

УДК 575.21 + 577.2

10.7124/visnyk.utgis.19.1-2.1434

**ФРАГМЕНТИ МОБІЛЬНИХ ГЕНЕТИЧНИХ ЕЛЕМЕНТІВ
РІЗНОГО ПОХОДЖЕННЯ У ГЕНОМІ КОРОНАВІРУСУ SARS-CoV-2**

О. В. ПІДПАЛА, Л. Л. ЛУКАШ

Інститут молекулярної біології і генетики НАН України
Україна, 03680, м. Київ, вул. Академіка Заболотного, 150
e-mail: pidpala@ukr.net

Мета. Проаналізувати нуклеотидну послідовність нового коронавірусу людини SARS-CoV-2 на наявність фрагментів мобільних генетичних елементів (МГЕ) про- та еукаріотного походження. **Методи.** Гомологію між нуклеотидними послідовностями визначали програмою BLAST 2.6.1. Результати пошуку та ідентифікації МГЕ здійснено за допомогою програм ISfinder та CENSOR. **Результати.** У геномі коронавірусу людини SARS-CoV-2 ідентифіковано фрагменти 11 бактеріальних IS-елементів (0,68 % вірусного геному) та 23 еукаріотних МГЕ (4,6 %). Із 11 генів SARS-CoV-2, фрагменти бактеріальних IS-елементів присутні у двох генах (ORF1ab і M) та у 3'UTR. Фрагменти МГЕ еукаріот виявлено у межах чотирьох генів (ORF1ab, S, N і ORF7b) та у 3'UTR. Найбільший відсоток МГЕ виявлено у гені ORF7b та у 3'UTR. Аналізуючи інсерційні профілі IS-фрагментів на прикладі відомих коронавірусів людини та у їхніх проміжних хазяїв, дійшли висновку про можливість використання фрагментів прокаріотних МГЕ для філогенетичних досліджень нового коронавірусу людини SARS-CoV-2. **Висновки.** На підставі одержаних результатів висловлюємо припущення, що фрагменти МГЕ про- та еукаріотного походження можуть відігравати роль в еволюції геному SARS-CoV-2, зокрема при формуванні вірусних генів. Також вони можуть бути інформативними філогенетичними маркерами.

Ключові слова: коронавіруси людини, SARS-CoV-2, МГЕ, IS-елементи, філогенетичні маркери.

Вступ. Віруси не лише найчисельніші та найрізноманітніші, але і найзагадковіші біологічні об'єкти на Землі (Mushegian, 2020). Упродовж десятиліть не вщухають суперечки не тільки щодо їхнього наукового визначення, але і щодо їхнього походження та еволюції (Holmes, 2007). На сьогодні існують три основні гіпотези про походження вірусів: регресивна гіпотеза, гіпотеза клітинного походження і гіпотеза коеволуції (Forterre, 2006), проте жодна із них не є загальноприйнятою і беззаперечною. Можливо, що віруси незалежно виникали кілька разів за одним чи кількома сценаріями, оскільки між різними групами вірусів є значні розбіжності в організації генетичного матеріалу. Еволюційну історію вірусів можна прослідкувати, аналізуючи сучасні вірусні геноми. Дослідження на молекулярному рівні виявили зв'язок між вірусами та інфікованими ними організмами. Зокрема, геноми багатьох організмів містять ендегенні вірусні елементи (EVEs), які є залишками давніх вірусних інфекцій. Наприклад, геноми більшості видів хребетних містять від сотні до тисячі таких послідовностей (Sotero-Caio et al., 2017). Ці ендегенні ретровірусні елементи є цінним джерелом ретроспективних даних про еволюційну історію вірусів і стали вагомим підґрунтям для палеовірусології (Emerman, Malik, 2010).

Мобільні генетичні елементи (МГЕ) є компонентами геномів усіх живих організмів — від бактерій до людини. Вони сприяють регуляції експресії генів при розвитку й адаптації та є основним джерелом генетичної варіації в еволюції геномів (Miller, Capu, 2004). МГЕ виявлено також і у складі вірусних геномів, зокрема у гігантського вірусу *Pandoravirus salinus* ідентифіковано MITE-елементи (miniature inverted-repeat transposable element) (Sun et al., 2015). Також у складі вірусів, які належать до родин *Astroviridae*, *Caliciviridae*, *Picornaviridae*, *Coronaviridae* виявлено мобільний елемент *s2m* із невідомою функцією, можливо подібною до РНК-інтерференції (Tengs et al., 2013).

На основі літературних даних можна зробити висновок, що упродовж тисячолітньої еволюції віруси і МГЕ неодноразово могли бути джерелом нових послідовностей один для одного (Skalka, 2014). Якщо це так, то і вірусні геноми можуть зберігати сліди своїх колишніх хазяїв у вигляді різноманітних фрагментів МГЕ. Тому метою даної роботи було проаналізувати нуклеотидну послідовність нового коронавірусу людини SARS-CoV-2 на наявність фрагментів МГЕ про- та еукаріотного походження.

