

НЕКАНОНІЧНІ СТРУКТУРИ В ГЕНОМІ СПУМАВІРУСУ ВЕЛИКОЇ РОГАТОЇ ХУДОБИ

О.Ю. ЛИМАНСЬКА^{1,*}, О.К. БАЛАК², О.П. ЛИМАНСЬКИЙ³

¹Національний науковий центр «Інститут експериментальної і клінічної ветеринарної медицини»,
вул. Г. Сковороди, 83, 61023, Харків, Україна

²Харківський національний медичний університет, пр. Науки, 4, 61022, Харків, Україна

³Інститут фізіологічно активних сполук, пр. Науки, 58, 61072, Харків, Україна

E-mail: olgaliman@ukr.net, ok.balak@knu.edu.ua, apliman@ukr.net

Автор для кореспонденції — О.Ю. Лиманська, e-mail: olgaliman@ukr.net

За допомогою біоінформатичних методів ідентифіковано потенційні досконалі G-квадруплекси (G4s) та потрібні з'єднання 3WJs (three-way junctions) в геномі спумавірусу великої рогатої худоби (СВ ВРХ). Штучний інтелект AlphaFold 3 використано для підтвердження потенційних G4s та 3WJs через побудову 3D моделей цих неканонічних структур. G4s — вторинні структури, які утворено G-багатими послідовностями. Мультиспіральне з'єднання 3WJ (яке утворено трьома з'єднаними у точці дуплексами), як і G4s, відносять до альтернативних структур в ДНК та РНК, які відрізняються від класичної дволанцюгової В-ДНК. В роботі створено карту локалізації на геномі СВ ВРХ потенційних консервативних внутрішньомолекулярних G4s, які утворено двома G-тетрадами. На підставі множинного вирівнювання 37 ізолятів СВ ВРХ з повним геномом в сенсовій нитці провірусної ДНК СВ ВРХ знайдено 7 потенційних консервативних G-квадруплексів, в антисенсовій нитці — 22 G4s, що утворено двома G-тетрадами, G-рахунок яких становить від 32 до 36. Для сенсової нитки провірусної ДНК СВ ВРХ щільність G4s становила 0,6 G4/тисяча нуклеотидів (т.н.), в той час як для антисенсової нитки — 1.8 G4/т.н. Для набору з 37 ізолятів СВ ВРХ знайдено один консервативний мотив 3WJ довжиною 73 н. зі 100%-вим рівнем подібності, який локалізовано в 5'-нетрансльованій області і частково на 5'-кінці гена gag. Структуру 3WJ в РНК СВ ВРХ стабілізовано 20 комплементарними п.н. з вільною енергією ΔG , що становить — 19,8 ккал/моль. Доведено значущість даної структури для функціонування спумавірусу ВРХ. Застосування AlphaFold 3 для побудови 3D моделей теоретично визначених досконалих потенційних G4s та 3WJs в геномі СВ ВРХ дозволило надійно визначати потенційні альтернативні структури в нуклеїнових кислотах.

Ключові слова: альтернативна структура, біоінформатика, AlphaFold 3, спумавірус великої рогатої худоби.

добі, G-квадруплекс (G4), потрібне з'єднання (3WJ), передбачення неканонічних структур.

Вступ. Неканонічні структури, які утворюються в молекулах нуклеїнових кислот (шпильки, хрести, триплекси, мультиспіральні з'єднання 3WJs (three-way junctions), 4WJs, 5WJs, G-квадруплекси (G4s), і-мотиви), є регуляторними елементами, які впливають на функціонування геному та контролюють різноманітні біологічні процеси (Kardmas et al., 1995).

Мультиспіральне з'єднання 3WJ — це характерний для вторинної структури молекул нуклеїнових кислот лабільний гнучкий фрагмент, що складається з трьох з'єднаних у точці дуплексів (Wu et al., 2004).

Встановлено, що 3WJs є консервативними мотивами внутрішнього сайту входу рибосоми вірусів гепатиту А та С і відіграють певну роль у внутрішній ініціації трансляції за механізмом, який є альтернативним кеп-залежній трансляції (Ouellet et al., 2010; Koirala et al., 2019). Неканонічному механізму трансляції сприяють також 3WJs, які локалізовано у послідовностях геномних РНК деяких вірусів рослин (Ojha et al., 2024).

Інформація щодо можливості формування 3WJ у геномі вірусів тварин на цей час є доволі обмеженою (Heng et al., 2024).

G-квадруплекси (G4s) — це неканонічні структури, фолдинг яких відбувається у збагачених гуанідіном послідовностях молекул ДНК або РНК за наявності сольватованих катіонів. Основною одиницею G-квадруплексу є G-тетрада, утворена чотирма гуанідінами через Хугстинівські взаємодії. G-квадруплекси стабілізовано стекінг-взаємодією плоских G-тетрад та додатково іонами фізіологічно значущих одновалентних металів, насамперед, K^+

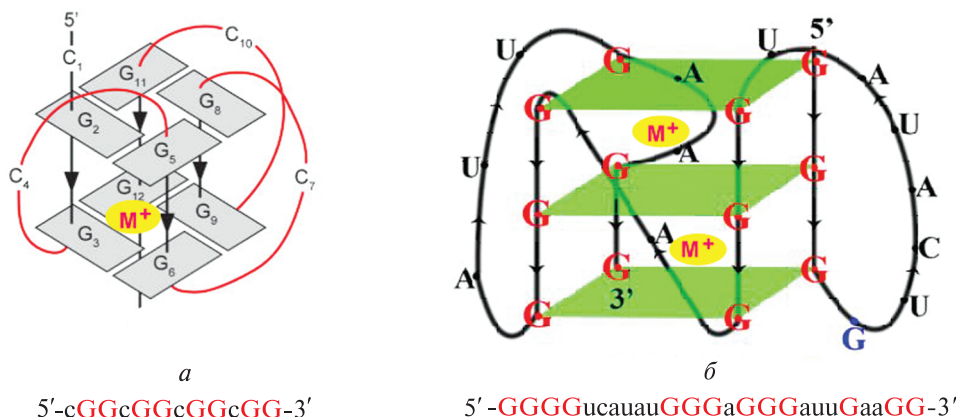


Рис. 1. Структура внутрішньомолекулярних паралельних G-квадруплексів, що утворено через стекінг двох G-тетрад з мотивом 5'-(CGG)₄-3' (Wolfe et al., 2014) (а) і стекінг трьох G-тетрад, які містять три фрагменти GGG, G та GG в гені L РНК вірусу Ебола (Wang et al., 2016), та стабілізовано центральним сольватованим моновалентним катіоном ($K^+ > Na^+ > Li^+$) (б). G-квартет містить чотири копланарних гуанідини, які стабілізовано водневими зв'язками

та в меншій мірі — Na^+ , а також дво валентних металів (зокрема, Sr^{2+} , Ba^{2+}) (Lu et al., 2015; Fay et al., 2017) (рис. 1).

G4s відіграють важливу роль у таких процесах, як мутагенез, патогенез, транскрипція, трансляція, реплікація, рекомбінація тощо (Belotserkovskii et al., 2013; Varshney et al., 2020), впливають на стабільність геному та можуть бути асоційованими з онкологічними, серцево-судинними, неврологічними захворюваннями, що дозволяє розглядати ці неканонічні структури як перспективні терапевтичні мішені та не виключає їх використання у молекулярній діагностиці. G4s виконують регуляторну функцію завдяки здатності G-мотивів за певних умов переходити з класичної В-ДНК в форму G4.

Спумавіруси (СВ) належать до родини *Reoviridae* підродини *Spumaretrovirinae*, зустрічаються у багатьох видів тварин, у тому числі у великої рогатої худоби (ВРХ), та є найбільш давніми серед усіх відомих ретровірусів (Rethwilm et al., 2013; Pinto-Santini et al., 2017).

Інформація щодо існування у геномі спумавірусів неканонічних структур з потенціалом утворення G4s та 3WJs у літературі на цей час є відсутньою.

В даній роботі знайдено консервативну 3WJ структуру в геномній РНК СВ ВРХ та визначено потенційні консервативні досконалі G-квадруплекси, які утворено двома G-тетрадами в (+)- та (-)-нитках провірусної ДНК СВ

ВРХ і підтверджено 3D моделями, отриманими за допомогою штучного інтелекту AlphaFold 3.

Матеріали і методи. Нуклеотидні послідовності 37 ізолятів спумавірусу ВРХ у форматі FASTA отримано через пошук таксономічного ідентифікатора (txid) 207343 в геномній базі даних GenBank Національного центру біотехнологічної інформації (США) станом на січень 2025 р.

Пошук мотивів GGG в провірусній ДНК СВ ВРХ, маніпулювання з нуклеотидними послідовностями (зворотна, комплементарна) та множинне вирівнювання проведено за допомогою програми Molecular Evolutionary Genetics Analysis MEGA12 (версія 12.0.9) (Kumar et al., 2024). Програму BioEdit (версія 7.2.5) (Hall, 1999) використано для побудови консенсусної повногеномної послідовності провірусної ДНК СВ ВРХ.

