

АНАЛІЗ НУКЛЕОТИДНИХ ПОСЛІДОВНОСТЕЙ БАКТЕРІАЛЬНИХ РЕТРОІНТРОНІВ НА НАЯВНІСТЬ ГОМОЛОГІЇ ДО ЕУКАРІОТНИХ МГЕ

О. В. ПІДПАЛА *orcid* 0000-0002-9446-1590, Л. Л. ЛУКАШ *orcid* 0000-0003-4522-1600

Інститут молекулярної біології і генетики НАН України,
Україна, 03143, м. Київ, вул. Академіка Заболотного, 150
e-mail: pidpala@ukr.net

Мета. Проаналізувати послідовності бактеріальних ретроінтронів на наявність гомології до представників різних класів еукаріотних мобільних генетичних елементів (МГЕ). **Методи.** Нуклеотидні послідовності бактеріальних інтронів II групи взято для аналізу із бази даних Database for Bacterial Group II Introns. Результати пошуку та ідентифікації МГЕ здійснено за допомогою програми SENSOR. **Результати.** У проаналізованих послідовностях бактеріальних ретроінтронів виявлено фрагменти послідовностей обох класів еукаріотних МГЕ у різноманітних комбінаціях: non-LTR; LTR; DNA; non-LTR+LTR; non-LTR+DNA; LTR+DNA; non-LTR+LTR+DNA. Найбільша група ретроінтронів містить МГЕ різних класів. **Висновки.** Бактеріальні ретроінтрони можуть бути не лише предками non-LTR ретроелементів, але і носіями фрагментарних нуклеотидних послідовностей для формування різноманітних еукаріотних МГЕ.

Ключові слова: ретроелементи, ретроінтрони, еукаріотні МГЕ.

Вступ. Мобільні генетичні елементи (МГЕ) — послідовності ДНК, які можуть переміщуватись у межах геному. Вони є невід'ємною складовою про- та еукаріотних геномів (Wells and Feschotte, 2020; Ross et al., 2021; Almojil et al., 2021). Залежно від структурної організації та механізмів транспозиції виокремлюють два основні класи МГЕ (Wicker et al., 2007). Ретротранспозони (елементи класу I — non-LTR і LTR) переміщуються за допомогою процесу зворотної транскрипції, тоді як ДНК-транспозони (елементи класу II) переміщуються у вигляді ДНК копій із використанням кількох різних механізмів. Серед ретроелементів особливу увагу привертають ретроінтрони (мобільні інтрони групи II, G2I) (Lambowitz and Belfort, 2014). Вони є найстарішою групою серед non-LTR ретроелементів. Це один із двох класів самосплайсуючих інтронів, які знаходяться у прокариотних геномах і геномах симбіотичних органел. Ймовірно, що перші мобільні інтрони групи II виникли саме у геномі бактерій. У процесі утворення еукаріотних клітин, ретроінтрони увійшли до їхнього складу разом із мітохондріями і хлоропластами (Zimmerly et al., 2001). Є припущення, що саме бактеріальні ретроінтрони є ключовими елементами в еволюції еукаріот і предками не лише сплайсосомних інтронів і теломер, але і non-LTR ретроелементів (Garavís et al., 2013; Lambowitz and Belfort, 2014; Novikova and Belfort, 2017). Проте ще багато лишається питань щодо еволюційних зв'язків між ретроінтронами і non-LTR ретроелементами та походженням еукаріотних ретротранспозонів (Novikova and Belfort, 2017). Метою даної роботи було проаналізувати послідовності бактеріальних ретроінтронів на наявність гомології до представників різних класів еукаріотних МГЕ.

Матеріали і методи

Нуклеотидні послідовності бактеріальних інтронів II групи взято для аналізу із бази даних Database for Bacterial Group II Introns (<http://webapps2.ualgary.ca/~groupii>) (Candales et al., 2012). Пошук та ідентифікацію МГЕ здійснено за допомогою програми SENSOR (<https://www.girinst.org/censor/>) (Kohany et al., 2006).

Результати та обговорення

Ретроінтрони є одночасно каталітичними РНК (рибозимами) і ретроелементами, які зустрічаються переважно в геномах бактерій і в мітохондріальних та хлоропластних геномах деяких еукаріот, зокрема, грибів і рослин, але рідко зустрічаються в археях і відсутні в еукаріотичних ядерних геномах (Zimmerly and Semper, 2015). Бактеріальні ретроелементи в основному діють як ретроелементи і складаються з каталітичної самосплайсуючої РНК і багатофункціонального інтрон-кодованого білка (IEP) із зворотньою транскриптазою, матуразною та ендонуклеазною активністю. Розмір повнорозмірних ретроінтронів становить 2–3 т.н.п. В органелах ретроінтрони зазвичай присутні в неактивній формі і часто не мають IEP. Бактеріальні ретроінтрони здебільшого розташовані в міжгенних регіонах або в різних типах мобільних ДНК, які включають острівки патогенності та плазміди вірулентності (Candales et al., 2012; Dai and Zimmerly, 2002; Dai et al., 2003; Toro and Martínez-Abarca, 2013). Більшість бактеріальних геномів містять лише кілька ретроінтронів (1–10 копій). Органелярні ретроінтрони розподілені щільніше в геномах мітохондрій та хлоропластів і присутні переважно в генах рРНК і тРНК (Zimmerly and Semper, 2015).

