

ЕВОЛЮЦІЯ ГЕНА *MGMT* У ПРИМАТІВ: ВІД СТРУКТУРИ ДО МОБІЛЬНИХ ГЕНЕТИЧНИХ ЕЛЕМЕНТІВ

О. В. ПІДПАЛА, Л. Л. ЛУКАШ

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

У статті зібрано дані щодо еволюції репаративного гена *MGMT* у приматів а також простежено за еволюційною долею низки мобільних генетичних елементів (МГЕ). Зроблено висновок, що еволюція гена *MGMT* у приматів зазнавала різноманітних змін від набуття до втрати екзон / інтронних послідовностей. Що стосується приматоспецифічних МГЕ, то вони мають різну еволюційну історію і шлях — від формування видо / родоспецифічних послідовностей до делеційної деградації і можуть бути компонентами не лише інтронних чи екзонних послідовностей але і складовими екзонів у вигляді фрагментованих послідовностей і не ідентифікуюватись як МГЕ.

Ключові слова: примати, ген *MGMT*, приматоспецифічні мобільні генетичні елементи.

Вступ. Генотип клітини зазнає різноманітних зовнішніх впливів, серед них УФ-випромінювання, іонізуюча радіація, хімічні мутагени, температурні коливання. Еволюційно клітина сформувала системи репарації для захисту і виправлень таких пошкоджень ДНК. Серед ферментів прямої репарації можна виокремити ензим O^6 -метилгуанін-ДНК метилтрансферазу (*MGMT*), який видаляє алкільні групи з O^6 -позиції гуаніну в ДНК і захищає клітини від їхнього токсичного та мутагенного впливів (Pegg, 2000). Відомо, що цей ензим також відіграє значну роль у алкілувальній хіміотерапії раку (Kaina et al., 2007; Pegg, 2011). *MGMT* широко розповсюджений у про- та еукаріотних організмів. Небагато відомо про еволюцію даного ензиму (Margison et al., 2007), ще менше даних про еволюцію цього гена. Ми проаналізували низку ортологів гена *MGMT* представників різних царств еукаріотних організмів, зокрема грибів (*S.cerevisiae*), тварин (*D.melanogaster*; *C.elegans*; *D.rerio*) (Підпала, Лукаш, 2017) а також протистів (*D.discoideum*; *D.lacteum*; *D.purpureum*; *G.niphandrodes*; *P.pallidum* PN500; *P.persalinus*; *S.arctica* JP610; *T.vaginalis* G3) (Підпала, Лукаш, 2018). Цікаво було прослідкувати за еволюцією даного гена на вищому еволюційному рівні — у приматів з акцентом на участі МГЕ у цьому процесі а також простежити за еволюційною долею низки приматоспецифічних МГЕ.

Філогенія приматів. Сучасні примати представлені двома великими еволюційними гілками — мокроносими (*Strepsirrhini*) і сухоносими (*Haplorhini*) приматами або мавпами. Обидві групи приматів мають багато спільних ознак і походять від єдиного предка, проте, розділились вони близько 87 млн років тому (Perelman et al., 2011). Мокроносі примати (нижчі примати або напівмаври) є еволюційно древнішим і примітивнішим підрядом. Вони об'єднують 3 надродина (інфраряди) — лемуриформиди (*Lemuriformes*), лорієвиди (*Lorisoidea*) та вимерлу групу ранніх приматів адапіди (*Adapiformes*). Лорієвиди відокремились від пов'язаних із ними лемурів, за різними даними, приблизно 50–80 млн років тому. Тоді як поділ лорієвидих на дві родини (лорієві та галагові) відбувся понад 40 млн років тому. Формування родів і видів лорієвих датують 37 млн років тому, а галагових — 19,9 млн років тому. Підряд сухоносих мавп об'єднує вищих або справжніх мавп (*Simiiformes*) і сестринську групу довгоп'ятів (*Tarsiiformes*), від яких вони відділились 77–80 млн років тому. На філогенетичному дереві приматів довгоп'яти (*Tarsiiformes*) є проміжною ланкою між мокроносими мавпами (*Strepsirrhini*) та мавповидими або вищими приматами (*Simiiformes*). У свою чергу вищих мавп поділяють на парворяд широконосих мавп, або мавп Нового Світу (*Platyrrhini*) та вузьконосих мавп, або мавп Старого Світу (*Catarrhini*).

Час поділу оцінюють 40–43 млн років тому. Мавпові (*Cercopithecidae*) — родина приматів з підряду сухоносих мавп (*Haplorhini*), яка нараховує близько 150 видів і є найрозгалуженішою родиною приматів (Perelman et al., 2011).

Разом із гоміноїдами (*Hominoidea*) вони входять до парворяду вузьконосих (*Catarrhini*).

Мавпові дуже різноманітні, але між крайніми формами є майже неперервні морфологічні переходи. Також, вони є однією із найеволюційніших груп мавп.

До родини мавпових належать дві підродини: мавпові (*Cercopithecinae*) та колобусові (*Colobinae*). Людиноподібні або гоміноїди чи антропоїди (*Hominoidea* або *Anthropomorphae*) це надродина вузьконосих мавп (*Catarrhini*), які мають будову тіла схожу до людської. Включає дві родини — людинових (*Hominidae*) та гібонових (*Hylobatidae*).

Хромосомна локалізація гена *MGMT* у приматів. Відомо, що ген *MGMT* людини локалізований на теломерній ділянці хромосоми 10 у положенні 10q26 у позитивній орієнтації (табл. 1). Що стосується інших представників гоміноїдів (*Hominoidea*), то відомо, що у представників родини людинових (*Hominidae*) ген *MGMT* також локалізова-

ний на теломерній ділянці хромосоми 10 у позитивній орієнтації, тоді як у гібона білощогого чубатого (*Nomascus leucogenys*), який належить до родини гібонових (*Hylobatidae*), зазначений ген локалізований на теломерній ділянці хромосоми 3 у зворотній орієнтації (табл. 1). Про хромосомну локалізацію ортологів гена *MGMT* у мавпових на сьогодні відомо лише для деяких представників. Досліджуваний ген локалізований на хромосомі 9 у позитивній орієнтації. Цікаво, що у тих випадках, коли для ортологів не ідентифіковано хромосомну локалізацію, останні мають зворотну орієнтацію (табл. 1). У широконосих мавп (*Platyrrhini*) про хромосомну локалізацію ортологів гена *MGMT* поки що відомо лише для *Callithrix jacchus*. Ген локалізований на хромосомі 12 також у позитивній орієнтації. Цікаво, що у *Saimiri boliviensis boliviensis* та у *Cebus capucinus imitator*, які належать до родини капуцинових, ген *MGMT* має зворотну орієнтацію. Не відомо про хромосомну локалізацію і для представника довгоп'ятових — довгоп'ята філіппінського (*Carlito syrichta*). У мокроносих приматів про хромосомну локалізацію гена *MGMT* на сьогодні відомо лише для *Microcebus murinus*. Він локалізований на теломерній ділянці хромосоми 15 (табл. 1).

Таблиця 1. Основна інформація про досліджувані ортологи гена *MGMT* приматів за даними *Ensembl*.

Організм	Хромосомна локалізація	Довжина гена, п.н.	Довжина мРНК, п.н.	Довжина білка, aa	Ек., n	Ін., n
<i>Homo sapiens</i>	10 (+)	300330	1256	238	5	4
<i>Pan troglodytes</i>	10 (+)	314257	717	238	5	4
<i>Pan paniscus</i>	10 (+)	303938	717	238	5	4
<i>Gorilla gorilla</i>	10 (+)	311423	717	238	5	4
<i>Pongo abelii</i>	10 (+)	311290	2411	238	5	4
<i>Nomascus leucogenys</i>	3 (-)	313929	717	238	5	4
<i>Chlorocebus sabaeus</i>	9 (+)	51414	2898	164	3	2
<i>Macaca nemestrina</i>	?	52539	684	227	4	3
<i>Macaca mulatta</i>	9 (+)	291453	827	238	5	4
<i>Macaca fascicularis</i>	9 (+)	292426	2864	238	5	4
<i>Theropithecus gelada</i>	9 (+)	290621	1894	238	5	4
<i>Mandrillus leucophaeus</i>	? (-)	65007	612	203	4	3
<i>Cercocebus atys</i>	? (-)	289992	670	207	5	4
<i>Papio anubis</i>	9 (+)	295177	1528	238	5	4
<i>Rhinopithecus bieti</i>	? (-)	247110	366	122	3	2
<i>Colobus angolensis palliatus</i>	? (+)	285732	717	238	5	4
<i>Ptilocolobus tephrosceles</i>	? (-)	284730	1955	241	6	5
<i>Aotus nancymaae</i>	? (+)	279053	717	238	5	4
<i>Saimiri boliviensis boliviensis</i>	? (-)	278372	717	238	5	4
<i>Cebus capucinus imitator</i>	? (-)	277434	740	238	5	4
<i>Callithrix jacchus</i>	? (+)	275654	717	238	5	4
<i>Carlito syrichta</i>	? (+)	61835	567	188	4	3
<i>Propithecus coquereli</i>	? (+)	286162	726	241	6	5
<i>Microcebus murinus</i>	15 (+)	286849	690	215	5	4
<i>Promomys simus</i>	? (+)	269603	729	242	5	4

Екзон-інтронна структура ортологів гена *MGMT* у приматів. У людини ген *MGMT* складається із п'яти кодуючих екзонів та чотирьох інтронів (табл. 1, 2). Така ж структура даного гена у досліджуваних представників гоміноїдів. У більшості мавпових ген *MGMT*, як і у людини, складається із п'яти екзонів та чотирьох інтронів. Хоча є випадки, коли ген *MGMT* має або меншу кількість структурних одиниць — чотири екзони і три інтрони як у випадку *Macaca nemestrina* та *Mandrillus leucophaeus* чи три екзони і два інтрони як у *Chlorocebus sabaeus* та *Mandrillus leucophaeus*, або ж більшу кількість

структурних одиниць на прикладі ортолога у *Ptilocolobus tephrosceles* (табл. 1, 2). У представників досліджуваних широконосих мавп ген *MGMT* теж складається із п'яти кодуючих екзонів та чотирьох інтронів. У *Carlito syrichta* даний ген має чотири екзони і три інтрони. Що стосується мокроносих приматів, то у *Microcebus murinus* і у *Prolemur simus* досліджуваний ген має п'ять кодуючих екзонів та чотири інтрони, тоді як у *Propithecus coquereli* ген складається із шести кодуючих екзонів та п'яти інтронів (табл. 1, 2).

