

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

АНАЛІЗ ОРТОЛОГІВ ГЕНА *MGMT* ЛЮДИНИ У НАЙПРОСТІШИХ ЕУКАРІОТ

Мета. Проаналізувати інтронні послідовності ортологів гена *MGMT* на ранніх етапах формування еукаріот на прикладі протистів. **Методи.** Гомологію між нуклеотидними послідовностями визначали програмою *BLAST 2.6.1*. Нуклеотидні послідовності бактеріальних та мітохондріальних інтронів II групи взято для аналізу із бази даних *Database for Bacterial Group II Introns*. Результати пошуку та ідентифікації мобільних генетичних елементів здійснено за допомогою програми *CENSOR*. **Результати.** Еволюція гена не завжди узгоджується з еволюцією організму, зокрема, це простежується у соціальних амеб. У межах роду досліджуваний ген набував і втрачав інтрони. Простежується фрагментарна гомологія між різними структурними одиницями ортологів. **Висновки.** На основі одержаних результатів показано ендегенний мозаїчний характер формування структурних одиниць гена

Ключові слова: ортологи гена O⁶-метилгуанін-ДНК метилтрансферази (*MGMT*), Протисти, еволюція гена, сплайсосомні інтрони, втрата і набуття інтрону.

Геном клітини постійно 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 O⁶-метилгуанін-ДНК метилтрансферазу (*MGMT*), який видаляє алкільні групи з O⁶-позиції гуаніну у ДНК і захищає клітини від їхнього токсичного та мутагенного впливів [1]. Цей ензим відіграє також значну роль в алкілувальній хімотерапії раку [2, 3]. *MGMT* широко розповсюджений у про- та еукаріотних організмах. Небагато відомо про еволюцію цього ензиму [4], ще менше про еволюцію його гена. Метою нашої роботи було дослідити структурні особливості та проаналізувати інтронні послідо-

вності названого гена на прикладі найпростіших еукаріотних організмів.

Матеріали і методи

Інформацію про ортологів гена *MGMT*, що кодує відповідний репаративний ензим, одержано із бази даних *Ensembl Protist* і наведено у табл. 1, 2. Гомологію між досліджуваними послідовностями визначали програмою *BLAST 2.6.1*. Нуклеотидні послідовності бактеріальних та мітохондріальних інтронів II групи взято для аналізу із бази даних *Database for Bacterial Group II Introns*. Результати пошуку та ідентифікації мобільних генетичних елементів (МГЕ) здійснено за допомогою програми *CENSOR*.

Результати та обговорення

Протисти (Найпростіші, Protista) – чисельна група різних за будовою і способом життя ядерних організмів. В еволюційному плані вони є перехідною групою (переважно одноклітинних еукаріот) від прокариот до еукаріот із багатоклітинною організацією [5].

Характерною ознакою еукаріотних геномів є екзон-інтронна організація генів. У геномах багатоклітинних модельних організмів, зокрема таких, як *D. melanogaster*, *C. elegans* і *A. thaliana*, більше 80 % генів мають сплайсосомні (ядерні) інтрони [6–8]. У геномах людини, щура і миші частка таких генів складає більше 90 % [9]. Що стосується нижчих еукаріот, то частка генів з інтронами у їхніх геномах коливається від 0,7 до 97 % [10]. За даними просеквенованих геномів протистів теж можна зробити висновок про великий діапазон зміни частки генів з інтронами. Зокрема, для геномів *Leishmania infantum*, *Trypanosoma brucei*, *Cryptosporidium parvum* характерно, що вони містять одиничні гени з інтронами, тоді як у геномах *Dicystelium discoideum*, *Plasmodium falciparum*, *Paramecium tetraurelia* і *Theileria parva* таких генів більше 50 %. Як правило, у протистів, порівняно з іншими еукаріотами, інтрони коротші.

