняно з кодуєчими ділянками ORF1, ORF2. Аналіз несинонімічних дистанцій в ORF3 дозволив виявити порядок еволюційних подій в історії ізолятів ВГЕ-3 свині. В якості контролю використано штами астровірусів птахів, філогенетичні відношення між якими відповідають порядку еволюції їхніх хазяїв. Топологія дерева для ORF3 ВГЕ-3 свині на підставі несинонімічних дистанцій значно відрізняється від структури дерев, побудованих на основі нуклеотидних та синонімічних дистанцій. Позитивне значення коефіцієнта D тесту Таджими для ORF1, ORF2 ВГЕ-3 свині вказує на дію позитивного відбору на ці відкриті рамки зчитування, а негативне значення для ORF3 — на очищаючий відбір

Ключові слова: позитивний відбір, тест Таджими, негативний відбір, ентропія Шеннона

Вірус гепатиту Е (ВГЕ, HEV, hepatitis E virus) є причиною небезпечного захворювання печінки — гострого та хронічного гепатиту Е, а також асоційований з низкою позапечінкових розладів: неврологічними синдромами, панкреатитом, гематологічними проблемами, нирковими травмами [1]. Геномна одноланцюгова молекула РНК ВГЕ містить три відкриті рамки зчитування (open reading frame, ORF), які кодують неструктурний поліпротеїн (ORF1); основний протеїн капсиду вірусу, асоційований зі складанням віріона, взаємодією з організмом хазяїна та імуногенністю (ORF2); фосфопропротеїн, залучений до процесу виходу вірусу з інфікованої клітини (ORF3) [2, 3]. Серед відомих гепатотропних вірусів тільки генотипи 3 та 4 ВГЕ, для яких основним природним резервуаром є свині та дикі кабани, характеризуються зоонозним потенціалом [4–6].

Основою молекулярної еволюції вірусів, як і інших організмів, є зміни їх геномів та макромолекул віріонів. Мутаційні зміни в молекулах РНК та ДНК (нуклеотидні заміни, інсерції, делеції, рекомбінації) постійно відбуваються в геномі та можуть бути причиною зміни рівня резистентності до лікарських препаратів, патогенності тощо [7, 8]. На рівні нуклеотидних послідовностей основними факторами еволюції є природний відбір, направлений мутаційний тиск та дрейф (вплив на генетичний склад популяції випадкового процесу вимирання та розмноження ліній).

Природний відбір спрямований на елімінацію шкідливих (несприятливих) мутацій, тобто проявляється у вигляді негативного (або очищаючого) відбору, та накопичення і фіксацію корисних (сприятливих) мутацій поряд зі збереженням функціональної активності відповідного білка, а також на підтримку поліморфізму (позитивний відбір). Згідно з нейтральною теорією еволюції, генетичні мутації, які швидко накопичуються, є селективно нейтральними, не

впливають на функціонування білків та, отже, не підпадають під дію природного відбору і, таким чином, їхня фіксація в популяції відбувається шляхом випадкових процесів. При цьому теорія нейтральності не заперечує, але мінімізує роль природного відбору як джерела генетичної варіабельності [9]. Мутаційний тиск, механізм якого вивчений для низки вірусів та бактерій, обумовлений підвищеною частотою та наступною фіксацією заміни аденіну та тиміну на гуанін та цитозин (ГЦ-тиск) або навпаки (АТ-тиск), що призводить до зміни рівня ГЦ-вмісту в ряду поколінь [10–12].