Таблиця 2. Дані про екзон-інтронну структуру ортологів гена *MGMT* у приматів за даними *Ensembl*.

Організм	Довжина структурних одиниць, п.н.												
	Ек.*	Ін.*	Ек.1	Ін.1	Ек.2	Ін.2	Ек.3	Ін.3	Ек.4	Ін.4	Ек.5	Ін.**	Ек.**
<i>Homo sapiens</i>			107	68944	137	171517	149	51158	140	7446	732		
<i>Pan troglodytes</i>			81	70867	137	179992	149	54574	140	8107	210		
<i>Pan paniscus</i>			81	70783	137	173267	149	51064	140	8107	210		
<i>Gorilla gorilla</i>			81	67193	137	184195	149	51224	140	8094	210		
<i>Pongo abelii</i>			221	74052	137	181911	149	44826	140	8090	1764		
<i>Nomascus leucogenys</i>			81	74381	137	187645	149	43141	140	8045	210		
<i>Chlorocebus sabaeus</i>							874	40446	140	8070	1884		
<i>Macaca nemestrina</i>					145	3480	189	40329	140	8046	210		
<i>Macaca mulatta</i>			88	70315	137	171424	149	40820	140	8067	313		
<i>Macaca fascicularis</i>			140	69342	137	171726	149	40424	140	8070	2298		
<i>Theropithecus gelada</i>			209	70834	137	169299	149	40525	140	8069	1259		
<i>Mandrillus leucophaeus</i>					80	137	182	56182	140	8076	210		
<i>Cercocebus atys</i>			34	69833	137	170639	149	40796	140	8054	210		
<i>Papio anubis</i>			150	70397	137	175014	149	40129	140	8109	952		
<i>Rhinopithecus bieti</i>			81	69550	137	177194	148						
<i>Colobus angolensis palliatus</i>			81	69881	137	166895	149	40187	140	8052	210		
<i>Ptilocolobus tephrosceles</i>	161	44077	107	26118	137	164709	149	39824	140	8047	1261		
<i>Aotus nancymaae</i>			81	67642	137	165528	149	37192	140	7974	210		
<i>Saimiri boliviensis boliviensis</i>			81	67547	137	165556	149	36504	140	8048	210		
<i>Cebus capucinus imitator</i>			104	65732	137	165698	149	37330	140	7934	210		
<i>Callithrix jacchus</i>			81	62227	137	168437	149	36369	140	7904	210		
<i>Carlito syrichta</i>					56	4226	161	48637	140	8405	210		
<i>Propithecus coquereli</i>			81	65354	137	169838	161	43154	140	7087	189	3	18
<i>Microcebus murinus</i>			30	77203	137	163997	161	38343	140	6166	222		
<i>Prolemur simus</i>			81	64529	137	157904	161	39629	140	6812	210		

Примітки. Ек.*, Ек.** — додаткові екзони; Ін.*, Ін.** — додаткові інтрони.

Різниця довжини гена у представників приматів варіює переважно через довжину інтронів. Сумарна довжина екзонів у більшості випадків суттєво не відрізняється. Найконсервативнішими є екзони 2 і 4. Що стосується екзону 3, то його довжина у мокроносих приматів і у *Carlito syrichta* більша (161 п.н.), а у інших представників приматів менша (149 п.н.). Найваріабельнішими за довжиною є екзони 1 і 5, переважно за рахунок набуття 5' і 3' нетрансльованих ділянок. Таким чином, у ортологів гена *MGMT* приматів структура гена переважно має п'ять екзонів та чотири інтрони. Є випадки як втрати екзонів/інтронів (*Chlorocebus sabaeus*, *Macaca nemestrina*, *Mandrillus leucophaeus*, *Rhinopithecus bieti* та *Propithecus coquereli*), так і набуття додаткових структурних одиниць (*Ptilocolobus tephrosceles* та *Propithecus coquereli*) (табл. 1, рис. 1).

Цікаво було простежити за еволюцією варіабельних екзонів у ортологів гена *MGMT* на прикладі мокроносих приматів. У *Microcebus murinus* довжина екзону 1 становить 30 п.н. і він є некоду-

ючим. У *Prolemur simus* і *Propithecus coquereli* цей екзон кодує і його довжина збільшилась на 51 п.н. (табл. 2). Виявлено, що некодуєчий екзон 1 *Microcebus murinus*, із незначними точковими мутаціями, є частиною кодуєчого екзону 1 *Prolemur simus* і *Propithecus coquereli*. Нуклеотидні послідовності екзону 1 у *Prolemur simus* і *Propithecus coquereli* гомологічні між собою на 93 % (Підпала, Лукаш, 2020).

У людини екзон 1 має кодуєчу (81 п.н.) і 5' нетрансльовану ділянку (26 п.н.). Якщо порівнювати його послідовність із послідовністю екзону 1 *Propithecus coquereli*, то вони гомологічні між собою у межах трансльованої ділянки на 77 %. Послідовність некодованого екзону 1 *Microcebus murinus* у ортолога гена людини зберегла лише часткову гомологію до цієї послідовності і теж у межах трансльованої ділянки екзону 1 (Підпала, Лукаш, 2020). Отже, на прикладі варіабельних екзонів показано, що некодуєчі послідовності можуть виконувати кодуєчу роль на різних етапах еволюції гена.

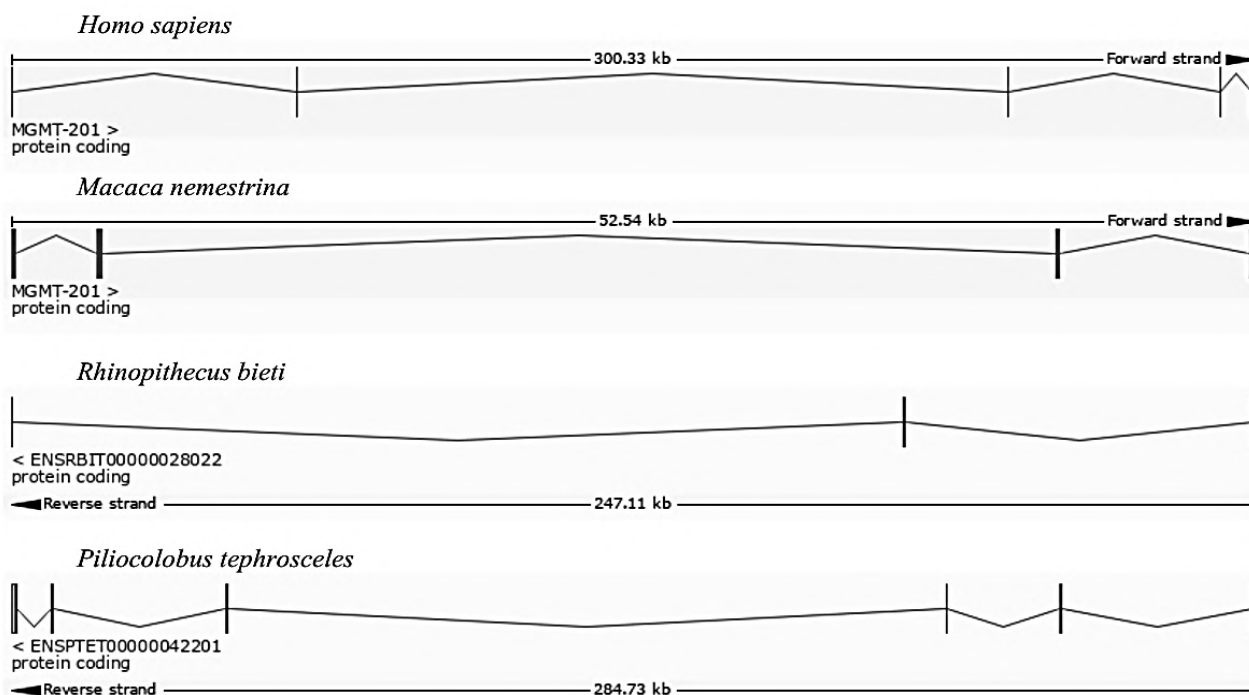


Рис. 1. Різноманітна екзон-інтронна структура гена *MGMT* у приматів за даними *Ensembl*.

Структурні особливості гена *MGMT* у широконосих мавп. У широконосих мавп ген *MGMT*, як і у людини, складається із п'яти кодуєчих екзонів та чотирьох інтронів. Різниця довжини

досліджуваного гена у представників мавп Нового Світу варіює, переважно, через довжину інтронів, тоді як довжина екзонів консервативніша. Сумарна довжина екзонів суттєво не відрізняється. Най-

консервативнішими є переважна більшість екзонів. Винятком є екзон 1 *Cebus capucinus imitator*, який має додаткові 23 п.н. 5'UTR, що гомологічні до послідовності 5'UTR людини. У межах кодуючої послідовності екзону 1 виявлено 10 нуклеотидних замін (7 одонуклеотидних і 3 динуклеотидних), тоді як послідовність 5'UTR не зазнала мутацій. Єдине, що у процесі еволюції довжина 5'UTR у гені *MGMT* людини збільшилась на 3 п.н. (Підпала та Лукаш, 2021).

Які послідовності гена *MGMT* капуцина білоголового могли бути потенційним джерелом 5'UTR? Як показали результати аналізу, фрагментарну гомологію довжиною 7-12 п.н. виявлено у двох послідовностях екзонів — у межах блоку кодуючої ділянки екзону 1 та екзону 5 і в усіх чотирьох інтронах. Загальна кількість частково гомологічних послідовностей у межах досліджуваного гена становить 233 одиниці. Фрагментарний характер гомології 5'UTR дає підстави для припущення щодо можливого ендегенного поход-

ження та рекомбіногенного характеру формування 5'UTR гена *MGMT Cebus capucinus imitator* (Підпала та Лукаш, 2021).

Мобільні генетичні елементи у гені *MGMT* приматів. МГЕ є вагомою складовою більшої еукаріотних геномів (Chenais et al., 2012). Для усіх просеквендованих на сьогодні геномів приматів частка МГЕ коливається від 42 до 50 % (Sotero-Caio et al., 2017), хоча є дані, що цей відсоток у людини може бути вищим (de Koning et al., 2011). У випадку гена *MGMT* серед різних представників приматів загальна кількість МГЕ становить від 23,73 до 37,72 % (табл. 3) із перевагою Non-LTR ретротранспозонів (23,73–31,18 %). Як видно із наведених даних, частка LTR ретротранспозонів найнижча серед МГЕ (1,55–3,12 %). Відсоток ДНК-транспозонів трохи вищий (1,34–5,19 %). Коливання кількості МГЕ у ортологів гена *MGMT* приматів можна пояснити втратою чи набуттям у деяких представників екзонних та інтронних послідовностей.