Завдяки самосплайсингу і ретромобільності ретроінтрони є ключовими елементами не лише у прокариотно-еукаріотному переході, але і відіграли важливу роль в еволюції еукаріот, зокрема дали початок багатьом різноманітним функціям і послідовностям, серед яких non-LTR ретроелементи. Проте, незважаючи на значний прогрес у розумінні структурної організації і роботи цих ретроелементів, ще багато лишається питань щодо їхніх еволюційних зв'язків (Novikova and Belfort, 2017).

Із проаналізованих 90 послідовностей бактеріальних інтронів II групи фрагменти non-LTR ретроелементів виявили у 16 ретроінтронах (рис. 1). Вони належать до 12 кладів (CRE, Daphne, Jockey, L1, Loa, Nimb, Penelope, Proto2, R1, RTE, Tad1, Vingi) 5 груп відповідних ретроелементів (CRE, L1, RTE, I, CR1) згідно класифікації RTclass1, яку використовують у базі даних Repbase (Bao et al., 2015). Переважають фрагменти L1 (6 випадків) і Nimb (3 випадки) елементів. Довжина фрагментів коливається від 37 п.н. до 152 п.н., що становить від 1,86 до 7,72 % загальної довжини інтрону. Рівні подібності становлять від 64,67 до 86,49 % (середнє значення — 77,40 %).

Варто зазначити, що елементи групи CRE є еволюційно найстарішими серед non-LTR ретроелементів (Malik et al., 1999; Novikova, Blinov,

2009), а час виникнення самих non-LTR ретроелементів збігається із виникненням еукаріотних організмів. Non-LTR ретротранспозони присутні у геномах рослин, грибів, безхребетних і хребетних тварин, але переважають у геномах тварин.

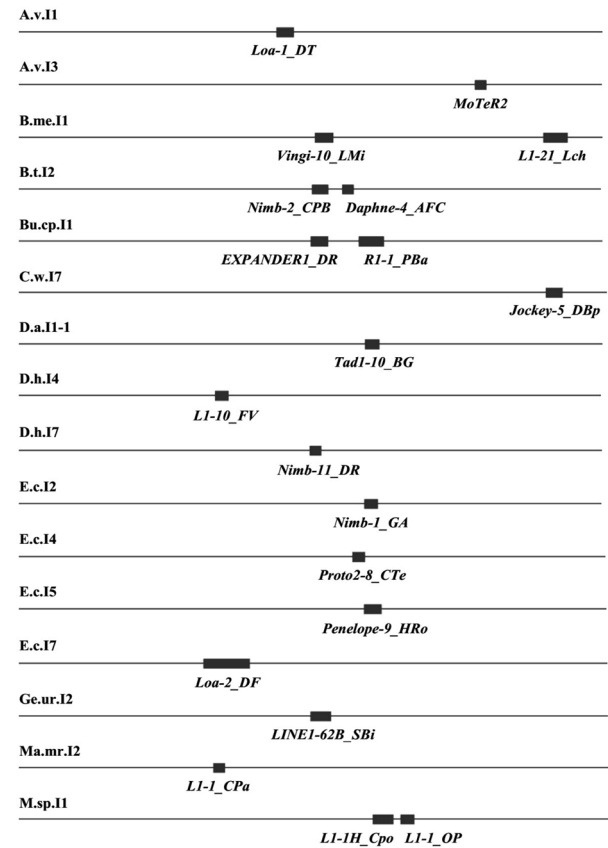


Рис. 1. Фрагменти non-LTR ретроелементів у бактеріальних ретроінтронах.

Фрагменти лише LTR ретроелементів ідентифіковано в 11 ретроінтронах (рис. 2). У двох послідовностях (B.f.I1 та Bu.xe.I2) вони ідентичні. У цих ретроінтронах виявлено по шість фрагментів LTR ретроелементів. У чотирьох послідовностях ідентифіковано по два різних фрагменти, а у п'яти ретроінтронах виявлено по одному фрагменту LTR-ретроелементів. Вони належать до 4 надродин: BEL, Soría, Gypsy, ERV. Послідовності BEL ретроелементів поширені у безхребетних тварин, Soría і Gypsy виявлені у геномах рослин, грибів і у безхребетних та хребетних, а ERV елементи поширені у хребетних. Серед фрагментів LTR ретроелементів у зазначених ретроінтронах переважають фрагменти Gypsy (11 випадків) і Soría (6 випадки) елементів. Довжина одиночних фрагментів коливається від 30 п.н. до 89 п.н., а загальний відсоток послідовностей становить від 1,08 до 20,24 % довжини ретроінтрону. Рівні подібності становлять від 69,66 до 90,00 % (середнє значення — 79,71 %).