Матеріали і методи

Інформацію про нуклеотидні послідовності коронавірусів одержано із GenBank: SARS-CoV-2 (NC_045512); SARS-CoV (AY274119); MERS-CoV (NC_019843); HCoV-229E (NC_002645); HCoV-NL63 (NC_005831); HCoV-OC43 (NC_006213); HCoV-HKU1 (NC_006577); Bat coronavirus RaTG13 (MN996532); bat-SL-CoVZXC21 (MG772934); bat-SL-CoVZC45 (MG772933); Pangolin coronavirus isolate MP789 (MT084071); SARS coronavirus civet007 (AY572034); MERS-CoV NRCE-HKU270 (KJ477103). Гомологію між досліджуваними послідовностями визначали програмою *BLAST 2.6.1*. Для пошуку та ідентифікації фрагментів бактеріальних IS-елементів використовували програму *ISfinder* а для фрагментів еукаріотних МГЕ — програму *Censor*.

Результати та обговорення

Коронавірус SARS-CoV-2 — оболонковий одноланцюговий (+)РНК-вірус, який належить до роду *Betacoronavirus* (Wu et al., 2020). Він спричиняє тяжке інфекційне захворювання COVID-19, вперше виявлене наприкінці 2019 року. На сьогодні це сьомий вірус людини із родини *Coronaviridae*.

Довжина геному коронавірусу SARS-CoV-2 становить 29903 нуклеотидів. До його складу входять дві нетрансльовані ділянки довжиною 265 (5'UTR) та 229 (3'UTR) н. і 11 генів, які кодують 29 білків. Дві третини геному коронавірусу SARS-CoV-19 із 5' кінця припадає на ген *ORF1ab*, що кодує два поліфункціональні білки, які розщеплюються на 16 неструктурних білків і беруть участь у транскрипції та реплікації вірусного геному (Wu et al., 2020). До складу 3'-кінця геному коронавірусу входять 4 гени, які кодують структурні білки, і є характерними для усіх коронавірусів (*S* кодує поверхневий глікопротеїн або білок шипа, *E* кодує білок оболонки вірусу, *M* кодує мембранний білок і *N* кодує білок нуклеокапсиду). Також до складу коронавірусного геному SARS-CoV-19 входять гени, які кодують так звані «допоміжні білки» (*ORF3a*, *ORF6*, *ORF7a*, *ORF7b*, *ORF8* і *ORF10*). Вони допомагають змінити середовище всередині зараженої клітини, таким чином, полегшуючи реплікацію вірусу. У межах 3'UTR із координатами 29728..29768 знаходиться мобільний елемент *s2m* довжиною 41 н.

Фрагменти прокаріотних IS-елементів у межах геному SARS-CoV-2. IS-елементи належать до класу ДНК-транспозонів і є найпростішими бактеріальними МГЕ (Siguier et al., 2014). Як структурні складові елементи, послідовності IS-елементів входять до складу інших прокаріотних МГЕ, таких як Tn-елементів та IStrons. Крім того, бактеріальні IS-елементи є складовими деяких еукаріотних МГЕ, зокрема, Helitrons (Thomas, Pritham, 2015), *IS1595/merlin*, *IS256/mutator*, *IS630/Tc1/mariner* та *IS1380/piggyBac* (Hickman et al., 2010).

У геномі коронавірусу людини SARS-CoV-2 ідентифіковано фрагменти 11 різних IS-елементів довжиною від 18 до 23 н., що становить 0,68 % вірусного геному. Із 11 генів SARS-CoV-2, фрагменти IS-елементів присутні у двох генах *ORF1ab* та *M* (із перевагою в гені *ORF1ab*) і у 3'UTR, де не перекриваються із мобільним елементом *s2m*.

Фрагменти еукаріотних МГЕ у межах геному SARS-CoV-2. МГЕ є вагомими компонентами більшості еукаріотних геномів. Їх поділяють на два основні класи: клас I — ретротранспозони (сюди входять LTR елементи та non-LTR елементи) і клас II — ДНК-транспозони (Wicker et al., 2007).

У геномі коронавірусу людини SARS-CoV-2 ідентифіковано фрагменти 23 різних еукаріотних МГЕ: 8 ДНК-транспозонів (1,53 %); 8 LTR

елементів (1,23 %) і 7 non-LTR елементів (1,91 %). Серед організмів-хазяїв МГЕ більшість становлять безхребетні тварини із значною перевагою членистоногих, зокрема, комарів: *Aedes aegypti* — комара жовтолихоманкового і *Culex quinquefasciatus* — південного домашнього комара, які є переносниками багатьох вірусів-збудників і серед них, вірусу Зіка. Довжина фрагментів МГЕ еукаріот більша, ніж довжина ідентифікованих бактеріальних IS-елементів — від 28 до 160 н. і їхня загальна довжина у геномі досліджуваного вірусу становить 4,6 %. Фрагменти еукаріотних МГЕ виявлено у межах чотирьох генів SARS-CoV-2 і у 3'UTR (рис. 1).