Таблиця 1. Основна інформація про досліджувані ортологи

Організм	Назва гена	Умовне позначення	Довжина гена, п. н.	Довжина мРНК, п. н.	Довжина білка, аа	Екзони, п	Інтрони, п
<i>Dictyostelium discoideum</i>	<i>mgmt-1</i>	pr1	618	618	205	1	-
<i>Dictyostelium lacteum</i>	<i>DLAC_00935</i>	pr2	573	573	190	1	-
<i>Dictyostelium purpureum</i>	<i>DICPU-DRAFT_75205</i>	pr3	689	606	201	2	1
<i>Gregarina niphandrodes</i>	<i>GNI_094800</i>	pr4	621	621	206	1	-
<i>Polysphondylium pallidum</i> PN500	<i>PPL_08681</i>	pr5	1900	1104	367	4	3
<i>Pseudocohnilembus persalinus</i>	<i>PPERSA_06800</i>	pr6	1040	753	250	4	3
<i>Sphaeroforma arctica</i> JP610	<i>SARC_02795</i>	pr7	888	888	186	1	-
<i>Trichomonas vaginalis</i> G3	<i>TVAG_401340</i>	pr8	459	459	152	1	-

Таблиця 2. Дані про екзон-інтронну структуру ортологів гена *MGMT*

Організм	Назва гена	Довжина структурних одиниць, п. н.						
		ек.1	ін.1	ек.2	ін.2	ек.3	ін.3	ек.4
<i>Dictyostelium discoideum</i>	<i>mgmt-1</i>	618						
<i>Dictyostelium lacteum</i>	<i>DLAC_00935</i>	573						
<i>Dictyostelium purpureum</i>	<i>DICPUDRAFT_75205</i>	454	83	152				
<i>Gregarina niphandrodes</i>	<i>GNI_094800</i>	621						
<i>Polysphondylium pallidum</i> PN500	<i>PPL_08681</i>	704	457	235	70	128	264	37
<i>Pseudocohnilembus persalinus</i>	<i>PPERSA_06800</i>	229	162	278	72	167	56	79
<i>Sphaeroforma arctica</i> JP610	<i>SARC_02795</i>	888						
<i>Trichomonas vaginalis</i> G3	<i>TVAG_401340</i>	459						

Крім варіювання кількості генів з інтронами у геномах протистів, виявлено різне співвідношення довжин екзонів та інтронів у кожному геномі [11].

Фрагментарний характер гомології у безінтронних ортологів гена *MGMT*. Досліджуваний ген у двох соціальних амеб *Dictyostelium discoideum* і *Dictyostelium lacteum* та у *Gregarina niphandrodes*, *Sphaeroforma arctica* JP610 і *Trichomonas vaginalis* G3 представлений одним екзоном. Відсоток гомології між генами коливається від 14 до 61 %. Аналізуючи нуклеотидні послідовності цих генів, ми виявили, що гомологія між ними має фрагментарний характер. На рис. 1, як приклад, наведено характер гомології між двома ортологами. Гомологічні послідовності довжиною 7–12 п. н. у безінтронних генів розкидані по гену та іноді

мають протилежний напрямок (рис. 2).

Аналіз інтронних послідовностей у досліджуваних генах. Серед досліджуваних ортологів інтрони присутні у трьох генах (табл. 1, 2 та рис. 1): гені *DICPUDRAFT_75205 Dictyostelium purpureum* (83 п.н.), гені *PPL_08681 Polysphondylium pallidum* PN500 (457, 70, 264 п. н.) та у гені *PPERSA_06800 Pseudocohnilembus persalinus* (162, 72, 56 п. н.). Гомологію досліджували як між єдиним інтроном гена *DICPUDRAFT_75205 D. purpureum* і відповідно між різними інтронами генів *PPL_08681 P. pallidum* PN500 та *PPERSA_06800 P. burspersalinus*, так і між інтронами у межах одного гена. Між усіма інтронами гомологія не була протяжною і коливалася від 12 % до 83 %, але навіть у випадку високої гомології її характер також був фрагментарним.