Для пошуку мотивів внутрішньомолекулярних G-квадруплексів в (+)- та (-)-нитках провірусної ДНК СВ ВРХ, а також для визначення G-рахунку (параметра, який характеризує стабільність G-квадруплексу у відносних одиницях) використано програму для аналізу квадруплексів QGRS Mapper на веб-сервері (<https://bioinformatics.ramapo.edu/QGRS/analyze.php>) (Kikin et al., 2006; Menendez et al., 2012) за наступних параметрів: максимальна довжина G-квадруплексу — 45 н., мінімальний розмір G-тетради — 2, петля — від 0 до 20 н.

Для пошуку потенційних досконалих G4s застосовано програму QGRS Mapper через найкращі показники порівняно з іншими програмами в частині помилково та правильно знайдених, помилково та правильно ненайдених G4s при тестуванні на одному й тому ж G-мотиві (Lombardi et al., 2020).

Послідовність, яку використовують для передбачення G4s за допомогою біоінформатичних методів, можна описати мотивом $G_{3-6}-N_x-G_{3-6}-N_x-G_{3-6}$, де N відповідає будь-якому нуклеотиду у складі петлі, x – довжина петлі (від 0 до 12 н.) (Lombardi et al., 2020).

Додатково потенційні G4s ідентифіковано за допомогою програми G4Hunter на веб-сервері <http://bioinformatics.ibp.cz> (Bedrat et al., 2016; Brazda et al., 2019) за наступних параметрів: вікно пошуку – 25 н., поріг – 1.2.

Для визначення термодинамічних параметрів шпилькової структури (температури плавлення (T_m) та вільної енергії Гіббса ΔG) використовували програму Mfold (Zucker, 2003).

В якості контролю коректності пошуку G4-мотивів використовували послідовності, що утворюють G-квадруплекси, які експериментально підтверджено для низки вірусів та бактерій. Для таких послідовностей, зокрема, для ізоляту AJ320484 вірусу імунodefіциту людини (ВІЛ) (Piekna-Przybylska et al., 2014), вірусу Ебола (Wang et al., 2016) нами розраховано G-рахунок, який становить 64, 71 та 33 відповідно. 13 G4s було знайдено в геномі вірусу гепатиту В (HBV), для 7 з яких значення G-рахунку варіювало від 19 до 35.

Значення G-рахунку для G4s в генах *murE*, *ftsB*, *mexC* *Pseudomonas aeruginosa*, які автори (Evans et al., 2023) визначали для довжини G-квадруплексу до 30 н., нами розраховано для довжини G-квадруплексу до 45 н. та становило 32, 35 та 35 відповідно.

Беручи до уваги експериментальні дані щодо параметрів G4s, які було детектовано в геномі HBV, ВІЛ, *Pseudomonas aeruginosa*, аналізували тільки консервативні фрагменти ізолятів СВ ВРХ, які можуть утворювати G4s з G-рахунком ≥ 32 . Для передбачення потенційних G4s використано мотив $G_x N_y G_x N_y G_x N_y G_x$ з максимальною довжиною 45 н. та $2 \leq x \leq 4$, $1 \leq y \leq 10$ (Kikin et al., 2006).

3D моделі G4s будували за допомогою штучного інтелекту AlphaFold 3 (Abramson et al., 2024) за 10 іонів K^+ . Для оцінки надійності передбачення моделі в AlphaFold 3 використано параметр pIDDT, значення якого для високого рівня достовірності моделі має становити понад 90 за шкалою від 0 до 100. Значення pIDDT від 70 до 90 відповідає середньому рівню надійності моделі, а $70 > pIDDT > 50$ – низькому рівню. Для візуалізації неканонічних структур використовували також програму RasWin Molecular Graphics (версія 2.7.5.2; www.openrasmol.org).

Для пошуку внутрішньомолекулярних 3WJs в геномній РНК СВ ВРХ використано програму Vfold2D (<http://rna.physics.missouri.edu>) (Xu et al., 2016). Для підтвердження вторинної структури та визначення термодинамічних параметрів 3WJs використовували програму Mfold (Zucker, 2003). Множинне вирівнювання нуклеотидних послідовностей та пошук консервативних мотивів 3WJ-структур для 37 ізолятів СВ ВРХ з повним геномом з бази даних GenBank проведено в пакеті програм MEGA (версія 6.06) за допомогою програми Clustal W (Higgins et al., 1994).

Пошук неканонічних структур у геномі СВ ВРХ та визначення їх температур плавлення проводили за фізіологічних умов: іонна сила $I = 0,2 \text{ M Na}^+$, $[Mg]^{2+} = 0,0 \text{ mM}$ або $I = 0,15 \text{ M Na}^+$, $[Mg]^{2+} = 0,2 \text{ mM}$.

Пошук неканонічних 3WJ-структур проводили для індивідуальної послідовності молекули мРНК ізоляту JX307861 СВ ВРХ (виділено в Польщі; довжина становить 12010 н.), яку було розрізано на 114 фрагментів довжиною 145 н., що перекриваються на 40 н. Далі набір з 37 ізолятів СВ ВРХ аналізували на наявність 3WJs.

Результати досліджень та їх обговорення. Незважаючи на численні дані щодо властивостей гуанідинових квадруплексів, наші знання про наявність цих структур в геномі патогенів все ще залишаються неповними.

Потенційні G4s у геномі снумавірусу ВРХ. G4s в ДНК та РНК мають різну топологію. G4s в ДНК можуть приймати паралельну, антипаралельну або змішану конформації. Топологія міжмолекулярних G4s в РНК обмежена паралельною конформацією, в якій всі чотири нитки орієнтовано в одному напрямку. Внутріш-

ньомолекулярні взаємодії в G4s в РНК ведуть до підвищення їхньої стабільності порівняно з G4s в ДНК (Fay et al., 2017).

G4s можуть впливати на експресію генів також за умови локалізації на (–)-нитці по відношенню до напрямку транскрипції. Беручи до уваги наведену вище обставину, пошук мотивів потенційних G4s здійснювали для мРНК, (+)- і (–)-нитки провірусної ДНК СВ ВРХ. Відносно недавно в експериментах *in vitro* та *in vivo* було підтверджено існування недосконалих G4s (De Nicola et al., 2016; Piazza et al., 2017). Для передбачення G4s, які мають випуклості, помилки в G-мотиві, недосконалі G-тетради, було розроблено нові біоінформатичні методи, які допускають існування до трьох недосконалостей в окремій послідовності (помилки, випуклості та/або петлі довжиною понад 9 н.). В даній роботі розглянуто тільки досконалі G4s в геномі СВ ВРХ.

На першому етапі на підставі множинного вирівнювання 37 ізолятів спумавірусу ВРХ з повним геномом в сенсовій нитці провірусної ДНК СВ ВРХ знайдено 19 потенційних консервативних G-квадруплексів з двох G-тетрад, G-рахунок яких становив від 32 до 36 (табл. 1). Для антисенсової нитки знайдено 31 G4s, які утворено двома G-тетрадами, з G-рахунком від 30 до 36 та 2 G4s, що утворено трьома G-тетрадами, G-рахунок яких становив 53. Ці два G4s є прямими повторами DR1, які локалізовано в 5'- та 3'-LTR.

На другому етапі використано AlphaFold 3 для побудови 3D моделей потенційних G4s в

геномі СВ ВРХ. В якості контролю коректності побудови цих моделей використовували 3D модель G4, утворення якого експериментально підтверджено в гені *texC* грампозитивної бактерії *P. aeruginosa*.

З теоретично визначених на першому етапі 19 консервативних G-квадруплексів в сенсовій нитці провірусної ДНК СВ ВРХ побудовано 3D моделі G4, які містять дві G-тетради, тільки для 7 (табл. 1). Для антисенсової нитки з 31-го G4s, які утворено двома G-тетрадами, та 2-х G4s, що утворено трьома G-тетрадами, побудовано 3D моделі G4 тільки для 22-х, які утворено двома G-тетрадами (табл. 2 та рис. 2).

Для мРНК СВ ВРХ існування G4s не підтверджено, оскільки для жодного з 23 теоретично визначених G4s в мРНК 3D моделі потенційних G4s AlphaFold 3 не побудував.

При визначенні потенційних G4s враховували (i) вплив на їхню структуру довжини петлі, збільшення якої веде до зменшення стабільності G4, (ii) дестабілізуючий ефект аденозина для одонуклеотидної петлі або в 5'-положенні короткої петлі порівняно з тимідіном та цитидіном (Jana et al., 2021).

