BIOINFORMATICS

Analysis of distribution of mobile genetic elements within the human TP53 gene and its 5 -flanking region

O. V. Pidpala, A. P. Iatsyshyna, L. L. Lukash

The Institute of Molecular Biology and Genetics, NAS of Ukraine
150 Acad. Zabolotny Str., Kyiv, 03143, Ukraine

E. mail: specrada@imbg.org.ua

Computational analysis of distribution of mobile genetic elements within the human TP53 gene and its 5 -flanking region has been performed. There was no difference revealed for SINE and LINE repeats, but it has been shown that the LINE elements are preferentially present within the TP53 gene and the LINE2 elements are preferentially distributed within 5 -flanking region of the TP53 gene. Alu repeats have been found to be the most common repeats within the TP53 gene and its 5 -flanking region. LTR repeats have been absent at all and DNA transposons have been determined only within the TP53 gene. It has been revealed that mobile genetic elements within TP53 gene and its 5 -flanking region preferentially form clusters, which contain mobile genetic elements from different repeat families and subfamilies.

Key words: human TP53 gene, mobile genetic elements, Alu-repeats, mosaic cluster structures.

Introduction According to the peculiarities of structure and the mode of moving, mobile genetic elements (MGE) can be divided into 4 basic classes: SINE (short interspersed nuclear elements), LINE (long interspersed nuclear elements), LTR elements (retrovirus-like elements with long terminal repeats) and DNA-transposons [1]. MGE are not only the factors of spontaneous and inducted mutations [2-4], but they are vAluable functional genome components [5-8]. Due to MGE, genome is considered to be a dynamic system which reacts in an active way to the environment changes [9-12]. Having the promoters of their own, they can change the activity of the gene adjacent to them, form promoters of vitally important genes [14, 15] or play the role of various regulatory sequences [16-18].

In human genome the MGE part is app. 45 % of nuclear DNA. They are spread randomly, mainly in intergenic or intron regions, rarely in exons [6, 19]. In many genes, MGE or sequences, originated from them, are present in promoter, 5'- and 3'- flanking regions, which points out to their important role in functioning of genome, in gene activity regulation, in particular [15, 20].

The human TP53 gene is the oncosuppressor, the mutations of which are the most widely spread disorders in all kinds of malignant tumors [21, 22]. Its basic function is cell genetic stability maintenance and it also participates in apoptosis, reparation, and angiogenesis [23-25]. This polyfunctionality demands complex regulation which possibly involves MGE. Some researches show the presence of SINE class representatives (Alu-repeats) in this gene [26] and in 3r-flanking region, in particular, [27] however it is not determined what subfamilies they belong

to, and other MGE are not mentioned as well. The given data on MGE distribution in TP53 gene in GenBank database (reg. No. AY838896) do not give information about the MGE class and subfamily, detailed analysis of the total amount of MGE in gene in general and in exons and in intorns separately is also absent. Therefore, the aim of our work was to analyze MGE distribution in the human TP53 gene and its 5'-flanking region in details.

Materials and Methods. Human gene TP53 nucleotide sequence was obtained from GenBank (reg. No.U94788, 843 – 20303bp). 5r-flanking sequence of this gene (10kb) was obtained from GENE DATABASE GENATLAS web-site (http://www.dsi.univ-paris5.fr/genatlas /struc_exon/TP53_1.html). Both sequences were analyzed for the presence of repeats using RepeatMasker software, available at BCM Search Launcher (http://searchlauncher.bcm.tcm.edu /seq-util.html).

Results and Discussion. Table 1 represents the results of the MGE distribution analysis in the human gene TP53. Summarized data on TP53 gene and its 5r-flanking region are represented in Table 2. It has been determined that general percentage of SINE and LINE elements in gene and in its 5r-flanking region do not differ essentially. However, LINE1 elements are located in within gene exceptionally, while LINE2 elements are located in 5r-flanking region preferentially. Among MGE Alu-repeats are observed the most frequently. LTR elements are not detected at all, and DNA-transposons in some minor quantity are represented in gene only.