Інсерції та делеції виникають у більшості випадків у некодуючому протеїн фрагменті гена. У білок-кодуючих послідовностях заміни нуклеотидів у третій позиції кодону є, як правило, синонімічними та не пов'язані з заміною амінокислоти на відміну від нуклеотидних заміни у першій та другій позиціях кодону, які є несинонімічними та асоційовані з заміщенням відповідної амінокислоти. Для визначення напрямку та сили відбору з застосуванням різних методів розраховують відношення числа несинонімічних заміни на кількість несинонімічних сайтів (dN) до числа синонімічних заміни на кількість синонімічних сайтів (dS) у білок-кодуючій послідовності. Про консервативність амінокислотної послідовності та, отже, дію очищаючого відбору свідчить значення $dN/dS < 1$. Навпаки, значення $dN/dS > 1$ означає, що відповідний протеїн знаходиться під дією позитивного відбору [13, 14]. Для ідентифікації схильних до мутацій позицій у послідовності протеїну розраховують ентропію Шеннона. Високий рівень ентропії Шеннона у певній позиції є індикатором підвищеної хаотичності та схильності до мутацій, у той час як низький рівень ентропії Шеннона означає підвищену стабільність та низький рівень безладності [15].

В даній роботі на основі філогенетичного аналізу визначено параметри трьох відкритих рамок читування ORF1–ORF3 в геномі ВГЕ-3 свині на макромолекулярному рівні та встановлено внутрішньовидові еволюційні відносини для 11 штамів ВГЕ-3.

Матеріали та методи. Ізоляти ВГЕ свині 3-го генотипу з повним геномом (таксономічний ідентифікатор (txid) 1678143), а також астровірусів курей (txid 336960, 336961) та індиків (txid 246343, 364370) 1-го і 2-го типів отримано через пошук в базі даних нуклеотидів Національного центру біотехнологічної інформації (США) NCBI. Параметри 11 штамів ВГЕ-3 та відомості щодо країн їхнього походження, дати отримання штамів наведено в табл. 1.

Таблиця 1 — Перелік штамів HEV-3, ізольованих з біологічного матеріалу свині, що проаналізовано в роботі

Штам	Номер у GenBank	Країна	Рік
swMN06-A1288	AB290312	Монголія	2006
JIO-swJ19-1	AB443623	Японія	2002
SWHEV75BO2012	KY766999	Італія	2012
CH_VW117	MG573193	Швейцарія	2017
HEV/13RS985-5	MG674164	Італія	2013
MXCDg3_B1_2016	MG833836	Мексика	2016
HEV_006_1	MN184579	Велика Британія	2013
HEVF46	MN450023	Таїланд	2015
HEVGB166	MN450031	Таїланд	2015
swX07-E1	EU360977	Швеція	2008
SAAS-JDY5	FJ527832	Китай	2008

Вирівнювання послідовностей ORF1–ORF3 проведено за допомогою програми Molecular Evolutionary Genetics Analysis (MEGA) (версія 6.06) [16].

Програму BioEdit (версія 7.2.5) [17] використано для маніпулювання з нуклеотидними та амінокислотними послідовностями, зокрема, для мутаційного аналізу протеїнів, що трансклюються з ORF1–ORF3, для визначення амінокислотних заміни ВГЕ свині. Тиск відбору на гени оцінено через проведення тесту Таджими на нейтральність, який вбудовано в програму MEGA 6.

Трансляцію нуклеотидних послідовностей в амінокислотні послідовності з наступним аналізом кодонів проведено за допомогою програми BioEdit.

Ентропію Шеннона використано для визначення різноманіття в кожній позиції відкритих рамок зчитування. Графіки ентропії для ORF1–ORF3, які представляють кількісну варіабельність для кожної амінокислотної (нуклеотидної) позиції вирівняних послідовностей 11 штамів ВГЕ-3, побудовано за допомогою програми BioEdit.

Філогенетичний аналіз виконано декількома методами. Для усіх методів позиції, що містили пробіли, виключено з попарного порівняння послідовностей. Бутстреп-аналіз виконано для 100 реплікацій. Нуклеотидні відстані проаналізовано з використанням методу приєднання сусідів (neighbour-joining, NJ), який базується на двопараметровому методі Кімури оцінки відстані (p-дистанції). Для ORF1–ORF3 аналіз нуклеотидних, несинонімічних та синонімічних дистанцій виконано за допомогою MEGA 6. Додатково аналіз несинонімічних та синонімічних замін виконано за допомогою корекції Jukes-Cantor для методу Nei-Gojobori. Для оцінки достовірності дерева бутстреп-аналіз проведено з використанням 100 випадкових вибірок. Кластери зі значенням, що перевищує 90, вважаються достовірними.