Таблиця 3. Частка МГЕ у гені *MGMT* приматів.

Організм, вид	Кількість МГЕ, %			
	Загальна (n*)	ДНК-транспозони	LTR ретро-транспозони	Non-LTR ретро-транспозони
<i>Homo sapiens</i>	30,11 (n = 249)	4,22	2,49	23,41
<i>Pan troglodytes</i>	29,76 (n = 262)	3,98	2,40	23,38
<i>Pan paniscus</i>	29,97 (n = 259)	4,07	2,44	23,46
<i>Gorilla gorilla gorilla</i>	29,99 (n = 264)	4,05	2,52	23,42
<i>Pongo abelii</i>	27,26 (n = 259)	4,06	2,31	20,90
<i>Nomascus leucogenys</i>	26,79 (n = 255)	3,97	2,26	20,56
<i>Chlorocebus sabaesus</i>	31,18 (n = 40)	1,34	3,12	26,46
<i>Macaca nemestrina</i>	31,12 (n = 41)	1,51	3,00	25,61
<i>Macaca mulatta</i>	29,58 (n = 244)	4,01	3,04	22,50
<i>Macaca fascicularis</i>	28,79 (n = 242)	4,03	2,87	21,37
<i>Theropithecus gelada</i>	28,37 (n = 235)	3,78	2,86	21,73
<i>Mandrillus leucophaeus</i>	24,15 (n = 39)	4,37	2,92	20,18
<i>Cercocebus atys</i>	27,17 (n = 242)	3,81	2,91	20,44
<i>Papio anubis</i>	29,74 (n = 246)	3,91	2,76	23,02
<i>Rhinopithecus bieti</i>	25,26 (n = 202)	4,62	2,60	18,04
<i>Colobus angolensis palliatus</i>	27,00 (n = 235)	3,99	2,88	20,12
<i>Ptilocolobus tephrosceles</i>	27,46 (n = 241)	4,12	2,83	20,06
<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
<i>Carlito syrichta</i>	37,72 (n = 49)	2,83	1,99	32,89
<i>Propithecus coquereli</i>	25,13 (n = 211)	4,98	2,17	17,98
<i>Prolemur simus</i>	23,73 (n = 178)	5,19	1,55	16,99
<i>Microcebus murinus</i>	25,92 (n = 199)	4,42	2,30	19,20

Примітка. n* — Кількість мобільних генетичних елементів.

Найбільше МГЕ в інтроні 3, а найменше в інтроні 4 (табл. 4). Хоча за кількістю МГЕ переважають інтрон 2 та інтрон 1 (табл. 2), як і за довжиною. Винятком є ген *MGMT Microcebus murinus*. Він не містить МГЕ в інтроні 4, але містить фрагмент ендегенного ретровірусу в екзоні 5 (Підпала та Лукаш, 2020). Довжина фрагме-

нта *HERV-Fc1* становить 110 п.н., а його частка, відповідно, 49,5 %. Цікаво зазначити, що ендегенний ретровірус *HERV-Fc1* належить до групи *ERV1* і є специфічним ендегенним ретровірусом для *Hominidae* (Bénit et al., 2003). Чи зберігся він у гені *MGMT* людини у такому ж вигляді чи зазнав еволюційних змін?

Таблиця 4. Розподіл МГЕ у інтронах ортологів гена *MGMT* приматів.

Організм, вид	Інтрони			
	1	2	3	4
<i>Homo sapiens</i>	28,02 (n=61)	27,72 (n=141)	44,75 (n=44)	8,26 (n=3)
<i>Pan troglodytes</i>	28,32 (n=64)	28,44 (n=151)	39,69 (n=43)	7,65 (n=3)
<i>Pan paniscus</i>	27,95 (n=66)	28,55 (n=146)	41,80 (n=42)	5,71 (n=3)
<i>Gorilla gorilla gorilla</i>	27,97 (n=59)	28,12 (n=156)	42,40 (n=47)	7,87 (n=3)
<i>Pongo abelii</i>	26,11 (n=72)	26,97 (n=147)	34,66 (n=35)	9,59 (n=5)
<i>Nomascus leucogenys</i>	26,62 (n=68)	25,96 (n=149)	35,03 (n=35)	4,72 (n=4)
<i>Chlorocebus sabaues</i>	-	-	37,52 (n=35)	7,66 (n=3)
<i>Macaca nemestrina</i>	-	8,25 (n=1)	36,65 (n=36)	9,38 (n=4)
<i>Macaca mulatta</i>	30,22 (n=69)	28,63 (n=134)	36,64 (n=36)	9,37 (n=5)
<i>Macaca fascicularis</i>	30,61 (n=68)	27,18 (n=132)	36,92 (n=37)	9,36 (n=4)
<i>Theropithecus gelada</i>	28,91 (n=64)	27,17 (n=130)	37,73 (n=36)	7,72 (n=5)
<i>Mandrillus leucophaeus</i>	-	0*	27,14 (n=36)	5,63 (n=3)
<i>Cercocebus atys</i>	30,82 (n=71)	24,34 (n=131)	37,38 (n=36)	5,65 (n=4)
<i>Papio anubis</i>	30,76 (n=71)	28,82 (n=135)	37,46 (n=36)	6,87 (n=4)
<i>Rhinopithecus bieti</i>	28,99 (n=66)	23,90 (n=137)	-	-
<i>Colobus angolensis palliatus</i>	29,82 (n=70)	24,60 (n=128)	36,60 (n=35)	6,93 (n=3)
<i>Ptilocolobus tephrosceles</i>	34,89 (n=48)** 23,06 (n=23)	25,33 (n=133)	36,70 (n=35)	6,93 (n=4)
<i>Aotus nancymae</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)
<i>Carlito syrichta</i>	-	6,53 (n=1)	45,49 (n=43)	10,95 (n=5)
<i>Propithecus coquereli</i> ***	27,16 (n=55)	23,79 (n=120)	31,91 (n=34)	2,55 (n=2)
<i>Prolemur simus</i>	24,17 (n=48)	24,15 (n=107)	25,77 (n=22)	2,73 (n=2)
<i>Microcebus murinus</i>	24,41 (n=64)	25,64 (n=110)	35,06 (n=26)	0*

Примітки. * — відсутні МГЕ; ** — МГЕ у додатковому інтроні.

Еволюційний шлях фрагмента *HERV-Fc1* у екзоні 5 ортологів гена *MGMT* приматів.

Проведено пошуки гомології між фрагментом ендегенного ретровірусу *HERV-Fc1* і послідовностями екзону 5 досліджуваних приматів (Підпала та Лукаш, 2020). У *Microcebus murinus* даний фрагмент гомологічний послідовності з координатами 33-142 відповідного екзону. У *Prolemur simus* послідовність фрагмента *HERV-Fc1* зазнала тринуклеотидної (у межах 70 п.н.) і дев'ятинуклеотидної (у межах 110 п.н.) делецій та її послідовність зменшилась до 98 п.н. У *Propithecus coquereli* простежується така ж кар-

тина, як і у *Prolemur simus*. У людини у межах кодуючої ділянки екзону 5 послідовність фрагмента *HERV-Fc1* зменшилась до 72 п.н. Крім того, присутні ще 13 фрагментів різних за довжиною (від 7 до 10 п.н.) і за напрямком: сім фрагментів у кодуючій ділянці екзону 5 і шість фрагментів у межах 3'UTR (рис. 2).

Ендегенні ретровіруси (ERV) — це залишки давніх ретровірусів, які були активними мільйони років тому, інфікували клітини, а потім зазнали різних мутацій, що призвело до втрати їхніх інфекційних властивостей (Magreg and Stoye, 2015). Більшість ERV втратили здатність

до ретротранспозиції, хоча і зберігають довгі термінальні повтори (*LTR*), що фланкують функціональні регуляторні послідовності і можуть використовуватись як альтернативні промотори для клітинних генів. *ERV* належать до родини *LTR*-ретроелементів, які широко представлені у хребетних.

Ендогенні ретровіруси відіграли активну роль у формуванні геномів. Більшість досліджень зосереджені переважно на геномах ви-

щих приматів і людини. Довгі кінцеві повтори (*LTR*), що фланкують геноми *ERV*, можуть відігравати роль альтернативних промоторів для клітинних генів. Крім того, ретровірусні білки можуть бути задіяні у процесах відтворення і розвитку. Рекомбінація між гомологічними послідовностями ретровірусних геномів також сприяє перетасуванню генів і генерації генетичної мінливості (Lee et al., 2015).

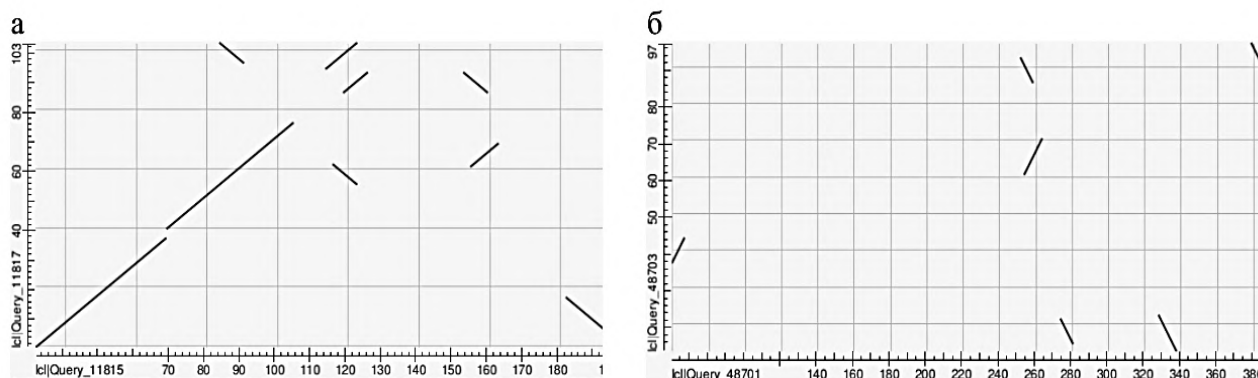


Рис. 2. Частка фрагмента *HERV-Fc1* у екзоні 5 гена *MGMT* людини: а) кодуюча ділянка та б) 3'UTR. По вісі абсцис — координати гомологічної послідовності у межах екзону, по вісі ординат — координати фрагмента ендогенного ретровірусу.