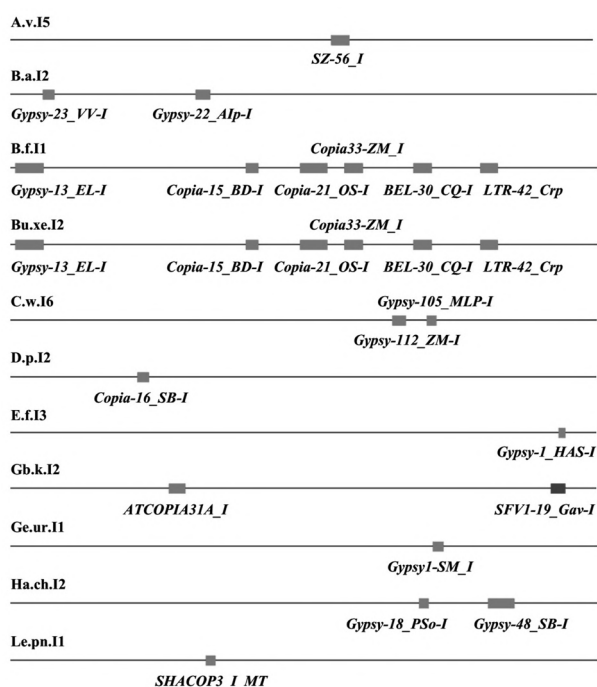


Рис. 2. Фрагменти LTR ретроелементів у бактеріальних інтронах II групи.

На основі філогенетичного аналізу ретроелементів було зроблено висновок, що non-LTR ретроелементи є прабатьками еукаріотних LTR ретроелементів, які у свою чергу дали початок кільком класам вірусів, включаючи ретровіруси хребетних (Malik and Eickbush, 2001). Є припущення, що предками LTR ретроелементів могли бути послідовності non-LTR ретроелементів, які гомологічні елементам TART і DRE, або HeT-A (Danilevskaya et al., 1999; Pardue and DeBaryshe, 1999).

Надродина Copia і Gypsy є найстарішими представниками LTR ретроелементів. Із них найближчими до non-LTR ретроелементи є послідовності надродина Copia, оскільки мають подібну доменну організацію гена *pol* (McClure, 1991; Kennell et al., 1993). У свою чергу послідовності надродина Gypsy є джерелом елементів надродина Bel/Pao, ретровірусів і каулімовірусів (Kim et al., 1994; Llorens et al., 2009).

Через еволюційно молодший вік LTR ретроелементів, їхню складну структуру та відсутність будь-яких прокаріотних попередників передбачено, що LTR ретроелементи виникли внаслідок набуття ДНК-транспозоном гена зворотної транскриптази non-LTR ретроелемента (Caru et al., 1997; Malik and Eickbush, 2001; Stocking and Kozak, 2008). Інтеграза LTR ретроелементів має подібність до транспозази деяких ДНК-транспозонів, особливо IS3, IS481, Ginger1, Ginger2 і Polinton (Bao et al., 2010).

Також ідентифікували фрагменти ретро-транспозонів обох підкласів. Їх виявлено у шістнадцяти ретроінтронах. Кількість виявлених фрагментів коливається від двох до п'яти послідовностей у різних ретроінтронах. Їхня довжина варіює від 27 до 182 п.н., що становить від 3,19 до 21,63 % загальної довжини ретроінтрону. У двох послідовностях (B.f.I2 і Bu.xe.I1) фрагменти ретро-транспозонів ідентичні але мають різний напрямок. Сумарна кількість фрагментів non-LTR ретроелементів становить 24 послідовності. Вони належать до 10 кладів (CR1, Crack, Jockey, L1, L2, R1, RTE, Tad1, Tx1, Vingi) 4 груп відповідних ретроелементів (L1, RTE, I, CR1). Переважають фрагменти L1 послідовностей. Кількість фрагментів LTR ретроелементів, які присутні разом із non-LTR ретроелементами, становить 26 послідовностей. Вони належать до 4 надродин: BEL, Copia, Gypsy, ERV. Переважають фрагменти Gypsy елементів (16 послідовностей). Варто зазначити, що у випадку ретроінтрону An.v.I1, послідовності фрагментів LTR (PRIMA41) і non-LTR ретроелементів (L1-7_ECa) утворюють кластер. Цікаво, що PRIMA41 є ендемічним ретровірусом людини.

