

КОНСЕРВАТИВНІ НУКЛЕОТИДНІ МОТИВИ В ПРОМОТОРАХ ГЕНОМНИХ РНК ТОБАМОВІРУСІВ

А. М. Кириченко, І. С. Щербатенко

*Інститут мікробіології і вірусології ім. Д.К. Заболотного НАН України,
вул. Академіка Заболотного, 154, Київ, 03143, Україна
e-mail: kirangel.07@meta.ua*

Мета. Реплікація тобамовірусів, як і інших вірусів рослин, що мають одноланцюгову смислову геномну РНК, включає синтез позитивних і негативних ланцюгів геномних РНК, а також транскрипцію двох субгеномних РНК. Первинні нуклеотидні послідовності та вторинні структурні елементи, необхідні для ініціації транскрипції субгеномних РНК, були охарактеризовані у багатьох видів вірусів, однак компоненти промоторів синтезу геномних вірусних РНК все ще залишаються майже недослідженими. Зважаючи на фрагментарність експериментальних даних і суперечливість уявлень щодо ініціації синтезу геномних вірусних РНК, метою нашої роботи було проведення порівняльного комп'ютерного аналізу консервативних компонентів у геномних промоторах тобамовірусів. **Методи.** Геномні сиквенси 17-ти штамів ВТМ та 21 штаму 18-ти інших видів тобамовірусів були завантажені з ГенБанку (<http://www.ncbi.nlm.nih.gov/>). Пошук і порівняльний аналіз консервативних компонентів промоторів проводили за консенсусною послідовністю ділянки ICR2 промотору генів tRNA та одноланцюговою розгорткою стеблової петлі D1L3S3 штаму ВТМ-L, використовуючи власні вузькоспеціалізовані комп'ютерні програми. **Результати.** Встановлено, що 3'-кінцеві ділянки геномних РНК 38-ми досліджених штамів тобамовірусів містять висококонсервативні нуклеотидні мотиви, подібні ICR2 та D1L3S3. У промоторах реплікації геномних (-)РНК ці мотиви локалізовані перед і поблизу стартових точок синтезу (-)ланцюгів, містять загальну паліндромну консенсусну послідовність – TTCGAA, оточену 2 – 4 парами комплементарних нуклеотидів і мають лише компенсаторні нуклеотидні заміни, за яких зберігається вторинна структура консервативних мотивів. Поблизу стартових точок синтезу (+) ланцюгів РНК, локалізованих на 3'-кінцях (-)ланцюгів, нуклеотидні ділянки, подібні ICR2 чи D1L3S3, а також інші консервативні мотиви зовсім не виявляються. За результатами аналізу можливих причин цього явища нами запропоновано гіпотетичні моделі ініціації синтезу геномних і субгеномних РНК тобамовірусів з використанням дволанцюгових і циркулярних матриць. **Висновки.** Запропоновані гіпотетичні моделі добре узгоджуються з отриманими результатами і даними літератури, зокрема: 1) наявність подібних консервативних мотивів і вторинних структур у різних типах промоторів – субгеномних, геномних, вірусних, еукаріотичних; 2) здатність вірусних РНК-полімераз до використання дволанцюгових і циркулярних матриць; 3) можливість ініціації синтезу протилежних ланцюгів РНК одним паліндромним сигнальним мотивом; 4) локалізація АТ-багатих нуклеотидних ділянок після стартових точок транскрипції та реплікації РНК, а GC-багатих ділянок – перед ними; 5) підвищення реплікації вірусів за зв'язування 5'- і 3'-кінців геномних РНК.

Ключові слова: Тобамовіруси, консервативні мотиви геномних промоторів, первинна і вторинна структура консервативних мотивів, гіпотетичні моделі синтезу геномних і субгеномних РНК.

На сьогодні переконливо показано, що циркуляція вірусів в агроценозах, перебіг інфекцій, симптоми враження та ступені стійкості рослин детермінуються тонко збалансованою взаємодією численних вірусних і клітинних компонентів [1, 2]. До найважливіших вірусних регуляторів інфекційного процесу, патогенезу і захисних реакцій рослин відносяться регулятори експресії вірусних генів: контексти стартових і термінальних кодонів транскрипції та трансляції; супресори посттранскрипційного мовчання генів; реплікаційні (геномні) і транскрипційні (субгеномні) промотори та енхансери [3, 4].

Структурно-функціональна організація промоторів субгеномних РНК тобамовірусів залишається маловивченою у порівнянні з іншими вірусними рослин, а експериментальні дані щодо властивостей їх геномних промоторів дуже фрагментарні і суперечливі. Вважається, що позитивні і негативні ланцюги вірусної РНК синтезуються одним ферментом – РНК-залежною РНК-полімеразою, проте однакові чи подібні компоненти в промоторах протилежних ланцюгів не виявляються [5]. У той же час встановлено наявність подібних нуклеотидних мотивів у промоторах (+) ланцюгів субгеномних і (-) ланцюгів геномних вірусних РНК [6]. Зважаючи на суперечливість наведених даних, наша робота була спрямована на комп'ютерний пошук консервативних нуклеотидних мотивів і вторинних структур у геномних промоторах тобамовірусів, подібних знайденим нами [7] в 76-ти субгеномних промоторах двох генів 38-ми вірусів [7].

Матеріали і методи. В роботі було використано геномні сиквенси 17-ти штамів ВТМ та 21 ізоляту/штаму 19-ти інших видів тобамовірусів, отриманих з ГенБанку (<http://www.ncbi.nlm.nih.gov/>). Оскільки об'єкти дослідження мають різні назви: види, ізоляти, штами, ізоляти штамів чи видів з певних рослин чи країн тощо, всі дослідні віруси ми називаємо штамми і використовуємо зручні в роботі їх короткі назви. Наведений перелік штамів включає коротку робочу назву, повну назву та номер доступу в ГенБанку.

