

ФРАГМЕНТИ МОБІЛЬНИХ ГЕНЕТИЧНИХ ЕЛЕМЕНТІВ ЯК СКЛАДОВІ ПОТЕНЦІЙНИХ АЛЬТЕРНАТИВНИХ ПРОМОТОРІВ ГЕНА *MGMT* ЛЮДИНИ

Геном еукаріотів є складною і динамічною структурою. Лише незначна частина генома людини (3–5%) представлена білок-кодуючими послідовностями, тоді як близько 50 % ДНК охоплено мобільними генетичними елементами (МГЕ) [1]. Тривалий час цю частину генома вважали «сміттєвою», «гоїстичною», «темною матерією», тобто непотрібним баластом. Проте все більше дослідників схилиються до думки, що МГЕ відіграють суттєву роль у мінливості та еволюції геномів, спричиняючи геномні перебудови, створюючи нові гени і модифікуючи регуляторні механізми існуючих генів [2–4]. Відомо, що близько 25 % генів людини містять МГЕ у промоторних ділянках, і на сьогодні є переконливі дані про їхнє залучення, зокрема промоторів ретроелементів, до регуляції транскрипційної активності генів [5–10]. Є випадки, коли МГЕ виконують роль альтернативних промоторів, що призводить або до збільшення рівня експресії відповідного гена, або до зміни тканинспецифічності його експресії [10].

У попередніх своїх дослідженнях ми вивчали розподіл МГЕ у гені репаративного ензиму O⁶-метилгуанін-ДНК метилтрансферази людини (*MGMT*) і у його промоторній ділянці. Виявили, що МГЕ присутні як в інтронних послідовностях [11], так і у промоторній ділянці досліджуваного гена [12]. В інтронних послідовностях МГЕ утворюють композиційні кластерні структури, які можуть бути джерелом різноманітних регуляторних послідовностей і мають потенціал для формування альтернативних промоторів [11]. У промоторі гена *MGMT* ідентифіковано три послідовності, які гомологічні фрагментам МГЕ: два *LTR*-повтори і фрагмент ДНК транспозону. Той факт, що один із *LTR*-повторів містить описані елементи відгуку на глюкокортикоїдні гормони (GRE) [13, 14], і те, що відомий мінімальний промотор і послідовність сайту SP1 розташовані у межах фрагмента ДНК-транспозону, свідчать на користь важливої ролі МГЕ у генній регуляції [12].

У базі даних AceView наведено інформацію про можливі варіанти мРНК та про нуклеотидні послідовності, які можуть містити потенційні альтернативні промоторні ділянки гена *MGMT* людини. Чи перекриваються вони із послідовністю реферованого промотора гена *MGMT* людини, чи містять у своєму складі МГЕ і чи несуть вони регуляторний потенціал? Ці питання визначили мету нашого дослідження.

Матеріали і методи

Дані про потенційні промоторні ділянки гена репаративного ензиму *MGMT* одержано із бази даних AceView і основну інформацію подано у табл. 1. Результати пошуку та ідентифікації МГЕ здійснено за допомогою програми *CENSOR*. Гомологію між досліджуваними послідовностями визначали програмою *BLAST 2.2.32*. Пошук потенційних регуляторних послідовностей здійснювали програмами *SITECON*, *SiteGA*, *BLASTN 2.2.26*, використовуючи ресурси бази даних TRRD (Transcription Regulatory Regions Database), а також програмами *Cister: Cis-element Cluster Finder*, *WWW Signal Scan* і *Tfsitescan*.

Результати та обговорення

Ген *MGMT* людини локалізований на теломерній ділянці хромосоми 10 у положенні 10q26 і складається з одного некодуючого і чотирьох кодуючих екзонів та чотирьох інтронів. Реферований промотор досліджуваного гена (номер у GeneBank X61657) має довжину 1157 п.н. і охоплює екзон 1 і частину інтрону 1 [13, 14]. У базі даних AceView представлено нуклеотидну послідовність довжиною 2 тис. п.н., яка охоплює реферований промотор (bAug10), а також вісім потенційних промоторних ділянок. Для чотирьох із них є інформація, що ці послідовності можуть містити промотор (табл. 1).