Ендогенні ретровіруси людини (*HERV*) складають значну частину геному людини. За різними підрахунками це приблизно 100000 *ERV*-елементів та фрагментів, що становлять близько 8 % (Lander et al., 2001). Вони найрізноманітніші серед МГЕ людини. На сьогодні, на жаль, не існує загальноприйнятої класифікації *HERV* (Blomberg et al., 2009; Vargiu et al., 2016). Їх об'єднують у 30-50 родин, які у свою чергу поділяють на більш, ніж 200 груп і підгруп. Труднощі з класифікацією *HERV* відображають їхню тривалу і складну еволюцію. Є дані, що *HERV*, можливо, притаманні не лише для лінії приматів. Це наводить на думку про те, що різниця між *HERV* та *ERV* може бути умовною і дозволяє припустити, що екзогенні аналоги деяких *HERV* інтегрували до геномів ссавців задовго до появи приматів та людини (Escalera-Zamudio and Greenwood, 2016).

Ендогенний ретровірус людини *HERV-Fc1* належить до родини *HERV-H/F*, яка є однією із найдавніших серед *HERV* (приблизний час інтеграції 60–70 млн років тому). Зазначається, що елемент *HERV-Fc1*, можливо, пройшов через позаклітинну інфекційну стадію, перш ніж потрапити до людинових (за різними підрахунками час

інтеграції від 35–20 до 15–10 млн років тому) (Bénil et al., 2003).

Розходження мокроносих і сухоносих приматів, до яких належить людина, відбулося близько 87 млн років тому (Perelman et al., 2011). Оскільки фрагмент специфічного для людини *HERV-Fc1* виявлено у представників мокроносих, можна припустити, що ендогенні ретровіруси приматів пройшли тривалий і неоднозначний процес ендегенізації. З одного боку, ретровірусні послідовності еволюціонували у нові *ERV*, а з іншого боку, зазнаючи фрагментації, ставали складовими елементами різних структур компонентів геному.

На основі наведених даних, можна висловити припущення про те, що у формуванні кодуючої ділянки екзону 5 і 3'UTR людини могли брати участь фрагментовані послідовності ендогенного ретровірусу *HERV-Fc1*. Таким чином, еволюційні зміни гена *MGMT* відбуваються на рівні різних структурних одиниць, причому МГЕ можуть бути не лише компонентами інтронів, але і складовими екзонів у вигляді фрагментованих послідовностей і не ідентифікуватись як МГЕ.

Формування ретроелемента *L1Hs* в інтроні 3 гена *MGMT* людиноподібних мавп. У людини переважна більшість МГЕ неактивна.

Активними є лише незначна кількість МГЕ (зокрема деякі *LINE-1* елементи, *Alu*-повтори, *SVA*-елементи і *HERV-K*) (Mills et al., 2007). Показовим є приклад із *LINE-1* елементами: із тисячі послідовностей на сьогодні свою активність зберегли не більше 100 послідовностей (Brouha et al., 2003). Усі вони належать до людиноспецифічних *L1Hs* елементів (Skowronski et al., 1988). Як вони виникли і що відомо про їхню еволюцію?

Вивчаючи розподіл МГЕ у гені *MGMT* людини ми показали, що вони присутні в інтронних послідовностях із перевагою *LINE-1* елементів (Pідпала and Lukash, 2012). У інтронах 2 та 3 МГЕ утворюють композиційні кластерні структури, до складу яких входять фрагменти *LINE1*-елементів самостійно чи у поєднанні із представниками інших класів МГЕ. В одному із трьох кластерів у межах інтрону 3 присутня майже повнорозмірна послідовність *L1Hs* елемента. Цікаво було простежити за формуванням людиноспецифічного *L1Hs* елемента в інтроні 3 гена *MGMT* на прикладі людиноподібних мавп (гоміноїдів).

Формування кластеру МГЕ із *L1Hs* елементом у інтроні 3 гена *MGMT* гоміноїдів. У всіх представників людиноподібних мавп МГЕ у інтроні 3 досліджуваного гена формують кластерні структури (Підпала та Лукаш, 2019). Детальніше простежили за зміною складових одного із кластеру, а саме того, де формується людиноспецифічний *L1Hs* елемент. Довжина повнорозмірного *L1Hs* елемента становить 6064 п.н. Він кодує дві відкриті рамки зчитування ORF1 (908..1921) і ORF2 (1988..5812). У випадку гена *MGMT* людини в інтроні 3 *L1Hs* елемент усічений з обох кінців, більше із 3'-кінця і має довжину 6049 п.н., що на 15 п.н. менше від повнорозмірної послідовності. З обох боків *L1Hs* елемент обрамлений *L1PA6* послідовностями. *L1PA6* елемент є 3'-кінцем *L1* повтору (підродини *L1PA6*) довжиною 901 п.н. Обидві *L1PA6* послідовності розташовані у зворотному до транскрипції напрямку. Вони різні за довжиною і за нуклеотидною послідовністю, перекриваються лише гомологічними 10 п.н. До складу даного кластера також входить фрагмент *L1* повтору (підродини *L1PA2*) довжиною 733 п.н., який є складовою ORF2 (2110..5403). Загальна довжина відповідного кластеру із чотирьох МГЕ становить 7280 п.н.

У шимпанзе присутні ті ж кластерні складові, але вони дещо відрізняються довжиною від попереднього. Більша і загальна довжина кластеру — 7284 п.н., і довжина людиноспецифічного *L1Hs* елемента. Щодо досліджуваного кластеру у боно-

бо, виявлено, що складові кластеру ті ж МГЕ, але істотна різниця у довжині *L1Hs* елемента (4243 п.н.) та однієї із *L1PA6* послідовностей при загальній довжині кластеру 5886 п.н.

У горили даний кластер ще не сформувався, а його потенційні складові, у переважній більшості, віддалені між собою і розділені іншими МГЕ. Цікаво, що у інтроні 3 горили ідентифіковано три різні фрагменти *L1Hs* елемента. Один із них гомологічний 5'-кінцю даної послідовності, а два інші фрагменти частково гомологічні ORF1 і ORF2.

У орангутана виявлено лише дві складові досліджуваного кластеру, а саме фрагменти послідовностей *L1PA6* та *L1*, тоді як фрагменти *L1Hs* елемента не ідентифікуються. У гібона ж відсутні навіть будь-які складові досліджуваного кластеру. Як формувався даний кластер і безпосередньо сам *L1Hs* елемент?

Різниця у загальній кількості МГЕ в інтроні 3 між послідовністю людини і гібона становить 10 %. Переважно за рахунок кількості *LINE1*-елементів і, саме складових досліджуваного кластеру. Це чітко простежується із характеру гомології між послідовністю інтрону 3 гена *MGMT* людини і гібона (Підпала та Лукаш, 2019).

Характер гомології між складовими елементами кластеру в орангутана та послідовністю інтрону 3 гена *MGMT* гібона. Оскільки в орангутана виявлено дві складові досліджуваного кластеру (фрагменти послідовностей *L1PA6* та *L1*) простежили, які саме послідовності МГЕ в інтроні 3 гібона могли бути джерелом формування цих двох елементів. За результатами BLAST аналізу показано, що в обох випадках у послідовностях *L1PA6* і *L1* є фрагменти, які гомологічні МГЕ із інтрону 3 гібона. У більшості випадків це послідовності *LINE1*-елементів. Крім того, для послідовності *L1PA6* елемента виявлено гомологію із фрагментом *AluY* повтору (Підпала та Лукаш, 2019), а у випадку *L1* послідовності, із ендегенними ретровірусами *MSTD* (ERV3) та *MLT1_Mim* (ERV1). Також гомологію мали фрагменти, не ідентифіковані як МГЕ. Були і такі послідовності, які не виявляли гомології із послідовністю інтрону 3 гібона.

Пошуки гомології між *L1Hs* елементом та послідовністю інтрону 3 гена *MGMT* гібона. Частка гомології між *L1Hs* елементом і послідовністю інтрону 3 досліджуваного гена у гібона становить 13 %. Аналіз виявлених гомологічних ділянок показав, що часткову гомологію, як і у випадку послідовностей *L1PA6* і *L1*, мають переважно фрагменти різних *LINE1*-елементів. Цікаво, що фрагменти гомологічних ділянок до *L1Hs* елемента

виявлено і в інших інтронних послідовностях гена *MGMT* гібона (Підпала та Лукаш, 2019).

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 (Kazazian and Moran, 1998; Furano, 2000; Boissinot et al., 2000). Зараз на частку геномної ДНК ссавців припадає приблизно 15–20 % *L1* елементів (Lander et al., 2001; Furano, 2000; Gu et al., 2000). Більшість із них усічені з 5' кінця, хоча існує і невелика кількість повнорозмірних *L1*, довжина яких становить 6–7 т.п.н. Родину *L1* поділяють на підродини: *L1Hs* (*L1PA1*), *L1PA2-16*, *L1PB1-3*, *L1MA1-10*, *L1MB1-8*, *L1MC*, *L1MD*, *L1ME*, де *L1Hs* наймолодша (і містить ще активні копії), а *L1ME* — найстаріша (Smit et al., 1995).

Ретротранспозони груп *L1PA(1-5)* специфічні для геномів мавп Старого Світу. Наймолодша група — *L1PA1* (або *L1Hs*, чи *L1Ta*) — виникла близько 4 млн. років тому, а пік ретропозицій її представників у геномі людини був приблизно 3 млн. років тому (Boissinot et al., 2000; Smit et al., 1995; Cantrell et al., 2000). Як зазначалось вище, деякі представники *L1Hs* (*L1PA1*) все ще активні. 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.

На основі наведених результатів можна підсумувати, що складовими кластера, де формувалася *L1Hs* елемент у людинових є фрагменти *L1PA6* елемента. У гібона ж у інтроні 3 досліджуваного гена серед груп *L1* присутні представники старіших підродин — *L1PB*, *L1MC*, *L1MD* і *L1ME*, тоді як елементи *L1PA6* групи відсутні, хоча ці елементи розповсюджені у мавп Старого і Нового Світу. Це питання лишається відкритим. Можливо подальші дослідження допоможуть зрозуміти узгодженість еволюції між різними представниками МГЕ серед приматів.