Штами ВТМ: **u1**, Звичайний штам U1, J02415; **x1**, Повної назви не знайдено, V01408; **kc**, Корейський штам NC 82, X68110; **im**, ІМ ізолят з бальзаміну, AV369276; **pet**, Петунієвий штам pet, AV369275; **pt**, Петунієвий штам pet-TW, EF392659; **ch**, Китайський штам, AF165190; **ch1**, Китайський ізолят Fujian, AF395127; **ch2**, Китайський ізолят TMV-017, AF395128; **ch3**, Китайський ізолят TMV-152, AF395129; **tsL**, Томатний штам L (BTM-L), X02144; **chm**, Китайський мутант штаму BTM-L, AF155507.2; **k1**, Казахський штам K1, AJ243571; **k2**, Казахський штам K2, Z92909; **b9**, Ізолят B935A з рослин *Vicia faba*, AJ011933; **cr**, BTM із хрестоцвітих, Z29370; **ob**, Штам BTM-Ob, L11665;

Штами 19-ти видів тобамовірусів: **1**, Вірус м'якої зеленої мозаїки тютюну, M34077; **2**, Вірус зеленої плямистої мозаїки огірка, D12505; **3**, Вірус м'якої плямистості перцю, M81413; **4**, Вірус мозаїки подорожника, ізолят Shanghai, AF254924; **5**, Вірус мозаїки томатів, AF332868; **6**, Тобамовірус хрестоцвітих, штам Wasabi, AB017503; **7**, Вірус м'якої плямистості плодів огірка, AF321057; **8**, Вірус кільцевої плямистості одонтоглосума, X82130; **9**, Вірус посвітління жилок турнепсу, U03387; **10**, Вірус перцю

обуда, D13438; **11**, Вірус зеленої плямистої мозаїки цукіні, AJ295949; **12**, Вірус зеленої плямистої мозаїки кіурі, AJ295948; **13**, Сінгапурський латентний вірус гібіскусу, AF395898; **14**, Вірус м'якої плямистості бругмансії, AM398436; **15**, Тобамовірус хрестоцвітих (некласифікований), AY318866; **16**, Вірус мозаїки китайської ріпи (Youcai), EU571218; **17**, Вірус мозаїки йокаї, U30944; **18**, Вірус мозаїки подорожника, ізолят *Impatiens*, DQ223770; **19**, Вірус плямистості болгарського перцю, DQ355023; **20**, Вірус мозаїки маракуї (пасифлори), DQ356949; **21**, Вірус м'якої плямистості кактуса, EU043335.

Комп'ютерний пошук консервативних нуклеотидних мотивів та вторинних структур проводили в повногеномних сиквенсах вірусних РНК та в промоторних ділянках довжиною 190 нуклеотидів від 5'- і 3'-кінців геномів.

Для порівняльного комп'ютерного аналізу консервативних елементів у промоторних ділянках вірусних РНК використовували власні вузько-спеціалізовані програми (утиліти), написані для конкретних задач дослідження, зокрема, вирізання з геномних сиквенсів нуклеотидних ділянок, подібних за довжиною, позиціями локалізації та кількістю збіжних нуклеотидів; пошук найбільших кластерів позиційно-зв'язаних нуклеотидних сайтів, які містяться в найбільшій кількості дослідних сиквенсів; пошук вторинних структур за їх лінійними розгортками та нуклеотидних мотивів за їх консенсусними послідовностями.

Встановлення консенсусних послідовностей і пошук подібних нуклеотидних мотивів проводили з використанням кодів нуклеотидів, рекомендованих Міжнародним біохімічним союзом (IUB): **A**, **T** (**U**), **G**, **C** – стандартні нуклеотиди; **R** – пуринові нуклеотиди (**puRine**, **A/G**); **Y** – піримідинові (**pYrimidine**, **C/T**); **S** – сильні (**Strong**, **G/C**); **W** – слабкі (**Weak**, **A/T**); **M** – аміновмісні (**aMine**, **A/C**); **K** – кетоновмісні (**Ketone**, **G/T**).

Геномні сиквенси та нуклеотидні ділянки штамів ВТМ у тексті подані двома-трьома символами, починаючи з латинської літери, а сиквенси та нуклеотидні сайти штамів інших видів тобамовірусів – їх порядковими номерами відповідно до наведеного переліку штамів. Позиції локалізації консервативних мотивів та вторинних структур, виявлених на плюс- і мінус-ланцюгах РНК, подані як порядкові номери їх початкових нуклеотидів відносно 5'- або 3'-кінців вірусних геномів.

Результати. *Консервативні сайти промоторів синтезу геномних (-)РНК.* Зважаючи на те, що геномна (-)РНК синтезується РНК-залежною РНК-полімеразою на матриці (+)РНК, нами проведено пошук у геномних промоторах нуклеотидного сайту TCGTT, який є комплементарним до консервативного мотиву субгеномних промоторів тобамовірусів.

Встановлено, що в промоторних ділянках 17-ти штамів ВТМ сайт TCGTT зустрічається лише один раз. У відповідних ділянках 21 штаму інших тобамовірусів він зустрічається 2 рази, тоді як у повних геномах штамів ВТМ його частота сягає 126, а в геномах інших штамів тобамовірусів – 161 (табл. 1). Отже, консенсусна послідовність субгеномних промоторів (TCGTT) не міститься в промоторних ділянках синтезу геномної (-)РНК.

Аналіз частоти сайтів, подібних до TCGTT, показав, що найдовшою послідовністю, загальною для промоторів синтезу (-)РНК тобамовірусів є TTCGAAT. Вона міститься у 3'-кінцевих 190-нуклеотидних ділянках (+)РНК всіх 38-ми досліджених вірусів (табл. 1) і дуже рідко зустрічається поза цими ділянками (у одного штаму ВТМ і п'яти штамів інших тобамовірусів). Сайт GGTTCGAATCC міститься в промоторних ділянках синтезу (-)РНК 13-ти штамів ВТМ і 15-ти штамів інших тобамовірусів, однак його варіант із заміною лише одного нуклеотиду – АТTCGAATCC зовсім відсутній як у промоторах, так і повних геномах, а подібні варіанти GATTCGAATTC та GGGATTCGAATCCC зустрічаються в промоторах лише чотирьох штамів ВТМ та трьох штамів інших видів тобамовірусів.

Отримані результати свідчать про подібність субгеномних і геномних промоторів тобамовірусів, а також геномних промоторів вірусів і еукаріотів, що проявляється як наявність у них дуже схожих або однакових нуклеотидних мотивів. Зокрема, всі нуклеотидні сайти, виявлені на 3'-кінцях (+)ланцюгів геномних РНК тобамовірусів, у ділянках промоторів синтезу (-)ланцюгів містять загальний мотив TTCGAAT. Мотиви GGTTCGAAT та GATTCGAAT, знайдені в ділянках геномних промоторів, є комплементарними раніше виявленим нами мотивам субгеномних промоторів, які мають загальну консенсусну послідовність – GRITTCGWWW з двома найчастішими варіантами WWW – ААТ чи ААА. Мотив GGGATTCGAATCCC являє собою стеблову петлю D1S3L3, яка міститься на 3'-кінці геному ВТМ-L і є важливою для синтезу (+)ланцюга геномної РНК цього штаму [8], а 11-нуклеотидний мотив GGTTCGAATCC відповідає консенсусній послідовності ділянки ICR2 промотору генів tRNA (GGTTCGAnTCC), яка знаходиться також у T psi C петлі tРНК і на 5'-кінці (-)РНК кукум-ітмовірусів та (+) і (-)РНК бромовірусів [9].