Найбільшу кількість фрагментів еукаріотних МГЕ (17 із 23) виявлено у гені *ORF1ab*, але, зважаючи на велику довжину гена (близько 70 % від загальної довжини геному), відсоток фрагментів еукаріотних МГЕ незначний і становить 4,6 %.

У гені *S* ідентифіковано фрагменти трьох МГЕ еукаріот, загальна довжина яких складає 5,1 % від довжини гена. Трішки менший відсоток еукаріотних МГЕ виявлено у гені *N* (3,49 %).

Рекордсменом за довжиною послідовностей, гомологічних до МГЕ еукаріот, є ген *ORF7b*. Про його функцію до сьогодні ще невідомо. Зазначається, що він перекривається із сусіднім геном *ORF7a* (Wu et al., 2020). Майже половину послідовності гена *ORF7b* (49,2 %) припадає на ділянку, яка гомологічна фрагменту ДНК-транспозону *ISL2EU-4_HM Hydra vulgaris* (рис. 1).

Друге місце за відсотком МГЕ належить 3'UTR ділянці (37,5 %). Вона містить послідовність, гомологічну фрагменту non-LTR елемента *CR1-66_HM*, організмом-хазяїном якого також є *Hydra vulgaris*. Фрагмент *CR1-66_HM* не перекривається із мобільним елементом s2m SARS-CoV-2, але перекривається із послідовністю бактеріального IS-елемента, а саме фрагментом *ISMth1* термофільної археї *Methanosaeta thermophila*.

Аналізуючи одержані результати виникає питання — чи можна МГЕ розглядати лише як свідків тривалих еволюційних процесів, чи вони також можуть бути і безпосередніми учасниками цих подій?

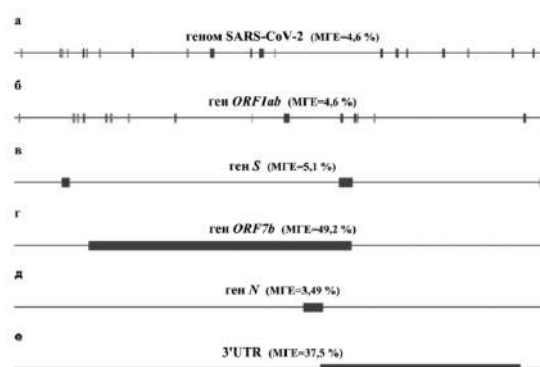


Рис. 1. Фрагменти еукаріотних МГЕ у геномі коронавірусу людини SARS-CoV-2 (а) та у його генах (б-д) чи неструктурних ділянках (е). У дужках зазначено відсоток МГЕ.

Інсерційні профілі прокаріотних IS-фрагментів у геномах коронавірусів людини. На сьогодні відомо про сім коронавірусів, які можуть інфікувати людину. Чотири із них — HCoV-229E, HCoV-NL63, HCoV-OC43 і HCoV-NKU1 є слабопатогенними і не спричиняють тяжкої симптоматики (Corman et al., 2018). На жаль, відомі і важкі форми коронавірусних інфекцій, збудниками яких є високопатогенні коронавіруси такі як SARS-CoV (Rota et al., 2003), MERS-CoV (Zaki et al., 2012) і SARS-CoV-2 (Wu et al., 2020).

Найбільшу гомологію до геномної послідовності SARS-CoV-2 серед коронавірусів людини має SARS-CoV (79 %) (Zhou et al., 2020). Аналізуючи інсерційні профілі IS-фрагментів у геномах коронавірусів людини, також бачимо, що саме SARS-CoV має найбільше гомологічних фрагментів IS-елементів із SARS-CoV-2 (рис. 2), що дає нам підстави розглядати фрагменти прокаріотних МГЕ як можливі філогенетичні маркери.

Коронавірусна інфекція є зооантропонозною, тобто можлива передача від тварин до людини. Виявлено, що джерелом SARS-CoV були дрібні хижачки гімалайської цивети (*Paguma larvata*) (Wang et al., 2005), а MERS-CoV — одногорбі верблюда або дромадери (*Camelus dromedarius*) (Chu et al., 2014). Як видно із результатів, наведених на рис. 3, інсерційні профілі IS-фрагментів у геномах відповідних коронавірусів людини і у їхніх проміжних хазяїв переконливо свідчать про можливість використання фрагментів прокаріотних МГЕ для філогенетичних досліджень.