Фрагменти ДНК-транспозонів ідентифіковано у 5 проаналізованих ретроінтронах (A.pt.I1, Ch.ph.I1-1, C.w.I1-1, D.h.I2, Kl.pn.I1) (рис. 3). По одному фрагменту ДНК-транспозонів ідентифіковано у трьох ретроінтронах. У двох послідовностях (A.pt.I1 та Kl.pn.I1) вони ідентичні. По два і по три фрагменти ідентифіковано ще у двох ретроінтронах. Виявлені фрагменти належать до 4 надродин ДНК-транспозонів: EnSpm/CASTA, Helitron, Mariner/Tc1 і MuDR, які присутні у геномах рослин, грибів, безхребетних і хребетних тварин. У проаналізованих ретроінтронах виявлено по два фрагменти послідовностей Helitron і MuDR та по одному фрагменту послідовностей EnSpm/CASTA і Mariner/Tc1. Довжина одиночних фрагментів ДНК-транспозонів коливається від 30 п.н. до 93 п.н., а загальний відсоток послідовностей становить від 1,87 до 7,47 % довжини інтрону. Рівні подібності становлять від 68,48 до 90,00 % (середнє значення — 78,91 %).

ДНК-транспозони виявлено у геномах про-і еукаріот (Feschotte and Pritham, 2007). Традиційно клас еукаріотних ДНК-транспозонів охоплює два підкласи, які відрізняються за способом транспозиції (Wicker et al., 2007). Перший підклас містить власне ДНК-транспозони (класичні ДНК-транспозони). Інший підклас охоплює хелітрони (Helitrons) і полінтони (Polintons). Власне ДНК-транспозони дуже різноманітні і знайдені в усіх вивчених таксонах еукаріот (вони об'єднують, за різними класифікаціями, від 7 до

21 надродини). У цих транспозонів відносно проста структура: вони мають одну відкриту рамку зчитування, що кодує фермент для переміщення — транспозазу, яка фланкована короткими інвертованими повторами (TIRs — terminal inverted repeats). Найрозповсюдженішими серед них є елементи Mariner/Tc1, MuDR, hAT і piggyBac.

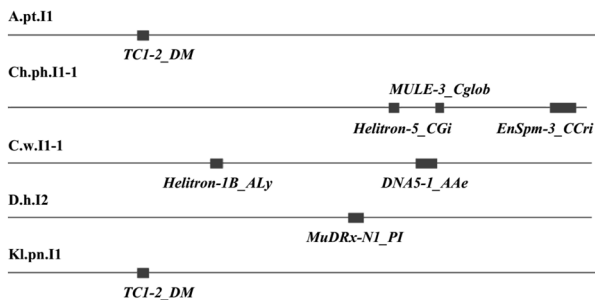


Рис. 3. Фрагменти еукаріотних ДНК-транспозонів у бактеріальних інтронах II групи.

Представники надродини Helitron переміщуються за допомогою механізму «кільця, що котиться» (rolling — circletransposition). Вони кодують ДНК-хеліказу і нуклеазу/лігазу і не мають TIRs (Kapitonov and Jurka, 2001). Хелітрони знайдено у більшості таксонів еукаріот, охоплюючи рослини, гриби, безхребетних і хребетних. За механізмом транспозиції вони нагадують бактеріальний інсерційний елемент IS91 і геміновіруси. Припускають, що такі елементи дали початок рослинним вірусам. Представники надродини Polinton (чи Maverick) об'єднують найбільші за розмірами транспозони (Kapitonov and Jurka, 2006; Pritham et al., 2007). Можливо вони походять від фагу Mavirus (Fischer and Suttle, 2011) чи лінійних плазмід, а древні Polintons, у свою чергу, дали початок аденовірусам. Деякі дослідники виокремлюють елементи типу Helitrons і Polintons в окремі підкласи (Kapitonov and Jurka, 2001; Pritham et al., 2007) або навіть у два класи (Hellen and Brookfield, 2013; Ray et al., 2007). Проте спільне походження їх від прокаріотних послідовностей (Hickey, 1992; Werren, 2011; Alzohairy et al., 2013) і переміщення без участі РНК-посередника дають всі підстави об'єднувати ДНК-транспозони еукаріот в один клас (Fedoroff, 2001).