Аналізуючи послідовності інтрону 3 гена *MGMT* у людиноподібних мавп можна зробити висновок, що формування людиноспецифічного *L1Hs* елемента відбувалось упродовж еволюції гоміноїдів (рис. 3). Утворення кластера із різних фрагментів *LINE1*-елементів і самого *L1Hs* елемента відбувалось паралельно. Також, на основі наведених даних, можна висловити припущення про те, що у формуванні *L1Hs* елемента могли брати участь не лише фрагменти різних МГЕ (переважно *LINE1*-елементів), але і послідовності, які не ідентифікуються як МГЕ. Таким чином, форму-

вання людиноспецифічного *L1Hs* елемента 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 (Підпала та Лукаш, 2019).

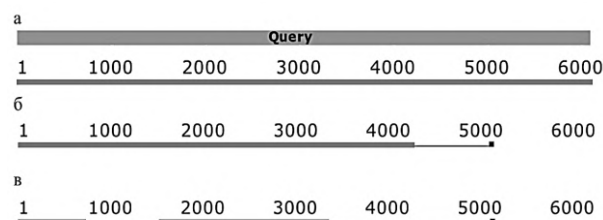


Рис. 3. Формування людиноспецифічного *L1Hs* елемента в інтроні 3 гена *MGMT* у людиноподібних мавп: а) людини (6049 п.н.) і шимпанзе (6057 п.н.); б) бонобо (4243 п.н.); в) горили (699 п.н. + 1812 п.н.).

Еволюційна історія видоспецифічних *Alu*-повторів на прикладі гена *MGMT* приматів. Умовно МГЕ приматів поділяють на дві групи: древні мобільні елементи, які є не лише у ссавців, але й у низки хордових (*MIR*, *L2*, *L3* і деякі ДНК-транспозони) і молоді приматоспецифічні повтори, до яких належать *Alu*-повтори.

Alu-повтори — неавтономні МГЕ класу SINE елементів, які походять від 7SL РНК і зустрічаються лише у приматів (Ullu and Tschudi, 1984). Вони є функціонально значимими елементами, які впливають на архітектуру геномів приматів та беруть активну участь у регуляції експресії генів (Häsler and Strub, 2006; Deininger, 2011). Також, *Alu*-повтори є важливим джерелом міжвидових і внутрішньовидових геномних варіацій приматів (Salem et al., 2003), тому їх широко використовують для еволюційних досліджень (Schmitz et al., 2016). На основі «діагностичних» мутацій *Alu*-повтори поділяють на три основні родини — древні (J), середні (S) і молоді (Y) повтори та на 14 основних підродин, які відрізняються еволюційним віком (Karitonov and Jurka, 1996). 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. Деякі підродини можуть бути спільними для низки близькоспорідних таксонів, тоді як інші є унікальними лише для певного виду або роду (McLain et al., 2013; Vaker et al., 2017). Ця паралельна еволюція підродин *Alu*-повторів призводить до того, що кожна лінія приматів має свою власну мережу інтегрованих

елементів (Konkel et al., 2010). Аналізуючи поширення видоспецифічних МГЕ у інтронних послідовностях ортологів гена *MGMT* широконосих мавп нами виявлено *Alu*-повтор довгоп'ятого філіппінського *Alu2_TS* (Підпала та Лукаш, 2021). Цікаво було простежити за еволюцією видоспецифічних *Alu*-повторів у ортологів гена *MGMT* приматів.

Alu-повтор є гетеродимером, який складається із двох протилежно направлених форм древнього *Alu*-мономера (FAM — fossil *Alu* monomer) довжиною приблизно 130 н.п. і розділених полі (A) вставкою (Quentin, 1992). Довжина *Alu* послідовності становить близько 300 п.н., із яких 282 п.н. — консенсусна послідовність, а інші нуклеотиди входять до складу полі (A) послідовності на 3' кінці. Більшість *Alu* послідовностей фланковані короткими прямими повторами (10–20 п.н.), які є дуплікаціями сайту ДНК-мішени і утворюються у процесі ретропозиції. Лівий мономер (FLAM — fossil left *Alu* monomer) довжиною 140 п.н. з'єднаний полі (A) вставкою із довшим правим мономером (FRAM — fossil right *Alu* monomer), який містить додаткові 31 п.н. (Quentin, 1992). Лівий мономер містить два промоторні елементи для РНК-полімерази III, блок А та блок В, кожен із яких має довжину близько 10 п.н. Більшість *Alu*-повторів мають димерну структуру. Зустрічаються «усічені» послідовності із частково делетованими 5' або 3' кінцями, а також, в поодиноких випадках, мономерні повтори FRAM та FLAM. Знайдено і ряд тетрамерних послідовностей. Крім того, зазначається, що неактивність більшості *Alu*-повторів може бути пов'язана із їхньою послідовною деградацією (Mighell et al., 1997).

Видоспецифічні *Alu*-повтори в інтронах ортологів гена *MGMT* мавпових. У проаналізованих послідовностях гена *MGMT* мавпових, крім представників основних родин *Alu*-повторів ідентифіковано фрагменти видоспецифічних повторів — *Otolemur garnettii Alu2_OG*, *Microcebus murinus AluJ_Mim* та *Tarsius syrichta Alu2_TS* (Підпала та Лукаш, 2022). Переважно вони присутні в інтронах 1 і 2. Не виявлено видоспецифічних *Alu*-повторів у *Chlorocebus sabaues*, оскільки у нього відсутні інтрони 1, 2. У *Macaca nemestrina* та *Mandrillus leucorhaeus* також не виявлено видоспецифічних *Alu*-повторів, можливо через те, що у них втрачено інтрон 1, а послідовності інтрону 2 делетовані.

Серед специфічних для представників мокроносих приматів *Alu*-повторів у досліджуваних ортологів гена *MGMT* мавпових виявлено два повтори: *Alu2_OG* та *AluJ_Mim*. Цікаво, що фрагменти *Alu*-повтору *Otolemur garnettii Alu2_OG*

виявлено у представників підродини *Cercopithecinae* і лише в одного представника підродини *Colobinae* — у *Rhinopithecus bieti*. Що стосується *Alu*-повтору *Microcebus murinus AluJ_Mim*, то його послідовності виявлено у представників підродини *Colobinae* — *Colobus angolensi spalliatu* та *Ptilocolobus tephrosceles*, тоді як у представників підродини *Cercopithecinae* їх не ідентифіковано. Видоспецифічний для одного із представників довгоп'ятів (Tarsiiformes) *Alu*-повтор *Tarsius syrichta Alu2_TS* виявлено у представників обох підродин мавпових (Підпала та Лукаш, 2022).

Видоспецифічний *Alu*-повтор *Alu2_OG*. *Alu*-повтор *Alu2_OG* є видоспецифічним повтором *Otolemur garnettii* (галаго Гарнетта). *Otolemur garnettii* належить до підряду мокроносих приматів із інфраряду лорієвидих, родини галагових, роду товстохвостих галаго. Повтор *Alu2_OG* має довжину 230 п.н. Якщо простежити еволюційну динаміку повтору *Alu2_OG* від мокроносих приматів до мавпових, то у мокроносих приматів, які належать до інфраряду лемуroidних, вони мають різну довжину і напрямок та локалізовані у різних інтронних послідовностях (Підпала та Лукаш, 2022). Зокрема, у *Microcebus murinus* довжина фрагмента повтору *Alu2_OG* становить 97 п.н. (координати послідовності 36–132, позитивний напрямок) і він локалізований у межах інтрону 1. У *Propithecus coquereli* фрагмент даного повтору локалізований у межах інтрону 3 і має довжину 56 п.н. (1–56, зворотній напрямок). У гені *MGMT* досліджуваних представників довгоп'ятових і широконосих мавп послідовності, гомологічні *Alu*-повтору *Alu2_OG* не виявлено. Що стосується фрагментів *Alu2_OG* у мавпових, то у представників підродини *Cercopithecinae* послідовності довжиною 51 п.н. (координати 1–51, зворотній напрямок) локалізовані у межах інтрону 1. У *Rhinopithecus bieti*, представника підродини *Colobinae*, фрагмент послідовності *Alu2_OG* має довжину 29 п.н. (координати 1–29, позитивний напрямок) і локалізований у межах інтрону 2. Серед досліджуваних представників людиноподібних приматів у межах гена *MGMT* послідовність *Alu*-повтору *Alu2_OG* не виявлено (Підпала та Лукаш, 2022).

На основі наведених даних, можна зробити висновок, що еволюційна історія *Alu*-повтору *Otolemur garnettii Alu2_OG* охоплює час формування мокроносих мавп і мавпових. Динаміка делетційної деградації пов'язана винятково із послідовністю мономеру FLAM. Чому у ортологів гена *MGMT* представників довгоп'ят і широконосих мавп послідовності *Alu2_OG* не виявлено, лиша-

ється відкритим питанням. Цікаво також, чому *Alu2_OG* у мавпових виявлено переважно у представників підродини *Cercopithecinae*. Ці питання потребують детальніших досліджень (Підпала та Лукаш, 2022).

Видоспецифічний *Alu*-повтор *AluJ_Mim*. Ще один специфічний для представників мокроносих приматів *Alu*-повтор *AluJ_Mim*, який виявлено у ортологів гена *MGMT* мавпових, є видоспецифічним повтором *Microcebus murinus* (мишачого мікроцебусу). *Microcebus murinus* належить до підряду мокроносих приматів із інфраряду лемуroidних, родини карликових лемуroidів, роду мишиних лемуroidів. Повтор *AluJ_Mim* має довжину 317 п.н. Серед мокроносих приматів у гені *MGMT* він присутній лише у представників інфраряду лемуroidних, а у *Otolemur garnettii* — представника інфраряду лорієвидних його не ідентифіковано (Підпала та Лукаш, 2022). Не виявлено *AluJ_Mim* і у досліджуваних представників довгоп'ятових та широконосих мавп, а у ортологів гена *MGMT* мавпових він присутній лише у представників підродини *Colobinae* (*Colobus angolensi spalliatu*s і *Ptilocolobus tephrosceles*). Крім усічених послідовностей із частково делетованими 5' чи 3'кінцями, мономерними повторами FRAM і FLAM, *AluJ_Mim* зазнає делеційної деградації (ідентифіковано фрагменти довжиною 64 п.н. із координатами 18-81) ще у інтронних послідовностях досліджуваного гена *Microcebus murinus*. У послідовностях гена *MGMT* *Colobus angolensi spalliatu*s та *Ptilocolobus tephrosceles* виявлено лише мономерні повтори FRAM. Цікаво, що послідовність такого ж мономеру ідентифіковано в інтроні 1 представника людиноподібних мавп - *Nomascus leucogenys*.