Результати BLAST аналізу потенційних промоторних ділянок гена *MGMT* людини. Виявлено, що послідовність bAug10 перекривається із реферованою послідовністю (hO⁶P1), за ви-

Таблиця 1

Потенційні альтернативні промоторні ділянки гена репаративного ензиму *MGMT* людини (за даними AceView)

Умовне позначення мРНК варіанту	Довжина мРНК, п.н.	Передбачуваний білок, а.з.	Довжина послідовності перед стартом транскрипції, тис.п.н.	Інформація щодо наявності промотора
aAug10	1814	351	2	МОЖЛИВО МІСТИТЬ
cAug10	886	207	2	МОЖЛИВО МІСТИТЬ
dAug10	514	171	2	*
eAug10	592	166	2	МОЖЛИВО МІСТИТЬ
eAug10	510	144	2	*
gAug10	736	133	2	*
hAug10	569	95	2	*
iAug10	999	73	2	МОЖЛИВО МІСТИТЬ

Примітка. * – Інформація відсутня.

нятком фланкуючої ділянки довжиною 156 п.н., яка частково охоплює енхансерну послідовність. Із восьми потенційних промоторних ділянок три послідовності (cAug10, hAug10 та iAug10) мають гомологію із реферованою послідовністю, а п'ять послідовностей гомології не виявили. Зокрема, послідовність dAug10 локалізована у межах інтрону 1. Чотири інші послідовності (aAug10, eAug10, fAug10 і gAug10) розташовані у межах інтрону 2.

Фрагменти МГЕ в альтернативних промоторних ділянках досліджуваного гена. Показано, що чотири із восьми досліджуваних послідовностей містять фрагменти МГЕ (табл. 2, рис.). Цікаво, що це саме ті послідовності, які, можливо, включають промотор, як зазначається у базі даних AceView. Тому детальніше проаналізували саме ці послідовності. Дві послідовності можливих альтернативних промоторів (cAug10 та iAug10), що перекриваються із реферованою послідовністю (hO^oP1), містять додаткові фрагменти МГЕ – *MIRc* (NonLTR-ретротранспозон) і *MER117* (ДНК-транспозон), тоді як послідовності двох фрагментів *LTR*-повторів ретровірусів ссавців, а саме *LTR*-повтор ретровірусоподібного елемента MaLR (*MLTIC*) і *LTR*-повтор ендегенного ретровірусу ERV3 (*MLTIC2*), ідентифіковані програмою CENSOR як один фрагмент ендегенного ретровірусу через більший розмір нуклеотидної послідовності. Цю особливість програми варто враховувати у подальших дослідженнях. У послідовностях aAug10 і fAug10, що розташовані у межах інтрону 2, ідентифіковано фрагменти МГЕ різних родин (табл. 2, рис.). На особливу увагу заслуговують дві послідовності: *Tigger15a* (ДНК-транспозон) і *AluSx1* (NonLTR-ретротранспозон), які є специфічними МГЕ для ссавців і приматів.

Потенційні регуляторні послідовності у межах ідентифікованих фрагментів МГЕ.

Усі досліджені нами фрагменти МГЕ у складі альтернативних промоторів гена *MGMT* людини збагачені різноманітними регуляторними послідовностями (табл. 3, 4). Зокрема, *AluSx1*-повтор містить низку нових, крім описаних [15–19], сайтів зв'язування транскрипційних факторів, а також тканинно-специфічні промоторні, енхансерні та сайленсерні послідовності (табл. 3). Відомо, що *Alu*-повтори можуть впливати на клітинні процеси, надаючи нові сайти термінації транскрипції або сайти сплайсингу чи виконуючи роль альтернативних промоторів [9, 20–22]. Подібні ситуації можуть порушувати нормальне функціонування гена, проте іноді можуть призводити до утворення білків із новими функціями [23]. ДНК-транспозони також збагачені регуляторними послідовностями [24]. Досліджуваний нами фрагмент елемента *Tigger15a* містить сайти зв'язування для низки транскрипційних факторів (ERE, HRE, MEF-2 і C/EBP), а також промоторні, енхансерні і сайленсерні послідовності та таку, що подібна до локус-контролюючої ділянки (Locus Control Region-like region).

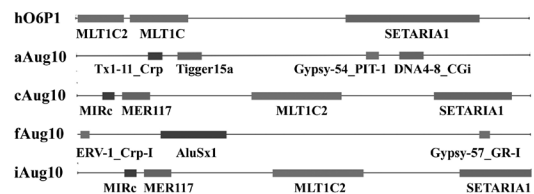


Рис. Фрагменти МГЕ у реферованому промоторі гена *MGMT* людини (hO^oP1) та у його потенційних промоторних послідовностях (aAug10, cAug10, fAug10, iAug10)

Таблиця 2

Послідовності МГЕ у потенційних альтернативних промоторних ділянках гена репаративного ензиму *MGMT* людини