			ISAsp24			
			ISGur4			
			ISPTo4			
			ISEc81			
			Tn4430			
			ISHma15			
			ISLhe11			
		ISAv8	ISDsp3			
	ISFop4	ISSto15	ISSto15	ISMth1		ISMth1
	Tn4430	ISCa7	ISCa7	ISSsu4	ISBF5	ISPLba3
	ISCig1	ISAbal1	ISAbal1	ISLhe10	ISTco2	ISFbr1
	MICBce6	ISFnu1	ISFnu1	ISMbov2	ISCKu8	ISSba20
	ISTph1	ISSha1	ISSha1	ISAbey9	ISHahy3	ISDeu11
	ISFac11	ISCbe3	ISCbe3	ISBusp7	ISAb42	ISButh7
ISMth1	ISSau3	ISCKu1	ISCKu1	ISGva2	ISTco1	ISNpu13
ISBf7	ISVsa17	ISSto10	ISSto10	ISAZo14	ISBeh1	ISKo13
ISDsac1	ISRme9	ISFnu5	ISFnu5	ISBce15	ISAtc5	ISPen1
ISLhe60	ISPin4	IS31831	ISMma3	ISTko1	ISBvu2	ISMae27
ISAZs29	IS1272	IS1207	ISLjo2	IS240F	ISNpu5	IS1271

(HCoV-229E) (HCoV-HKU1) (HCoV-NL63) (HCoV-OC43) (MERS-CoV) (SARS-CoV) (SARS-CoV-2)

Рис. 2. Інсерційні профілі прокариотних IS-фрагментів у геномах коронавірусів людини. Жирним шрифтом виділено гомологічні фрагменти.

У випадку коронавірусу SARS-CoV-2 можливим джерелом інфекції також є тварини. За результатами генетичного аналізу виявлено гомологію із коронавірусами Підковикових (кажанів) (Zhou et al., 2020), проте на даний час достовірно невідомо чи є вони лише природним резервуаром чи, можливо, і проміжним хазяїном інфекції.

а		б	
IS1470	IS1470	ISBF5	ISBF5
ISMth1	ISMth1	ISCB09	ISCB09
ISFbr1	ISFbr1	ISTco2	ISTco2
ISRel3	ISRel3	ISCKu8	ISCKu8
MITEPlu2	MITEPlu2	ISHahy3	ISHahy3
ISBco1	ISBco1	ISAb42	ISAb42
ISBvu2	ISBvu2	ISTco1	ISTco1
ISNpu5	ISNpu5	ISBeh1	ISBeh1
(SARS-CoV)	(SARS coronavirus civet007)	ISAtc5	ISAtc5
		ISTko1	ISTko1
		IS240F	IS240F
		(MERS-CoV)	(MERS-CoV NRCE-HKU270)

Рис. 3. Інсерційні профілі прокариотних IS-фрагментів у геномах коронавірусів людини та їхніх проміжних хазяїв: а — SARS-CoV і SARS coronavirus civet007; б — MERS-CoV і MERS-CoV NRCE-HKU270. Жирним шрифтом виділено гомологічні фрагменти.

Аналіз інсерційних профілів прокариотних IS-фрагментів у вірусних геномах потенційних переносників SARS-CoV-2. Рукокрилих — природні хазяї і переносники різних вірусів, бактерій та патогенних грибів (Schountz, 2014). Зокрема, вони є резервуаром вірусів сказу, Марбурга, Ебола та коронавірусів, які входять до групи особливо небезпечних інфекцій.

Наразі відомо, що у циркуляцію коронавірусів задіяно щонайменше 30 видів із 11 різних родин Рукокрилих (Drexler et al., 2014).

Відомо, що природним резервуаром SARS-подібних коронавірусів є Підковикові. Показано, що всі складові частини геному SARS-CoV присутні у SARS-подібних коронавірусів, які циркулюють серед підковикових кажанів *Rhinolophus sinicus* в одній із печер китайської провінції Юньнань (Quan et al., 2010). Припускають, що коронавірус SARS-CoV міг формуватися внаслідок рекомбінації генів між даними вірусами (Hu et al., 2017).

Що стосується коронавірусу MERS-CoV, то на думку багатьох дослідників, його природним резервуаром могли бути кажани, які належать до різних родів, зокрема, *Rousettus*, *Tylonycteris* і *Pipistrellus* (Memish et al., 2013). Висловлено гіпотезу, що предок MERS-CoV міг існувати у комахоїдних кажанів Старого Світу, і має африканське походження.

Людина може заразитись напямучу від Рукокрилих, але найчастіше спочатку заражаються тварини, які є проміжними хазяями, а вже через них вірус потрапляє до людини.