У тридцяти із проаналізованих послідовностей бактеріальних інтронів II групи ідентифіковано фрагменти МГЕ різних класів/підкласів у різних комбінаціях: non-LTR+LTR+DNA; non-LTR+DNA; LTR+DNA. Кількість фрагментів МГЕ коливається від двох до семи, а довжина — від 29 п.н. до 144 п.н., що становить від 2,14 до 21,06 % загальної довжини відповідного ретроінтрону. Ідентифіковані фрагменти non-LTR ретроелементів належать до 12 кладів (CR1, Crack, Daphne, I, Jockey, L1, L2, Loa, Nimb, R1, RTE,

Tx1) 4 груп відповідних ретроелементів (L1, RTE, I, CR1) а фрагменти LTR ретроелементів належать до 4 надродин: BEL, Soria, Gypsy, ERV. Виявлені фрагменти ДНК-транспозонів належать до 10 надродин: EnSpm/CACTA, Harbinger, hAT, Helitron, ISL2EU, Mariner/Tc1, MuDR, P, Polinton, Sola (Sola2). На рис. 4 наведено найцікавіші варіанти транспозонного ландшафту у даній групі. Цікаво, що у двох випадках (E.fm.II і L.II) виявлено формування кластерів із фрагментів LTR ретроелементів і ДНК-транспозонів. Також, у послідовності ретроінтрону C.a.II ідентифіковано фрагмент LTR повтору (LTR24C) довжиною 60 п.н. ендегенного ретровірусу людини.

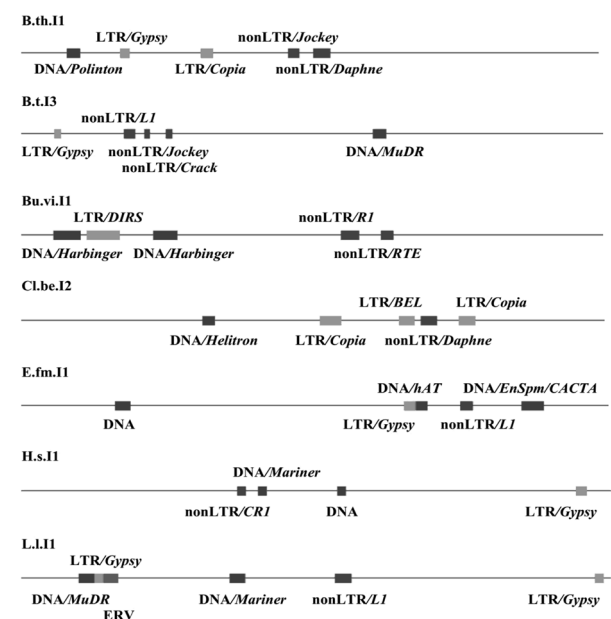


Рис. 4. Фрагменти різноманітних еукаріотних мобільних генетичних елементів у бактеріальних інтронах II групи.

У чотирьох проаналізованих бактеріальних інтронів II групи (Al.bo.II, B.c.II, Cl.sp.II, G.v.II) поряд із фрагментами еукаріотних МГЕ виявили фрагменти каулімовірусів (рис. 5). Довжина одиначних фрагментів каулімовірусів коливається від 54 п.н. до 69 п.н., а загальний відсоток їхніх послідовностей становить від 1,68 до 3,26 % довжини відповідних ретроінтронів. Рівні подібності становлять від 74,07 до 80,36 % (середнє значення — 77,55 %).

Каулімовіруси (Caulimoviridae) — родина ДНК-вмістних вірусів рослин із механізмом зворотної транскрипції і дволанцюговою ДНК, тобто є вірусами, які містять стадію зворотної транскрипції у своєму реплікативному циклі. Оскільки ці віруси мають багато спільного із ретровірусами, їх часто об'єднують у групу параретровірусів. Проте тут є важлива відмінність між ретровірусами і вірусами родини Caulimoviridae. На відміну від ретровірусів вони не інтегрують свій геном до

хромосоми, оскільки їхній геном не кодує ферменту інтегрази (Vassiliev et al., 2023).

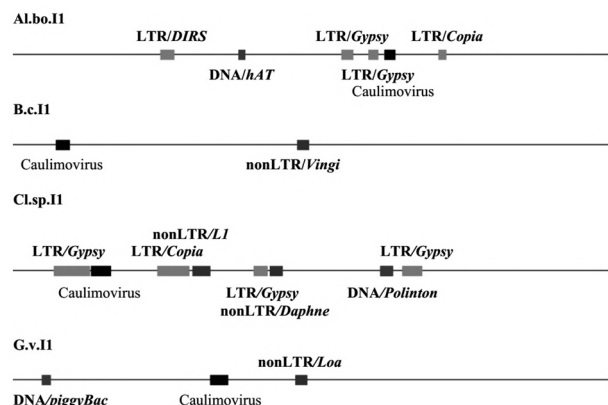


Рис. 5. Фрагменти каулімовірусів у бактеріальних інтронах II групи.