Проведено порівняння ефективності пошуку G4s за допомогою двох програм – QGRS Mapper та G4Hunter за рекомендованих авторами параметрів. За допомогою програми QGRS Mapper виявлено 7 потенційних G4s для мРНК та (+)-нитки провірусної ДНК СВ ВРХ (табл. 1), в той час як за допомогою програми G4Hunter не виявлено жодного G4. Для (–)-нитки провірусної ДНК СВ ВРХ за

Таблиця 1. Консервативні мотиви потенційних G-квадруплексів у сенсовій нитці провірусної ДНК спумавірусу великої рогатої худоби. G-багаті фрагменти послідовностей підкреслено. Позиції генів наведено для сенсової нитки ізоляту JX307861 СВ ВРХ. DR3, DR4, DR5 – прямі повтори. Параметр pIDDT характеризує надійність передбачення моделі в AlphaFold 3

Положення	Ген/елемент	Мотив G-квадруплексу			G-рахунок
56-69	U3	<u>GGAGGATTGGCTGG</u>	pIDDT > 90	DR3	34
10754-10767	U3	<u>GGAGGATTGGCTGG</u>	pIDDT > 90	DR3	34
486-499	U3	<u>GGTTCGGAGGATGG</u>	90 > pIDDT > 70	DR4	34
11184-11197	U3	<u>GGTTCGGAGGATGG</u>	90 > pIDDT > 70	DR4	34
574-590	U3	<u>GGTTCGAGGTCAGGCGG</u>	90 > pIDDT > 70	DR5	32
11272-11288	U3	<u>GGTTCGAGGTCAGGCGG</u>	90 > pIDDT > 70	DR5	32
7309-7349	<i>env</i>	<u>GGcactgttacgtGGtacatattccgGGGgaaagaatGG</u>	70 > pIDDT > 50		36

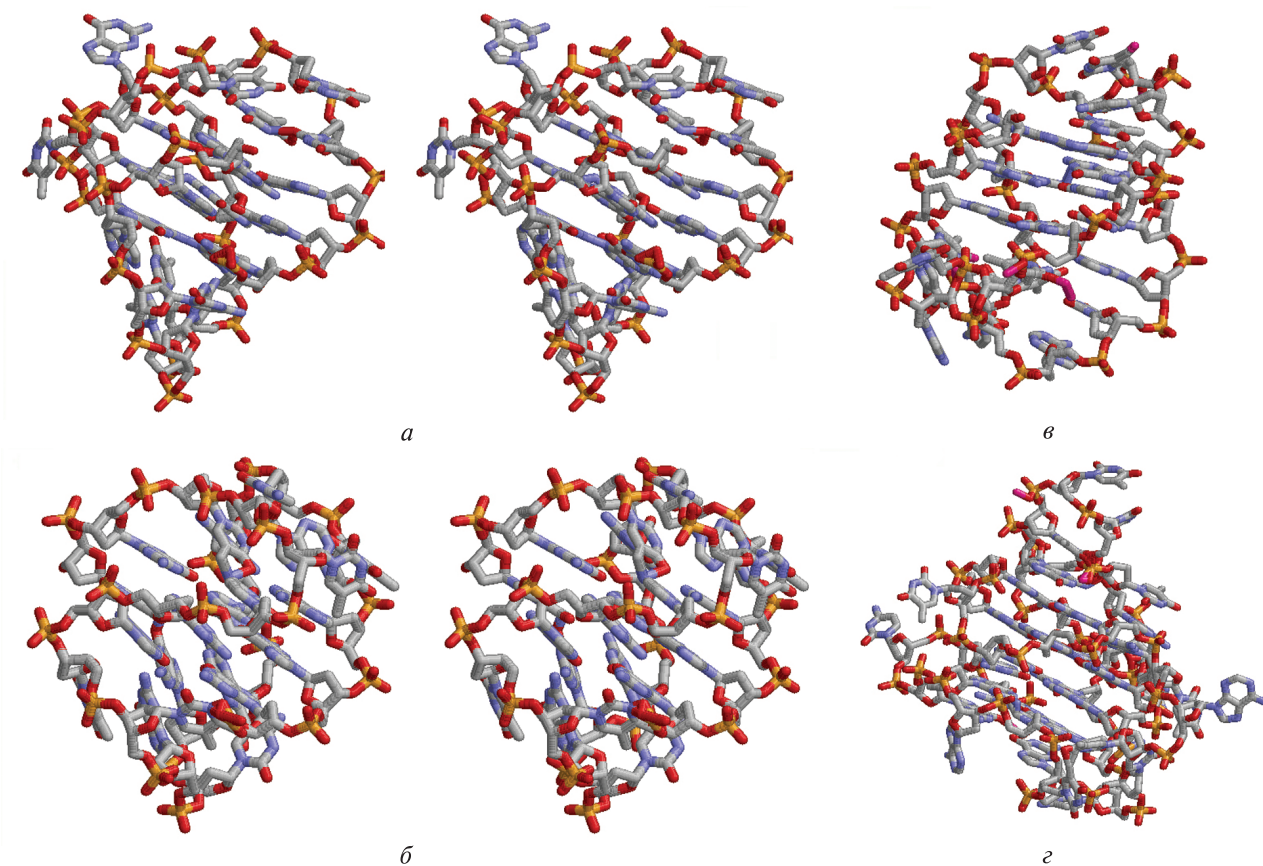


Рис. 2. 3D моделі потенційних G-квадруплексів в антисенсовій нитці провірусної ДНК ізоляту JX307861 СВ ВРХ (а, б, в) та експериментально визначеного G4 в гені *texC* *P. aeruginosa* (б), які побудовано за допомогою AlphaFold 3. а, б – стереозображення. Положення на геномі ізоляту JX307861: а – 7061-7082 (90 > pIDDT > 70); б – 4250-4274 (90 > pIDDT > 70); в – 2986-3020 (70 > pIDDT > 50)

допомогою програми QGRS Mapper виявлено 22 потенційні G4s, в той час як за допомогою програми G4Hunter ідентифіковано чотири G4s (табл. 2, мотиви виділено), які є фрагментами G4s, що виявлено за допомогою програми QGRS Mapper.

На геномній карті СВ ВРХ, яка показує позиції структурних генів, регуляторних елементів, генетичних маркерів та відстань між ними, нами зазначено сайти потенційних G4s (рис. 3). В генах *bel1*, *bel2*, *gag*, *pol*, *env* (–)-нитки ідентифіковано 1, 1, 9, 1 та 3 G4s, в той час як (+)-нитки – тільки один G4 в гені *gag*. Наведені дані свідчать, що утворення G4s в (+)- та (–)-нитках провірусної ДНК СВ ВРХ відбувається асиметрично, що може вести до різного впливу G4s, які ідентифіковано в (+)- та (–)-нитках, на транскрипцію.

Відомо, що G4s в (–)-нитці, один з яких наведено на рис. 3 (б, в), істотним чином інгібують транскрипцію на відміну від G4s в (+)-нитці, які не впливають на транскрипцію (Agarwal et al., 2014). Проведений комп'ютерний аналіз 37 ізолятів СВ ВРХ з повним геномом показав, що G4s нерівномірно розподілено як серед структурних генів, так і в (+)- та (–)-нитках провірусної ДНК СВ ВРХ.

Створена геномна карта сайтів потенційних G4s дозволила виявити особливості структурної організації (+)-нитки порівняно з (–)-ниткою провірусної ДНК СВ ВРХ. Для (+)-нитки виявлено 3 прямих повтори (DR3 – DR5), для (–)-нитки – 1 прямий повтор DR1, які містять G-квадруплекси, що утворено двома G-тетрадами, G-рахунок яких становить 32 – 34.

Таблиця 2. Консервативні мотиви потенційних G-квадруплексів в антисенсовій нитці провірусної ДНК спумавірусу великої рогатої худоби. G-багаті фрагменти послідовностей підкреслено. Позиції генів наведено для сенсової нитки ізоляту JX307861 CB BPX. W = a/t, y = c/t. DR1 – прямий повтор. Мотиви, які ідентифіковано програмою G4Hunter, виділено. G-рахунок визначено за допомогою програм QGRS Mapper та G4Hunter (в круглих дужках). Параметр pIDDT характеризує надійність передбачення моделі в AlphaFold 3

Положення	Ген/елемент	Мотив G-квадруплексу	G-рахунок
10715-10753	U3	<u>GGGCAAGTAAAGGGGGGCTGG</u> cyyttgctwatttcy <u>GGG</u> DR1	32
17-55	U3	<u>GGGCAAGTAAAGGGGGGCTGG</u> cyyttgctwatttcy <u>GGG</u> DR1	32
10671-10685	bel2	<u>GGAGGGCATGGCTGG</u> pIDDT > 90	35
10092-10121	bel1	<u>GGTCATCCCTGGCACATTGGCACAAAGAGG</u> 90 > pIDDT > 70	34
8234-8250	env	<u>GGAGGGAGAAGGCTCGG</u> 90 > pIDDT > 70	34
7561-7586	env	<u>GGAGTATAGGGTAGGGCTACAGAAGG</u> 90 > pIDDT > 70	32
7061-7082	env	<u>GGGTTTTGGGTGAGGTTCCAGG</u> 90 > pIDDT > 70	35
6193-6218	pol	<u>GGGCAGTGGTCTGTGAAGGAGAAGGG</u> 70 > pIDDT > 50	33
5743-5773	pol	<u>GGTTGCGTTGGCAGTTTGGGCCTTGGTGGGG</u> (33) 70 > pIDDT > 50	36
5669-5683	pol	<u>GGAGGGAGGGAGTG</u> pIDDT > 90	35
5165-5178	pol	<u>GGTTGGTGGCCTGG</u> 90 > pIDDT > 70	34
4250-4274	pol	<u>GGATACATTGGCCTTTGGCTTTGGG</u> 90 > pIDDT > 70	34
2986-3020	pol	<u>GGCGGAGTGCCGGCATCTGTGGGGTATGGGTGG</u> (34) 70 > pIDDT > 50	36
2874-2889	gag	<u>GGGGGCGGGACCTCGG</u> 90 > pIDDT > 70	32
2746-2771	gag	<u>GGATTGGGCCTGAGGGGATAACGAGG</u> (31) 70 > pIDDT > 50	33
1987-2006	gag	<u>GGCCCTGGAGCAGGTGCAGG</u> 90 > pIDDT > 70	36
1969-1979	gag	<u>GGTGGAGGAGG</u> pIDDT > 90	36
1939-1958	gag	<u>GGAGCAGGCAATGGAGCTGG</u> 70 > pIDDT > 50	36
1915-1932	gag	<u>GGCTGGAGGTGGTGGAGG</u> (36) 90 > pIDDT > 70	35
1891-1918	gag	<u>GGAGGAGCAGGTCCTTGTGGAGCCGCTGG</u> 70 > pIDDT > 50	36
1633-1652	gag	<u>GGGTAAGGTTGGGTCCGAGG</u> pIDDT > 90	34
1573-1610	gag	<u>GGTCCCATTGAGGAATAGGCAAGGGTTGCCCTGCAGG</u> 70 > pIDDT > 50	36

Структурні гени (+)-нитки порівняно з (–)-ниткою провірусної ДНК CB BPX характеризуються різною щільністю G4s. Для гена *gag* (–)-нитки щільність G4s становила 5.5 G4/т.н., для гена *pol* – 1.6 G4/т.н., для гена *env* – 1.0 G4/т.н., в той час як для гена *gag* (+)-нитки – 0.6 G4/т.н.