Результати досліджень. Мутаційний процес — це такі випадкові зміни послідовності ДНК як заміни, делеції (випадіння), вставки, складні події. Швидкість мутацій визначають як ймовірність зміни на нуклеотид за покоління.

Під відбором розуміють диференційне розмножування генотипів (штамів патогена), а під пристосованістю — очікувану ефективність розмноження особини (штаму) з певним генотипом (геномом патогена).

Аналіз мутацій в відкритих рамках зчитування ORF1–ORF3. Нуклеотидні послідовності ORF1, ORF2, ORF3 отримано з повновимірних геномів 11 штамів ВГЕ-3 свині. Для аналізу еволюційних відношень між ізолятами ВГЕ-3 спочатку нуклеотидні послідовності ORF1–ORF3 було вирівняно. Несинонімічні (табл. 2) та синонімічні заміни ідентифіковано в трьох відкритих рамках зчитування ізолятів ВГЕ-3 свині. Такі зміни ведуть до зміни поверхневих властивостей ORF.

Визначення еволюційних відстаней (p-відстаней) для пар послідовностей відкритих рамок зчитування (число нуклеотидних або амінокислотних замін, що має місце між ними) дозволило виявити співвідношення синонімічних та несинонімічних замін для цих пар. P-відстань визначають як пропорцію (p) нуклеотидних сайтів, в яких послідовності, що порівнюють, відрізняються.

Зокрема, для пари ізолятів AB443623 (найраніше отриманий штам за відомостями з бази даних GenBank) та MG573193 (має найраніше еволюційно відгалужену гілку ORF3, як впливає з філогенетичного аналізу) визначено наступне. ORF3 ізоляту MG573193 характеризується більшим числом як несинонімічних, так і синонімічних замін порівняно з ORF3 ізоляту AB443623. Переважання несинонімічних замін вказує на ненейтральний характер еволюції протеїну. Більше число синонімічних замін є характерним для внутрішніх протеїнів.

Аналіз ентропії Шеннона для протеїнів ORF1–ORF3. Ентропія є мірою варіабельності всередині кожного сайту та має більш високе значення для варіабельних сайтів порівняно з менш варіабельними сайтами.

У випадку, якщо для всіх послідовностей в вирівнюванні в певній позиції існує тільки один нуклеотид, можна зробити висновок, що існує максимум інформації про зазначену позицію та в цьому випадку ентропія дорівнює нулю. І навпаки, якщо для всіх послідовностей в вирівнюванні в певній позиції наявними є чотири нуклеотиди з частотою 0,25, то інформація щодо цієї позиції в гомологічній послідовності зменшується до нуля, а ентропія набуває максимального значення.

За допомогою аналізу ентропії Шеннона 91, 23 та 25 амінокислотних варіабельних сайтів було ідентифіковано для протеїнів, що трансклюються з ORF1, ORF2 та ORF3 ВГЕ-3 відповідно (рис. 1). Варіабельність розраховано як ентропію для позиції кожного амінокислотного залишку. Кількість варіабельних сайтів для відкритих рамок зчитування становила: ORF1 — 9,2 % (157/1710), ORF2 — 8,0 % (53/662) та ORF3 — 21,7 % (25/115). Аналогічні відсоткові дані було раніше отримано автором для відкритих рамок зчитування ORF1–ORF3 ВГЕ-4 кабана та свині.

Аналіз ентропії показав, що серед відкритих рамок зчитування ВГЕ-3 свині ORF3 характеризується найбільшою варіабельністю, як і ORF-3 ВГЕ-4 свині та кабана (як показано автором раніше).