a)

```

1 ATGAGCAAAA GACAAAGAGA CTCAGACAAA AATCCTTCAC AAAAAAAAAA ATTCAAAGAA 60
61 ACACAGATTT ATTTTGATCA ATTTGAAACT GGAATTGGTG TAATAGGGTT STATTCSTAAT 120
121 GAAAAGGAGT TATTGCAAAT TGAAATTGAC AGTACTGATG ATAAATCCCT GTACAAATGT 180
181 AATATCATTA CGAATCAATT CAAAAAGAA TTAGAAAGAT ACTTTGGTGG ACTATTTGAT 240
241 GAAAATGAGT GGT TAAAAC ACCAATTGAC TATAACAAG GAAGTGAATT CCAACAAGCT 300
301 GTTTGGAAC AAATGTTAAA GATACCTTAT GGAAAACTA CTTCCSTATGG TACCATGGCT 360
361 AAAGAAATGG GATATACAAG TTGCTATTCA AGAGCTATTG CATCTGCTTG CAGACTTAAT 420
421 AAATTCCTCAA TTATCATACC TTGTCATAGA GTAGTTTCTT CCACTGGTGA GATTGTTGGC 480
481 TATATGGGTA AACATGGTAT TTCAAAACAA AAATTCCTAT TTAATTTAGA AACACCCT 540
541 TTATCAACAC CAAAAAAGA AATTGATACA TAG

```

б)

```

301 GTTTGGAAC AAATGTTAAA GATACCTTAT GGAAAACTA CTTCCSTATGG TACCATGGCT 360
      166 TAAA GATA 174
      248 AAAC AAAT 255      302 AT GGA-AACTA C 313
367 GTTTGGA 373

361 AAAGAAATGG GATATACAAG TTGCTATTCA AGAGCTATTG CATCTGCTTG CAGACTTAAT 420
      420 AG TTGCTA 427      184 CATCTGCT 177
      582 AAT

421 AAATTCCTCAA TTATCATACC TTGTCATAGA GTAGTTTCTT CCACTGGTGA GATTGTTGGC 480
      69 A TTATCAT 62      521 TTTCTT C 527
      AAAT 575

```

Рис. 1. Характер гомології між ортологами pr2 та pr1: а – загальний вигляд гомології між послідовностями генів *DLAC_00935 D. lacteum* та *mgmt-1 D. discoideum* із зазначенням гомологічних ділянок жирним шрифтом; б – розгорнута інформація на прикладі ділянки гена *DLAC_00935 D. lacteum* (у межах координат 301–480 п. н.) та фрагментів гомології до гена *mgmt-1 D. discoideum* із координатами у межах відповідного гена.

```

121 GAAAAGGAGT TATTGCAAAT TGAAATTGAC AGTACTGATG ATAAATCCCT GTACAAATGT 180
      83 AAAT TGAAA 91      181 GATG ATAAATC 191
      422 ACTGATG A 415
      42 GAC AGTAC 35      269 ATCCCT GT 276
      881 TGAAATTGAC 872
      172 TACTGATG 165
      133 TGCAAAT T 140

```

pr1
pr1
pr4
pr7
pr7
pr8

Рис. 2. Фрагменти гомології між безінтронними ортологічними генами до ділянки послідовності гена *DLAC_00935 Dictyostelium lacteum* (у межах координат 121–180 п. н.).

Варто зазначити, що гомологічні послідовності трапляються не лише у різних інтронах, але і у екзонах ортологів (рис. 3). На прикладі гена *DICPUDRAFT_75205 D. purpureum* бачимо, що одні й ті ж фрагменти нуклеотидних послідовностей можуть бути складовими елементами різних структурних частин гена.