На підставі молекулярно-генетичних досліджень послідовності SARS-CoV-2, припускають, що його природним резервуаром також є кажани. Наприклад, геном коронавірусу RaTG13, який виділено з організму азіатського підковоноса (*Rhinolophus affinis*), збігається з SARS-CoV-2 на 96 % (Zhou et al., 2020). Але хто ж міг бути потенційним проміжним хазяїном і переносником SARS-CoV-2 від кажанів до людини? Серед претендентів на цю роль розглядають панголінів, оскільки

послідовності геномів вірусів від яванських панголінів (*Manis javanica*) схожі на 91,02 % із SARS-CoV-2 (Lam et al., 2020). При цьому в одному із зразків виявлено ділянку геному, що кодує S-білок, яка на 99 % подібна до аналогічної послідовності SARS-CoV-2 (гомологія у межах відповідної ділянки у RaTG13 — 77 %). Автори пояснюють це або паралельною еволюцією, або рекомбінацією. Про рекомбінацію серед коронавірусів відомо давно (Su et al., 2016). Цей механізм нерідко призводить до появи гібридних вірусів, які здатні інфікувати новий вид хазяїв, проте для цього необхідно, щоб два коронавіруси одночасно інфікували один організм. Де і як це сталося у випадку SARS-CoV-2, поки невідомо. Крім того, структура геному панголін-асоційованого коронавірусу має низку специфічних особливостей, що не дозволило дослідникам зробити остаточний висновок про безпосередню роль цих тварин у розповсюдженні коронавірусу серед людей (Liu et al., 2020).

Порівнюючи інсерційні профілі IS-фрагментів у геномах SARS-CoV-2 та у коронавірусів кажанів (RaTG13 і bat-SL-CoVZXC21) чи у коровавірусу панголіну, бачимо, що послідовність коронавірусу RaTG13 має найбільшу кількість гомологічних IS-фрагментів, тоді як у коронавірусного ізоляту від *Manis javanica* вони неідентифіковані (рис. 4). Цікаво, що у геномі одного із SARS-подібного коронавірусу кажана, а саме, bat-SL-CoVZXC21 присутній IS-фрагмент ISPen1, який є у SARS-CoV-2 але відсутній у геномі RaTG13. Можливо, що ці два коронавіруси кажанів одночасно інфікували потенційного проміжного хазяїна SARS-CoV-2 і сформували рекомбінантний варіант вірусу, який надалі, подолавши видовий бар'єр, інфікував людину.

ISMth1	ISMba1	ISMth1	ISEfa7
ISPIba3	ISMth1	ISFbr1	ISAs17
ISFbr1	ISFbr1	ISAs17	MITEPlu2
ISSba20	ISLho3	ISCSa6	ISEc66
ISDeul1	ISButh7	ISRel5	ISVat1
ISButh7	ISKol3	ISPa52	ISDec1
ISNpu13	ISAzo31	ISButh7	ISEfa10
ISKol3	ISMAe27	ISDce1	ISPPr12
<u>ISPen1</u>	ISClbu2	ISVpa2	IS2311
ISMAe27	IS1500B	<u>ISPen1</u>	ISClbu2
IS1271	IS1271	ISTfe1	
(SARS-CoV-2)	(RaTG13)	(bat-SL-CoVZXC21)	(Pangolin coronavirus)

Рис. 4. Інсерційні профілі прокаріотних IS-фрагментів у геномах коронавірусів: людини — SARS-CoV-2; кажанів — RaTG13 і bat-SL-CoVZXC21 та панголінового коронавірусу (ізолят MP789). Жирним шрифтом виділено гомологічні фрагменти. Підкреслено послідовність, яка відсутня у геномі коронавірусу RaTG13, але присутня у послідовності одного із SARS-подібного коронавірусу кажана.

Участь еукаріотних МГЕ у формуванні коронавірусних генів. Віруси пройшли тривалий час власної еволюції разом із своїми хазяями на різних рівнях живого — від прокаріотних через низку еукаріотних організмів до людини. Інфікуючи клітини, віруси могли не лише лишати свої послідовності (як у випадку ендегенних ретровірусів), але і запозичувати фрагменти МГЕ. Розподіл і склад МГЕ та вірусів специфічні для різних доменів живого, що дозволяє зробити припущення про їхнє взаємоперетворення упродовж еволюції, при якому МГЕ можуть бути джерелом вірусів (Koonin et al., 2015). Також є думка, що екзогенні ретровіруси утворились із Ту3/Gypsy LTR-ретротранспозонів, оскільки їхні послідовності виявлено у геномах більшості живих організмів (Skalka, 2014).

На прикладі коронавірусів, досліджуючи їхні геноми, простежили участь МГЕ, зокрема, еукаріотних, у формуванні вірусних генів. Спочатку проаналізували відповідні коронавірусні послідовності на наявність фрагментів еукаріотних МГЕ. Кількість МГЕ різна і коливається від 23 у SARS-CoV-2 до 38 у RaTG13. Є фрагменти, які присутні у більшості проаналізованих геномах, а є й такі, які виявлено лише для певних вірусів.