Таблиця 1
Консервативні нуклеотидні сайти в геномних РНК тобамовірусів

Консервативні сайти ¹⁾	Штами ВТМ		Штами 18-ти видів тобамовірусів	
	Кількість штамів, що містять сайт	Кількість сайтів у геномах	Кількість штамів, що містять сайт	Кількість сайтів у геномах
TCGTT	1	126	2	161
TCGT	1	306	5	429
TCGA	17	513	21	619
TCGAA	17	179	21	208
TCGAAT	17	42	21	47
TTCGAAT	17	18	21	26
GTTCGAAT	13	14	18	20
GTTCGAATC	13	14	18	18
GGTTCGAATC	13	14	15	15
GGTTCGAATCC	13	13	15	15
GATTCGAATCC	0	0	0	0
GATTCGAATT	4	4	3	3
GGGATTCGAATCCC	4	4	3	3

¹⁾ Сірим фоном виділені: мотиви геномних промоторів, комплементарні консервативним мотивам субгеномних промоторів генів ТБ (GGTTCGAAT) і БО (GATTCGAAT), консенсусна послідовність ділянки ICR2 промотору генів tRNA (GGTTCGAATCC) та стеблова петля D1S3L3 штаму ВТМ-U1 – GGGATTCGAATCCC.

Зважаючи на наведені дані про подібність геномних і субгеномних промоторів вірусів, а також промоторів вірусів і еукаріотів, ми провели пошук схожих консервативних мотивів та/або вторинних структур у промоторах синтезу (-) і (+)ланцюгів геномних РНК тобамовірусів. В роботі було використано одноланцюгову розгортку стеблової петлі D1L3S3 та модифіковану консенсусну послідовність ділянки ICR2 промотору генів tRNA – GGTTCGAATCC. Модифікація була зумовлена наявністю нуклеотиду А замість довільного нуклеотиду n у восьмій позиції всіх 11-нуклеотидних мотивів, знайдених у геномних промоторах (табл. 1).

Консервативні мотиви в промоторах синтезу (-) і (+)ланцюгів геномних РНК тобамовірусів. Встановлено, що повну збіжність нуклеотидів з розгорткою стеблової петлі D1L3S3 (19/19) мають консервативні мотиви промоторів синтезу (-)РНК семи тобамовірусів – чотирьох штамів ВТМ (tsL, chm, k1, k2), а також вірусів 5, 8 та 19 (табл. 2). Решта тобамовірусів (31 штам з 38-ми досліджених) мають по дві еквівалентні заміни нуклеотидів – А на G і Т на С у позиції 4, за яких не лише зберігається, а ще й підсилюється взаємодія комплементарних нуклеотидів стеблових петель за рахунок трьох водневих зв'язків у пари GC замість двох у AT. По чотири однакові еквівалентні заміни нуклеотидів у близьких позиціях мають консервативні мотиви промоторів двох груп тобамовірусів, які містять три (7, 11, 12) та десять (cr, 6, 9, 17, 18, 1, 4, 14, 15, 16) штамів. Нееквівалентні одонуклеотидні заміни – С або G на А, які спричиняють некомплементарність нуклеотидів у другій позиції стебла петлі D1L3S3, мають лише два із 38-ми штамів тобамовірусів. Ці дані, а також

Таблиця 2

Консервативні 15-нуклеотидні мотиви, локалізовані на (+) і (-) ланцюгах геномних РНК 38-ми штамів тобамовірусів

РНК	Збіг ¹⁾	Мотиви ²⁾		Штами ³⁾	
		1 2 3 4	4 3 2 1		
(+) ланцюг	15	GGGATTCGAATTC CC		tsL, chm, k1, k2, 5, 8, 19	7
	13	. . . g	c	S11, 2, 3, 10, 20, 21	16
	12	. . . g	c . . a .	im	1
		. a . g	c	13	1
	11	. a . g	c . t .	cr, 1, 4, 6, 9, 14, 15, 16, 17, 18	10
. . t g		c a . .	7, 11, 12	3	
(-) ланцюг	12 c	c . . . t	S9, 2, 3, 10, 20, 21	14
		. . . a	c . . . t	tsL, chm, k1, k2, 5, 8	6
	 c	c t . . .	13	1
		. . . a	c . . a .	19	1
	11	t c	c . . . t	im	1
		. t c	a . . . t	7, 11, 12	3
	10	a c	c t . . t	cr, 1, 4, 6, 9, 14, 15, 16, 17, 18	10

¹⁾ Кількість нуклеотидів у мотивах, збіжних з послідовністю стеблової петлі.

²⁾ Нуклеотидна послідовність стеблової петлі D1S3L3 штаму ВТМ-U1 (жирні літери) і відповідні їй консервативні мотиви. Цифрами (1 – 4) позначені комплементарні нуклеотиди, точками – збіжні нуклеотиди, малими літерами – компенсаційні заміни комплементарних нуклеотидів, червоним фоном – некомпенсаційні заміни, що порушують вторинну структуру стеблових петель, сірим фоном – спільний консервативний мотив.

³⁾ Назви та кількість штамів, що мають однаковий мотив. S11 – 11 штамів ВТМ (u1, b9, b9, kc, ch, ch1, ch2, ch3, pet, pt, ob), S9 – штами S11 без pt і ob.

ідентичність неспарених ділянок стеблових петель в промоторах всіх 38-ми досліджених штамів свідчать про важливість як первинної послідовності, так і вторинної структури консервативних мотивів для забезпечення функціональної активності вірусних промоторів. Отже, результати наших досліджень суперечать висновку щодо «більш важливої ролі вторинних структур, а не первинних послідовностей у забезпеченні функціональної активності промоторів» [10] і цілком узгоджуються з даними про залежність різних функцій промоторів від різних первинних та/або вторинних структурних елементів [11, 12, 13, 14].

