

ПІДПАЛА О. В.✉, ЛУКАШ Л. Л.

Інститут молекулярної біології та генетики НАН України,
Україна, 03680, м. Київ, вул. Заболотного, 150, e-mail: pidpala@ukr.net
✉ pidpala@ukr.net, (063) 433-92-78

ВИДОСПЕЦИФІЧНІ МОБІЛЬНІ ГЕНЕТИЧНІ ЕЛЕМЕНТИ У ГЕНІ РЕПАРАТИВНОГО ЕНЗИМУ *MGMT* ШИРОКОНОСИХ МАВП

Мета. Проаналізувати поширення видоспецифічних мобільних генетичних елементів (МГЕ) в ортологів гена *MGMT* широконосих мавп. **Методи.** Гомологію між нуклеотидними послідовностями визначали програмою *BLAST 2.6.1*. Результати пошуку та ідентифікації МГЕ здійснено за допомогою програми *CENSOR*. **Результати.** На прикладі ортологів гена *MGMT* широконосих мавп з'ясовано, що різні видоспецифічні МГЕ, ідентифіковані у їхніх інтронних послідовностях, можуть мати різну еволюційну хронологію. У випадку елемента *Alu2_TS*, який виник у представника довгоп'ятових, виявлено, що в еволюційно найближчих приматів він знає делеційної деградації, тоді як фрагменти людино-специфічного *L1Hs* елемента траплялися у геномах еволюційно віддалених приматів ще задовго до формування і виникнення цього ретроелемента. **Висновки.** Хронологія еволюційних змін гена та його видоспецифічних МГЕ може мати різний характер і відбуватися паралельно та незалежно.

Ключові слова: Platyrrhina, ген *MGMT*, МГЕ, *Alu2_TS*, *L1Hs*.

Ензим O⁶-метилгуанін-ДНК метилтрансфераза (*MGMT*) видаляє алкільні групи з O⁶-позиції гуаніну у ДНК, захищає клітини від їхнього токсичного та мутагенного впливів і належить до ферментів прямої репарації [1]. Небагато відомо про еволюцію такого ензиму [2]. Ще менше даних про еволюцію його гена, тому ми сконцентрували увагу на вивченні ортологів гена *MGMT* від протистів (найпростіших еукаріот) до людини. Аналізуючи структурні особливості та нуклеотидні послідовності ортологів гена *MGMT* у протистів, виявлено, що еволюційне походження інтронів, як і інших структурних одиниць гена, може мати ендегенний мозаїчний характер [3]. Особливий інтерес викликає дослідження еволюції цього гена названого у приматів із акцентом на участі у цьому процесі мобільних генетичних елементів (МГЕ). У своїх

попередніх роботах ми вивчали утворення людино-специфічного *L1Hs* елемента в інтроні 3 гена *MGMT* гоміноїдів [4] та аналізували ортологів гена *MGMT* на прикладі найдавніших приматів, а саме мокроносих (*Strepsirrhini*) [5]. Встановлено, що формування людино-специфічного *L1Hs* елемента відбувалось упродовж еволюції гоміноїдів паралельно із утворенням кластерної структури МГЕ у людинових із різних підродів *LINE1*-елементів, складові компоненти якого, очевидно, також задіяні в утворенні *L1Hs* елемента [4]. На основі порівняння послідовностей мокроносих приматів і *H.sapiens* висловлено припущення, що еволюційні зміни гена *MGMT* можуть відбуватися на рівні різних структурних одиниць, причому МГЕ можуть бути не лише компонентами інтронів, але і складовими екзонів у вигляді фрагментованих послідовностей [5]. Метою нашої роботи було проаналізувати поширення видоспецифічних МГЕ на прикладі ортологів гена *MGMT* широконосих мавп, або мавп Нового Світу (*Platyrrhina*).

Матеріали і методи

Інформацію про ген репаративного ензиму *MGMT* і про його нуклеотидні послідовності у людини (*Homo sapiens*) та широконосих мавп (*Platyrrhina*) – західноамазонської мірікіни (*Aotus nancymaae*), болівійського саймірі (*Saimiri boliviensis boliviensis*), капуцина білоголового (*Cebus capucinus imitator*) та ігрунки звичайної (*Callithrix jacchus*) – одержано із баз даних *Ensembl* та *OrthoDB*. Гомологію між досліджуваними послідовностями визначали програмою *BLAST 2.6.1*. Результати пошуку та ідентифікації МГЕ здійснено за допомогою програми *CENSOR*.