Для (–)-нитки провірусної ДНК CB BPX є характерною надзвичайно висока щільність G4s для окремих фрагментів генів, що не спостерігається для (+)-нитки. Зокрема, у фрагменті гена *gag* дожиною 116 н. (позиції 1891–2006 н.) ідентифіковано 5 G4s, і щільність G4s становила 43.1 G4/т.н. Для фрагмента гена *gag* більшої довжини (1448 н., позиції 1573–3020 н.) з урахуванням додатково виявлених чотирьох G4s їхня щільність становила 6.2 G4/т.н. Для сенсової нитки провірусної ДНК CB BPX щільність G4s становила 0.6 G4/т.н., в той час як для антисенсової нитки – 1.8 G4/т.н.

Ізолят JX307861 CB BPX, як і інший наявний в базі даних GenBank ізолят JX307862 (знайдено в Німеччині), містить два промотори – канонічний в довгому кінцевому повторі (LTR) та внутрішній (в позиції 9261–9360 гена *env*) (Wu et al., 2010; Hechler et al., 2012). Транскрипція спумавірусу регулюється на ранній та пізній фазах, які розділено використанням двох промоторів. 5'- та 3'-LTR сенсової нитки містять шість G4s, які є прямими повторами (DR3 – DR5, табл. 1), а 5'- та 3'-LTR антисенсової нитки містить два G4s, які також є прямими повторами (DR1, табл. 2), для яких побудовано 3D моделі з двома G-тетрадами за допомогою AlphaFold 3 (результат не показано). Для раніше теоретично визначених двох G4, які є прямими повторами DR2 в 5'- та 3'-LTR сенсової нитки та задовольняли критеріям визначення G4, не побудовано тривимірні альтернативні структури.

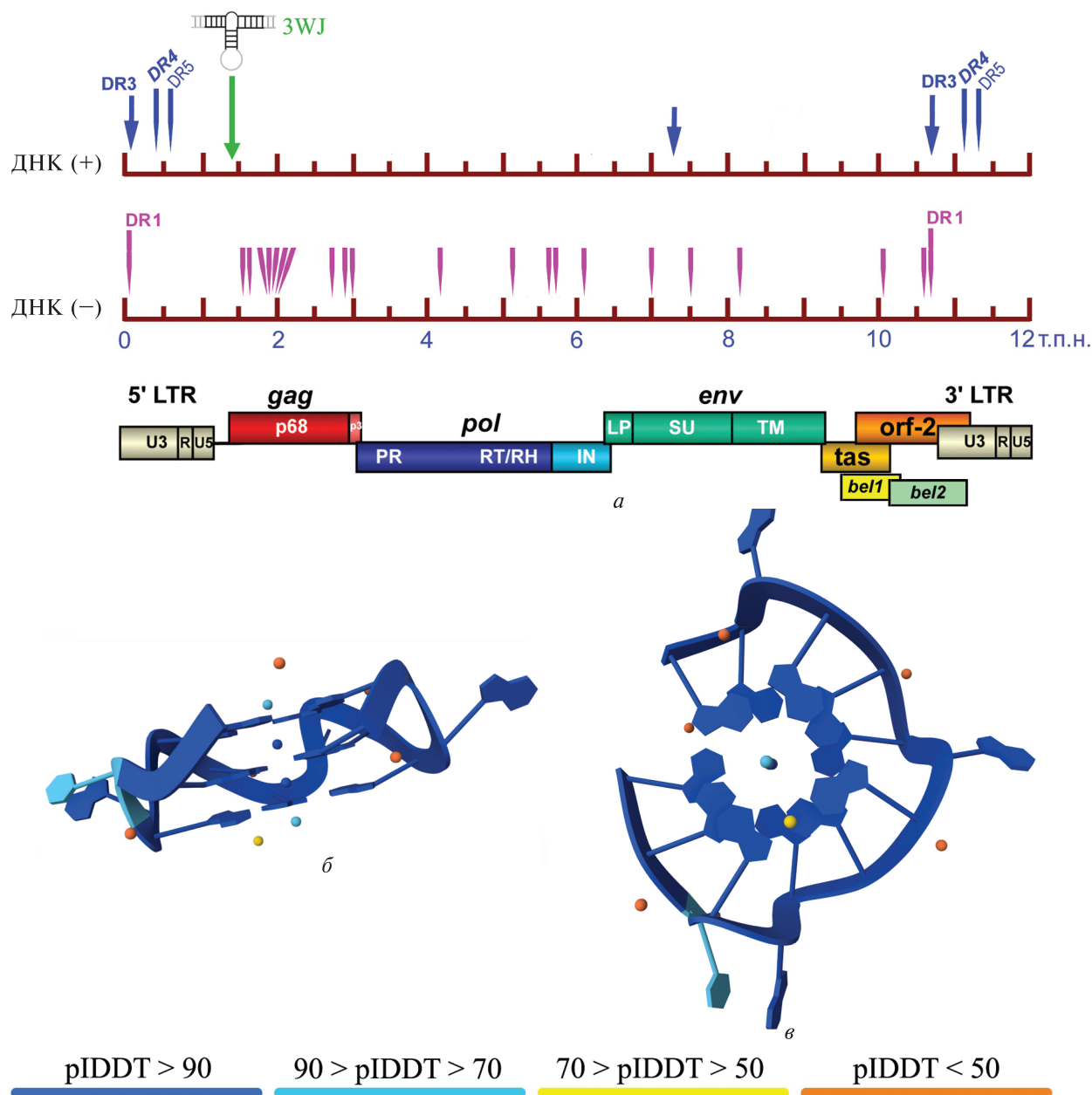


Рис. 3. Фізична карта геному спумавірусу BPX з наведеними позиціями відомих (Hamann et al., 2016) та додаткових генів *bel1* і *bel2* (Mekata et al., 2021). Стрілками та вертикальними лініями показано мотиви G-квадруплексів в (+)- та (-)-нитках провірусної ДНК відповідно. DR1 – прямий повтор в (-)-нитці, DR3 – DR5 – прямі повтори в (+)-нитці провірусної ДНК СВ BPX. 3WJ – в положенні 1388-1460 ізоляту JX307861 СВ BPX (а). G4 в положенні 5669-5683 провірусної ДНК СВ BPX (пIDDT>90): вид збоку (б) та вид зверху (в)

В геномі табака (*Nicotiana tabacum*) було визначено області промоторів як області вище генів на 2000 н.п., 1500 н.п., 1000 н.п. та 500 н.п., розраховано кількість G4s для кожної

області та встановлено, що промотори 60,4 % генів містять, принаймні, один G4 (Song et al., 2024). В роботі (Stefos et al., 2022) промоторну область визначили як область, яка локалізу-

вана на 500 н. вище сайту старту транскрипції. Всі промотори безпосередньо ранніх генів вірусу простого герпеса першого типу містять консервативні послідовності, які можуть утворювати G4s (Frasson et al., 2019), що підкреслює можливу роль G4s в транскрипційному та трансляційному регулюванні вірусних генів, а також в пошкодженні та відновленні провірусної ДНК (Monsen et al., 2022).

Порівняння визначених параметрів – G-рахунків – для G4s, відомих з літературних експериментально визначених даних, та для СВ ВРХ показує, що G-рахунок для потенційних квадруплексів в геномі СВ ВРХ (32–36) за порядком величини збігається з G-рахунком для вірусу гепатиту В, ВІЛ (Piekna-Przybylska et al., 2014; Wang et al., 2023).

Створена карта розподілу потенційних G4s підіймає декілька питань, оскільки біологічна функція переважної більшості знайдених G4s є невизначеною. Відомо, що G4s, які локалізовані в 5'- та 3'-LTR, мають різний вплив на трансляцію (Whisnant et al., 2014). G4s в антисенсовій нитці на відміну від G4s в сенсовій нитці можуть створювати блоки для РНК-полімерази II і таким чином зменшувати рівень транскрипції (Agarwal et al., 2014).