Пошук консервативних мотивів у промоторах синтезу геномних (+) РНК показав відсутність у них нуклеотидних мотивів, подібних стебловій петлі D1L3S3, а також суттєву відмінність цієї петлі від комплементарної їй послідовності на 3'-кінцях геномних (-)РНК. По-перше, на негативних ланцюгах РНК, на відміну від позитивних, не виявляються сайти, які мають 15, 14 чи 13 нуклеотидів, збіжних з розгорткою петлі D1L3S3. По-друге, за низької збіжності з розгорткою переважна більшість подібних мотивів виявляється не в промоторних ділянках. Зокрема, кількість мотивів, збіжних з розгорткою петлі за 10-ма нуклеотидами, становить 36 у промоторних ділянках і 180 поза ними. Поряд з цим мотиви промоторних ділянок мають численні некомпенсаційні нуклеотидні заміни в стеблах петлі D1L3S3, що порушують її вторинну структуру, а також заміни першого нуклеотиду мотиву TTCGAAT, які не зустрічаються у промоторах синтезу (-)ланцюгів.

Таблиця 3

Консервативні 11-нуклеотидні мотиви, локалізованих на (+) і (-) ланцюгах геномних РНК 38-ми штамів тобамовірусів

РНК	Збіг ¹⁾	Мотиви ²⁾		Штами ³⁾	
		1 2	2 1		
(+) ланцюг	11	GGTTCGAATCC		S13,1,2,3,4,6,9,10,13,14,15,16,17,18,20,21	28
	9	. a t .		tsL, chm, k1, k2, 5, 8, 19	7
		t a		7,11,12	3
(-) ланцюг	9	. a c . .		S12, 2, 3, 10, 20, 21	17
		a a		tsL, chm, k1, k2, 5, 8, 19	7
	8	. a c . t		cr, 1, 4, 6, 9, 13, 14, 15, 16, 17, 18	11
		. a c a .		7,11,12	3

¹⁾ Кількість нуклеотидів у мотивах, збіжних з консенсусною послідовністю ділянки ICR2 промотору генів tRNA.

²⁾ Консенсусна послідовність ділянки ICR2 (жирні літери) і відповідні їй консервативні мотиви. Цифрами (1 – 2) позначені комплементарні нуклеотиди, точками – збіжні нуклеотиди, малими літерами – компенсаційні заміни комплементарних нуклеотидів, червоним фоном – некомпенсаційні заміни, що порушують вторинну структуру стеблових петель, сірим фоном – спільний консервативний мотив.

³⁾ Назви та кількість штамів, що мають однаковий мотив. S13 – 13 штамів BTM: u1, x1, b9, kc, im, ch, ch1, ch2, ch3, pet, pt, ob, cr. S12 – штами S13 без cr.

Наступним етапом роботи був пошук у геномних промоторах консервативних мотивів, подібних консенсусній послідовності 11-ти нуклеотидної ділянки ICR2 промотору генів tRNA – GGTTCGAATCC, яка знайдена у промоторах різних вірусів та еукаріотів [9]. Результати досліджень показали, що схожі на ICR2 11-нуклеотидні мотиви в геномних промоторах синтезу (+)РНК (на 3'-кінцях (-)ланцюгів) не виявляються, однак їх комплементарні послідовності на 5'-кінцях позитивних ланцюгів РНК мають високу консервативність (табл. 3). Так, 28 штамів тобамовірусів на 3'-кінцях (+)РНК містять загальний консервативний мотив, ідентичний ділянці ICR2. Мотиви решти 10-ти штамів відрізняються від послідовності ICR2 за двома нуклеотидними замінами, які не порушують компліментарності 1-го нуклеотиду з 11-тим і 2-го з 12-тим. Отже, всі 3 консервативні мотиви у промоторах синтезу геномних (-)РНК мають подібну структурну організацію – загальний 7-нуклеотидний сайт TTCGAAT, оточений двома парами комплементарних нуклеотидів.

Консервативні мотиви, знайдені в промоторах синтезу (+)ланцюгів РНК суттєво відрізняються від виявлених у геномних промоторах (-)ланцюгів. Вони не відповідають консенсусній послідовності ICR2, а також мають численні нуклеотидні заміни, які порушують як первинну, так і вторинну структуру стеблової петлі D1L3S3. Ці дані свідчать про можливість використання різних сигнально-регуляторних мотивів у промоторах синтезу позитивних і негативних РНК тобамовірусів. З іншого боку, подібна структурно-функціональна організація виявлена у зовсім різних типів промоторів – субгеномних, геномних, вірусних та еукаріотичних [9]. Враховуючи ці дані та отримані результати, ми пропонуємо гіпотетичні моделі синтезу геномних і субгеномних РНК тобамовірусів з використанням первинних і вторинних структур подібних консервативних мотивів.

Обговорення. Паліндромність мотиву TTCGAA наводить на думку про можливість його участі в ініціації синтезу обох ланцюгів геномних РНК, однак це припущення суперечить тому, що сигнально-регуляторні елементи промоторів зазвичай розташовані поблизу стартових точок синтезу РНК, тоді як відстань від стартової точки реплікації (кінця (-)РНК) до TTCGAA перевищує 6000 нуклеотидів (рис. 1). Одним із способів зближення стартових точок з сигналами ініціації синтезу може бути циркуляризація РНК шляхом утворення пар між далекими комплементарними нуклеотидами, яка виявлена у кількох вірусів [15, 16, 17, 18]

На моделі ВТМ як типового представника РНК-геномних вірусів було встановлено, що реплікація цих вірусів включає синтез негативного ланцюга геномної РНК на матриці віріонної (+)РНК, а потім синтез позитивного ланцюга на матриці геномної (-)РНК [19, 20]. Продукти синтезу вірусних РНК у клітинах включають повністю дволанцюгові реплікативні форми (RF), частково дволанцюгові проміжні реплікативні форми (RI), а також одностанцюгові геномні та субгеномні молекули різної довжини [13, 19]. Дані стосовно ролі одно- і дволанцюгових форм у матричному синтезі РНК все ще залишаються дуже фрагментарними і суперечливими [20, 21, 22, 23]. Припущення про можливість транскрипції субгеномних РНК тобамовірусів з використанням дволанцюгових матриць було

висуно в нашій лабораторії [24] на основі виявлення АТ-вмісних ділянок після стартового нуклеотиду транскрипції РНК. Такі ділянки сприяють розплітання дволанцюгових молекул за рахунок слабших водневих зв'язків між комплементарними АТ-парами у порівнянні з GC-парами. Для синтезу геномних (-)РНК на матриці одноланцюгових віріонних (+) РНК розплітання молекул не потрібне, але, несподівано, ми виявили чітко виражені АТ-вмісні ділянки після стартових нуклеотидів синтезу геномних (-)РНК у всіх досліджених 38-ми тобамовірусів (рис. 2).