Результати та обговорення

Сучасні примати представлені двома великими еволюційними гілками – сухоносими (*Haplorrhini*) і мокроносими (*Strepsirrhini*) при-

Як видно із наведених даних (табл. 1), частка ДНК-транспозонів і LTR ретротранспозонів від широконосих мавп до людини не зазнає суттєвих змін (3,93–4,22 % і 1,92–2,49 %, відповідно), тоді як частка Non-LTR ретротранспозонів та загальний відсоток МГЕ дещо зростає.

Найбільший відсоток МГЕ в інтроні 3, а найменший в інтроні 4 (табл. 4), хоча за кількістю МГЕ переважають інтрон 2 та інтрон 1 (табл. 2), як і за довжиною. У межах інтронних послідовностей коливається не лише кількість і відсоток МГЕ, але і приналежність їх до різних класів мобільних елементів. Найлегше це простежити на прикладі інтрона 4 (табл. 2, рис. 2).

В усіх досліджених послідовностях ортологів гена *MGMT* як широконосих мавп, так і людини, у межах цього інтрона присутні фрагменти LTR-повторів ендегенного ретровірусу ERV3 – *LTR79*. Різні фрагменти ДНК-транспозонів hAT надродина у інтроні 4 широконосих мавп утворюють кластер, тоді як у людини присутня послідовність однієї із них – *CHESHIRE_A*. Як видно із наведених даних, у ігрунки звичайної (*Callithrix jacchus*) у межах окресленого інтрону відсутній фрагмент послідовності Non-LTR ретротранспозону *L2*. Цікаво, що у деяких представників широконосих мавп у інтроні 4 присутні фрагменти видоспецифічних послідовностей: у капуцина білоголового (*Cebus capucinus imitator*) – *L1B_Mim*, Non-LTR ретротранспозону *L1 Microcebus murinus*; у болівійського саймірі (*Saimiri boliviensis boliviensis*) – *6kbHsap*, тандемний повтор людини.

МГЕ у різних видів приматів умовно поділяють на дві групи. Перша – молоді видоспецифічні повтори, серед них *Alu*-повтори приматів. Друга група – древні мобільні елементи, які є не лише у ссавців, але й у низки хордових. До них належать послідовності *MIR*, *L2*, *L3* і деякі ДНК-транспозони. Цікаво простежити за долею деяких видоспецифічних МГЕ на прикладі ортологів гена *MGMT*.

Видоспецифічні МГЕ в інтронах ортологів гена MGMT широконосих мавп. У проаналізованих послідовностях гена *MGMT* широконосих мавп ідентифіковано фрагменти МГЕ, які специфічні не лише для представника цього парвораду *Callithrix jacchus*, але і для представників різних приматів – від найпростіших мокроносих мавп (лемуроподібних і лоріподібних), довгоп'ятоподібних до вузьконосих мавп (рис. 3).

Наприклад, в ігрунки звичайної (*Callithrix jacchus*) в інтронах виявлено послідовність *AluY*-повтору (інтрон 1), який виник 20 млн років тому і характерний для вузьконосих мавп, та ідентифіковано фрагмент людино-специфічного *L1Hs* елемента (інтрон 2). Також в інтроні 3 капуцина білоголового (*Cebus capucinus imitator*) виявлено фрагмент ендегенного ретровірусу *MacERV1_int*, який характерний для представників родини мавпових (Cercopithecidae). Крім того, як зазначалося вище (рис. 2), в інтроні 4 болівійського саймірі (*Saimiri boliviensis boliviensis*) ідентифіковано фрагмент тандемного повтору людини *6kbHsap*.

Таблиця 1. Частка МГЕ у гені *MGMT* людини і *Platyrrhina*

Організм, вид	Кількість МГЕ, %			
	Загальна (n*)	ДНК-транспозони	LTR ретротранспозони	Non-LTR ретротранспозони
<i>Homo sapience</i>	30,10 (n=250)	4,22	2,49	23,39
<i>Aotus nancymaae</i>	27,07 (n=243)	4,28	1,90	16,98
<i>Saimiri boliviensis boliviensis</i>	25,65 (n=226)	4,18	2,09	19,39
<i>Cebus capucinus imitator</i>	25,44 (n=227)	4,25	2,09	19,09
<i>Callithrix jacchus</i>	25,73 (n=224)	3,93	1,92	19,87

Примітка. *У дужках наведено кількість мобільних генетичних елементів.