Таблиця 2 — Позиції кодонів з несинонімічними замінами в ORF3 ВГЕ-3 свині (амінокислотну послідовність ORF3 ізоляту АВ443623 використано як референсну)

Позиція кодона	Номер у GenBank	Заміна амінокислоти
5	EU360977, FJ527832, MG573193, MH184579	S (F)
6	MG573193 MG833836, MH450031	C (Y) C (F)
8	AB290312, EU360977, FJ527832 MG573193	S (L) S (V)
16	EU360977	L (I)
36	EU360977	A (V)
41	KY766999, MG833836	A (V)
43	AB290312	A (V)
64	MH184579	S (L)
75	EU360977, KY766999, MG674164, MH450023, MH450031	S (L)
78	EU360977, KY766999, MG674164, MH450023, MH450031	M (T)
79	AB290312	S (L)
80	EU360977, MH184579, MH450023, MH450031	F (Y)
81	AB290312	H (R)
90	MG573193	D (G)
91	AB290312, FJ527832, MG573193, MG833836, MH450023, MH450031	N (S)
92	AB290312	R (H)
93	AB290312	P (H)
94	MH450023	A (V)
96	AB290312, EU360977, KY766999, MG573193, MG674164, MH184579, MH450023, MH450031	L (S)
97	AB290312, EU360977, FJ527832, KY766999, MG573193, MG674164, MG833836, MH184579, MH450023, MH450031	V (A)
98	AB290312, EU360977, FJ527832, KY766999, MG573193, MG674164, MG833836, MH184579, MH450023, MH450031	Q (P)
101	EU360977, FJ527832	V (A)
109	EU360977	P (L)
110	MG573193	P (L)
112	MG573193, MH450031	V (A)

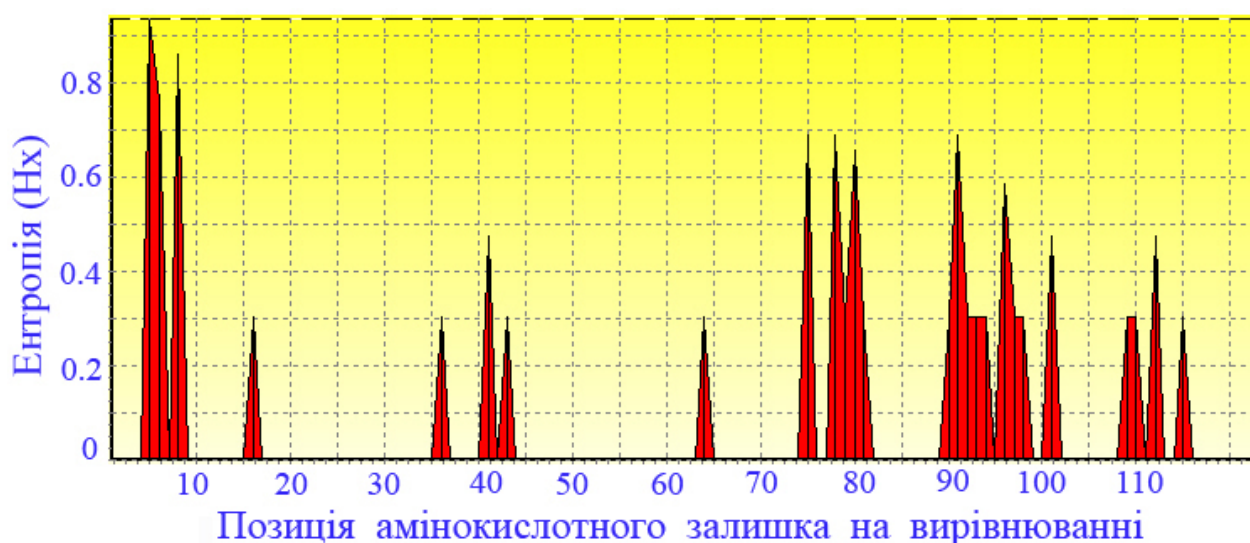


Рис. 1. Графік ентропії Шеннона для протеїну, який транслюється з ORF3 ВГЕ-3 свині, — залежність міри мінливості амінокислотних залишків в заданому положенні від їхньої позиції, — що отримано на підставі вирівнювання та наступної трансляції для 11 ізолятів.