Еволюційна історія *Alu*-повтору *Microcebus murinus AluJ_Mim*, як і у випадку *Alu*-повтору *Otolemur garnettii Alu2_OG*, охоплює час формування мокроносих мавп і мавпових. На відміну від *Alu2_OG*, послідовність якого у ортологів гена *MGMT* мавпових виявлено переважно у представників підродини *Cercopithecinae*, *Alu*-повтор *AluJ_Mim* ідентифіковано у представників підродини *Colobinae*. Що стосується динаміки деградації *AluJ_Mim*, то найрізноманітніші варіанти зафіксовано ще у *Microcebus murinus*, а у послідовностях ортологів гена *MGMT* мавпових (представників підродини *Colobinae*) і представника людино-

подібних мавп (*Nomascus leucogenys*) виявлено лише мономерні повтори FRAM (Підпала та Лукаш, 2022).

Видоспецифічний *Alu*-повтор *Alu2_TS*. *Alu*-повтор *Alu2_TS* є видоспецифічним повтором *Tarsius syrichta* (довгоп'ята філіппінського). Разом із широконосими і вузьконосими мавпами довгоп'ят зараховано до підряду сухоносих приматів (Haplorhini). Довжина вихідної послідовності *Alu2_TS* становить 283 п.н. У гені *MGMT* довгоп'ята філіппінського присутні дві майже повнорозмірні послідовності *Alu2_TS* елемента із координатами 1-278 і 12-283 п.н. у межах інтрону 2 у позитивній орієнтації. Аналізуючи розподіл видоспецифічних МГЕ у ортологів гена *MGMT* широконосих мавп, фрагменти *Alu2_TS* виявлено у послідовностях інтронів 1 та 2 (Підпала та Лукаш, 2021). У послідовностях інтрону 1 фрагмент *Alu2_TS* має координати 1-187, 12-107 (зворотній та позитивний напрямки), а у межах інтрону 2 фрагмент *Alu2_TS* має ще менші розміри — від 1-97 до 1-93 п.н. (зворотній напрямки).

У гені *MGMT* мавпових видоспецифічний *Alu*-повтор *Tarsius syrichta Alu2_TS* присутній у представників обох підродин і переважно в інтроні 2 (Підпала та Лукаш, 2022). Серед фрагментів *Alu2_TS* зустрічаються переважно делеційні фрагменти мономеру FLAM, хоча є і послідовності усіченого мономеру FRAM. У ортологів гена *MGMT* людиноподібних мавп послідовність повтору *Alu2_TS* не ідентифіковано. Таким чином, еволюційна історія *Alu*-повтору *Tarsius syrichta Alu2_TS* охоплює час формування сухоносих мавп (за винятком людиноподібних). Динаміка делеційної деградації і у широконосих мавп, і у мавпових переважно пов'язана із послідовністю мономеру FLAM.

Підсумовуючи наведені дані, можна зробити висновок, що еволюційна історія видоспецифічних *Alu*-повторів та особливості їхньої деградації мають індивідуальний характер (рис. 4). В той же час відсутність досліджених видоспецифічних *Alu*-повторів мокроносих (*Alu2_OG*, *AluJ_Mim*) у представників довгоп'ятових і широконосих мавп та їхнє вибіркоче розповсюдження у мавпових, спонукають до детальніших філогенетичних досліджень (Підпала та Лукаш, 2022).

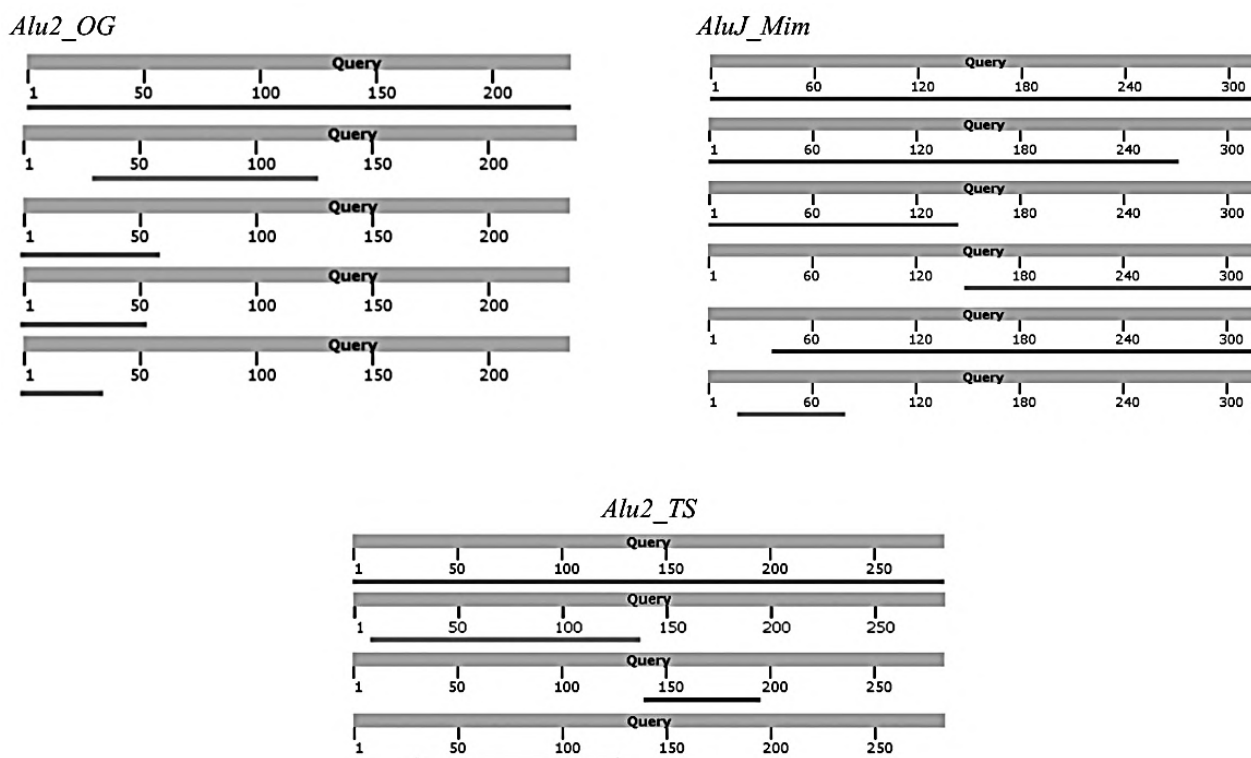


Рис. 4. Еволюційна динаміка деградації видоспецифічних *Alu*-повторів в інтронних послідовностях гена *MGMT*.

Поширення родоспецифічного *Alu*-повтору макак *AluMacYa3* у ортологів гена *MGMT* мавпових. Мавпові (*Cercopithecidae*) це родина приматів з підряду сухоносих мавп (*Haplorhini*) (Perelman et al., 2011). До цієї родини належать дві підродини: мавпові (*Cercopithecinae*) та колобусові (*Colobinae*), які розійшлися 18 млн років тому (Perelman et al., 2011). Підродину мавпові (*Cercopithecinae*) поділяють на дві триби — *Cercopithecini* та *Papionini*. В той же час класифікація колобусових (*Colobinae*) залишається спірною: їх поділяють або на дві триби — *Colobini* та *Presbytini*, або на три групи — африканська, лангура та дивноносі.

У переважній більшості мавпових ген *MGMT*, як і у людини, складається із п'яти екзонів та чотирьох інтронів (табл. 1). Хоча є й випадки, коли ген *MGMT* має або меншу кількість структурних одиниць — чотири екзони і три інтрони як у випадку *M. nemestrina* та *M. leucophaeus* чи три екзони і два інтрони як у *Ch. sabaues* та *M. leucophaeus*, або ж більшу

кількість структурних одиниць на прикладі ортолога у *Ptilocolobus tephrosceles* (табл. 1).

У ортологів гена *MGMT* мавпових загальний відсоток МГЕ коливається від 24,15 до 31,18 %. Переважають *Non-LTR* ретро-транспозони (18,04–26,46 %) (Підпала та Лукаш, 2022). Зокрема, відсоток *Alu*-повторів становить від 2,06 % у *Mandrillus leucophaeus* до 4,09 % у *Papio anubis*. У більшості випадків виявлено найбільший відсоток середніх повторів *AluS*, тоді як кількість древніх (J) і молодих (Y) повторів у двічі менша (Підпала та Лукаш, 2023).

Крім видоспецифічних *Alu*-повторів, у приматів зустрічаються і родоспецифічні повтори. Зокрема у макак описано три родоспецифічні *Alu*-повтори: *AluMacYa3*, *AluMacYb2* і *AluMacYb4* (Mighell et al., 1997). Усі вони належать до *AluSp* повторів.

*Частка родоспецифічного *Alu*-повтору *AluMacYa3* у макак.* Рід макак належить до підродини мавпові (*Cercopithecinae*), триби *Papionini* і нараховує, за різними даними, від 21 до 24 видів. Згідно молекулярних даних, мака-

ки відокремились від інших представників триби *Papionini* приблизно 9–10 мільйонів років тому (Raaijmakers et al., 2005). Високні дані вказують на те, що рід макак виник приблизно 7 млн років тому у Північній Африці, а потім родовід азіатських макак почав з'являтися близько 5,5 млн років тому (Perelman et al., 2011).

У проаналізованих ортологів гена *MGMT* макак *Alu*-повтор *AluMacYa3* виявлено у послідовностях інтронів 1 і 2 *Macaca mulatta* та *Macaca fascicularis*, за винятком *Macaca nemestrina* (Підпала та Лукаш, 2023). Відомо, що ген *MGMT* у *Macaca nemestrina* складається із 4 екзонів та 3 інтронів. У нього відсутня послідовність інтрону 1, а послідовність інтрону 2 усічена і гомологічна кінцевому фрагменту послідовностей відповідних інтронів *Macaca mulatta* та *Macaca fascicularis*.

Довжина вихідного повтору *AluMacYa3* становить 285 п.н. У гені *MGMT* *Macaca mulatta* та *Macaca fascicularis* ідентифіковано по чотири послідовності *Alu*-повтору *AluMacYa3*, по одній в інтроні 1 і по три в інтроні 2 (Підпала та Лукаш, 2023). Усі вони мають усічену на кілька нуклеотидів послідовність із 3'кінця (1-283; 1-281). Напрямок послідовності різний, переважає зворотний.

Щодо нуклеотидного поліморфізму послідовностей *Alu*-повторів *AluMacYa3* у *Macaca fascicularis* та *Macaca mulatta*, домінують точкові мутації. На прикладі повторів, які ідентифіковано в інтроні 1 гена *MGMT*, серед 11 точкових мутацій переважають транзиції (6 С-Т; 4 G-A), а трансверсію (С-А) виявлено тільки в одному випадку (Підпала та Лукаш, 2023).