Пошуки гомології між інтронами II групи та інтронами досліджуваних ортологів. Питання походження сплайсосомних інтронів ще й досі лишається відкритим. Розглядається зв'язок між сплайсосомними інтронами і бактеріальними самосплайсуючими інтронами групи II, які, згідно із сценарієм еукаріогенезу, спочатку потрапили до органел (мітохондрій і хлоропластів), а згодом до ядра [12].

Аналізуючи інтрони досліджуваних генів (pr3, pr5 і pr6), ми не виявили протяжної гомології між ними і бактеріальними чи мітохондріальними інтронами групи II. Відсоток гомології коливався від 12 % до 100 %. Але навіть за 100 % гомології він мав фрагментарний характер. Це проілюстровано на рис. 4 на прикладі найменшого із інтронів (ін.3, 56 п. н.) гена *PPERSA_06800 Pseudocohnilem buspersalinus*.

Мобільні генетичні елементи у послідовностях ортологів гена *MGMT*. Мобільні генетичні елементи широко представлені в еукаріотних геномах. У протистів вони становлять від 1 % до 60 % геному, зокрема у *Trypanosoma* 2–5 %, *Amoeba proteus* 4 %, *Paramecium tetraurelia* 8,5 %, *Entamoeba histolytica* 24 % [13–15].



Рис. 3. Гомологія між інтронними та екзонними послідовностями (вказано жирним шрифтом) гена *pr3*: *a* – загальний вигляд гомології; *б* – фрагменти гомології до екзонних послідовностей. Наведено координати послідовностей у межах відповідних екзонів та вказано їхні номери.

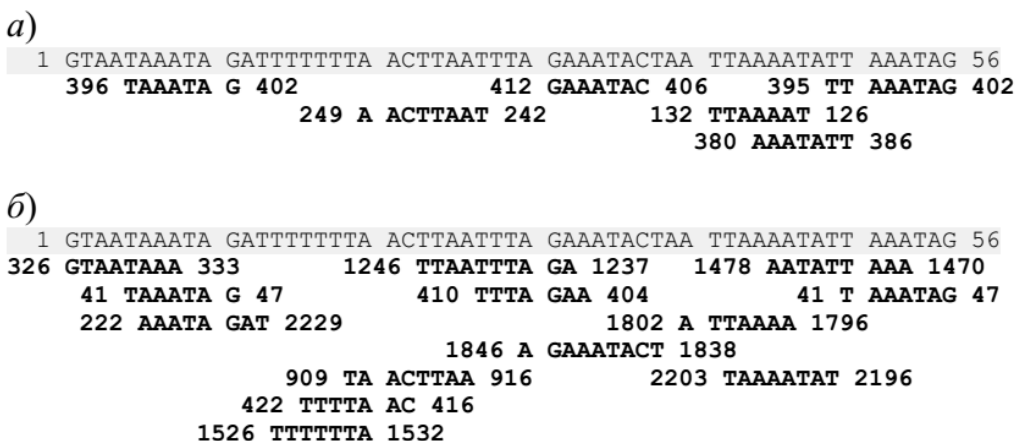


Рис. 4. Характер гомології між послідовністю інтрону 3 гена *PPERSA_06800 Pseudocohnilem buspersalinius* та інтронами II групи: *a* – бактеріального інтрону II групи *A.pt.II* (AF369871); *б* – мітохондріального інтрону II групи *LSU.I2* (AF114794). Гомологічні ділянки позначено жирним шрифтом і вказано координати у межах відповідних послідовностей.

Виявлено представників усіх класів МГЕ. Переважають LINE і SINE-елементи, менше ДНК-транспозонів і LTR-елементів. Рекордсменом серед одноклітинних мікроорганізмів є трихомонада вагінальна (*Trichomonas vaginalis*), геном якої на дві третини складається із МГЕ, переважно (37 %) із ДНК-транспозонів, зокрема таких, як *Maverick* [16]. У соціальних амеб відсоток МГЕ коливається від 1 % до 10 % [17].