Враховуючи дані про здатність РНК-полімерази використовувати одно- і дволанцюгові матриці, можливість ініціації синтезу обох ланцюгів РНК одним паліндромним сигнальним мотивом за умов зближення кінців матричних РНК, виявлення такого зближення у різних вірусів; локалізацію GC-багатих нуклеотидних ділянок перед стартовими точками транскрипції та реплікації РНК, а АТ-багатих сайтів – після них, ми пропонуємо гіпотетичні моделі синтезу субгеномних і геномних РНК тобамовірусів, які узгоджуються з наведеними даними (рис. 3).

Геномна (-)РНК, яка використовується лише як компонент дволанцюгової матриці для синтезу геномної та субгеномних (+)РНК, синтезується в дуже обмеженій кількості на матриці одноланцюгової віріонної РНК (рис. 3А). Синтез починається з передостаннього 3'-кінцевого стартового нуклеотиду С. Початок сигналу ініціації синтезу – ААGСТТ знаходиться на відстані 28 нуклеотидів від стартової точки (С). Збереження інтактності 3'-кінцевого фрагмента СССА, важливого для зв'язування

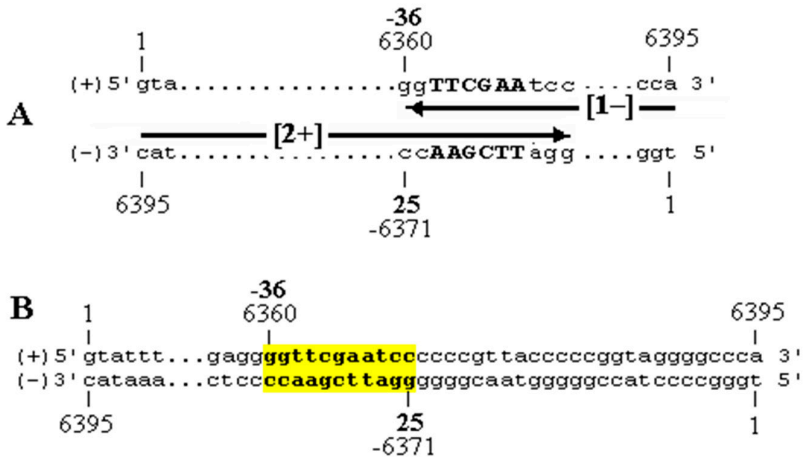


Рис. 1. Локалізація висококонсервативного 11-нуклеотидного мотиву на позитивних і негативних ланцюгах геномної РНК ВТМ.

А – схема синтезу ланцюгів. **В** – локалізація консервативного мотиву у геномі штаму ВТМ-U1. [1-] – початкова позиція (асс) і напрямком (стрілка) синтезу (-) ланцюга геномної РНК на матриці віріонної (+)РНК. [2+] – початкова позиція (cat) і напрямком синтезу (+)ланцюга вірусної РНК на матриці геномної (-)РНК. 6360 – позиція висококонсервативного мотиву на (+)ланцюгу і його відстань (-36) від стартової точки реплікації РНК. 25 – позиція консервативного мотиву на (-)ланцюгу і його відстань (-6360) від стартової точки реплікації. ААGСТТ – паліндромний сайт консервативного мотиву, орієнтований за напрямком синтезу протилежного ланцюга РНК.

вірусної полімерази зі стартовою точкою синтезу РНК, забезпечується шляхом приєднання тимідилату до 5'-кінцевого фрагмента GGG за допомогою нуклеотиділтрансферази хазяїна [25]. Синтезована дволанцюгова РНК може розглядатись як давно відома реплікативна форма RF [20].

Субгеномні РНК капсидного і транспортного білків синтезуються на матриці геномної (-)РНК у складі дволанцюгового комплексу (рис. 5В). Сигнал ініціації транскрипції AAGCAA дуже схожий на такий у промоторах геномних РНК (AAGCTT), ідентичний у обох субгеномних РНК і входить до складу консервативних мотивів промоторів гена транспортно-го (ССААGCAA) і капсидного білка (СТААGCAA). Початок сигналу ініціації транскрипції локалізований за 3 нуклеотиди від її стартової точки. Комплекс дволанцюгової матриці з субгеномними РНК схожий на проміжні реплікативні форми RI, описані раніше [26].

У гіпотетичній моделі синтезу геномної (+)РНК на лінійній дволанцюговій матриці (рис. 5С) сигнал ініціації синтезу локалізований не перед стартовим нуклеотидом, а після нього і до того ж на дуже великій

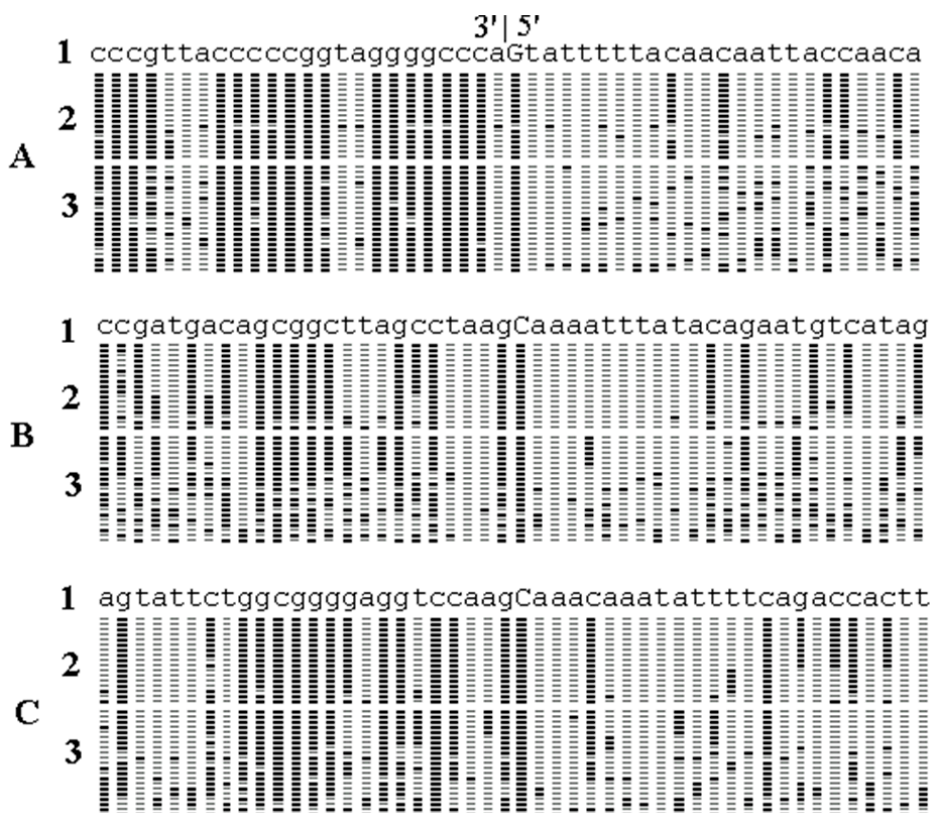


Рис. 2. GC- та АТ-багаті сайти, що оточують стартові кодони реплікації і трансляції РНК тобамовірусів.