Таблиця 2. Розподіл МГЕ (%) у інтронах гена *MGMT* людини і *Platyrrhina*

Організм, вид	Інтрони			
	1	2	3	4
<i>Homo sapience</i>	28,02 (n=61)	27,74 (n=142)	44,75 (n=44)	8,26 (n=3)
<i>Aotus nancymaae</i>	28,62 (n=64)	26,12 (n=140)	33,52 (n=34)	6,29 (n=4)
<i>Saimiri boliviensis boliviensis</i>	27,66 (n=60)	24,04 (n=131)	33,50 (n=30)	6,47 (n=5)
<i>Cebus capucinus imitator</i>	29,21 (n=62)	23,24 (n=126)	33,65 (n=33)	8,52 (n=5)
<i>Callithrix jacchus</i>	29,46 (n=57)	24,45 (n=131)	32,35 (n=35)	6,55 (n=3)

Примітка. У дужках наведено кількість мобільних генетичних елементів.

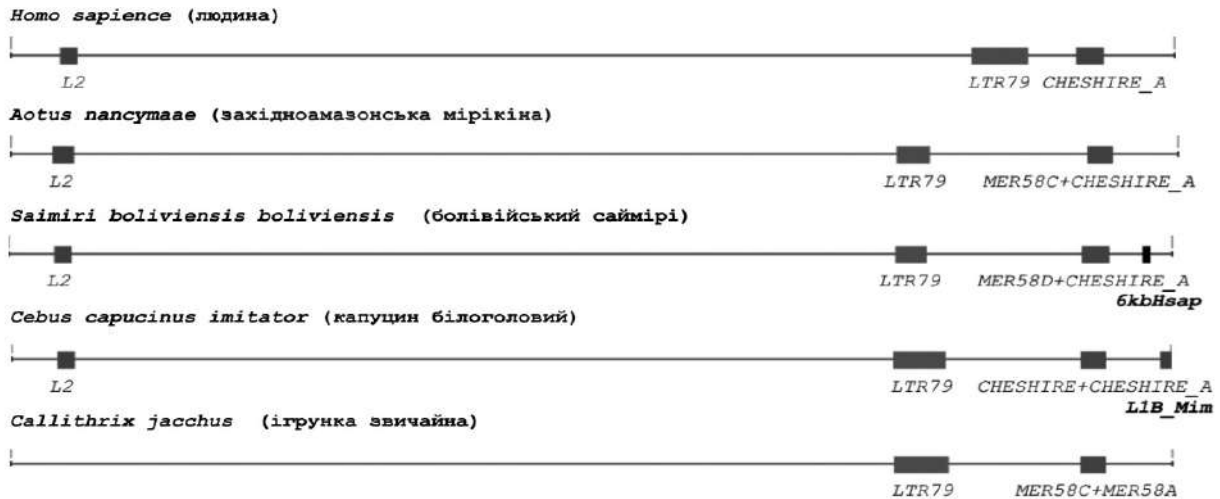


Рис. 2. Повторювані послідовності в інтроні 4 ортологів гена *MGMT* людини і широконосих мавп. Жирним шрифтом зазначено видоспецифічні послідовності.

специфічні для			
микроносих	довгоп'ят	широконосих	вужконосих
<i>LIB_Mim</i>	<i>Alu2_TS</i>	<i>L1-1_Cja</i>	<i>AluY</i>
<i>MLT1_Mim</i>	<i>L1-2_TS</i>	<i>L1-2_Cja</i>	<i>L1HS</i>
<i>LTR9_OG</i>	<i>L1-3_TS</i>	<i>HAL1-1B_Cja</i>	<i>MacERV1_int</i>
<i>hAT-2N1_MM</i>	<i>L1-4_TS</i>	<i>HAL1-1C_Cja</i>	
	<i>ERV2-2_TSy-I</i>	<i>HAL1-1_Cja</i>	
	<i>hAT-2N2_TS</i>	<i>ERV1-2_CJa-I</i>	
	<i>hAT-2N3_TS</i>	<i>ERV1-3_CJ-I</i>	
		<i>ERV2-1_CJ-I</i>	

Рис. 3. Примато-специфічні МГЕ в інтронах гена *MGMT* широконосих мавп.