Отже, ми проаналізували 90 послідовностей бактеріальних інтронів групи II. У восьми ретроінтронах не виявили фрагментів еукаріотних МГЕ (Ba.fr.I1, D.h.I5, D.p.I1, E.a.I1, E.c.I3, Ha.ch.I1, La.re.I1, Ma.mr.I3). Найбільша група ретроінтронів ($n = 30$) містить МГЕ різних класів. Фрагменти ретротранспозонів обох підкласів ідентифікували у шістнадцяти ретроінтронах. Також виявили послідовності бактеріальних інтронів II групи, які містять фрагменти або лише nonLTR-ретроелементів ($n = 16$), або лише LTR-ретроелементів ($n = 11$). Фрагменти лише ДНК-транспозонів ідентифікували у п'яти ретроінтронах. У чотирьох бактеріальних інтронів II групи виявили фрагменти каулімовірусів.

У проаналізованих послідовностях бактеріальних ретроінтронів виявлено фрагменти послідовностей обох класів еукаріотних МГЕ у різноманітних комбінаціях (non-LTR, LTR, DNA, non-LTR+LTR, non-LTR+DNA, LTR+DNA, non-LTR+LTR+DNA). Серед non-LTR ретроелементів переважають фрагменти L1 елементів, серед LTR ретроелементів фрагменти Gypsy послідовностей, а серед фрагментів ДНК-транспозонів у досліджених ретроінтронах домінують фрагменти hAT, MuDR і Helitron елементів. Середня довжина фрагментів становить 54 п.н. (мінімальна 27 п.н., максимальна 182 п.н.), а середній відсоток щодо загальної довжини ретроінтрону — 6,8 (мінімальний 1,08 %, максимальний 21,63 %). Щодо рівнів подібності, то їхній відсоток суттєво не відрізняється у різних класах МГЕ. У елементів класу I, зокрема у non-LTR ретроелементів — 77,40 %, у LTR ретроелементів — 79,71 % а у елементів класу II (ДНК-транспозонів) — 78,91 %. Цікаво, що відсоток фрагментів non-LTR ретроелементів збігається із відсотком фрагментів ДНК-транспозонів. А максимальні відсотки еукаріотних фрагментів у ретроінтронах виявили для LTR ретроелементів,

у комбінаціях non-LTR+LTR та non-LTR+LTR+DNA, non-LTR+DNA, LTR+DNA.

Висновки

На основі одержаних результатів можна зробити висновок, що бактеріальні ретроінтрони можуть бути не лише предками non-LTR ретроелементів, але і носіями фрагментарних нуклеотидних послідовностей для формування різноманітних еукаріотних мобільних генетичних елементів.

Перелік літератури:

1. Almojil D., Bourgeois Y., Fali M., Hariyani I., Wilcox J., Boissinot S. The structural, functional and evolutionary impact of transposable elements in eukaryotes. *Genes*. 2021. Vol. 12, № 6. P. 918. doi: 10.3390/genes12060918.
2. Alzohairy A. M., Gyulai G., Jansen R. K., Bahieldin A. Transposable elements domesticated and neofunctionalized by eukaryotic genomes. *Plasmid*. 2013. Vol. 69, № 1. P. 1–15. doi: 10.1016/j.plasmid.2012.08.001.
3. Bao W., Kapitonov V. V., Jurka J. Ginger DNA transposons in eukaryotes and their evolutionary relationships with long terminal repeat retrotransposons. *Mob. DNA*. 2010. Vol. 1, № 1. P. 3. doi: 10.1186/1759-8753-1-3.
4. Bao W., Kojima K. K., Kohany O. Repbase Update, a database of repetitive elements in eukaryotic genomes. 2015. *Mobile DNA*. Vol. 6. P. 11–17. doi: 10.1186/s13100-015-0041-9.
5. Candaes M. A., Duong A., Hood K. S., Li T., Neufeld R. A., Sun R., McNeil B. A., Wu L., Jarding A. M., Zimmerly S. Database for bacterial group II introns. *Nucleic Acids Res.* 2012. Vol. 40, № 1. D187-90. doi: 10.1093/nar/gkr1043.
6. Capy P., Langin T., Higuete D., Maurer P., Bazin C. Do the integrases of LTR-retrotransposons and class II element transposases have a common ancestor? *Genetica*. 1997. Vol. 100. P. 63–72. doi: 10.1023/A:1018300721953.
7. Dai L., Zimmerly S. Compilation and analysis of group II intron insertions in bacterial genomes: evidence for retroelement behavior. *Nucleic Acids Res.* 2002. Vol. 30, № 5. P. 1091–1102. doi: 10.1093/nar/30.5.1091.
8. Dai L., Toor N., Olson R., Keeping A., Zimmerly S. Database for mobile group II introns. *Nucleic Acids Research*. 2003. Vol. 31, № 1. P. 424–426. doi: 10.1093/nar/gkg049.
9. Danilevskaya O. N., Traverse K. L., Hogan N. C., DeBaryshe P. G., Pardue M. L. The two *Drosophila* telomeric transposable elements have very different patterns of transcription. *Mol. Cell. Biol.* 1999. Vol. 19, № 1. P. 873–881. doi: 10.1128/mcb.19.1.873.
10. Fedoroff N. V. How transposition was discovered. *Nat. Struct.* 2001. Vol. 8, № 4. P. 300–301. doi: 10.1038/86161.
11. Feschotte C., Pritham E. J. DNA transposons and the evolution of eukaryotic genomes. *Annu. Rev. Genet.* 2007. Vol. 41. P. 331–368. doi:10.1146/annurev.genet.40.110405.090448.
12. Fischer M. G., Suttle C. A. A virophage at the origin of large DNA transposons. *Science*. 2011. Vol. 332, № 6026. P. 231–234. doi: 10.1126/science.1199412.
13. Garavís M., González C., Villasante A. On the origin of the eukaryotic chromosome: the role of noncanonical DNA structures in telomere evolution. *Genome Biol. Evol.* 2013. Vol. 5, № 6. P. 1142–1150. doi: 10.1093/gbe/evt079.
14. Hellen E. H. B., Brookfield J. F. Y. Transposable element invasions. *Mob. Genet. Elements*. 2013. Vol. 3, № 1. e23920. doi: 10.4161/mge.23920.
15. Hickey D. A. Evolutionary dynamics of transposable elements in prokaryotes and eukaryotes. *Genetica*. 1992. Vol. 86, № 1–3. P. 269–274. doi: 10.1007/BF00133725.