А, В, С – 48-нуклеотидні (48-нт) сайти, що містять: 3'-кінцеві та 5'-початкові 24-нт ділянки геномних РНК (А), 24-нт ділянки, локалізовані перед і після стартового кодону транскрипції субгеномних РНК генів БО (В) та ТБ (С). 1 – 3 – 48-нт сайти: штаму ВТМ-U1 (1), 17-ти штамів ВТМ (2), 21 штаму/ізоляту інших 18-ти видів тобамовірусів (3). Темні горизонтальні лінії показують нуклеотиди G і C, світлі лінії – А і Т.

відстані (6351 нуклеотид) у порівнянні з розмірами картованих промоторів тобамовірусів [10]. У моделі синтезу геномної (+)РНК з використанням дволанцюгової кільцевої РНК (рис. 5D) сигнал ініціації синтезу, як і в усіх інших моделях, знаходиться перед стартовою точкою С і поблизу неї. За зближення протилежних кінців обох ланцюгів РНК у циркулярних структурах РНК-залежна РНК-полімераза впізнає сигнал ініціації синтезу РНК – мотив ААGСТТ, оточений комплементарними парами нуклеотидів. Паліндромність цього мотиву забезпечує ідентичність орієнтації сигналу у напрямку синтезу РНК. Суттєву залежність селективності та інтенсивності синтезу (+)- і (-)-ланцюгів РНК від орієнтації сигнальних компонентів у промоторах показано на моделі шпильок Н4+ та Н4– Вірусу зморшкватості ріпи [27], а також на моделі прямої – ТАТАААААА і оберненої – ТТТТТАТА орієнтації канонічного ТАТА боксу в еукаріотичних промоторах [28]. Знайшовши сигнал ініціації синтезу РНК, РНК-полімераза

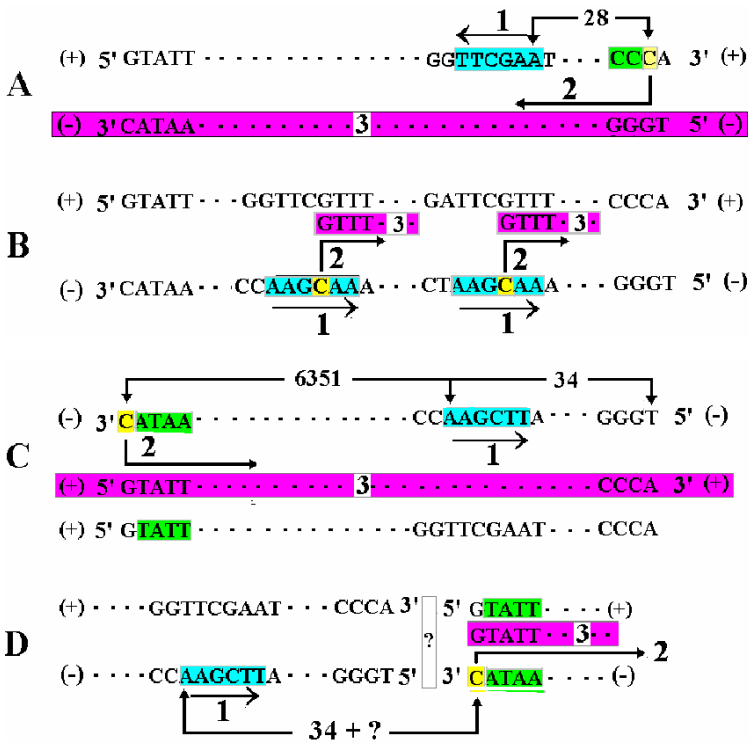


Рис. 3. Гіпотетичні моделі синтезу субгеномних і геномних РНК тобамовірусів
 А – реплікація геномної (-)РНК на матриці віріонної (+)РНК. В – транскрипція субгеномних РНК транспортного і капсидного білків на матриці (-)РНК з використанням дволанцюгової геномної РНК. С, D – реплікація геномної (+) РНК з використанням дволанцюгової лінійної (С) або кільцевої (D) віріонної РНК. 1 - сигнали ініціації синтезу РНК і їх орієнтація в напрямку елонгації: від 5'- до 3'-кінців РНК (стрілки). 2 – стартовий нуклеотид С у матричному ланцюгу РНК і напрямок синтезу протилежного ланцюга (стрілки). 3 – ланцюг РНК, що синтезується на матриці. 28 і 6351 – відстань між стартовим нуклеотидом реплікації РНК (С) і сигналом ініціації синтезу в матричному ланцюгу (ААGСТТ). 34 – відстань між початком сигналу ініціації синтезу (+)ланцюга РНК і стартовою точкою реплікації: 3'-кінцевим нуклеотидом С (-)ланцюга. «?» – невідома відстань між 3'- і 5'-кінцями однакових ланцюгів кільцевої РНК.

сканує матрицю в напрямку орієнтації сигналу (від 5'- до 3'-кінця) у пошуку стартової точки реплікації – першого нуклеотиду С перед ділянкою слабо зв'язаних АТ-багатих комплементарних нуклеотидних пар. У такий спосіб синтез (-)РНК ефективно блокується міцно зв'язаними GC-багатими ділянками і реалізується асиметричний конвеєрний синтез великої кількості головної вірусної РНК – віріонної (+)РНК.