Аналіз позитивного та негативного відбору на ORF1–ORF3. Для ідентифікації фрагментів ORF, які перебувають під дією відбору, проводили тест Таджими на нейтральність, в якому визначають, в тому числі, коефіцієнт D. Для послідовностей, які аналізують, D показує відповідність нуклеотидних заміни гіпотези нейтральності. Тест Таджими дозволяє визначити відхилення нуклеотидної різноманітності, що спостерігають, від моделі нейтральної еволюції.

Фрагменти ORF1 та ORF2 ВГЕ-3 перебувають під тиском позитивного відбору, на що вказує позитивне значення коефіцієнту D (табл. 3). На фрагмент ORF3 діє негативний (очищаючий) відбір, який зменшує варіабельність.

Таблиця 3 — Параметри мутаційного аналізу відкритих рамок зчитування ORF1–ORF3 ВГЕ генотипу 3 свині

	m	S	p_s	Θ	π	D
ORF1	11	1863	0,366372	0,125086	0,171975	1,822164
ORF2	11	583	0,295789	0,100987	0,135427	1,654770
ORF3	11	54	0,158358	0,054066	0,052359	-0,149395

Примітки: m — число послідовностей для аналізу, n — загальне число сайтів, S — число поліморфних сайтів в послідовностях, $p_s = S/n$, $\Theta = p_s/a_1$, π — нуклеотидна різноманітність, D — коефіцієнт теста Таджими (оцінка відповідності характеру нуклеотидних заміни гіпотези нейтральності).

При негативному відборі найприспосованіший генотип (штам патогена) є частим, він підтримує вихідний стан та перешкоджає змінам.

При позитивному відборі найприспосованіший генотип (штам патогена) є рідким. Позитивний відбір сприяє змінам. Після того, як позитивний відбір зробив свою справу і найбільш пристосований генотип (штам патогенна) став частим, відбір стає негативним за незмінності адаптивного ландшафту (який є графіком пристосованості від генотипу (штам патогена). Балансуючий відбір сприяє будь-якому рідкому генотипу, що є можливим тільки за зміни адаптивного ландшафту.

Відповідно до теорії нейтральності ймовірність того, що мутація буде селективно нейтральною (нешкідливою) тим більше, чим меншою мірою вона відбивається на структурі та функції молекули. Інакше кажучи, молекули, які підлягають відносно слабким функціональним обмеженням, еволюціонують швидше молекул, що характеризуються більшими обмеженнями [18].

Внутрішньовидовий філогенетичний аналіз ізолятів ВГЕ-3. Понад 30 років минуло після ізоляції першого штаму ВГЕ свині [19, 20].

Для вивчення еволюційних сил, що діють у внутрішньовидовій популяції ВГЕ-3 свині, окремо проаналізовано нуклеотидні, синонімічні та несинонімічні дистанції для трьох рамок зчитування.

Топологія дерева, отриманого як результат філогенетичного аналізу ORF1 (рис. 2А) та ORF2 (рис. 2Б) на підставі несинонімічних дистанцій, не втрачається, а зберігається та є подібною до структури дерев ORF1–ORF3 ВГЕ-3, отриманих на підставі нуклеотидних, синонімічних дистанцій (результати не наведено).

Топологія дерева для ORF3 на підставі несинонімічних дистанцій (рис. 2В) значно відрізняється від структури зазначених вище дерев.

Аналіз несинонімічних дистанцій в ORF3 дозволив виявити порядок еволюційних подій в історії ізолятів ВГЕ-3 свині. В якості контролю використано астровіруси птахів, філогенетичні відношення між якими відповідають порядку еволюції їхніх хазяїв. Природним хазяїном вірусу нефриту птахів є кури, для яких встановлено, зокрема, філогенетичні відносини з індиками [21].