За різними оцінками, родина *LINE1*-елементів (або *L1*) виникла 100–170 млн років тому перед поділом ссавців на порядки, і її представники розповсюдились по їхніх геномах [11–13]. Родину *L1* поділяють на підродини: *L1Ns* (*L1PA1*), *L1PA2-16*, *L1PB1-3*, *L1MA1-10*, *L1MB1-8*, *L1MC*, *L1MD*, *L1ME* [14]. Наймолодша група – *L1PA1* (або *L1Ns* чи *L1Ta*) – виникла близько 4 млн років тому, а пік ретропозицій її представників у геномі людини був приблизно 3 млн років тому [13–15]. Ретротранспозони груп *L1PA* (6–15) розповсюджені і у мавп Старого Світу, і у мавп Нового Світу, а представники груп *L1PA* (15–16), *L1MA* (1–3) – в усіх приматів.

Формування людино-специфічного *L1HS* елемента, очевидно, відбувалось упродовж еволюції гоміноїдів паралельно із утворенням кластерної структури МГЕ у людинових із різних підродин *LINE1*-елементів, складові компоненти якого, очевидно, також задіяні в утворенні *L1HS* елемента [4]. Ідентифікація фрагмента *L1HS* елемента у представника широконосих мавп дозволяє припустити, що хронологія виникнен-

ня фрагментів людино-специфічного *L1HS* елемента може мати ширші межі і відбуватись упродовж еволюції мавпоподібних (*Simiiformes*).

Серед МГЕ, які специфічні для широконосих мавп, в ортологів гена *MGMT* виявлено лише фрагменти послідовностей, специфічні для *Callithrix jacchus* (рис. 3). Цікаво, що інтронні послідовності досліджуваного гена західноамазонської мірікіни (*Aotus nancymaе*), болівійського саймірі (*Saimiri boliviensis boliviensis*) і капуцина білоголового містять не лише МГЕ, які присутні в інтронах ігрунки звичайної (*Callithrix jacchus*), але і такі, які не ідентифіковані у межах її гена (рис. 4).

Переважна більшість виявлених МГЕ має гомологічні послідовності однакової орієнтації, схожі розміри та положення серед інших МГЕ. Винятком є видоспецифічний *Alu*-повтор довгоп'ята філіппінського (*Tarsius syrichta*, *Carlito syrichta*) *Alu2_TS*. *Alu*-повтори (довжина близько 300 п. н.) є пасивними МГЕ та належать до класу SINE елементів, які походять від 7SL РНК і трапляються лише у приматів [16].

- Callithrix jacchus*
- інтрон 1 **L1-1_Cja**
 - інтрон 3 **ERV2-1_CJ-I**
- Aotus nancymaae*
- інтрон 1 **L1-1_Cja**
 - інтрон 2 **L1-1_Cja** HAL1-1B_Cja
HAL1-1_Cja L1-2_Cja
- Saimiri boliviensis*
- інтрон 1 **L1-1_Cja** ERV1-2_CJa-I
 - інтрон 2 **L1-1_Cja** L1-2_Cja
- Cebus capucinus imitator*
- інтрон 1 **L1-1_Cja** HAL1-1C_Cja
HAL1-1_Cja
 - інтрон 2 **L1-1_Cja** ERV1-3_CJ-I
L1-2_Cja
- L1-1_Cja**
L1-2_Cja
HAL1-1B_Cja
HAL1-1C_Cja
HAL1-1_Cja
ERV1-2_CJa-I
ERV1-3_CJ-I
ERV2-1_CJ-I

Рис. 4. Фрагменти видоспецифічних для *Callithrix jacchus* МГЕ в інтронах ортологів гена *MGMT* широконосих мавп. Жирним шрифтом виділено МГЕ, які присутні у досліджуваному гені *Callithrix jacchus*.

Їх поділяють на три основні родин – древні (J), середні (S) і молоді (Y) та на 14 основних підродин, які відрізняються еволюційним віком [17]. Вік найстаріших *AluJ*-повторів оцінюють у 50–80 млн років, тоді як час виникнення *AluS* і *AluY* відповідно становить приблизно 35 і 20 млн років тому. *Alu*-повтори містять значну кількість точкових мутацій, так званих «діагностичних» мутацій, спектр яких характерний для певної підродини [18].

У гені *MGMT* довгоп'ята філіппінського присутні дві майже повнорозмірні послідовності *Alu2_TS* елемента із координатами 1–278 і 12–283 п. н. у позитивній орієнтації. Довжина вихідної *Alu2_TS* послідовності становить 283 п. н. У послідовностях інтрона 1 досліджуваних широконосих мавп він змінює не лише своє положення серед інших МГЕ, розмір послідовності, але й орієнтацію (рис. 5). Фрагменти *Alu2_TS* менших розмірів (93–97 п. н.) виявлено також і у проаналізованих послідовностях інтрона 2.