Еволюційна історія родоспецифічного *Alu*-повтору макак *AluMacYa3* у представників підроду мавпових. Підроду мавпових (*Cercopithecinae*) поділяють на дві триби — *Cercopithecini* та *Papionini*, які виникли від спільного предка приблизно 11,5 млн років тому (Perelman et al., 2011). До триби *Papionini* зараховують сім родів. Ми змогли проаналізувати ортологи гена *MGMT* п'яти родів (крім трьох представників роду *Macaca*, по одному представнику родів *Papio*, *Theropithecus*, *Cercocobus* і *Mandrillus*) (Підпала та Лукаш, 2023).

У проаналізованих послідовностях гена *MGMT* представників триби *Papionini* послідовність *Alu*-повтору *AluMacYa3* виявлено у *Papio anubis* і *Theropithecus gelada* в інтронах 1 і 2 та у *Cercocobus atys* в інтронах 1, 2 і 3. Не виявлено *Alu*-повтору *AluMacYa3* у послідов-

ності ортолога *Mandrillus leucophaeus*. У нього відсутня послідовність інтрону 1, а послідовність інтрону 2 становить лише 137 п.н. (за даними *Ensembl*). Цікаво, що у *Papio anubis* також ідентифіковано ще одну родоспецифічну послідовність *Alu*-повтору макак *AluMacYb2* (Підпала та Лукаш, 2023).

Для досліджених послідовностей *AluMacYa3* у інтронних послідовностях ортологів гена *MGMT* представників триби *Papionini* також характерний поліморфізм і усіченість послідовностей із частково делетованими 5' та / або 3'кінцями (1-283; 2-283; 5-283; 198-283). Послідовність *AluMacYa3* у гені *MGMT* *Cercocobus atys* виявлено також в інтроні 3 (Підпала та Лукаш, 2023).

Щодо представників триби *Cercopithecini*, до складу якої входять п'ять родів (*Erythrocebus*, *Chlorocebus*, *Cercopithecus*, *Miopithecus* і *Allenopithecus*), проаналізувати наявність послідовності *AluMacYa3* у гені *MGMT* змогли лише для одного представника роду *Chlorocebus* — *Chlorocebus sabaeus*. Родоспецифічний *Alu*-повтор макак *AluMacYa3* у *Chlorocebus sabaeus* виявлено лише у межі інтрону 1, який гомологічний послідовності інтрону 3, оскільки послідовності інтронів 1 і 2 у нього делетовані (Підпала та Лукаш, 2023).

Порівнюючи розповсюдження послідовності *AluMacYa3* у гені *MGMT* представників підроду мавпових (*Cercopithecinae*) виявили, що даний родоспецифічний *Alu*-повтор макак переважно локалізований в інтронах 1 і 2. У послідовностях інтрону 3 *Alu*-повтор *AluMacYa3* присутній у одного із представників триби *Papionini* — *Cercocobus atys* та в одного із представників триби *Cercopithecini* — *Chlorocebus sabaeus*. Обидві послідовності частково усічені з 5' кінця і мають довжину 282 п.н. Також, вони мають різне положення серед інших МГЕ (Підпала та Лукаш, 2023).

Еволюційна історія родоспецифічного *Alu*-повтору макак *AluMacYa3* у представників підроду колобусових. Підроду колобусових (*Colobinae*) охоплює 10 родів і 59 видів, які поділяють на дві триби — *Colobini* та *Presbytini*, що розділились 10–13 млн років тому (Perelman et al., 2011).

У проаналізованих послідовностях гена *MGMT* представників триби *Colobini* послідовність *Alu*-повтору *AluMacYa3* виявили у *Colobus angolensis palliatus* та *Ptilocolobus tephrosceles*, а у представників триби *Pres-*

bytini — у *Rhinopithecus bieti* (Підпала та Лукаш, 2023). Послідовності *AluMacYa3* у *Colobus angolensis palliatus* та *Rhinopithecus bieti* ідентифіковані у інтроні 2, а у *Ptilocolobus tephrosceles* крім інтрону 2, *Alu*-повтори *AluMacYa3* виявлено у додатковій інтронній послідовності. У жодному випадку не було виявлено повної нуклеотидної послідовності *AluMacYa3*. Усі вони мають усічені послідовності із частково делетованими 5' та / або 3'кінцями (1-283; 3-283; 35-283). Крім того, у двох випадках виявлено мономерні повтори FLAM.

Узагальнюючи еволюційну історію родоспецифічного *Alu*-повтору макак *AluMacYa3* у представників родини мавпові (*Cercopithecidae*) бачимо, що крім нуклеотидного поліморфізму для послідовності зазначеного повтору також характерна делеційна деградація (Підпала та Лукаш, 2023). Зокрема, в одного із представників підродино *Cercopithecinae* — *Cercocebus atys* виявлено послідовність із частково делетованою послідовністю мономеру FLAM (198-283) а у представника підродино *Colobinae* — *Rhinopithecus bieti* виявлено повтори *AluMacYa3* із частково делетованими послідовностями як мономеру FLAM (35-283), так і FRAM (2-199).

Відомо, що дві підродино мавпових — мавпові (*Cercopithecinae*) та колобусові (*Colobinae*) розійшлися 18 млн. років тому (Perelman et al., 2011), але схеми класифікації дискусивні, оскільки є невідповідності між морфологічними і генетичними даними, а також є розбіжності між різними генетичними дослідженнями. Зазначається, що колобусові почали формуватися приблизно 12 млн років тому, а мавпові виникли від спільного предка близько 11,5 млн років тому (Perelman et al., 2011). Що стосується представників триби *Papionini*, то час їхньої появи оцінюють від 6 до 8 млн років тому, зокрема представники роду макак почали формуватися приблизно 5,1 млн років тому (Perelman et al., 2011). Якщо орієнтуватися на ці дані, то *Alu*-повтор *AluMacYa3* міг бути присутній, лише у представників роду макак. Пояснити ж присутність послідовності *Alu*-повтору *AluMacYa3* не лише у представників триби *Papionini* (*Theropithecus gelada*, *Cercocebus atys*, *Papio anubis*) і *Cercopithecini* (*Chlorocebus sabaeus*), але й у представників підродино колобусові (*Rhinopithecus bieti*, *Colobus angolensis palliatus*, *Ptilocolobus tephrosceles*) на основі загальноприйнятих

філогенетичних даних потребує подальших досліджень.

Таким чином, в еволюційній історії родоспецифічного *Alu*-повтору макак *AluMacYa3* переважає не делеційна деградація, як у випадку видоспецифічних *Alu*-повторів *Alu2_OG*, *AluJ_Mim* та *Alu2_TS* (Підпала та Лукаш, 2022), а нуклеотидний поліморфізм. На прикладі гена *MGMT* показано, що родоспецифічний *Alu*-повтор *AluMacYa3* присутній не лише у макак, а й у інших представників мавпових. Аналіз нуклеотидної послідовності цього *Alu*-повтору свідчить про те, що його еволюційна історія поєднує нуклеотидний поліморфізм та делеційну деградацію (Підпала та Лукаш, 2023).

Отже, на основі наведених даних можна зробити висновок, що еволюція гена *MGMT* у приматів зазнавала різноманітних змін від набуття до втрати екзон/інтронних послідовностей. Що стосується приматоспецифічних МГЕ, то вони мають різну еволюційну історію і шлях — від формування видо/родоспецифічних послідовностей до делеційної деградації і можуть бути компонентами не лише різноманітних МГЕ але й інтронних чи екзонних послідовностей.

Висновки

Упродовж еволюції приматів ген *MGMT* зазнавав змін як на рівні екзонних, так і інтронних послідовностей. Крім набуття чи втрати екзон/інтронних послідовностей у ортологів гена *MGMT* показано, що некодуючі послідовності екзонів могли ставати кодуєчими. Аналізуючи еволюційну долю фрагмента ендемічного ретровірусу *HERV-Fc1*, висловлено припущення, що МГЕ можуть бути не лише компонентами інтронів, але і складовими екзонів у вигляді фрагментованих послідовностей і не ідентифікуватись як МГЕ. Виявлено, що формування людноспецифічного *L1Hs* елемента відбувалось упродовж еволюції гоміноїдів паралельно із утворенням кластерної структури МГЕ у людюних із різних підродин *LINE1*-елементів, складові компоненти якого, очевидно, також задіяні в утворенні *L1Hs* елемента. Показано, що еволюційна історія видоспецифічних *Alu*-повторів (*Alu2_OG*, *AluJ_Mim*, *Alu2_TS*) та родоспецифічного повтору *AluMacYa3* має індивідуальний характер і свої особливості щодо делеційної деградації послідовностей.