З рис. 2А та 2Б випливає, що найстарішими ORF3 є відкриті рамки зчитування ізолятів MG573193, FJ527823, AB443623.

Несинонімічні заміни, оскільки вони змінюють амінокислоту, перебувають під впливом сильного позитивного чи негативного тиску. Навпаки, синонімічні заміни, які зберігають амінокислоту, перебувають під слабким тиском відбору або тиск відбору відсутній.

Структура філогенетичного дерева для ORF1 та ORF2 для ізолятів ВГЕ-3, яку отримали при аналізі несинонімічних дистанцій (рис. 2), відрізняється порядком гілок від такої, отриманої при аналізі синонімічних дистанцій (результати не показано).



Рис. 2. Філогенетичні відношення між 11 ізолятами ВГЕ-3 свині на основі несинонімічних дистанцій в ORF1 (А), ORF2 (Б), ORF3 (В). Наведено значення бутстреп-аналізу (100 випадкових вибірок) для оцінки достовірності дерева, а також номери ізолятів в GenBank. Дерева побудовано з використанням р-дистанцій та методу приєднання сусідів. АВ033998 — вірус нефриту птахів типу 1, АВ046864 — вірус нефриту птахів типу 2, АF206663 — астровірус індика типу 1, Y15936 — астровірус індика типу 2.

Висновки. 1. У роботі визначено особливості структурної організації трьох відкритих рамок зчитування ORF1–ORF3 вірусу гепатиту Е свині генотипу 3.

2. Позитивне значення коефіцієнта D тесту Таджими для ORF1, ORF2 ВГЕ-3 вказує на дію позитивного відбору на ці відкриті рамки зчитування, а негативне значення для ORF3 вказує на очищаючий відбір, що зменшує її варіабельність.

3. ORF3 характеризується найбільшим числом варіабельних амінокислотних залишків (21,7 %) порівняно з кодуєчими ділянками ORF1, ORF2.

4. Топологія дерева для ORF3 на підставі несинонімічних дистанцій значно відрізняється від структури дерев, побудованих на основі нуклеотидних та синонімічних дистанцій. За допомогою аналізу несинонімічних дистанцій виявлено порядок еволюційних подій в історії ізолятів ВГЕ-3 свині.

Перспективи використання отриманих результатів. Молекулярне моделювання вторинної та третинної структури протеїнів з урахуванням визначеного порядку еволюційних подій в історії ізолятів ВГЕ-3 свині на основі несинонімічних дистанцій в відкритих рамках зчитування ORF1–ORF3 дозволить дослідити вплив тиску відбору на поверхневі властивості та конформацію протеїнів.

Фінансування. Роботу виконано за фінансової підтримки гранта 34.02.01.01Ф Національної академії аграрних наук України.