На філогенетичному дереві приматів довгоп'яти є проміжною ланкою між мокроносими мавпами (*Strepsirhini*, наприклад, лемурами, лорі, галаго) та мавпоподібними (*Simiiformes*, зокрема людиноподібні, включаючи і людей). Разом із широконосими і вузьконосими мавпами не так давно довгоп'ятів зараховано до підзагону сухоносих приматів (*Haplorhini*). Велику роль у вирішенні цього питання відіграли геномні дослідження, в першу чергу аналіз *Alu*-повторів, які є інформативними філогенетичними маркерами і широко використовуються для еволюційного вивчення приматів [19].

В інтронних послідовностях західноамазонської мірікіни (*Aotus nancymaae*) також ідентифіковано фрагменти видоспецифічних *LINE1*-

елементів довгоп'ята філіппінського *L1-4_TS* (в інтроні 1), *L1-3_TS* (в інтроні 2), *L1-2_TS* (в інтроні 3). Серед видоспецифічних МГЕ довгоп'ятових, крім фрагментів Non-LTR ретротранспозонів, ідентифіковано послідовності видоспецифічного ретровірусу *ERV2-2_TSy-I* (у *Cebus capucinus imitator*, інтрон 1) та ДНК-транспозонів *hAT-2N3_TS* (у *Saimiri boliviensis boliviensis*, інтрон 2) і *hAT-2N2_TS* (у *Saimiri boliviensis boliviensis*, інтрон 2). Щодо послідовностей МГЕ, характерних для представників мокроносих мавп (рис. 3), то в інтронах ортологів гена *MGMT* у широконосих мавп виявлено фрагменти МГЕ, які специфічні для мишачого мікроцебусу (*Microcebus murinus*) – *L1B_Mim*, *MLT1_Mim* та *hAT-2N1_MM* і для галаго Гарнетта (*Otolemur garnettii*) – *LTR9_OG*.

Підсумовуючи одержані дані, можна висловити припущення щодо ендегенного походження та рекомбіногенного характеру формування 5'UTR у досліджуваного гена *Cebus capucinus imitator*. Також на прикладі двох видоспецифічних послідовностей МГЕ – *Alu*-повтору довгоп'ята філіппінського *Alu2_TS* і людино-специфічного *L1Hs* елемента, можна простежити їхню еволюційну хронологію. Елемент *Alu2_TS* виник у представника довгоп'ятових, а у широконосих мавп зазнає делеційної деградації. Що стосується людино-специфічного *L1Hs* елемента, то його окремі фрагменти трапляються у геномах еволюційно віддалених приматів ще задовго до формування і виникнення цього ретроелемента. Отже, еволюція видоспецифічних МГЕ, можливо, може відбуватися паралельно і незалежно від еволюції гена.



Рис. 5. Положення фрагментів видоспецифічного *Alu*-повтору довгоп'ята філіппінського *Alu2_TS* серед інших МГЕ в інtronі 1 гена *MGMT* широконосих мавп.

Висновки

Хронологія еволюційних змін гена *MGMT* та його видоспецифічних МГЕ може мати різний характер і відбуватися паралельно та незалежно.