Зв'язок між МГЕ та інтронами обговорюється давно. Є навіть експериментальні підтвердження, що МГЕ можуть генерувати нові ін-

тронні послідовності у різних організмів [18–21].

У випадку досліджуваних генів серед протистів не виявлено видоспецифічних МГЕ, лише фрагменти МГЕ вищих тварин. У гені *mgmt-1 D. discoideum* (*pr1*) ідентифіковано дві послідовності, гомологічні ДНК-транспозонам (*Polinton 3_SM*; *hAT-N26_DR*), які становлять 42 % від загальної довжини гена. Фрагмент ДНК-транспозону (*EnSpm-21_HMa*) виявлено і у гені *DLAC_00935 D.lacteum* (*pr2*). Його відсоток значно менший і становить майже 8 %. У гені

DICPUDRAFT_75205 ще однієї соціальної амеби *D. purpureum* (pr3) також присутні фрагменти ДНК-транспозонів, дві послідовності в екзоні 1 (*EnSpm-20_HMa*; *CryptonV-N2_DR*), і одна послідовність (*Academ-8_SK*) на 72 % перекривається із інтроном і охоплює початок екзону 2. Загальна ж частка МГЕ становить 27 % від довжини гена. У гені *PPL_08681 P. pallidum PN500* (pr5), крім двох фрагментів ДНК-транспозонів (*DNA-8-29_DR*; *Chapaev-8_HM*), є послідовність LTR-елемента (*Gypsy-1-I_LG*). У цьому гені фрагменти МГЕ переважають в інтронній послідовності (29 %), тоді як у екзонах їхня частка не перевищує 1 %. Цікаво, що у гені *PPERSA_06800 P. buspersalinus* (pr6), на відміну від інших ортологічних генів у *Protist*, серед МГЕ виявлено фрагменти Non-LTR елементів (*Kiri-6_DK*; *L1_Canid_5end*), які переважно локалізовані в екзонних послідовностях. У трьох генах (pr4, pr7 і pr8) МГЕ не виявлено.

Отже, на прикладі ортологів гена *MGMT* у найпростіших еукаріотних організмів бачимо, що еволюція гена не завжди узгоджується із еволюцією організму, зокрема, це простежується у соціальних амеб (pr1, pr2, pr3 і pr5). У межах роду досліджуваний ген набував і втрачав

інтрони. Це узгоджується із літературними даними, що *Dictyostelium*, як тварини і гриби, незалежно від дивергенції, мають схожі механізми втрати інтронів [22]. Питання ж гомології інтронів ортологічних генів досить складне. Показано, що втрата і поява інтронів і навіть зміна їхнього положення у гені відбувається доволі часто, що може із великою ймовірністю змінювати хід їхньої еволюції [23]. Щодо походження інтронів, то і у випадку ортологів гена *MGMT* у протистів простежується фрагментарна гомологія між різними структурними одиницями гена (екзонами, інтронами, МГЕ), що підтверджує припущення про можливий рекомбіногенний характер формування інтронів [24].

Висновки

На основі аналізу структурних особливостей та нуклеотидних послідовностей ортологічних генів *MGMT* у протистів показано, що і на нижчих рівнях організації ген може зазнавати появи і втрати інтронів і що еволюційне походження інтронів, як і інших структурних одиниць гена, може мати ендегенний мозаїчний характер і здійснюється на основі коротких гомологічних фрагментів.