References

1. Baker J. N., Walker J. A., Vanchiere J. A., Phillippe K. R., St. Romain C.P., Gonzalez- Quiroga P., Denham M. W., Mierl J. R., Konkel M. K., Batzer M. A. Evolution of *Alu* subfamily structure in the Saimiri lineage of new world monkeys. *Genome Biol. Evol.* 2017. Vol. 9, № 9. P. 2365–2376. doi: 10.1093/gbe/evx172.
2. Bénéit L., Calteau A., Heidmann T. Characterization of the low-copy HERV-Fc family: evidence for recent integrations in primates of elements with coding envelope genes. *Virology.* 2003. Vol. 312, № 1. P. 159–168.
3. Blomberg J., Benachenhou F., Blikstad V., Sperber G., Mayer J. Classification and nomenclature of endogenous retroviral sequences (ERVs): problems and recommendations. *Gene.* 2009. Vol. 448, № 2. P. 115–123. doi:10.1016/j.gene.2009.06.007.
4. 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.
5. Brouha B., Schustak J., Badge R. M., Lutz-Prigge S., Farley A. H., Moran J. V., Kazazian H. H. Jr. Hot L1s account for the bulk of retrotransposition in the human population. *Proc. Natl. Acad. Sci. USA.* 2003. Vol. 100, № 9. P. 5280–5285.
6. 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.
7. 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.
8. Deininger P. Alu elements: know the SINEs. *Genome Biol.* 2011. Vol.12, № 12. P. 236.
9. de Koning A. P., Gu W., Castoe T. A., Batzer M. 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.
10. Escalera-Zamudio M., Greenwood A. D. On the classification and evolution of endogenous retrovirus: human endogenous retroviruses may not be 'human' after all. *APMIS.* 2016. Vol. 124, № 1–2. P. 44–51. doi: 10.1111/apm.12489.
11. 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.
12. Gu Z., Wang H., Nekrutenko A., Li W. H. Densities, length proportions, and other distributional features of repetitive sequences in the human genome estimated from 430 megabases of genomic sequence. *Gene.* 2000. Vol. 259, № 1–2. P. 81–88.
13. Häsler J., Strub K. Alu elements as regulators of gene expression. *Nucleic Acids Res.* 2006. Vol. 34, № 19. P. 5491–5497. doi: 10.1093/nar/gkl706.
14. Kaina B., Christmann M., Naumann S., Roos W. P. MGMT: key node in the battle against genotoxicity, carcinogenicity and apoptosis induced by alkylating agents. *DNA Repair (Amst).* 2007. Vol. 8. P. 1079–1099. doi:10.1016/j.dnarep.2007.03.008.
15. Kapitonov V., Jurka J. The age of Alu subfamilies. *J. Mol. Evol.* 1996. Vol. 42, № 1. P. 59–65.
16. Kazazian H. H. Jr, Moran J. V. Mobile DNA in Health and Disease. *N. Engl. J. Med.* 2017. Vol. 377, № 4. P. 361–370. doi: 10.1056/NEJMr1510092.
17. Konkel M. K., Walker J. A., Batzer M. A. LINEs and SINEs of primate evolution. *Evol. Anthropol.* 2010. Vol. 19. P. 236–249.
18. Lander E., Linton L. M., Birren B., Nusbaum C., Zody M. et al. Initial sequencing and analysis of the human genome. *Nature.* 2001. 409, № 6822. P. 860–921.
19. Lee H.-E., Eo J., Kim H.-S. Composition and evolutionary importance of transposable elements in humans and primates. *Genes Genomics.* 2015. Vol. 37, № 2. P. 135–140.
20. Mager D. L., Stoye J. P. Mammalian Endogenous Retroviruses. *Microbiol. Spectr.* 2015. Vol. 3, № 1. MDNA3-0009-2014. doi: 10.1128/microbiolspec.MDNA3-0009-2014.
21. 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., Santibáñez-Koref M. F. Alkyltransferase-like Proteins. *DNA Repair (Amst).* 2007. Vol. 6. P. 1222–1228. doi:10.1016/j.dnarep.2007.03.014.
22. McLain A. T., Carman G. W., Fullerton M. L., Beckstrom T. O., Gensler W., Meyer T. J., Faulk C., Batzer M. A. Analysis of western lowland gorilla (*Gorilla gorilla gorilla*) specific *Alu* repeats. *Mob. DNA.* 2013. Vol. 4, № 1. P. 26. doi: 10.1186/1759-8753-4-26.
23. Mighell A. J., Markham A. F., Robinson P. A. Alu sequences. *FEBS Lett.* 1997. Vol. 417, № 1. P. 1–5. doi: 10.1016/s0014-5793(97)01259-3.
24. Mills R. E., Bennett E. A., Iskow R. C., Devine S. E. Which transposable elements are active in the human genome? *Trends Genet.* 2007. Vol. 23, № 4. P. 183–191.
25. Pegg A. E. Repair of O⁶-alkylguanine by alkyltransferases. *Mutat. Res.* 2000. Vol. 262. P. 83–100.
26. Pegg A. E. Multifaceted roles of alkyltransferase and related proteins in DNA repair, DNA

- damage, resistance to chemotherapy, and research tools. *Chem. Res. Toxicol.* 2011. Vol. 24. P. 618–639. doi:10.1021/tx200031q.
27. Perelman P., Johnson W. E., Roos C., Seuánez H. N., Horvath J. E., Moreira M. A., Kes-sing B., Pontius J., Roelke M., Rumpler Y., Schneider M. P., Silva A., O'Brien S. J., Pecon-Slattey J. A molecular phylogeny of living primates. *PLoS Genet.* 2011. Vol. 7, № 3. e1001342. doi:10.1371/journal.pgen.1001342.
 28. Pidpala O. V., Lukash L. L. Distribution of mobile genetic elements in the human O(6)-methyl-guanine-DNA methyltransferase gene. XII International Scientific and Practical Conference «Current problems of modern biology and human health». 2012. Issue 12. P. 150–154 [in Ukrainian] / Підпала О. В., Лукаш Л. Л. Розподіл мобільних генетичних елементів у гені O⁶-метилгуанін-ДНК-метил-трансфераза людини. XII Міжнародна науково-практична конференція «Актуальні проблеми сучасної біології та здоров'я людини». 2012. Вип. 12. С. 150–154.
 29. Pidpala O. V., Lukash L. L. Recombination origin of the nuclear introns. *Factors Experimental Evol. Organisms.* 2017. Vol. 20. P. 329–334 [in Ukrainian] / Підпала О. В., Лукаш Л. Л. Реконбінаційне походження ядерних інтронів. Фактори експериментальної еволюції організмів. 2017. Т. 20. С. 329–334.
 30. Pidpala O. V., Lukash L. L. The analysis of human MGMT gene orthologous in Protists. *Factors Experimental Evol. Organisms.* 2018. Vol. 22. P. 345–351 [in Ukrainian] / Підпала О. В., Лукаш Л. Л. Аналіз ортологів гена GMT у найпростіших еукаріот. Фактори експериментальної еволюції організмів. 2018. Т. 22. С. 345–351.
 31. Pidpala O. V., Lukash L. L. Formation of the L1Hs retroelement in the intron of the MGMT gene of Hominoidea. *Factors Experimental Evol. Organisms.* 2019. Vol. 24. P. 338–344. [in Ukrainian] / Підпала О. В., Лукаш Л. Л. Формування ретроелемента L1Hs в інтроні гена MGMT людиноподібних мавп. Фактори експериментальної еволюції організмів. 2019. Т. 24. С. 338–344.
 32. Pidpala O. V., Lukash L. L. In silico analysis of MGMT gene orthologous in the most ancient Strepsirrhini primates. *Factors Experimental Evol. Organisms.* 2020. Vol. 26. P. 305–310 [in Ukrainian] / Підпала О. В., Лукаш Л. Л. In silico аналіз ортологів гена MGMT у найдревніших приматів Strepsirrhini. Фактори експериментальної еволюції організмів. 2020. Т. 26. С. 305–310.
 33. Pidpala O. V., Lukash L. L. Species-specific mobile genetic elements in the gene of repair enzyme MGMT in New World monkeys. *Factors Experimental Evol. Organisms.* 2021. Vol. 28. P. 128–134 [in Ukrainian] / Підпала О. В., Лукаш Л. Л. Видоспецифічні мобільні генетичні елементи у гені репаративного ензиму MGMT широконосих мавп. Фактори експериментальної еволюції організмів. 2021. Т. 28. С. 128–134.
 34. Pidpala O. V., Lukash L. L. Evolutionary history of species-specific Alu repeats on the example of the MGMT gene of Old World monkey. *Factors Experimental Evol. Organisms.* 2021. Vol. 28. P. 128–134 [in Ukrainian] / Підпала О. В., Лукаш Л. Л. Еволюційна історія видоспецифічних Alu-повторів на прикладі гена MGMT мавпових. Фактори експериментальної еволюції організмів. 2022. Т. 30. С. 126–132.
 35. Pidpala O. V., Lukash L. L. Distribution of the macaques genus-specific Alu repeat AluMacYa3 in the MGMT gene orthologs of Old World monkeys. *Factors Experimental Evol. Organisms.* 2023. Vol. 32. P. 148–154 [in Ukrainian] / Підпала О. В., Лукаш Л. Л. Поширення родоспецифічного Alu-повтору макак AluMacYa3 у гені MGMT мавпових. Фактори експериментальної еволюції організмів. 2023. Т. 32. С. 148–154.
 36. Quentin Y. Origin of the Alu family: a family of Alu-like monomers gave birth to the left and the right arms of the Alu elements. *Nucleic Acids Res.* 1992. Vol. 20, № 13. P. 3397–3401.
 37. Raam R. L., Sterner K. N., Noviello C. M., Stewart C. B., Disotell T. R. Catarrhine primate divergence dates estimated from complete mitochondrial genomes: concordance with fossil and nuclear DNA evidence. *J. Hum. Evol.* 2005. Vol. 48. P. 237–257.
 38. Salem A. H., Ray D. A., Xing J., Callinan P. A., Myers J. S., Hedges D. J., Garber R. K., Witherspoon D. J., Jorde L. B., Batzer M. A. Alu elements and hominid phylogenetics. *Proc. Natl. Acad. Sci. USA.* 2003. Vol. 100, № 22. P. 12787–12791. doi:10.1073/pnas.2133766100.
 39. 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 haplorrhine primate *Tarsius syrichta* reveals unusual insertions. *Nat. Commun.* 2016. Vol. 7. P. 12997. doi: 10.1038/ncomms12997.
 40. Skowronski J., Fanning T. G., Singer M. F. Unit-length line-1 transcripts in human teratocarcinoma Cells. *Mol. Cell. Biol.* 1988. Vol. 8, № 4. P. 1385–1397.
 41. Smit A. F., Tóth 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.
 42. 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.

43. Ullu E., Tschudi C. Alu sequences are processed 7SL RNA genes. *Nature*. 1984. Vol. 312, № 5990. P. 171–172.
44. Vargiu L., Rodriguez-Tomé P., Sperber G. O., Cadeddu M., Grandi N., Blikstad V., Tramontano E., Blomberg J. Classification and characterization of human endogenous retroviruses: mosaic forms are common. *Retrovirology*. 2016. Vol. 13. P. 7. doi: 10.1186/s12977-015-0232-y.

*Стаття надійшла до редакції 3.09.2023,
прийнята до друку 20.09.2023*

**EVOLUTION OF THE *MGMT* GENE
IN PRIMATES: FROM STRUCTURE
TO MOBILE GENETIC ELEMENTS**

Pidpala O. V., Lukash L. L.

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

The article collects data on the evolution of the reparative *MGMT* gene in primates and traces the evolutionary fate of a number of mobile genetic elements (MGEs). It was concluded that the evolution of the *MGMT* gene in primates underwent various changes from acquisition to loss of exon / intron sequences. As for primate-specific MGEs, they have a different evolutionary history and path — from the formation of species / genus-specific sequences to deletional degradation and can be components not only of intronic or exonic sequences, but also components of exons in the form of fragmented sequences and not be identified as MGEs.

Keywords: primates, gene *MGMT*, primate-specific mobile genetic elements.