MGE distribution in introns and exons was analyzed (Fig.1). MGE were found in five out of ten human TP53 gene introns. The highest percentage of MGE is in introns 1, 6, and 9 (73.03, 68.96, and 85.45 % respectively). The percentage is lower in introns 4 and 10 (40.66 and 32.83 %). Two SINE class MGE, namely one MIR element and one Alu-repeat, are present in exon 11. It is worth mentioning that within TP53 gene only Alu-repeats out of 39 MGE sequences are full-sized elements, all the rest are represented in fragments.

Alu-repeats are known to form clusters [28-30], which according to [31, 32] may consist of different subfamilies. The exception is young ALU-repeats subfamilies that are mostly distant from clusters [33]. The involving of Alu-repeats into chromosome reconstruction is discussed [34].

Within TP53 gene there are three big Alu-repeats clusters: two in intron 1 and one in intron 9. One cluster, which is located in intron 1, consists of eleven Alu-repeats and one Alu-monomer (two of which belong to young subfamilies). It is of interest that Alu-repeats in this cluster are surrounded or border with fragments of L1 element, developing composite structures (Fig. 2). The second intron 1 cluster consists of five Alu-repeats. There are cases when one Alu-repeat is inserted into another one (Fig. 2). Alu-repeats that belong to intron 9, form one big cluster, which consists of seven Alu-repeats and one Alu-monomer. In this case the cluster is surrounded with MER2 sequences and one more MER2 sequence is located within this cluster. Analyzing 10 kb of 5'-flanking region, within which there is WDR79 gene (previously FLJ10385), which codes hypothetical protein LOC55135 and has opposite transcription direction, we revealed that overwhelming majority of Alu-repeats is located in three clusters, the specific feature of which is framing with L2 sequences. Therefore, MGE within both human TP53 gene and its 5'-flanking region mainly (80% and 73%) form mosaic cluster structures, which are contain MGE of different families and subfamilies (among them young Alu-repeats subfamilies are also represented).

As the majority of human TP53 gene mutations are the missence-mutations that take place in exons 5-8 [35] one can discuss MGE involving in mutational processes only in some cases [36-37]. However, regarding large MGE representation in the human TP53 gene (59.96 %), the question about the role, which they play in genome functioning, arises.

V. A. Ratner defined a special role of MGE as 'movable cassettes of regulatory elements' and later as 'movable cassettes of functional sites' that having various sites of external signals reception, may influence genes expression substantially [38, 7]. As for Alu-repeats, the presence of functional binding sites for retinoid acid receptors was found [39] in consensus sequence, and additional presence of hormone-acceptor elements was shown for Alu-repeats located in promoter regions of some genes [40].

There are known cases of MGE participation in the regulation of cellular genes expressions [17]. Alu-repeats can be enhancers (e.g. in the case of adenosidiamenase gene) [41], transcriptional modulators (c-myc gene) [42], or transcriptional silencers (PCNA gene) in particular [43]. Alu-repeats can inactivate or change the functions of gene products, creating alternative sites of splicing or interfere into its mechanism (e.g. in the case of subunit of в1С-2 integrin) [44]. They can also act as insulators (KRT 18 gene) [45] and evidently implement other functions[46-48]. The presence of Alu-repeats and other retroposons in pre-mRNA affect polyadenilation of transcripts as well as influence translation effectiveness [49-51]. Alu-repeats contribute to the methylation of neighboring loci providing one more mechanism of control over genes expression [52-53].

High abundance of MGE in the analyzed human TP53 gene and its 5'-flanking region is probably connected with

Table 1.
Distribution Analysis of mobile genetic elements and their fragments in human TP53 gene.