Література

1. Pegg A.E. Repair of O⁶-alkylguanine by alkyltransferases. *Mutat. Res.* 2000. Vol. 262. P. 83–100.
2. 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.
3. 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.
4. 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.
5. Adl S.M., Simpson A.G., Farmer M.A., Andersen R.A., Anderson O.R., Barta J.R., Bowser S.S., Brugerolle G., Fensome R.A., Fredericq S., James T.Y., Karpov S., Kugrens P., Krug J., Lane C.E., Lewis L.A., Lodge J., Lynn D.H., Mann D.G., McCourt R.M., Mendoza L., Moestrup O., Mozley-Standridge S.E., Nerad T.A., Shearer C.A., Smirnov A.V., Spiegel F.W., Taylor M.F. The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. *J. Eukaryot. Microbiol.* 2005. Vol. 52. P. 399–451. doi: 10.1111/j.1550-7408.2005.00053.x.
6. Arabidopsis Genome Initiative. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature*. 2000. Vol. 408. P. 796–815.
7. C. elegans Sequencing Consortium. Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science*. 1998. Vol. 282. P. 2012–2018.
8. Adams M.D., Celniker S.E., Holt R.A. The genome sequence of *Drosophila melanogaster*. *Science*. 2000. Vol. 287. P. 2185–2195.
9. Venter J.C., Adams M.D., Myers E.W. The sequence of the human genome. *Science*. 2001. Vol. 291. P.1304–1351. doi: 10.1126/science.1058040.
10. Иващенко А.Т., Тауасарова М.К., Атамбаева Ш.А. Экзон-интронная структура генов полных геномов грибов. *Молекулярная биология*. 2009. Т. 43. С. 28–35.
11. Иващенко А.Т., Кабдуллина А.А. Экзон-интронная организация генов некоторых видов протистов. *Генетика и биотехнология XXI века. Фундаментальные и прикладные аспекты*. Минск: Изд. центр БГУ, 2008. С. 241–243.
12. Rogozin I.B., Carmel L., Csuros M., Koonin E.V. Origin and evolution of spliceosomal introns. *Biol. Direct*. 2012. Vol. 7. P. 11. doi: 10.1186/1745-6150-7-11.
13. Bringaud F., Ghedin E., El-Sayed N.M., Papadopoulou B. Role of transposable elements in trypanosomatids. *Microbes. Infect.* 2008. Vol. 10. P. 575–581. doi: 10.1016/j.micinf.2008.02.009.

14. Lorenzi H., Thiagarajan M., Haas B., Wortman J., Hall N., Caler E. Genome wide survey, discovery and evolution of repetitive elements in three *Entamoeba* species. *BMC Genomics*. 2008. Vol. 10. P. 595. doi: 10.1186/1471-2164-9-595.
15. Thomas M.C., Macias F., Alonso C., López M.C. The biology and evolution of transposable elements in parasites. *Trends Parasitol.* 2010. Vol. 26. P. 350–362. doi: 10.1016/j.pt.2010.04.001.
16. Bradic M., Warring S.D., Low V., Carlton J.M. The Tc1/mariner transposable element family shapes genetic variation and gene expression in the protist *Trichomonas vaginalis*. *Mob. DNA*. 2014. Vol. 5. P. 12. doi: 10.1186/1759-8753-5-12.
17. Glockner G. Genome analysis of social Amoebae. In: Dictyostelids: Evolution, Genomics and Cell Biology. Eds M.Romeralo, S. Baldanf, R. Escalante. *Springer Science and Business Media*. 2013. P. 35–48. doi: 10.1007/978-3-642-38487-5.
18. Menssen A., Höhmann S., Martin W., Schnable P.S., Peterson P.A., Saedler H., Gierl A. The En/Spm transposable element of *Zea mays* contains splice sites at the termini generating a novel intron from a dSpm element in the A2 gene. *EMBO J.* 1990. Vol. 9. P. 3051–3057.
19. Giroux M.J., Clancy M., Baier J., Ingham L., McCarty D., Hannah L.C. De novo synthesis of an intron by the maize transposable element Dissociation. *Proc. Natl. Acad. Sci. USA*. 1994. Vol. 91. P. 12150–12154.
20. Nouaud D., Boëda B., Levy L., Anxolabéhère D. A P element has induced intron formation in *Drosophila*. *Mol. Biol. Evol.* 1999. Vol. 16. P. 1503–1510.
21. Huff J.T., Zilberman D., Roy S.W. Mechanism for DNA transposons to generate introns on genomic scales. *Nature*. 2016. Vol. 538. P. 533–536. doi: 10.1038/nature20110.
22. Ma M.Y., Che X.R., Porceddu A., Niu D.K. Evaluation of the mechanisms of intron loss and gain in the social amoebae *Dictyostelium*. *BMC Evol. Biol.* 2015. Vol. 15. P. 286. doi: 10.1186/s12862-015-0567-y.
23. Roy S.W., Gilbert W. The evolution of spliceosomal introns: patterns, puzzles and progress. *Nat. Rev. Genet.* 2006. Vol. 7. P. 211–221. doi: 10.1038/nrg1807.
24. Підпала О.В., Лукаш Л.Л. Рекомбінаційне походження ядерних інтронів. *Фактори експериментальної еволюції організмів*. 2017. Т. 20. С. 329–334.