Оскільки кількість МГЕ еукаріотного походження значно більша, ніж прокаріотного, порівнювали коронавірусні геноми на прикладі конкретних генів. Вище ми навели дані про те, що майже половина послідовності гена *ORF7b* SARS-CoV-2 (49,2 %) припадає на ділянку, яка гомологічна фрагменту ДНК-транспозону *ISL2EU-4_HM Hydra vulgaris* (рис. 1). Ген *ORF7b* кодує допоміжний білок, роль якого ще не з'ясовано (Wu et al., 2020). Цей ген також присутній у геномі коронавірусу кажана RaTG13 (Zhou et al., 2020), який на сьогодні має найбільшу гомологію до нового коронавірусу людини, але неідентифікований у його еволюційно найближчих SARS-подібних коронавірусів кажанів (bat-SL-CoVZXC21, bat-SL-CoVZC45) (Hu et al., 2018) і у панголінового коронавірусу (Lam et al., 2020). Проаналізували ступінь гомології відповідних послідовностей. Цікаво, що виявлений у вірусних геномах SARS-CoV-2 і RaTG13 фрагмент ДНК-транспозону *ISL2EU-4_HM* перекривається із фрагментом LTR ретротранспозону *Gypsy-614_AA-I* у геномі bat-SL-CoVZXC21. У послідовності панголінового коронавірусу відповідного МГЕ не ідентифіковано.

На підставі одержаних результатів, висловлюємо припущення, що фрагменти МГЕ можуть бути задіяні у формуванні вірусних генів. Але яким чином фрагменти МГЕ еукаріотного походження потрапили до вірусних геномів? Можемо лише припустити, що це відбувалось не лише за рахунок рекомбінаційних подій між вірусними геномами, яке широко розповсюджене, але, і на певних етапах еволюції вірусів, за рахунок інтеграції / деінтеграції чи рекомбінації між ними і геномом хазяїна.

Висновки

На підставі одержаних результатів висловлено припущення, що фрагменти МГЕ про- та еукаріотного походження можуть відігравати роль в еволюції геному SARS-CoV-2, зокрема при формуванні вірусних генів. Також вони можуть бути інформативними філогенетичними маркерами для пошуку потенційного еволюційного попередника.