16. Hickman A. B., Chandler M., Dyda F. Integrating prokaryotes and eukaryotes: DNA transposases in light of structure. *Biochem. Mol. Biol.* 2010. ol. 45, № 1. P. 50–69. doi: 10.3109/10409230903505596.
17. Kapitonov V. V., Jurka J. Rolling circle transposons in eukaryotes. *Proc. Natl. Acad. Sci. USA.* 2001. Vol. 98, № 15. P. 8714–8719. doi: 10.1073/pnas.151269298.
18. Kapitonov V. V., Jurka J. Self-synthesizing DNA transposons in eukaryotes. *Proc. Natl. Acad. Sci. USA.* 2006. Vol. 103, № 12. P. 4540–4545. doi: 10.1073/pnas.0600833103.
19. Kennell J. C., Moran J. V., Perlman P. S., Butow R. A., Lambowitz A. M. Reverse transcriptase activity associated with maturase-encoding group II introns in yeast mitochondria. *Cell.* 1993. Vol. 73, № 1. P. 133–146. doi: 10.1016/0092-8674(93)90166-n.
20. Kim A., Terzian C., Santamaria P., Pe'lisson A., Purd'homme N., Bucheton A. Retroviruses in invertebrates: the gypsy retrotransposon is apparently an infectious retrovirus of *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA.* 1994. Vol. 91, № 4. P. 1285–1289. doi: 10.1073/pnas.91.4.1285.
21. Kohany O., Gentles A. J., Hankus L., Jurka J. Annotation, submission and screening of repetitive elements in Repbase: RepbaseSubmitter and Censor. *BMC Bioinformatics.* 2006. Vol. 7. P. 474. doi: 10.1186/1471-2105-7-474.
22. Kojima K. K. Structural and sequence diversity of eukaryotic transposable elements. *Genes and Genetic Systems.* 2019. Vol. 94, № 6. P. 233–252. doi: 10.1266/ggs.18-00024.
23. Lambowitz A. M., Belfort M. Mobile bacterial group II introns at the crux of eukaryotic evolution. *Microbiol Spectr.* 2015. Vol. 3, № 1. MDNA3–MDNA0050. doi:10.1128/microbiolspec.MDNA3-0050-2014.
24. Liorens C., Muñoz-Pomer A., Bernad L., Botellan H., Moya A. Network dynamics of eukaryotic LTR retroelements beyond phylogenetic trees. *Biol. Direct.* 2009. Vol. 4. P. 41. doi: 10.1186/1745-6150-4-41.
25. Malik H. S., Burke W. D., Eickbush T. H. The age and evolution of non-LTR retrotransposable elements. 1999. *Mol. Biol. Evol.* Vol. 16, № 6. P. 793–805. doi: 10.1093/oxfordjournals.molbev.a026164.
26. Malik H. S., Eickbush T. H. Phylogenetic analysis of ribonuclease H domains suggests a late, chimeric origin of LTR retrotransposable elements and retroviruses. *Genome Res.* 2001. Vol. 11, № 7. P. 1187–1197. doi: 10.1101/gr.185101.
27. McClure M. A. Evolution of retroposons by acquisition or deletion of retrovirus-like genes. *Mol. Biol. Evol.* 1991. Vol. 8, № 6. P. 835–856. doi: 10.1093/oxfordjournals.molbev.a040686.
28. Novikova O. S., Blinov A. G. Origin, evolution, and distribution of different groups of non-LTR retrotransposons among eukaryotes. *Genetika.* 2009. Vol. 45, № 2. P. 149–159. doi:10.1134/S102279540902001X.
29. Novikova O., Belfort M. Mobile group II introns as ancestral eukaryotic elements. *Trends Genet.* 2017. Vol. 33, № 11. P. 773–783. doi: 10.1016/j.tig.2017.07.009.
30. Pardue M. L., DeBaryshe P. G. *Drosophila* telomeres: two transposable elements with important roles in chromosomes. *Genetica.* 1999. Vol. 107, № 1–3. P. 189–196. doi: 10.1023/A:1003905210770.
31. Pritham E. J., Putliwala T., Feschotte C. Mavericks, a novel class of giant transposable elements widespread in eukaryotes and related to DNA viruses. *Gene.* 2007. Vol. 390, № 1–2. P. 3–17. doi: 10.1016/j.gene.2006.08.008.
32. Ray D. A., Walker J. A., Batzer M. A. Mobile element-based forensic genomics. *Mutat. Res.* 2007. Vol. 616, № 1–2. P. 24–33. doi: 10.1016/j.mrfmmm.2006.11.019.
33. Ross K., Varani A. M., Snesrud E., Huang H., Alvarenga D. O., Zhang J., Wu C., McGann P., Chandler M. TnCentral: a prokaryotic transposable element database and web portal for transposon analysis. *mBio.* 2021. Vol. 12, № 5. e0206021. doi: 10.1128/mBio.02060-21.
34. Stocking C., Kozak C. A. Endogenous retroviruses. *Cell. Mol. Life Sci.* 2008. Vol. 65, № 21. P. 3383–3398. doi: 10.1007/s00018-008-8497-0.
35. Toro N., Martínez-Abarca F. Comprehensive phylogenetic analysis of bacterial group II intron-encoded ORFs lacking the DNA endonuclease domain reveals new varieties. *PLoS One.* 2013. Vol. 8, № 1. e55102. doi:10.1371/journal.pone.0055102.
36. Vassiljeff H., Geering A. D. W., Choisne N., Teycheney P. Y., Maumus F. 2023. Endogenous caulimovirids: fossils, zombies, and living in plant genomes. *Biomolecules.* 2023. Vol. 13, № 7. P. 1069. doi: 10.3390/biom13071069.
37. Wells J. N., Feschotte C. A field guide to eukaryotic transposable elements. *Annu. Rev. Genet.* 2020. Vol. 54. P. 539–561. doi: 10.1146/annurev-genet-040620-022145.
38. Werren J. H. Selfish genetic elements, genetic conflict, and evolutionary innovation. *Proc. Natl. Acad. Sci. USA.* 2011. Vol. 108, № 2. P. 10863–10870. doi: 10.1073/pnas.110234311.
39. Wicker T., Sabot F., Hua-Van A., Bennetzen J. L., Capy P., Boulos C., Flavell A., Leroy P., Morgante M., Panaud O., Paux E., SanMiguel P., Schulman A. H. A unified classification system for eukaryotic transposable. *Nat. Rev. Genet.* 2007. Vol. 8, № 12. P. 973–982. doi: 10.1038/nrg2165.
40. Zimmerly S., Hausner G., Wu X. Phylogenetic relationships among group II intron ORFs. *Nucleic Acids Res.* 2001. Vol. 29, № 5. P. 1238–1250. doi: 10.1093/nar/29.5.1238.
41. Zimmerly S., Semper C. Evolution of group II introns. *Mobile DNA.* 2015. Vol. 6. P. 7. doi: 10.1186/s13100-015-0037-5.