Потенційні 3WJ-структури у геномі спумавірусу ВРХ. На першому етапі було знайдено 17 потенційних 3WJ-структур для послідовності мРНК ізоляту JX307861 (Hechler et al., 2012). Послідовність, що утворює 3WJ в РНК бактеріофага phi29 у *Bacillus subtilis* (Zhang et al., 2013), ми використовували як контроль для коректності біоінформатичного пошуку мотивів 3WJ (рис. 4).

Зазначена мала молекула РНК довжиною приблизно 120 н. відіграє істотну роль у компактизації ДНК phi29 *in vitro* (Guo et al., 1987). Три дуплекси цього 3WJ-мотиву, кристалічна структура якого була раніше підтверджена експериментально з роздільною здатністю 0,3 нм (Zhang et al., 2013; PDB: 4KZ2), містять 20 комплементарних пар нуклеотидів (рис. 4, а). Вільна енергія ΔG цієї 3WJ-структури, яку визначено за допомогою програми Mfold (але лише з примусовим утворенням комплементарних пар за наступних параметрів F 1 78 1), становить – 20,9 ккал/моль (рис. 4, б).

На другому етапі серед ідентифікованих 3WJs для подальшого аналізу вибирали такі, що за параметрами відповідали 3WJ-структурі, яку експериментально досліджено в роботі (Zhang et al., 2013) (рис. 4, а). Варто зазначити, що 3D модель цієї 3WJ, яку побудував AlphaFold 3, не підтверджує існування класичної 3WJ з трьома дуплексами. Замість трьох дуплексів для цієї структури AlphaFold 3 визначає чотири дуплекси (рис. 4 в, г), що зайвий раз підкреслює потужні можливості AlphaFold 3. Ці відомості збігаються з результатами проведеного нами біоінформатичного аналізу 3WJ з роботи (Zhang et al., 2013) (рис. 4, б), який також вказує на наявність чотирьох дуплексів в цій структурі, яка до того ж є енергетично не вигідною, оскільки її неможливо отримати без примусового утворення комплементарних пар (рис. 4, б).

Вільна енергія вибраних 3WJs становила не менше – 18 ккал/моль, а загальна кількість комплементарних пар становила не менше 20.

Нами знайдено 6 потенційних 3WJs в мРНК ізоляту JX307861 СВ ВРХ, які стабілізовано 20–26 комплементарними нуклеотидами та локалізовано в 5'-LTR (2 структури), генах *gag* (2 структури), *env*, *bel2* (табл. 3), а також три 4WJ-структури, дві 5WJ-структури (результати не показано). Однак лише дві 3WJ-структури (у генах *gag* (позиція 1388–1460) та *env* (позиція 8112–8174)) серед 6-ти вищезгаданих, які змодельовано за допомогою програми Vfold2D, збігаються з 3WJ-структурами, які визначено за допомогою програми Mfold (без примусового утворення комплементарних пар) (табл. 3).

П'ять із 6 ідентифікованих 3WJs не є консервативними для набору з 37 ізолятів СВ ВРХ. Визначено тільки один висококонсервативний мотив 3WJ (позиція 1388–1460 для ізоляту JX307861 СВ ВРХ, рис. 5) довжиною 73 н. зі 100%-вим рівнем подібності (рис. 5), який локалізовано між 5'-LTR (позиція 1–1312) та 5'-кінцем гена *gag* (позиція 1420–3054) і частково на 5'-кінці гена *gag*. Крім того, довший фрагмент довжиною 103 н. (позиція 1382–1484 для ізоляту JX307861 СВ ВРХ) також має 100 %-ву подібність для 37 ізолятів СВ ВРХ.

Ми використали 3WJ у РНК ВІЛ-1 (Song et al., 2021) як контроль для визначення потенційних 3WJs, оскільки цю неканонічну структуру

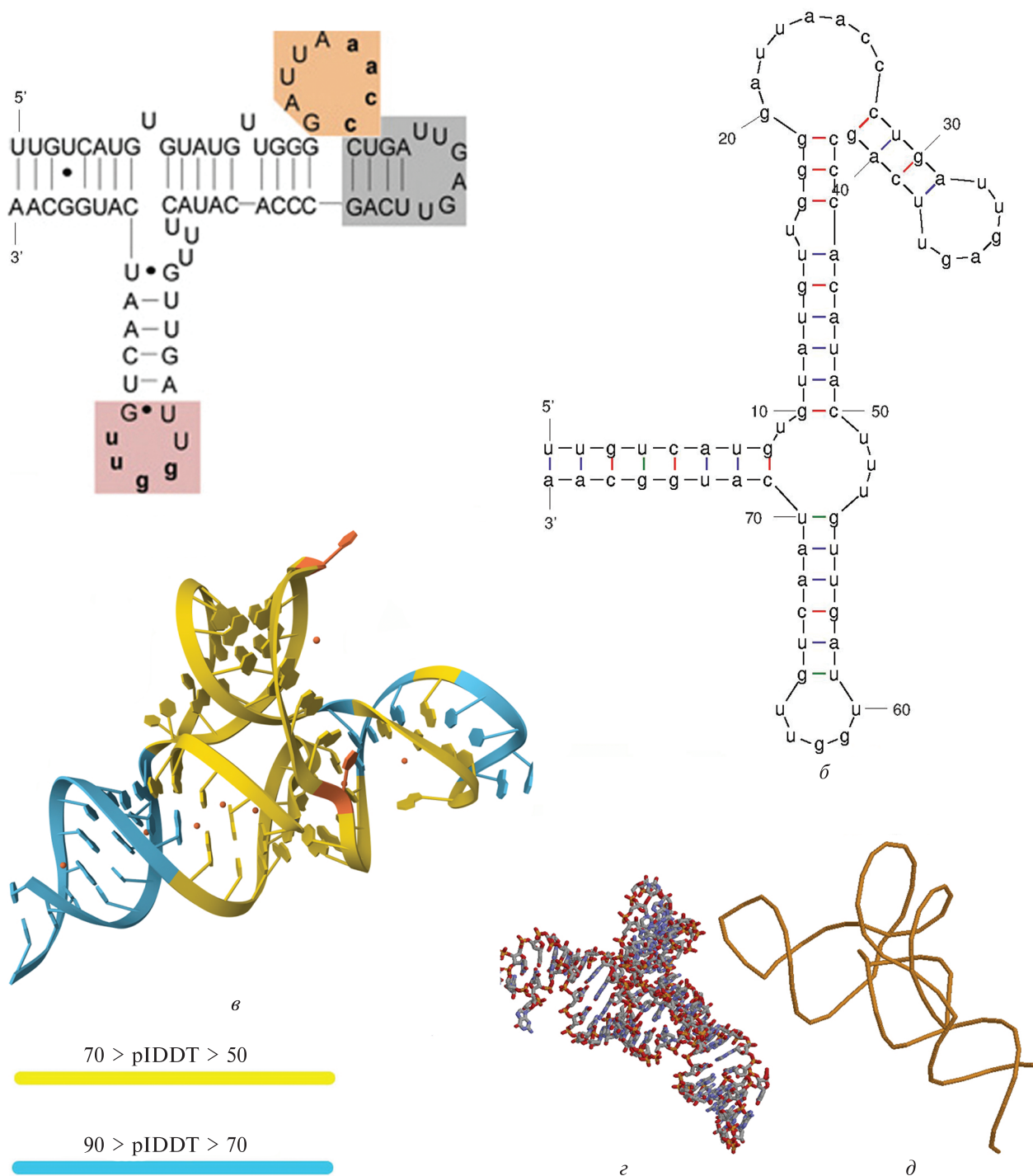
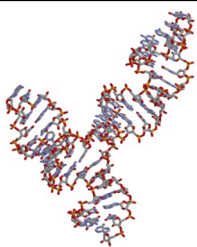
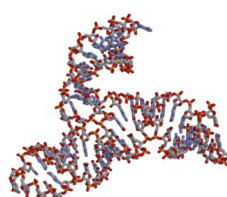
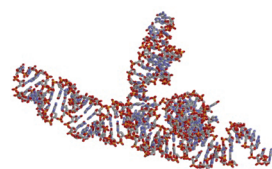
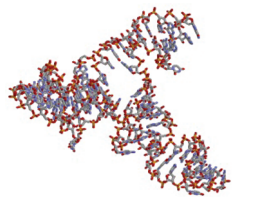
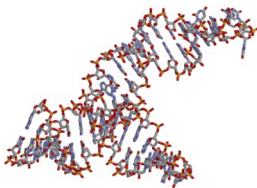
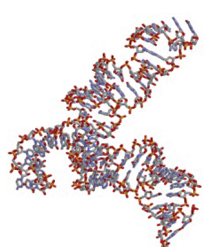


Рис. 4. Вторинна структура – 3WJ-структура, – яку знайдено в РНК пакувального мотору бактеріофага phi29, що зчеплений з пакуванням ДНК (Zhang et al, 2013) (а). Вторинну структуру 3WJ побудовано за допомогою програми Mfold (RNA Folding Form) (б), 3D модель – штучним інтелектом AlphaFold 3 з візуалізацією AlphaFold 3 (в) та програми RasWin Molecular Graphics (г – зв'язки, д – остов)

Таблиця 3. 3D моделі, які отримано за допомогою AlphaFold 3. Параметри 3WJ-структур в геномі ізоляту JX307861 СВ ВРХ і число комплементарних пар нуклеотидів у напрямку 5'→3' для трьох дуплексів, які утворюють 3WJ. Підкреслено параметри консервативної 3WJ-структури (№ 2) для всіх 37 ізолятів СВ ВРХ, що досліджено