Перелік літератури

1. Corman V. M., Muth D., Niemeyer D., Drosten C. Hosts and sources of endemic human coronaviruses. *Adv. Virus Res.* 2018. Vol. 100. P. 163–188. doi: 10.1016/bs.aivir.2018.01.001.
2. Chu D. K., Poon L. L., Goma M. M., Shehata M. M., Perera R. A., Abu Zeid D., El Rifay A. S., Siu L. Y., Guan Y., Webby R. J., Ali M. A., Peiris M., Kayali G. MERS coronaviruses in dromedary camels, Egypt. *Emerging Infect. Dis.* 2014. Vol. 20, № 6. P. 1049–1053. doi: 10.3201/eid2006.140299.
3. Drexler J. F., Corman V. M., Drosten C. Ecology, evolution and classification of bat coronaviruses in the aftermath of SARS. *Antiviral Res.* 2014. Vol. 101. P. 45–56. doi: 10.1016/j.antiviral.2013.10.013.
4. Emerman M., Malik H. S. Paleovirology — modern consequences of ancient viruses. *PLoS Biology.* 2010. Vol. 8, № 2. e1000301. doi:10.1371/journal.pbio.1000301.
5. Forterre P. The origin of viruses and their possible roles in major evolutionary transitions. *Virus Res.* 2006. Vol. 117, № 1. P. 5–16. doi: 10.1016/j.virusres.2006.01.010.
6. Hickman A. B., Chandler M., Dyda F. Integrating prokaryotes and eukaryotes: DNA transposases in light of structure. *Crit. Rev. Biochem. Mol. Biol.* 2010. Vol. 45, № 1. P. 50–69. doi: 10.3109/10409230903505596.
7. Holmes E. C. Viral evolution in the genomic age. *PLoS Biol.* 2007. Vol. 5, № 10. e278. doi:10.1371/journal.pbio.0050278.
8. Hu B., Zeng L.P., Yang X. L., Ge X. Y., Zhang W., Li B., Xie J. Z., Shen X. R., Zhang Y. Z., Wang N., Luo D. S., Zheng X. S., Wang M. N., Daszak P., Wang L.-F., Cui J., Shi Z. L. Discovery of a rich gene pool of bat SARS-related coronaviruses provides new insights into the origin of SARS coronavirus. *PLoS Pathogens.* 2017. Vol. 13, № 11. e1006698.
9. Koonin E. V., Dolja V. V., Krupovic M. Origins and evolution of viruses of eukaryotes: The ultimate modularity. *Virology.* 2015. Vol. 479–480. P. 2–25.
10. Lam T. T., Shum M. H., Zhu H. C., Tong Y. G., Ni X. B., Liao Y. S., Wei W., Cheung W. Y., Li W. J., Li L. F., Leung G. M., Holmes E. C., Hu Y. L., Guan Y. Identifying SARS-CoV-2 related coronaviruses in Malayan pangolins. *Nature.* 2020. doi:10.1038/s41586-020-2169-0.
11. Liu P., Jiang J. Z., Wan X. F., Hua Y., Li L., Zhou J., Wang X., Hou F., Chen J., Zou J., Chen J. Are pangolins the intermediate host of the 2019 novel coronavirus (SARS-CoV-2)? *PLoS Pathog.* 2020. Vol. 16, № 5. e1008421. doi: 10.1371/journal.ppat.1008421.
12. Memish Z. A., Mishra N., Olival K. J., Fagbo S. F., Kapoor V., Epstein J. H., Alhakeem R., Durosinloun A., Al Asmari M., Islam A., Kapoor A., Briese T., Daszak P., Al Rabeeah A. A., Lipkin W. I. Middle East respiratory syndrome coronavirus in bats, Saudi Arabia. *Emerg. Infect. Dis.* 2013. Vol. 19, № 11. P. 1819–1823. doi:10.3201/eid1911.13117.
13. Miller W. J., Cappy P. Mobile genetic elements as natural tools for genome evolution. *Methods Mol. Biol.* 2004. Vol. 260. P. 1–20. doi: 10.1385/1-59259-755-6:001.
14. Mushegian A. R. Are there 10^{31} virus particles on Earth, or more, or fewer? *J. Bacteriol.* 2020. Vol. 202, № 9. pii: e00052-20. doi: 10.1128/JB.00052-20.
15. Quan P. L., Firth C., Street C., Henriquez J. A., Petrosov A., Tashmukhamedova A., Hutchison S. K., Egholm M., Osinubi M. O. V., Niezgodna M., Ogunkoya A. B., Briese T., Rupprecht C. E., Lipkin W. I. Identification of a Severe Acute Respiratory Syndrome Coronavirus Like Virus in a leaf-nosed bat in Nigeria. *mBio.* 2010. Vol. 1, № 4. e00208-10.
16. Rota P. A., Oberste M. S., Monroe S. S., Nix W. A., Campagnoli R., Icenogle J. P., Peñaranda S., Bankamp B., Maher K., Chen M.-H., Tong S., Tamin A., Lowe L., Frace M., DeRisi J. L., Chen Q., Wang D., Erdman D. D., Peret T. C. T., Burns C., Ksiazek T. G., Rollin P. E., Sanchez A., Liffick S., Holloway B., Limor J., McCaustland K., Olsen-Rasmussen M., Fouchier R., Günther S., Osterhaus A. D. M. E., Drosten C., Pallansch M. A., Anderson L. J., Bellini W. J. Characterization of a novel coronavirus associated with severe acute respiratory syndrome. *Science.* 2003. Vol. 300, № 5624. P. 1394–1399. doi: 10.1126/science.1085952.