Список літератури

1. Kamar N. et al. Hepatitis E virus infections. *Clinical Microbiology Reviews*. 2014. Vol. 27, No 1. P. 116–138. DOI: <https://doi.org/10.1128/cmr.00057-13>.
2. Lhomme S. et al. Clinical manifestations, pathogenesis and treatment of Hepatitis E virus infections. *Journal of Clinical Medicine*. 2020. Vol. 9, No 2. P. 331–355. DOI: <https://doi.org/10.3390/jcm9020331>.
3. He B. et al. Functional epitopes on hepatitis E virions and recombinant capsids are highly conformation dependent. *Human Vaccines and Immunotherapeutics*. 2020. Vol. 16, No 7. P. 1554–1564. DOI: <https://doi.org/10.1080/21645515.2019.1703454>.
4. Doceul V. et al. Zoonotic hepatitis E virus: classification, animal reservoirs and transmission routes. *Viruses*. 2016. Vol. 8, No 10. P. 270. DOI: <https://doi.org/10.3390/v8100270>.
5. Hoan N. X. et al. High hepatitis E virus (HEV) positivity among domestic pigs and risk of HEV infection of individuals occupationally exposed to pigs and pork meat in Hanoi, Vietnam. *Open Forum Infectious Diseases*. 2019. Vol. 6, No 9. P. ofz306. DOI: <https://doi.org/10.1093/ofid/ofz306>.
6. Kenney S. P. The current host range of hepatitis E viruses. *Viruses*. 2019. Vol. 11, No 5. P. 452. DOI: <https://doi.org/10.3390/v11050452>.
7. Goto A. et al. Understanding the genetics of viral drug resistance by integrating clinical data and mining of the scientific literature. *Scientific Reports*. 2022. Vol. 12, No 1. P. 14476–14486. DOI: <https://doi.org/10.1038/s41598-022-17746-3>.
8. Stern A. et al. The evolutionary pathway to virulence of an RNA virus. *Cell*. 2017. Vol. 169, No 1. P. 35–46. DOI: <https://doi.org/10.1016/j.cell.2017.03.013>.
9. Kimura M. The neutral theory of molecular evolution: a review of recent evidence. *Japanese Journal of Genetics*. 1991. Vol. 66, No 4. P. 367–386. DOI: <https://doi.org/10.1266/jjg.66.367>.
10. Sueoka N. Directional mutation pressure and neutral molecular evolution. *Proceedings of the National Academy of Sciences*. 1988. Vol. 85. P. 2653–2657. DOI: <https://doi.org/10.1073/pnas.85.8.2653>.
11. Blazej P. et al. Optimization of mutation pressure in relation to properties of protein-coding sequences in bacterial genomes. *PLoS One*. 2015. Vol. 10, No 6. P. e0130411. DOI: <https://doi.org/10.1371/journal.pone.0130411>.
12. Nasrullah I. et al. Genomic analysis of codon usage shows influence of mutation pressure, natural selection, and host features on Marburg virus evolution. *BMC Evolutionary Biology*. 2015. Vol. 15. P. 174. DOI: <https://doi.org/10.1186/s12862-015-0456-4>.
13. Dasmeh P. et al. The influence of selection for protein stability on dN/dS estimation. *Genome Biology and Evolution*. 2014. Vol. 6, No 10. P. 2956–2967. DOI: <https://doi.org/10.1093/gbe/evu223>.
14. Aziz R. et al. Incorporation of transition to transversion ratio and nonsense mutations, improves the estimation of the number of synonymous and non-synonymous sites in codons. *DNA Research*. 2022. Vol. 29, No 4. P. 1–8. DOI: <https://doi.org/10.1093/dnares/dsac023>.
15. Valdar W. S. J. Scoring residue conservation. *Proteins*. 2002. Vol. 48. P. 227–241. DOI: <https://doi.org/10.1002/prot.10146>.
16. Tamura K. et al. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution*. 2013. Vol. 30, No 12. P. 2725–2729. DOI: <https://doi.org/10.1093/molbev/mst197>.
17. Hall T. A. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*. 1999. Vol. 41. P. 95–98.
18. Kimura M. The neutral theory of molecular evolution. *Cambridge: Cambridge University Press*, 1983. 367 p. DOI: <https://doi.org/10.1017/CBO9780511623486>.
19. Reyes G. R. et al. Isolation of a cDNA from the virus responsible for enterically transmitted non-A, non-B hepatitis. *Science*. 1990. Vol. 247, No 4948. P. 1335–1339. DOI: <https://doi.org/10.1126/science.2107574>.

20. Кордюм В. А. Эволюция вирусов — попытка нелинейного прогноза. *Биополимеры и клетина*. 2001. Т. 17, № 6. С. 467–486. DOI: <https://doi.org/10.7124/bc.0005D6>.
21. Lukashov V., Goudsmit J. Evolutionary relationships among Astroviridae. *Journal of General Virology*. 2002. Vol. 83. P. 1397–1405. DOI: <https://doi.org/10.1099/0022-1317-83-6-1397>.

CHARACTERIZATION OF SWINE HEPATITIS E VIRUS GENOTYPE 3 OPEN READING FRAMES

Lymanska O. Yu.