ANALYSIS OF NUCLEOTIDE SEQUENCES OF BACTERIAL RETROINTRONS FOR THE PRESENCE OF HOMOLOGY TO EUKARYOTIC MGE

O. V. Pidpala, L. L. Lukash

Institute of Molecular Biology and Genetics of Natl. Acad. Sci. of Ukraine, Ukraine, 03143, Kyiv, Akad. Zabolotnogo str., 150 e-mail: pidpala@ukr.net

Aims. To analyze the sequences of bacterial retrointrons for the presence of homology to representatives of various classes of eukaryotic mobile genetic elements (MGE). **Methods.** Nucleotide sequences of bacterial group II introns were taken for analysis from the Database for Bacterial Group II Introns. The results of the search and identification of MGE have been realized by CENSOR program. **Results.** In the analyzed sequences of bacterial retrointrons, fragments of sequences of both classes of eukaryotic MGEs were found in various combinations: non-LTR; LTR; DNA; non-LTR+LTR; non-LTR+DNA; LTR+DNA; non-LTR+LTR+DNA. The largest group of retrointrons contains MGEs of various classes. **Conclusions.** Bacterial retrointrons can be not only ancestors of non-LTR retroelements, but also carriers of fragmentary nucleotide sequences for the formation of various eukaryotic MGEs.

Keywords: retroelements, retrointrons, eukaryotic MGEs.