17. Schountz T. Immunology of bats and their viruses: challenges and opportunities. *Viruses*. 2014. Vol. 6, № 12. P. 4880–4901. doi: 10.3390/v6124880.
18. Siguier P., Gourbeyre E., Chandler M. Bacterial insertion sequences: their genomic impact and diversity. *FEMS Microbiol. Rev.* 2014. Vol. 38, № 5. P. 865–891. doi:10.1111/1574-6976.12067.
19. Skalka A. M. Retroviral DNA transposition: themes and variations. *Microbiol. Spectr.* 2014. Vol. 2, № 5. MDNA300052014. doi: 10.1128/microbiolspec.MDNA3-0005-2014.
20. Sotero-Caio C. G., Platt R. N., Suh A., Ray D. A. Evolution and diversity of transposable elements in vertebrate genomes. *Genome Biol. Evol.* 2017. Vol. 9, № 1. P. 161–177. doi: 10.1093/gbe/evw264.
21. Su S., Wong G., Shi W., Liu J., Lai A. C. K., Zhou J., Liu W., Bi Y., Gao G. F. Epidemiology, genetic recombination, and pathogenesis of coronaviruses. *Trends Microbiol.* 2016. Vol. 24, № 6. P. 490–502. doi: 10.1016/j.tim.2016.03.003.
22. Sun C., Feschotte C., Wu Z., Mueller R. L. DNA transposons have colonized the genome of the giant virus *Pandoravirus salinus*. *BMC Biol.* 2015. Vol. 13. P. 38. doi: 10.1186/s12915-015-0145-1.
23. Tengs T., Kristoffersen A. B., Bachvaroff T. R., Jonassen C. M. A mobile genetic element with unknown function found in distantly related viruses. *Virology*. 2013. Vol. 10. P. 132. doi: 10.1186/1743-422X-10-132.
24. Thomas J., Pritham E. J. Helitrons, the eukaryotic rolling-circle transposable elements. *Microbiol. Spectr.* 2014. Vol. 3, № 4. MDNA3-0049-2014. doi:10.1128/microbiolspec.MDNA3-0049-2014.
25. Wang M., Yan M., Xu H., Liang W., Kan B., Zheng B., Chen H., Zheng H., Xu Y., Zhang E., Wang H., Ye J., Li G., Li M., Cui Z., Liu Y. F., Guo R. T., Liu X. N., Zhan L. H., Zhou D. H., Zhao A., Hai R., Yu D., Guan Y., Xu J. SARS-CoV infection in a restaurant from palm civet. *Emerging Infect. Dis.* 2005. Vol. 11, № 12. P. 1860–1865.
26. Wicker T., Sabot F., Hua-Van A., Bennetzen J. L., Capy P., Chalhoub B., Flavell A., Leroy P., Morgante M., Panaud O., Paux E., SanMiguel P., Schulman A. H. A unified classification system for eukaryotic transposable elements. *Nat. Rev. Gen.* 2007. Vol. 8, № 12. P. 973–982. doi: 10.1038/nrg2165.
27. Wu F., Zhao S., Yu B., Chen Y. M., Wang W., Song Z. G., Hu Y., Tao Z. W., Tian J. H., Pei Y. Y., Yuan M. L., Zhang Y. L., Dai F. H., Liu Y., Wang Q. M., Zheng J. J., Xu L., Holmes E. C., Zhang Y. Z. A new coronavirus associated with human respiratory disease in China. *Nature*. 2020. Vol. 579, № 7798. P. 265–269. doi: 10.1038/s41586-020-2008-3.
28. Zaki A. M., van Boheemen S., Bestebroer T. M., Osterhaus A. D., Fouchier R. A. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *N. Engl. J. Med.* 2012. Vol. 367, № 19. P. 1814–1820. doi: 10.1056/NEJMoa1211721.
29. Zhou P., Yang X.-L., Wang X.-G., Hu B., Zhang L., Zhang W., Si H.-R., Zhu Y., Li B., Huang C.-L., Chen H.-D., Chen J., Luo Y., Guo H., Jiang R.-D., Liu M.-Q., Chen Y., Shen X.-R., Wang X., Zheng X.-S., Zhao K., Chen Q.-J., Deng F., Liu L.-L., Yan B., Zhan F.-X., Wang Y.-Y., Xiao G.-F., Shi Z.-L. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*. 2020. Vol. 579. P. 270–273. doi:10.1038/s41586-020-2012-7.

FRAGMENTS OF DIFFERENT ORIGINS MOBILE GENETIC ELEMENTS IN THE GENOME OF CORONAVIRUS SARS-CoV-2

O. V. Pidpala, L. L. Lukash

Institute of Molecular Biology and Genetics
of Natl. Acad. Sci. of Ukraine
Ukraine, 03680, Kyiv, Akad. Zabolotnogo str., 150
e-mail: pidpala@ukr.net

Aims. To analyze the presence of fragments of mobile genetic elements (MGE) of pro- and eukaryotic origin in the nucleotide sequence of the new human coronavirus SARS-CoV-2. **Methods.** The homology between nucleotide sequences was determined by BLAST 2.6.1. The results of the search and identification of MGE were performed using the ISfinder and CENSOR programs. **Results.** In the genome of the human coronavirus SARS-CoV-2, fragments of 11 bacterial IS elements (0,68 % of the viral genome) and 23 eukaryotic MGEs (4,6 %) have been identified. Of the 11 SARS-CoV-2 genes, fragments of IS-elements are present in two genes (ORF1ab and M) and in 3'UTR. Fragments of MGE eukaryotes have been detected within four genes (ORF1ab, S, N and ORF7b) and in 3'UTR. The highest percentage of MGE was found in the ORF7b gene and in 3'UTR. Analyzing the insertion profiles of IS fragments on the example of known human coronaviruses and their intermediate hosts, we have been concluded that the possibility of using fragments of prokaryotic MGE for phylogenetic studies of the new human coronavirus SARS-CoV-2. **Conclusions.** Based on the obtained results, we have been suggested that fragments of MGE of pro- and eukaryotic origin may play a role in the evolution of the SARS-CoV-2 genome, in particular in the formation of viral genes. They could also be informative phylogenetic markers.

Keywords: human coronaviruses, SARS-CoV-2, MGE, IS-elements, phylogenetic markers.