AlphaFold 3		Положення на геномі; pIDDT	Ген/елемент	Число комплементарних пар
зв'язки	остов			
		72-135 70>pIDDT>50	5'-LTR	6-7-4
		1388-1460 70>pIDDT>50	5'-HTO <i>gag</i>	9-6-5
		2100-2245 70>pIDDT>50	<i>gag</i>	—
		2138-2224 70>pIDDT>50	<i>gag</i>	4-12-10
		8112-8174 90>pIDDT>70	<i>env</i>	9-5-6
		10644-10717 70>pIDDT>50	<i>bel2</i>	5-9-10

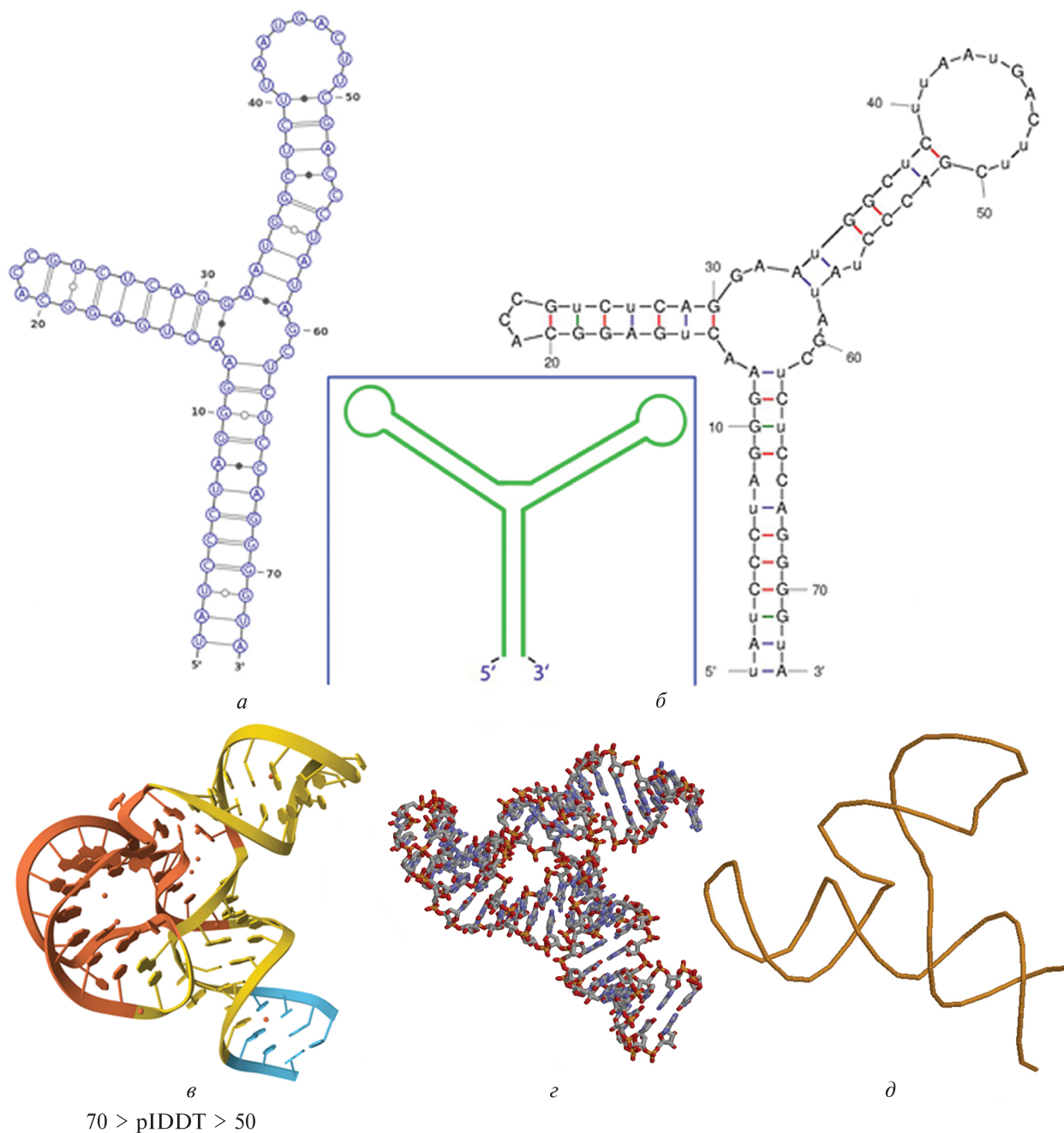


Рис. 5. Консервативна для 37 ізолатів СВ ВРХ структура 3WJ (позиція 1388-1460 для ізоляту JX307861), яку знайдено в 5'-НТО та 5'-кінці гена *gag* СВ ВРХ, що визначено за допомогою програм Vfold2D (а) та Mfold (б). 3D модель побудовано за допомогою AlphaFold 3 з візуалізацією AlphaFold 3 (в) та програми RasWin Molecular Graphics (г – зв'язки, д – остов). 3WJ характеризується 20 комплементарними парами нуклеотидів, які показано відрізками з крапкою (9-6-5 комплементарних пар у напрямку 5'→3' для трьох дуплексів відповідно). На вставці показано схему 3WJ-структури, яка складається з трьох дуплексів – двох шпильок та стебла

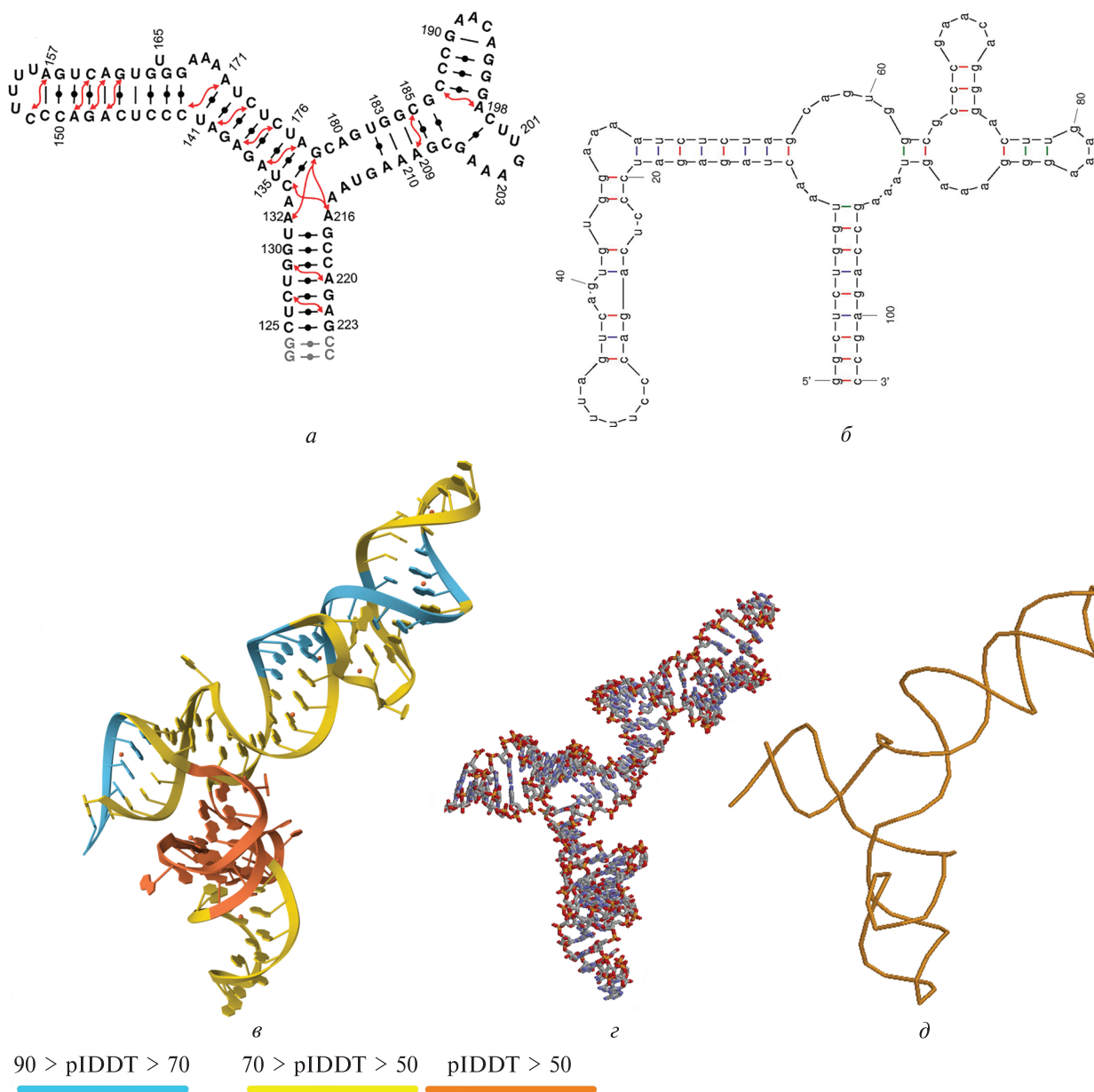


Рис. 6. 3WJ структура (довжиною 103 н.), що містить сайт зв'язування праймера, з якого відбувається ініціація зворотної транскрипції ВІЛ-1 (Song et al, 2021; PDB: 7lva) (а), яку розраховано за допомогою програми Mfold за наступних параметрів: примусове утворення комплементарних пар: F 12 56 8, F 20 45 2, F 65 75 3 (б). 3D модель побудовано за допомогою AlphaFold 3 з візуалізацією AlphaFold 3 (в) та програми RasWin Molecular Graphics (г – зв'язки, д – остов)

підтверджено побудовою 3D моделі за допомогою AlphaFold 3, незважаючи на те, що зазначена 3WJ може бути сформованою лише за умов примусового утворення комплементарних пар та є енергетично відносно не вигідною (рис. 6). Вільна енергія ΔG цього 3WJ-

мотиву дорівнює – 28,8 ккал/моль, тоді як ΔG більш вигідної вторинної структури, яка не утворює 3WJ для цього фрагмента РНК ВІЛ-1, дорівнює – 37,8 ккал/моль.