References

1. Pegg A.E. Repair of O⁶-alkylguanine by alkyltransferases. *Mutat. Res.* 2000. Vol. 262. P. 83–100.
2. 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.
3. 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.
4. 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., Santibañez-Koref M.F. Alkyltransferase-like proteins. *DNA Repair (Amst)*. 2007. Vol. 6. P. 1222–1228. doi: 10.1016/j.dnarep.2007.03.014.
5. Adl S.M., Simpson A.G., Farmer M.A., Andersen R.A., Anderson O.R., Barta J.R., Bowser S.S., Brugerolle G., Fensome R.A., Fredericq S., James T.Y., Karpov S., Kugrens P., Krug J., Lane C.E., Lewis L.A., Lodge J., Lynn D.H., Mann D.G., McCourt R.M., Mendoza L., Moestrup O., Mozley-Standridge S.E., Nerad T.A., Shearer C.A., Smirnov A.V., Spiegel F.W., Taylor M.F. The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. *J. Eukaryot. Microbiol.* 2005. Vol. 52. P. 399–451. doi: 10.1111/j.1550-7408.2005.00053.x.
6. Arabidopsis Genome Initiative. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature*. 2000. Vol. 408. P. 796–815.
7. C. elegans Sequencing Consortium. Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science*. 1998. Vol. 282. P. 2012–2018.
8. Adams M.D., Celniker S.E., Holt R.A. The genome sequence of *Drosophila melanogaster*. *Science*. 2000. Vol. 287. P. 2185–2195.
9. Venter J.C., Adams M.D., Myers E.W. The sequence of the human genome. *Science*. 2001. Vol. 291. P.1304–1351. doi: 10.1126/science.1058040.
10. Ivashchenko A.T., Tauasarova M.K., Atambayeva S.A. Exon-intron structure of genes of fungi genomes. *Molecular biology*. 2009. Vol. 43. S. 28–35.
11. Ivashchenko A.T., Kabdullina A.A. Exon-intron organization of genes of some Protista species. *Genety`ka y` by`otexnologiy`ya XXI veka. Fundamental`nye y` pry`kladnye aspekty*. Mu`nsk: Y`zd. centr BGU, 2008. С. 241–243.
12. Rogozin I.B., Carmel L., Csuros M., Koonin E.V. Origin and evolution of spliceosomal introns. *Biol. Direct*. 2012. Vol. 7. P. 11. doi: 10.1186/1745-6150-7-11.
13. Bringaud F., Ghedin E., El-Sayed N.M., Papadopoulou B. Role of transposable elements in trypanosomatids. *Microbes. Infect.* 2008. Vol. 10. P. 575–581. doi: 10.1016/j.micinf.2008.02.009.
14. Lorenzi H., Thiagarajan M., Haas B., Wortman J., Hall N., Caler E. Genome wide survey, discovery and evolution of repetitive elements in three *Entamoeba* species. *BMC Genomics*. 2008. Vol. 10. P. 595. doi: 10.1186/1471-2164-9-595.
15. Thomas M.C., Macias F., Alonso C., López M.C. The biology and evolution of transposable elements in parasites. *Trends Parasitol.* 2010. Vol. 26. P. 350–362. doi: 10.1016/j.pt.2010.04.001.
16. Bradic M., Warring S.D., Low V., Carlton J.M. The Tc1/mariner transposable element family shapes genetic variation and gene expression in the protist *Trichomonas vaginalis*. *Mob. DNA*. 2014. Vol. 5. P. 12. doi: 10.1186/1759-8753-5-12.
17. Glockner G. Genome analysis of social Amoebae. In: Dictyostelids: Evolution, Genomics and Cell Biology. Eds M.Romeralo, S. Baldanf, R. Escalante. *Springer Science and Business Media*. 2013. P. 35–48. doi: 10.1007/978-3-642-38487-5.