National Scientific Center “Institute of Experimental and Clinical Veterinary Medicine”, Kharkiv, Ukraine

The goal of this study was a determination of structural organization features of swine hepatitis E virus genotype 3 open reading frames. Gene selective pressure for genes was estimated using the Tajima's neutrality test as implemented in MEGA 6. Shannon entropy was used for diversity detection at each position of open reading frames. Shannon entropy plots representing quantitative diversity at each amino acid (nucleotide) position of 11 swine HEV-3 isolates aligned sequences were calculated by software BioEdit. For phylogenetic analysis, nucleotide distances were analyzed using the neighbour-joining algorithm. Analysis of nucleotide, non-synonymous and synonymous distances was performed using MEGA6. Parameters of mutational analysis for swine hepatitis E virus genotype 3 open reading frames ORF1 – ORF3 were determined using Tajima's neutrality test. Shannon entropy technique revealed that ORF3 is characterized by the most number of diverse amino acid residues (21,7 %) comparing with ORF1, ORF2 coding regions. Analysis of non-synonymous distances in ORF3 allowed to determine an order of evolutionary events in the history of swine HEV-3 isolates. Avian astroviruses strains were used as control because phylogenetic relationships among them correspond to the evolution history of their host species. Tree topology for swine HEV-3 ORF3 based on non-synonymous distances differs from tree structures based on nucleotide and synonymous distances. Positive values of *D* Tajima's factor for ORF1, ORF2 of swine HEV-3 show on operating positive selection onto these coding regions. Negative value of *D* Tajima's factor for ORF3 coding region indicates onto purifying selection

Keywords: positive selection, Tajima's neutrality test, negative selection, Shannon entropy

УДК 619:616.98:578.828:578.2'21:636.22/.28

DOI 10.36016/VM-2023-109-5

СТРУКТУРНИЙ АНАЛІЗ ПРОТЕЇНІВ ВІДКРИТИХ РАМОК ЗЧИТУВАННЯ ВІРУСУ ІМУНОДЕФІЦИТУ ВЕЛИКОЇ РОГАТОЇ ХУДОБИ

Балак О. К.

Харківський національний медичний університет, Харків, Україна

Лиманська О. Ю.

Національний науковий центр «Інститут експериментальної і клінічної ветеринарної медицини», Харків, Україна, e-mail: olgaliman@ukr.net

Метою роботи було визначення особливостей структурної організації протеїнів ORF2 та ORF3 відкритих рамок зчитування вірусу імунодефіциту великої рогатої худоби (ВІ ВРХ). За допомогою програми ATGgr, яка дозволяє ефективно передбачення кодонів ініціації трансляції з точністю до нуклеотида, визначено ORFs для двох ізолятів ВІ ВРХ з повним геномом. Для передбачення, аналізу вторинної структури та функції протеїнів використовували програму Rhyre2. Для пошуку фрагментів протеїнів в неупорядкованому, або нативно-розгорнутому, стані використовували програму PONDR-FIT. Аналіз амінокислотного складу протеїнів ORF2 та ORF3 ізолятів ВІ ВРХ щодо наявності неполярних, полярних, ароматичних, гідрофобних амінокислотних залишків проведено за допомогою програми PSIPRED. Моделі просторової структури протеїнів отримано за допомогою сервера I-TASSER. Для протеїну ORF3 передбачено 14 % α -спіралей, 17 % β -тяжів, 43 % неупорядкованої структури. Для поліпротеїну Gag, який транскрибується з відкритої рамки зчитування ORF2, визначено 37 % α -спіралей, 0 % β -тяжів, 41 % неупорядкованої структури. Розподіл заряджених амінокислотних залишків характеризує поверхневі властивості протеїнів. Для протеїну ORF2 їхня кількість сягає 23,9 %. Кількість Arg становить 5,2 %, Lys — 8,0 %, Glu — 7,3 %, Asp — 3,4 %. Загальна кількість заряджених амінокислотних залишків ORF3 становить 23,3 %. Кількість Arg становить 12,6 %, Lys —