Недавно 43 послідовності гена *gag* депоновано в базу даних GenBank (Bao et al.,

2020). Фрагмент гена *gag* цих 43 ізолятів СВ ВРХ на 40 н. перекривається з нуклеотидною послідовністю визначеної нами консервативної 3WJ-структури (довжиною 73 н., табл. 3).

Додатково нами проведено аналіз впливу числа пасажів лінії клітин нирки ВРХ, що перевивається, MDBK (Madin Darby Bovine Kidney), інфікованих штамом Riems СВ ВРХ (Bao et al., 2020), на конформацію знайденої консервативної 3WJ-структури. Для дослідження на додаток до 37 ізолятів СВ ВРХ з повним геномом нами використано 43 послідовності гена *gag* довжиною від 1620 н. до 1665 н., що отримано з провірусної ДНК 43 ізолятів СВ ВРХ, яку виділено після серії пасажів лінії клітин MDBK (Bao et al., 2020).

Виконаний нами аналіз одонуклеотидних поліморфізмів, які виникли після проведення серії пасажів лінії клітин MDBK, інфікованих штамом Riems СВ ВРХ (Bao et al., 2020), показав, що знайдена 3WJ-структура зберігається, що вказує на її значущість для функціонування СВ ВРХ. Визначений консервативний структурний 3WJ-мотив в РНК СВ ВРХ є енергетично вигіднішим порівняно з іншими потенційними структурами 3WJ для цієї послідовності, оскільки він був розрахований без примусового утворення комплексарних пар у трьох дуплексах, що утворюють цю вторинну структуру. Представлені дані підтверджують висновок про те, що СВ ВРХ є унікальним та нестандартним вірусом, а молекулярна структура гена *gag* через наявність 10 неканонічних структур (9 G4s та 3WJ) значно відрізняється порівняно з такою інших ретровірусів (Lindemann et al., 2021). Подібно до тих арбовірусів, геном яких, як і геном СВ ВРХ, представлено позитивною одноланцюговою РНК (Nicoletto et al., 2023), спумавірус ВРХ характеризується значною кількістю потенційних G4s, які локалізовано в кодуючих послідовностях та НТО (UTRs).

Автори висловлюють щире подяку д-ру J. Chariker (Університет Луїсвілля, США; University of Louisville, USA) за корисні коментарі, зауваження та пропозиції щодо визначення G-квадруплексів та рецензенту за конструктивні зауваження, які направлено на істотне покращення статті.

Дотримання етичних стандартів. Ця стаття не містить будь-яких досліджень з використанням людей і тварин як об'єктів.

Конфлікт інтересів. Автори заявляють, що не мають конфлікту інтересів.

Фінансування. Роботу частково підтримано грантом 34.02.01.01Ф Національної академії аграрних наук України.

NON-CANONICAL STRUCTURES IN THE GENOME OF BOVINE FOAMY VIRUS

O.Yu. Limanskaya, O.K. Balak, A.P. Limanskii

National Scientific Center «Institute of Experimental and Clinical Veterinary Medicine»,
83 G. Skovorody St., Kharkiv, 61023, Ukraine
Kharkiv National Medical University,
4 Nauka Ave., Kharkiv, 61022, Ukraine
Institute of Physiological Active Compounds,
58 Nauka Ave., Kharkiv, 61072, Ukraine
E-mail: olgaliman@ukr.net

Bioinformatics methods have been used to identify putative perfect G-quadruplexes (G4s) and three-way junctions (3WJs) in bovine foamy virus (BFV) genome. Artificial intelligence (AI) AlphaFold 3 was used to confirm putative G4s and 3WJs by building 3D models of these non-canonical structures. G4s are secondary structures formed by G-rich sequences. Multihelical 3WJs formed by three duplexes connected at the binding point and G4s are considered as alternative structures in DNA and RNA that differ from the classical double-stranded B-DNA. In the present paper, the localization map of putative conservative intramolecular G4s formed by two G-tetrads on the BFV genome was created. 7 putative conservative G-quadruplexes in the sense strand of BFV proviral DNA and 22 G4s in the antisense strand formed by two G-tetrads with G-score from 32 to 36 were found by the multiple alignment of 37 BFV isolates with complete genome. The density of G4s was 0.6 G4/kb for the sense strand of the BFV proviral DNA, while it was 1.8 G4/kb for the antisense strand. One conservative 3WJ motif with length of 73 bp with 100 % homology localized in the 5'-untranslated region and partially on the 5'-end of the *gag* gene was found for a set of 37 BFV isolates. The 3WJ structure in BFV RNA is stabilized by 20 complementary bp with a free energy ΔG of -19.8 kcal/mol. The significance of this structure for BFV functioning has been proven. The use of AI AlphaFold 3 to build 3D models of theoretically determined perfect G4s and 3WJs in the BFV genome allowed us to reliably determine putative alternative structures in nucleic acids.

СПИСОК ЛІТЕРАТУРИ

- Abramson, J., Adler, J., Dunger, J., et al., Accurate structure prediction of biomolecular interactions with AlphaFold 3, *Nature*, 2024, vol. 630, pp. 493–500. <https://doi.org/10.1038/s41586-024-07487-w>
- Agarwal, T., Roy, S., Kumar, S., et al., In the sense of transcription regulation by G-quadruplexes: asymmetric effects in sense and antisense strands, *Biochemistry*, 2014, vol. 53, pp. 3711–3718. <https://doi.org/10.1021/bi401451q>
- Bao, Q., Hotz-Wagenblatt, A., Betts, M.J., et al., Shared and cell type-specific adaptation strategies of Gag and Env yield high titer bovine foamy virus variants, *Infect. Genet. Evol.*, 2020, vol. 82, pp. 104287. <https://doi.org/10.1016/j.meegid.2020.104287>
- Bedrat, A., Lacroix, L., and Mergny, J.L., Re-evaluation of G-quadruplex propensity with G4Hunter, *Nucleic Acids Res.*, 2016, vol. 44, pp. 1746–1759. <https://doi.org/10.1093/nar/gkw006>
- Belotserkovskii, B.P., Neil, A.J., Saleh, S.S., et al., Transcription blockage by homopurine DNA sequences: role of sequence composition and single-strand breaks, *Nucleic Acids Res.*, vol. 201341, pp. 1817–1828. <https://doi.org/10.1093/nar/gks1333>
- Brazda V., Kolomaznik J., Lysek J et al. (2019) G4Hunter web application: a web server for G-quadruplex prediction. *Bioinformatics* 35:3493–3495. <https://doi.org/10.1093/bioinformatics/btz087>
- De Nicola, B., Lech, C.J., Heddi, B., et al., Structure and possible function of a G-quadruplex in the long terminal repeat of the proviral HIV-1 genome, *Nucleic Acids Res.*, 2016, vol. 44, pp. 6442–6451. <https://doi.org/10.1093/nar/gkw432>
- Evans, L., Kotar, A., Valentini, M., et al., Identification and characterisation of G-quadruplex DNA-forming sequences in the *Pseudomonas aeruginosa* genome, *RSC Chem. Biol.*, 2023, vol. 4, pp. 94–100. <https://doi.org/10.1039/d2cb00205a>
- Fay, M.M., Lyons, S.M., and Ivanov, P., RNA G-quadruplexes in biology: principles and molecular mechanisms, *J. Mol. Biol.*, 2017, vol. 429, pp. 2127–2147. <https://doi.org/10.1016/j.jmb.2017.05.017>
- Frasson, I., Nadai, M., and Richter, S.N., Conserved G-quadruplexes regulate the immediate early promoters of human alphaherpesviruses, *Molecules*, vol. 201924, pp. 2375. <https://doi.org/10.3390/molecules24132375>
- Guo, P., Erickson, S., and Anderson, D., A small viral RNA is required for in vitro packaging of bacteriophage phi 29 DNA, *Science*, 1987, vol. 236, pp. 690–694. <https://doi.org/10.1126/science.3107124>
- Hall, T.A., BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT, *Nucleic Acids Symp. Ser.*, 1999, vol. 41, pp. 95–98
- Hamann, M.V., Lindemann, D., Foamy virus protein-nucleic acid interactions during particle morphogenesis, *Viruses*, 2016, vol. 8, pp. 243. <https://doi.org/10.3390/v8090243>
- Hechler, T., Materniak, M., Kehl, T. et al., Complete genome sequences of two novel European clade bovine foamy viruses from Germany and Poland, *J. Virol*, 2012, vol. 86, pp. 10905–10906. <https://doi.org/10.1128/JVI.01875-12>
- Heng, X., Herrera, A.P., Song, Z., et al., Retroviral PBS-segment sequence and structure: Orchestrating early and late replication events, *Retrovirology*, 2024, vol. 21, pp. 12. <https://doi.org/10.1186/s12977-024-00646-x>
- Higgins, D., Thompson, J., Gibson, T. et al., 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, pp. 4673–4680. <https://doi.org/10.1093/nar/22.22.4673>
- Jana, J., Weisz, K., Thermodynamic stability of G-quadruplexes: impact of sequence and environment, *ChemBioChem* 2021, vol. 22, pp. 2848–2856. <https://doi.org/10.1002/cbic.202100127>
- Kardmas, J.L., Ravin, A.J., and Leontis, N.B., Relative stabilities of DNA three-way, four-way and five-way junctions (multi-helix junction loops): unpaired nucleotides can be stabilizing or destabilizing, *Nucleic Acids Res.*, 1995, vol. 23, pp. 2212–2222. <https://doi.org/10.1093/nar/23.12.2212>
- Kikin, O., D'Antonio, L., Bagga, P.S., QGRS Mapper: a web-based server for predicting G-quadruplexes in nucleotide sequences, *Nucleic Acids Res*, 2006, vol. 34, pp. W676–W82. <https://doi.org/10.1093/nar/gkl253>
- Koirala, D., Shao, Y., Koldobskaya, Y., et al., A conserved RNA structural motif for organizing topology within picornaviral internal ribosome entry sites, *Nat. Commun.*, 2019, vol. 10, pp. 3629. <https://doi.org/10.1038/s41467-019-11585-z>
- Kumar, S., Stecher, G., Suleski, M., Sanderford, M., Sharma, S., and Tamura, K., MEGA12: Molecular Evolutionary Genetics Analysis Version 12 for adaptive and green computing. *Mol. Biol. Evolut.*, 2024, vol. 41, pp. 1–9. <https://doi.org/10.1093/molbev/msae263>
- Lindemann, D., Hütter, S., Wei, G., et al., The unique, the known, and the unknown of spumaretrovirus, *Viruses*, 2021, vol. 13, pp. 105. <https://doi.org/10.3390/v13010105>
- Lombardi, E.P., Londono-Vallejo, A., A guide to computational methods for G-quadruplex prediction, *Nucleic Acids Res.*, 2020, vol. 48, pp. 1–15. <https://doi.org/10.1093/nar/gkz1097>
- Lu, H., Li, S., Chen, J., et al., Metal ions modulate the conformation and stability of a G-quadruplex with or without a small-molecular ligand, *Metallomics*,