18. Menssen A., Höhmann S., Martin W., Schnable P.S., Peterson P.A., Saedler H., Gierl A. The En/Spm transposable element of *Zea mays* contains splice sites at the termini generating a novel intron from a dSpm element in the A2 gene. *EMBO J.* 1990. Vol. 9. P. 3051–3057.
19. Giroux M.J., Clancy M., Baier J., Ingham L., McCarty D., Hannah L.C. De novo synthesis of an intron by the maize transposable element Dissociation. *Proc. Natl. Acad. Sci. USA.* 1994. Vol. 91. P. 12150–12154.
20. Nouaud D., Boëda B., Levy L., Anxolabéhère D. A P element has induced intron formation in *Drosophila*. *Mol. Biol. Evol.* 1999. Vol. 16. P. 1503–1510.
21. Huff J.T., Zilberman D., Roy S.W. Mechanism for DNA transposons to generate introns on genomic scales. *Nature.* 2016. Vol. 538. P. 533–536. doi: 10.1038/nature20110.
22. Ma M.Y., Che X.R., Porceddu A., Niu D.K. Evaluation of the mechanisms of intron loss and gain in the social amoebae *Dictyostelium*. *BMC Evol. Biol.* 2015. Vol. 15. P. 286. doi: 10.1186/s12862-015-0567-y.
23. Roy S.W., Gilbert W. The evolution of spliceosomal introns: patterns, puzzles and progress. *Nat. Rev. Genet.* 2006. Vol. 7. P. 211–221. doi: 10.1038/nrg1807.
24. Pidpala O.V., Lukash L.L. Recombinational origin of the nuclear introns. *Factors in Experimental Evolution of Organisms.* 2017. Vol. 20. S. 329–334.

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

THE ANALYSIS OF HUMAN *MGMT* GENE ORTHOLOGOUS IN PROTISTS

Aim. The intron sequences of orthologous O⁶-methylguanin-DNA methyltransferase (*MGMT*) genes in Protists on the early stages of their formation in eukaryotic organisms have been analysed. **Methods.** Homologous regions have been defined by the program *BLASTN 2.6.1*. Nucleotide sequences of the bacterial and mitochondrial group II introns have been taken from *Database for Bacterial Group II Introns*. Searching and identifying the MGEs have been realized by using *CENSOR*. **Results.** It has been shown that the evolution of the gene does not always coincide with the evolution of the organism. This is shown on the example of intron loss and gain in social amoebae *Dictyostelium*. Also it has been found the fragmentary nature of homology between various introns and exons of the orthologous genes. **Conclusions.** The obtained results allow offer a suggestion about the endogenous mosaic character of the evolutionary formation of the gene structural units.

Keywords: O⁶-methylguanin-DNA methyltransferase (*MGMT*) gene orthologous, Protists, gene evolution, spliceosomal introns, intron loss and gain.