Наведена модель синтезу геномної (+)РНК тобамовірусів з використанням дволанцюгової кільцевої матриці узгоджується з даними про реплікацію РНК Вірусу куцистої карликовості томатів шляхом заміщення новосинтезованою РНК позитивного ланцюга дволанцюгової матриці [29]. При цьому дволанцюгову матрицю можна розглядати як RF, а субгеномні РНК, що синтезуються на цій матриці – як RI [26]. На можливість реплікації вірусних РНК на кільцевих матрицях свідчить широко розповсюджене утворення кільцевих вірусних РНК шляхом взаємодії різноманітних вірусних і клітинних компонентів [16 – 18]. Таким чином, запропоновані гіпотетичні моделі синтезу субгеномних і геномних РНК тобамовірусів узгоджуються з отриманими нами і літературними даними і можуть бути використані для подальшого вивчення природи і механізмів взаємодії вірусних і клітинних компонентів, причетних до реплікації та транскрипції вірусних РНК.

КОНСЕРВАТИВНЫЕ НУКЛЕОТИДНЫЕ МОТИВЫ В ПРОМОТОРАХ ГЕНОМНЫХ РНК ТОБАМОВИРУСОВ

А.Н. Кириченко, И.С. Щербатенко

*Институт микробиологии и вирусологии им. Д.К.Заболотного НАН Украины,
ул. Академика Заболотного, 154, Киев, 03143, Украина*

Резюме

Цель. Репликация тобамовирусов, как и других вирусов растений, имеющих одноцепочечную смысловую геномную РНК, включает синтез позитивных и негативных цепочек геномных РНК, а также транскрипцию двух субгеномных РНК. Первичные нуклеотидные последовательности и вторичные структурные элементы, необходимые для инициации транскрипции субгеномных РНК, были охарактеризованы у многих видов вирусов, однако компоненты промоторов синтеза геномных РНК все еще остаются почти неисследованными. Учитывая фрагментарность экспериментальных данных и противоречивость мнений относительно инициации синтеза геномных вирусных РНК, целью нашей работы было проведение сравнительного компьютерного анализа консервативных компонентов в геномных промоторах тобамовирусов. **Методы.** Геномные сиквенсы 17-ти штаммов ВТМ и 21 штамма других видов тобамовирусов были получены из ГенБанка (<http://www.ncbi.nlm.nih.gov/>). Для поиска и сравнительного анализа консервативных компонентов промоторов использовали консенсусные последовательности участка ICR2 промотора генов tRNA и одноцепочечной развертки стеблевой петли D1L3S3 штамма ВТМ-L, а также собственные узкоспециализированные компьютерные программы. **Результаты.** Установлено, что все 3'-концевые участки геномных РНК 38-ми исследованных штаммов тобамовирусов содержат высококонсервативные нуклеотидные мотивы, подобные ICR2 и

D1L3S3. В промоторах репликации геномных (-)РНК эти мотивы локализируются перед и вблизи стартовых точек синтеза (-)цепочек, содержат общую палиндромную консенсусную последовательность – ТТСГАА, окруженную 2 – 4 парами комплементарных нуклеотидов и имеют только компенсаторные нуклеотидные замены, при которых сохраняется вторичная структура консервативных мотивов. Вблизи стартовых точек синтеза (+)цепочек РНК, локализованных на 3'-концах (-)цепочек, нуклеотидные участки, подобные ICR2 или D1L3S3, а также другие консервативные мотивы совсем не выявляются. По результатам анализа возможных причин этого явления мы предлагаем гипотетические модели синтеза геномных и субгеномных РНК тобамовирусов на двухцепочечных и кольцевых матрицах. **Выводы.** Предложенные модели хорошо согласуются с полученными результатами и данными литературы, в частности: 1) наличие подобных консервативных мотивов и вторичных структур в разных типах промоторов – субгеномных, геномных, вирусных, еукариотических; 2) способность вирусных РНК-полимераз использовать двухцепочечные и кольцевые матрицы; 3) возможность инициации синтеза противоположных цепочек РНК одним палиндромным сигнальным мотивом; 4) локализация АТ-богатых нуклеотидных участков после стартовых точек транскрипции и репликации РНК, а GC-богатых участков – перед ними; 5) увеличение интенсивности репликации вирусов при связывании 5'- и 3'-концов геномных РНК.

Ключевые слова: Тобамовирусы, консервативные мотивы геномных промоторов, первичная и вторичная структура консервативных мотивов, гипотетические модели синтеза геномных и субгеномных РНК.

CONSERVE NUCLEOTIDE MOTIFS WITHIN TOBAMOVIRUS GENOMIC PROMOTERS

A.N. Kyrychenko, I.S. Shcherbatenko

*Zabolotny Institute of Microbiology and Virology, NAS of Ukraine,
154 Acad. Zabolotny Str., Kyiv, 03143, Ukraine*

Summary

Aim. Replication of tobamoviruses, like that of other single-stranded positive sense RNA viruses involves the production of both genomic minus-strand and plus-strand RNAs as well as the transcription of two subgenomic positive-strand RNAs. The sequence elements and secondary structures required for promotion of the subgenomic RNAs have been characterized for many viruses, but only little is known about the promoter elements essential for genomic RNAs synthesis. Taking into consideration the insufficiency of experimental data and contradictoriness of notions regarding the promotion of viral genomic RNA replication, a comparative computational analysis of a conservative components in tobamovirus replicative promoters was the goal of our study. **Methods.** Complete genomic sequences of 38 tobamoviruses were downloaded from the NCBI website (<http://www.ncbi.nlm.nih.gov/>). A computer search and comparative analysis of conservative promoter components was carried out by means of a set of own tightly specialized computer programs using consensus sequence of the ICR2 motif of tRNA genes and the unfolded sequence of D1L3S3 stem-loop of BTM-L strain. **Results.** We show that the 3'-end regions of genomic RNA of all 38 tobamovirus tested contain a highly conserved

ICR2- and D1L3S3-like nucleotide motifs. In replicative promoters of the genomic minus-strand RNA this motifs are localized upstream and hereabouts of the start positions of (-)RNA synthesis, contain the common palindromic consensus sequence – TTCGAA, surrounding by 2 – 4 pairs of complementary nucleotides, and have only compensatory nucleotide exchange, which do not disruption of the complementary base pairing. On the contrary, no ICR2-, D1L3S3-like or others highly conserved nucleotide motifs were found hereabouts of the start point of (+)RNA synthesis localized at the 3'-end of genomic (-) RNA. By analyses of a possible cause of this appearance we propose the hypothetic models for subgenomic and genomic RNA synthesis on double-stranded and circular templates. **Conclusions.** The hypothetic models suggested are well consistent with our results and literature data, in particular: i) presence a similar conservative motifs in different promoter types – subgenomic, genomic, viral and eukaryotic; ii) ability of virus RNA polymerases for using of double stranded and circular templates; iii) possibility to initiate synthesis of opposite RNA strands by one palindromic signaling motif; iv) localization of viral AT-reach sequence regions downstream of replication and transcription start points of virus RNA, and localization of GC-reach sequences upstream of such points; v) enhancing of virus replication through promoting linkage between 5'- and 3'-untranslated regions.