- 2015, vol. 7, pp. 1508–1514. <https://doi.org/10.1039/c5mt00188a>
- Mekata, H., Okagawa, T., Konnai, S., et al., Molecular epidemiology and whole-genome analysis of bovine foamy virus in Japan, *Viruses* 2021, vol. 13, pp. 1017. <https://doi.org/10.3390/v13061017>
- Menendez, C., Frees, S., Bagga, P.S., QGRS-H Predictor: a web server for predicting homologous quadruplex forming G-rich sequence motifs in nucleotide sequences. *Nucleic Acids Res.*, 2012, vol. 40 (Web Server issue), pp. W96–W103. <https://doi.org/10.1093/nar/gks422>
- Monsen, R.C., Trent, J.O., and Chaires, J.B., G-quadruplex DNA: A longer story, *Acc. Chem. Res.*, 2022, vol. 55, pp. 3242–3252. <https://doi.org/10.1021/acs.accounts.2c00519>
- Nicoletto, G., Richter, S.N., Frasson, I., Presence, location and conservation of putative G-quadruplex forming sequences in arboviruses infecting humans, *Int. J. Mol. Sci.*, 2023, vol. 24, pp. 9523. <https://doi.org/10.3390/ijms24119523>
- Ojha, M., Vogt, J., Das, N.K., et al., Structure of saguaro cactus virus 3' translational enhancer mimics 5' cap for eIF4E binding, *Proc. Natl. Acad. Sci. USA*, 2024, vol. 121, pp. e2313677121. <https://doi.org/10.1073/pnas.2313677121>
- Ouellet, J., Melcher, S., Iqbal, A., et al., Structure of the three-way helical junction of the hepatitis C virus IRES element, *RNA*, 2010, vol. 16, pp. 1597–1609. <https://doi.org/10.1261/rna.2158410>
- Piazza, A., Cui, X., Adrian, M. et al., Non-Canonical G-quadruplexes cause the hCEB1 minisatellite instability in *Saccharomyces cerevisiae*, *eLife*, 2017, vol. 6, pp. e26884. <https://doi.org/10.7554/eLife.26884>
- Piekna-Przybylska, D., Sullivan, M.A., Sharma, G., et al., U3 region in the HIV-1 genome adopts a G-quadruplex structure in its RNA and DNA sequence, *Biochemistry*, 2014, vol. 53, pp. 2581–2593. <https://doi.org/10.1021/bi4016692>
- Pinto-Santini, D.M., Stenbak, C.R., and Linial, M.L., Foamy virus zoonotic infections, *Retrovirology*, 2017, vol. 14, pp. 55. <https://doi.org/10.1186/s12977-017-0379-9>
- Rethwilm, A., Bodem, J., Evolution of foamy viruses: the most ancient of all retroviruses. *Viruses* 2013, vol. 5, pp. 2349–2374. <https://doi.org/10.3390/v5102349>
- Song, Z., Gremminger, T., Singh, G., et al., The three-way junction structure of the HIV-1 PBS-segment binds host enzyme important for viral infectivity, *Nucleic Acids Res.*, 2021, vol. 49, pp. 5925–5942. <https://doi.org/10.1093/nar/gkab342>
- Song, K., Li, B., Li, H., et al., The characterization of G-quadruplexes in Tobacco genome and their function under abiotic stress, *Int. J. Mol. Sci.*, 2024, vol. 25, pp. 4331. <https://doi.org/10.3390/ijms25084331>
- Stefos, G.C., Theodorou, G., and Politis, I., Genomic landscape, polymorphism and possible LINE-associated delivery of G-quadruplex motifs in the bovine genes, *Genomics*, 2022, vol. 114, pp. 110272. <https://doi.org/10.1016/j.ygeno.2022.110272>
- Varshney, D., Spiegel, J., Zyner, K., et al., The regulation and functions of DNA and RNA G-quadruplexes, *Nat. Rev. Mol. Cell Biol.*, 2020, vol. 21, pp. 459–474. <https://doi.org/10.1038/s41580-020-0236-x>
- Wang, J., Huang, H., Zhao, K., et al., G-quadruplex in hepatitis B virus pregenomic RNA promotes its translation, *J. Biol. Chem.*, 2023, vol. 299, pp. 105151. <https://doi.org/10.1016/j.jbc.2023.105151>
- Wang, S.R., Zhang, Q.Y., Wang, J.Q., et al., Chemical targeting of a G-quadruplex RNA in the Ebola virus L gene, *Cell Chem. Biol.*, 2016, vol. 23, pp. 1113–1122. <https://doi.org/10.1016/j.chembiol.2016.07.01>
- Whisnant, A.W., Kehl, T., Bao, Q., et al., Identification of novel, highly expressed retroviral microRNAs in cells infected by bovine foamy virus. *J Virol* 2014, vol. 88, pp. 4679–4686. <https://doi.org/10.1128/JVI.03587-13>
- Wolfe, A.L., Singh, K., Zhong, Y., et al., RNA G-quadruplexes cause eIF4A-dependent oncogene translation in cancer, *Nature*, 2014, vol. 513, pp. 65–70. <https://doi.org/10.1038/nature13485>
- Wu, B., Girard, F., van Buuren, B., et al., Global structure of a DNA three-way junction by solution NMR: towards prediction of 3H fold, *Nucleic Acids Res.*, 2004, vol. 32, pp. 3228–3239. <https://doi.org/10.1093/nar/gkh645>
- Wu, Y.F., Tan, J., Su, Y., et al., Transcription factor API modulates the internal promoter activity of bovine foamy virus, *Virus Res.*, 2010, vol. 147, pp. 139–144. <https://doi.org/10.1016/j.virusres.2009.10.005>
- Xu, X.J., and Chen, S.J., A method to predict the structure and stability of RNA/RNA complexes. *Methods Mol. Biol.*, 2016, vol. 1490, pp. 63–72. https://doi.org/10.1007/978-1-4939-6433-8_5
- Zhang, H., Endrizzi, J.A., Shu, Y., et al., Crystal structure of 3WJ core revealing divalent ion-promoted thermostability and assembly of the Phi29 hexameric motor pRNA, *RNA*, vol. 19, pp. 1226–1237. <https://doi.org/10.1261/rna.037077.112>
- Zuker, M., Mfold web server for nucleic acid folding and hybridization prediction, *Nucleic Acids Res.*, 2003, vol. 31, pp. 3406–3415. <https://doi.org/10.1093/nar/kg595>

Надійшла в редакцію 11.02.2025
Після доопрацювання 03.05.2025
Прийнята до друку 18.09.2025