Умовне позначення промотору	Дані про мобільні генетичні елементи				
	Елемент	Клас, родина	Довжина, п.н.	Напрямок	Координати у межах промотору
aAug10	Tx1-11_Crp	NonLTR/Tx1	60	c	320-379
	Tigger15a	DNA/Mariner	105	c	448-552
	Gypsy-54_PIT-I	LTR/Gypsy	54	d	1278-1331
	DNA4-8_CGi	DNA	104	d	1425-1528
cAug10	MIRc	NonLTR/SINE/SINE2	53	c	115-167
	MER117	DNA/hAT	120	d	202-321
	MLT1C2	ERV/ERV3	397	c	771-1167
	SETARIA1	DNA	341	d	1576-1916
fAug10	ERV-1_Crp-I	ERV	38	c	19-56
	AluSx1	NonLTR/SINE/SINE1	285	c	374-658
	Gypsy-57_GR-I	LTR/Gypsy	44	c	1777-1820
iAug10	MIRc	NonLTR/SINE/SINE2	53	c	208-260
	MER117	DNA/hAT	120	d	295-414
	MLT1C2	ERV/ERV3	397	c	864-1260
	SETARIA1	DNA	314	d	1685-1998

Примітки: c – комплементарний, d – прямий.

Таблиця 3

Регуляторний потенціал послідовностей МГЕ, властивих геному людини, у межах потенційних промоторів гена *MGMT* людини

Умовне позначення промотору	Мобільні генетичні елементи		Регуляторні послідовності	
	Назва	Видова приналежність	Сайти зв'язування транскрипційних факторів	Регуляторні елементи
aAug10	Tigger15a	Mammalia	ERE; HRE; MEF-2; C/EBP	promoter; enhancer; silencer; Locus Control Region-like region
fAug10	AluSx1	Primates	TATA box ; YY1; GATA; SF1; Sp1; CAAT box; OCT1; SREBP; PAX2; NF-kappaB; PBX2; SOX9; SRY; WT1; AP-2; C/EBP; P53; CRE; CREB; MEF-2; nCaRE; TRalpha; RARalpha/RXR; IFN-stimulated response element	promoter; myeloid-specific promoter; erythroid-specific promoter; enhancer; lymphocyte-specific enhancer; H ₂ O ₂ -inducible enhancer; non-renal silencer cytokine responsive element; Ets-responsive region PMA and cAMP responsive region

Геном людини несе не лише специфічні для ссавців, приматів чи людини МГЕ або їхні фрагменти чи копії, але і фрагменти МГЕ інших організмів: тварин, рослин і навіть бактерій [12, 25–27]. Як видно із відомостей, наведених у табл. 4, такі послідовності також збагачені різноманітними регуляторними елементами і можуть вплива-

ти на транскрипційну активність репаративного гена *MGMT* людини. Варто зауважити, що наявність ТАТА боксу у межах МГЕ в обох потенційних промоторних ділянках (зазначено жирним шрифтом у табл. 3 та табл. 4) може бути передумовою для існування альтернативного промотора.

Таблиця 4

Регуляторний потенціал неспецифічних для геному людини фрагментів МГЕ у межах альтернативних промоторів гена *MGMT*

Умовне позначення промотору	Мобільні генетичні елементи		Регуляторні послідовності	
	Назва	Видова приналежність	Сайти зв'язування транскрипційних факторів	Регуляторні елементи
aAug10	Tx1-11_Crp	<i>Crocodylus porosus</i>	Sp1; NF-kappaB; HNF-3; TATA box ; UF1-H3beta	promoter; erythroid-specific promoter; housekeeping promoter; enhancer; neuronal specific enhancer region
	Gypsy-54_PIT-I	<i>Phytophthora infestans</i>	SRE; YY-1; Pit1; GATA; muE; TSE	promoter; T-cell specific enhancer
	DNA4-8_CGi	<i>Crassostrea gigas</i>	Sp1; NF-kappaB; AP-1; SOX4; HNF4; ASP; GRE-I; GRE-II	promoter; lymphocyte-specific enhancer
fAug10	ERV-1_Crp-I	<i>Crocodylus porosus</i>	GATA-1	promoter; IFN-gamma responsive negative regulatory region
	Gypsy-57_GR-I	<i>Gossypium raimondii</i>	SREBP; ERE; p53	promoter

Не викликає сумніву, що МГЕ можуть впливати на експресію еукаріотних генів. Перш за все, у цих процесах беруть участь їхні промотори та асоційовані з ними регуляторні послідовності. Глобальний аналіз експресії ретротранспозонів у геномі людини виявив ~275 000 TSS, що становить ~31 % від усіх відомих TSS геному людини, хоча рівень їхньої активності значно нижчий, ніж TSS звичайних генів [28]. Транскрипція послідовностей МГЕ також впливає на транскриптом генів, що кодують білки [28]. Показано, що 576 промоторів ретротранспозонів людини або їхніх фрагментів використовуються як альтернативні під час транскрипції відомих генів. Описано також випадки залучення енхансерів та інсуляторів МГЕ до участі в транскрипційних мережах геномів людини, тварин і рослин [29]. Усі ці

факти свідчать про важливу роль МГЕ у філогенезі та онтогенезі еукаріот, проте їхнє глобальне значення ще потребує вагоміших підтверджень [30].