References

- Pegg A.E. Repair of O(6)-alkylguanine by alkyltransferases. *Mutat.Res.* 2000. Vol. 462, № 2–3. P. 83–100.
- Margison G.P., Butt A., Pearson S.J., Wharton S., Watson A.J., Marriott A., Caetano C.M., Hollins J.J., Rukazenkova N., Begum G., Santibanez-Koref M.F. Alkyltransferase-like proteins. *DNA Repair (Amst)*. 2007. Vol. 6, № 8. P. 1222–1228.
- Pidpala O.V., Lukash L.L. Analysis of human *MGMT* gene orthologous in Protists. *Factors in Experimental Evolution of Organisms*. 2018. Vol. 22. P. 345–351. [in Ukrainian]
- Pidpala O.V., Lukash L.L. Formation of the *L1Hs* retroelement in the intron of the *MGMT* gene of Hominoidea. *Factors in Experimental Evolution of Organisms*. 2019. Vol. 24. P. 338–344. [in Ukrainian]
- Pidpala O.V., Lukash L.L. *In silico* analysis of *MGMT* gene orthologous in the most ancient Mammals Strepsirrhini. *Factors in Experimental Evolution of Organisms*. 2020. Vol. 26. P. 305–310. [in Ukrainian]
- Perelman P., Johnson W.E., Roos C., Seanez H.N., Horvath J.E., Moreira M.A., Kessing B., Pontius J., Roelke M., Rumpler Y., Schneider M.P., Silva A., O'Brien S.J., Pecon-Slattery J. A molecular phylogeny of living primates. *PLoS Genet.* 2011. Vol. 7, № 3. e1001342. doi: 10.1371/journal.pgen.1001342.
- Chenais B., Caruso A., Hiard S., Casse N. The impact of transposable elements on eukaryotic genomes: from genome size increase to genetic adaptation to stressful environments. *Gene*. 2012. Vol. 509, № 1. P. 7–15. doi: 10.1016/j.gene.2012.07.042.
- 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. P. 161–177. doi: 10.1093/gbe/evw264.
- de Koning A.P., Gu W., Castoe T.A., Batzer H.A., Pollock D.D. Repetitive elements may comprise over two-thirds of the human genome. *PLoS Genet.* 2011. Vol. 7, № 12. e1002384. doi: 10.1371/journal.pgen.1002384.
- Pidpala O., Lukash L. Regulatory potential of mobile genetic elements in the human *MGMT* gene. *J. Genet. Genomic Sci.* 2018. Vol. 3. P. 008–015. doi: 10.24966/GGS-2485/100008.
- Kazazian H.H.Jr., Moran J.V. The impact of L1 retrotransposons on the human genome. *Nat. Genet.* 1998. Vol. 19, № 1. P. 19–24.
- Furano A.V. The biological properties and evolutionary dynamics of mammalian LINE-1 retrotransposons. *Prog. Nucleic Acid Res. Mol. Biol.* 2000. Vol. 64. P. 255–294.
- Boissinot S., Chevret P., Furano A.V. L1 (LINE-1) retrotransposon evolution and amplification in recent human history. *Mol. Biol. Evol.* 2000. Vol. 17, № 6. P. 915–928.
- Smit A.F., Tyth G., Riggs A.D., Jurka J. Ancestral, mammalian-wide subfamilies of LINE-1 repetitive sequences. *J. Mol. Biol.* 1995. Vol. 246, № 3. P. 401–417.
- Cantrell M.A., Grahn R.A., Scott L., Wichman H.A. Isolation of markers from recently transposed LINE-1 retrotransposons. *Biotechniques*. 2000. Vol. 29, № 6. P. 1310–1316.
- Batzer M.A., Deininger P.L. Alu repeats and human genomic diversity. *Nature Rev. Genet.* 2002. Vol. 6, № 3. P. 370–379.
- Kapitonov V., Jurka J. The age of Alu subfamilies. *J. Mol. Evol.* 1996. Vol. 42. P. 59–65.
- Roy-Engel A.M., Carroll M.L., El-Sawy M., Salem A.H., Garber R.K., Nguyen S.V., Deininger P.L., Batzer M.A. Non-traditional Alu evolution and primate genomic diversity. *J. Mol. Biol.* 2002. Vol. 316, № 5. P. 1033–1040.
- Schmitz J., Noll A., Raabe C.A., Churakov G., Voss R., Kiefmann M., Rozhdestvensky T., Brosius J., Baertsch R., Clawson H., Roos C., Zimin A., Minx P., Montague M.J., Wilson R. K., Warren W. C. Genome sequence of the basal haplorhine primate *Tarsius syrichta* reveals unusual insertions. *Nat. Commun.* 2016. Vol. 7. P. 12997. doi: 10.1038/ncomms12997.

PIDPALA O.V., LUKASH L.L.

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

SPECIES-SPECIFIC MOBILE GENETIC ELEMENTS IN THE GENE OF REPAIR ENZYME *MGMT* IN NEW WORLD MONKEYS

Aim. To analyze the distribution of species-specific mobile genetic elements (MGE) in orthologs of the *MGMT* gene in Platyrrhina. **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 *CENSOR* program. **Results.** On the example of orthologs of the *MGMT* gene in New World monkeys, it has been shown that different species-specific MGE identified in their intron sequences may have different evolutionary chronologies. In the case of the element *Alu2_TS*, which originated in the Tarsiiformes representative, it was found that in evolutionarily close primates it undergoes deletion degradation, while fragments of the human-specific *L1Hs* element are found in the genomes of evolutionarily distant primates long before the formation and emergence of this retroelement. **Conclusions.** The chronology of evolutionary changes in the gene *MGMT* and its species-specific MGE can be of different nature and occur in parallel and independently.
Keywords: Platyrrhina, *MGMT* gene, MGE, *Alu2_TS*, *L1Hs*.