Keywords: Tobamoviruses, conserved motifs in genomic promoters, primary and secondary structures of conserved motifs, hypothetic models for subgenomic and genomic RNA synthesis.

1. Van der Heijden MW, Bol JF. Composition of alphavirus-like replication complexes: involvement of virus and host encoded proteins. *Arch. Virol.* 2002; 147:875–98.
2. Diez J, Ishikawa M, Kaido M, Ahlquist P. Identification and characterization of a host protein required for efficient template selection in viral RNA replication. *Proc. Natl. Acad. Sci. U.S.A.* 2000; 97:3913–18.
3. Ranjith-Kumar CT, Zhang X, Kao CC. Enhancer-Like Activity of a Brome Mosaic Virus RNA Promoter. *J. Virol.* 2003; 77:1830–39.
4. Olsthoorn RCL, Haasnoot PCJ, Bol JF. Similarities and Differences between the Subgenomic and Minus-Strand Promoters of an RNA Plant Virus. *J. Virol.* 2004; 78:4048–53.
5. Ayllon MA, Gowda S, Satyanarayana T, Dawson WO. cis-Acting elements at opposite ends of the Citrus tristeza virus genome differ in initiation and termination of subgenomic RNAs. *Virology.* 2004; 322:41–50
6. Sivakumaran K, Choi SK, Hema M, Kao CC. Requirements for brome mosaic virus subgenomic RNA synthesis in vivo and replicase-core promoter interactions in vitro. *J. Virol.* 2004; 78:6091–101
7. Kyrychenko AN, Shcherbatenko IS. [Conserve nucleotide motifs and secondary structures within tobamovirus subgenomic promoters]. *Microbiol. Zh.* 2019; 81(1):106–23. Ukrainian.
8. Osman TA, Hemenway CL, Buck KW. Role of the 3' tRNA-like structure in tobacco mosaic virus minus-strand RNA synthesis by the viral RNA-dependent RNA polymerase In vitro. *J. Virol.* 2000; 74:11671–80.
9. Marsh LE, Pogue GP, Hall TC. Similarities among plant virus (+) and (-) RNA termini imply a common ancestry with promoters of eukaryotic tRNAs. *Virology.* 1989; 172:415–27.

10. Grdzlishvili VZ, Chapman SN, Dawson WO, Lewandowski DJ. Mapping of the tobacco mosaic virus movement protein and coat protein subgenomic RNA promoters in vivo. *Virology*. 2000; 275:177–92.
11. Tilgner M, Shi PY. Structure and function of the 3' terminal six nucleotides of the west Nile virus genome in viral replication. *J Virol*. 2004; 78:8159–71.
12. Man M, Epel BL. Characterization of regulatory elements within the coat protein (CP) coding region of Tobacco mosaic virus affecting subgenomic transcription and green fluorescent protein expression from the CP subgenomic RNA promoter. *J Gen Virol*. 2004; 85:1727–38.
13. Gunawardene CD, Jaluba K, White KA. Conserved motifs in a tombusvirus polymerase modulate genome replication, subgenomic transcription, and amplification of defective interfering RNAs. *J Virol*. 2015; 89:3236–46.
14. Frolov I, Hardy R, Rice CM. Cis-acting RNA elements at the 5' end of Sindbis virus genome RNA regulate minus- and plus-strand RNA synthesis. *RNA*. 2001 Nov; 7:1638–51.
15. Filomatori CV, Lodeiro MF, Alvarez DE et al. A 5' RNA element promotes dengue virus RNA synthesis on a circular genome. *Genes Dev*. 2006; 20:2238–49.
16. Kumar A, Ray U, Das S. Human La protein interaction with GCAC near the initiator AUG enhances hepatitis C Virus RNA replication by promoting linkage between 5' and 3' untranslated regions. *J Virol*. 2013; 87:6713–26.
17. Villordo SM, Alvarez DE, Gamarnik AV. A balance between circular and linear forms of the dengue virus genome is crucial for viral replication. *RNA*. 2010; 16:2325–35.
18. López-Manríquez E, Vashist S, Ureña L, Goodfellow I. et al. Norovirus genome circularization and efficient replication are facilitated by binding of PCBP2 and hnRNP A1. *J Virol*. 2013; 87:11371–87.
19. Young N, Forney J, Zaitlin M. Tobacco mosaic virus replicase and replicative structures. *J Cell Sci Suppl*. 1987; 7:277–85.
20. Buck KW Comparison of the replication of positive-stranded RNA viruses of plants and animals. *Adv. Virus Res*. 1996; 47:159–251.
21. Panavas T, Stork J, Nagy PD. Use of double-stranded RNA templates by the tombusvirus replicase in vitro: Implications for the mechanism of plus-strand initiation. *Virology*. 2006; 352:110–20.
22. Kovalev N, Pogany J, Nagy PD. Template role of double-stranded RNA in tombusvirus replication. *J Virol*. 2014; 88:5638–51.
23. Osman TA, Olsthoorn RC, Livieratos IC. In vitro template-dependent synthesis of Pepino mosaic virus positive- and negative-strand RNA by its RNA-dependent RNA polymerase. *Virus Res*. 2012; 167:267–72.
24. Gordejchuk OI, Oleshchenko LT, Shcherbatenko IS. [Similar nucleotide blocks in tobamoviral subgenomic promoters]. *Microbiol. Zh*. 2007; 69(1):42–51. Ukrainian.
25. Dreher TW. Role of tRNA-like structures in controlling plant virus replication. *Virus Res*. 2009; 139:217–29.
26. Rongxiang F, Xiaofeng W, Bingyi Q, Faxing C, Ming B. The nature of the RNA products synthesized in vitro with a cell-free extract from TMV-infected tobacco leaves. *Sci Sin*. 1980; 23:357–70.

27. Wang Y, Jensen RC, Stumph WE. Role of TATA box sequence and orientation in determining RNA polymerase II/III transcription specificity. *Nucleic Acids Res.* 1996; 24:3100–6.
28. Sun X, Simon AE. A cis-replication element functions in both orientations to enhance replication of Turnip crinkle virus. *Virology.* 2006; 352:39–51.
29. Kovalev N, Pogany J, Nagy PD. Template role of double-stranded RNA in tombusvirus replication. *J Virol.* 2014; 88:5638–51.

Отримано 6.04.2019