Висновки

Аналізуючи нуклеотидні послідовності потенційних альтернативних промоторних ділянок гена *MGMT* людини із бази даних AceView, виявлено, що дві послідовності (сAug10 та іAug10) перекриваються із реферованим промотором, а дві інші послідовності (аAug10 і fAug10) розташовані у межах інтрону 2. Усі вони містять у своєму складі фрагменти МГЕ, які збагачені різноманітними регуляторними послідовностями і можуть впливати на регуляцію транскрипційної активності гена *MGMT* людини.

ЛІТЕРАТУРА

- Lander E.S., Linton L.M., Birren B. et al. International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome // *Nature*. – 2001. – 409, № 6822. – P. 860–921.
- Makaiowski W. Genomic scrap yard: how genomes utilize all that junk // *Gene*. – 2000. – 259, № 1–2. – P. 61–67.
- Sverdlov E.D. Retroviruses and primate evolution // *Bioessays*. – 2000. – 22, № 2. – P. 161–171.
- Kazazian H.H. Jr. Mobile elements: drivers of genome evolution // *Science*. – 2004. – 303, № 5664. – P. 1626–1632.
- Hamdi H.K., Nishio H., Tavis J. et al. Alu-mediated phylogenetic novelties in generegulation and development // *J. Mol. Biol.* – 2000. – 299, № 4. – P. 931–939.
- Medstrand P., Landry J.R., Mager D.L. Long terminal repeats are used as alternative promoters for the endothelin B receptor and apolipoprotein C-I genes in humans // *J. Biol. Chem.* – 2001. – 276, № 3. – P. 1896–1903.
- Murnane J.P., Morales J.F. Use of a mammalian interspersed repetitive (MIR) element in the coding and processing sequences of mammalian genes // *Nucleic Acids Res.* – 1995. – 23, № 15. – P. 2837–2839.
- Nekrutenko A., Li W.H. Transposable elements are found in a large number of human protein-coding genes // *Trends Genet.* – 2001. – 17, № 11. – P. 619–621.
- Jordan I.K., Rogozin I.B., Glazko G.V., Koonin E.V. Origin of a substantial fraction of human regulatory sequences from transposable elements // *Trends Genet.* – 2003. – 19, № 2. – P. 68–72.
- Faulkner G.J., Carninci P. Altruistic functions for selfish DNA // *Cell Cycle*. – 2009. – 8, № 18. – P. 2895–2900.

11. Підпала О.В., Лукаш Л.Л. Композиційні кластерні структури мобільних генетичних елементів у інтронах гена *MGMT* як джерело регуляторних послідовностей // Фактори експериментальної еволюції організмів зб. наук. праць [Під ред. В.А. Кунаха та ін.]. – 2014. – 14. – С. 220–224.
12. Підпала О.В., Лукаш Л.Л. Мотиви регуляторних послідовностей у промоторних ділянках гена *MGMT* людини в межах мобільних елементів // Фактори експериментальної еволюції організмів: зб. наук. праць [Під ред. В.А. Кунаха та ін.]. – 2015. – 16. – С. 230–235.
13. Harris L.C., Potter P.M., Tano K. et al. Characterization of the promoter region of the human O6-methylguanine-DNA methyltransferase gene // *Nucleic Acids Res.* – 1991. – 19, № 22. – P. 6163–6167.
14. Harris L.C., Remack J.S., Brent T.P. Identification of a 59 bp enhancer located at the first exon/intron boundary of the human O⁶-methylguanine DNA methyltransferase gene // *Nucleic Acids Res.* – 1994. – 22, № 22. – P. 4614–4619.
15. Vansant G., Reynolds W.F. The consensus sequence of a major Alu subfamily contains a functional retinoic acid response element // *Proc. Natl. Acad. Sci. – USA*, 1995. – 92, № 18. – P. 8229–8233.
16. Piedrafita F.J., Molander R.B., Vansant G. et al. An Alu element in the myeloperoxidase promoter contains a composite SP1–thyroid hormone–retinoic acid response element // *J. Biol. Chem.* – 1996. – 271, № 24. – P. 4412–4420.
17. Babich V., Aksenov N., Alexeenko V. et al. Association of some potential hormone response elements in human genes with the Alu family repeats // *Gene.* – 1999. – 239, № 2. – P. 341–349.
18. Polak P., Domany E. Alu elements contain many binding sites for transcription factors and may play a role in regulation of developmental processes // *BMC Genomics.* – 2006. – 7. – P. 133.
19. Laperriere D., Wang T.T., White J.H., Mader S. Widespread Alu repeat–driven expansion of consensus DR2 retinoic acid response elements during primate evolution // *BMC Genomics.* – 2007. – 8. – P. 23.
20. Sorek R., Ast G., Graur D. Alu-containing exons are alternatively spliced // *Genome Res.* – 2002. – 12, № 7. – P. 1060–1067.
21. Shankar R., Grover D., Brahmachari S.K., Mukerji M. Evolution and distribution of RNA polymerase II regulatory sites from RNA polymerase III dependant mobile Alu elements // *BMC Evol. Biol.* – 2004. – 4, № 1. – P. 37.
22. Callinan P.A., Batzer M.A. Retrotransposable elements and human disease // *Genome Dyn.* – 2006. – 1. – P. 104–115.
23. Romanish M.T., Nakamura H., Lai C.B. et al. A novel protein isoform of the multicopy human NAIP gene derives from intragenic Alu SINE promoters // *PLoS One.* – 2009. – 4, № 6. – e5761.
24. Bolotin E., Chellappa K., Hwang-Verslues W. et al. Nuclear receptor HNF4b binding sequences are widespread in Alu repeats // *BMC Genomics.* – 2011. – 12. – P. 560.
25. Mayorov V.I., Rogozin I.B., Elisaphenko E.A., Adkison L.R. B2 elements present in the human genome // *Mamm. Genome.* – 2000. – 11, № 2. – P. 177–179.
26. Підпала О.В., Яцишина А.П., Лукаш Л.Л. Мобільні генетичні елементи у гені-онкосупресорі p53: еволюційний аспект // Фактори експериментальної еволюції організмів: зб. наук. праць [Під ред. В.А. Кунаха та ін.]. – 2006. – 3. – С. 43–48.
27. Підпала О.В., Яцишина А.П., Лукаш Л.Л. Фрагменти бактеріальних IS-елементів і мобільних генетичних елементів еукаріотів у мтДНК людини // Досягнення і проблеми генетики, селекції та біотехнології. – 2007. – 1. – С. 498–502.
28. Faulkner G.J., Kimura Y., Daub C.O. et al. The regulated retrotransposon transcriptome of mammalian cells // *Nature Genet.* – 2009. – 41. – P. 563–571.
29. De Souza, F.S.J., Franchini L.F., Rubinstein M. Exaptation of Transposable Elements into Novel Cis-Regulatory Elements: Is the Evidence Always Strong? // *Mol. Biol. Evol.* – 2013. – 30, № 6. – P. 1239–1251.
30. Патрушев Л.И., Коваленко Т.Ф. Функции некодирующих последовательностей генома млекопитающих // Успехи биологической химии. – 2014. – 54. – С. 39–102.

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

FRAGMENTS OF THE MOBILE GENETIC ELEMENTS AS COMPONENTS FOR POTENTIAL ALTERNATIVE PROMOTERS OF HUMAN *MGMT* GENE

Aim. It has been carried out analysis of the MGEs in the alternative promoters of human *MGMT* gene for the sequences homologous to different regulatory elements and binding sites of the transcription factors. **Methods.** Searching and identifying these MGEs have been realized by using CENSOR. The functional sites have been defined by the programs *SITECON*, *SiteGA*, *BLASTN 2.2.26*, *Cister: Cis-element Cluster Finder*, *WWW Signal Scan* and *Tfsitescan*. **Results.** It has been found that two sequences (cAug10 and iAug10) are overlapping with the known referenced promoter (GeneBank X61657). The different sequences (aAug10 і fAug10) are located within intron 2. They compose fragments of the MGEs which are enriched by various regulatory elements and binding sites of the transcription factors. **Conclusions.** The obtained results allow considering the MGEs identified in potential alternative promoters of *MGMT* gene as carriers for the potential regulatory sequences which might affect the regulation of transcriptional activity of this gene.

Keywords: human O⁶-methylguanine-DNA methyltransferase (*MGMT*) gene, alternative promoters, mobile genetic elements (MGEs), regulatory elements.