MGE	Class/family	Coordinates in gene borders	Length(b.p.)	Chain	Localization
L1M2	LINE/L1	1516-1749	234	+	Intron 1
AluSq	SINE/Alu	1750-2043	294	-	Intron 1
L1ME2	LINE/L1	2044-3031	988	+	Intron 1
AluJo	SINE/Alu	3082-3379	298	+	Intron 1
L1	LINE/L1	3426-3485	60	-	Intron 1
AluSx	SINE/Alu	3486-3787	302	-	Intron 1
L1	LINE/L1	3787-3959	172	-	Intron 1
AluSx	SINE/Alu	3960-4095	136	-	Intron 1
AluSq	SINE/Alu	4096-4385	290	-	Intron 1
AluSx	SINE/Alu	4386-4560	175	-	Intron 1
L1	LINE/L1	4561-4618	58	-	Intron 1
FLAM_c	SINE/Alu	4621-4737	117	-	Intron 1
L1	LINE/L1	4749-4974	226	-	Intron 1
AluY	SINE/Alu	4975-5288	314	-	Intron 1
L1	LINE/L1	5289-5395	107	-	Intron 1
AluY	SINE/Alu	5396-5701	306	-	Intron 1
AluSq	SINE/Alu	5710-5851	142	-	Intron 1
AluSq	SINE/Alu	5852-6147	296	-	Intron 1
AluSq	SINE/Alu	6148-6327	180	-	Intron 1
L1	LINE/L1	6328-6401	74	-	Intron 1
AluSq	SINE/Alu	6402-6664	263	-	Intron 1
AluSx	SINE/Alu	6665-6984	320	-	Intron 1
L1	LINE/L1	6992-7052	387	-	Intron 1
MER2	DNA/MER2	7062-7281	220	-	Intron 1
MIR	SINE/MIR	7776-7853	58	+	Intron 1
AluSq	SINE/Alu	7861-8192	332	-	Intron 1
AluSp	SINE/Alu	8257-8557	301	-	Intron 1
AluJo	SINE/Alu	8668-8719	52	-	Intron 1
AluSx	SINE/Alu	8720-9036	317	-	Intron 1
AluJo	SINE/Alu	9037-9202	166	-	Intron 1
AluSq	SINE/Alu	9210-9520	311	-	Intron 1
L2	LINE/L2	9666-9863	198	+	Intron 1
AluSx	SINE/Alu	10235-10530	296	-	Intron 1
L2	LINE/L2	10532-10710	179	+	Intron 1

MGE	Class/family	Coordinates in gene borders	Length(b.p.)	Chain	Localization
AluJb	SINE/Alu	11754-12060	307	-	Intron 4
MIR	SINE/Alu	12683-12774	92	+	Intron 6
AluY	SINE/Alu	12789-13087	299	-	Intron 6
MER47A	DNA/MER2	14178-14307	130	+	Intron 9
AluSq	SINE/Alu	14308-14624	317	+	Intron 9
AluJo	SINE/Alu	14637-14854	218	+	Intron 9
AluJb	SINE/Alu	14863-15028	166	+	Intron 9
AluSg	SINE/Alu	15033-15343	311	+	Intron 9
AluJo	SINE/Alu	15346-15637	292	+	Intron 9
MER47A	DNA/MER2	15638-15792	155	+	Intron 9
AluSx	SINE/Alu	15855-16151	297	-	Intron 9
FLAM_A	SINE/Alu	16155-16279	125	-	Intron 9
AluSx	SINE/Alu	16281-16571	291	-	Intron 9
MER47A	DNA/MER2	16587-16692	106	+	Intron 9
AluSp	SINE/Alu	17250-17551	302	-	Intron 10
MIR	SINE/MIR	18206-18308	102	-	Exon 11
AluJb	SINE/Alu	18601-18900	300	+	Exon 11

Table 2.
Mobile genetic elements in human TP53 gene and its 5'-flanking region

Type of the element	Human TP53 gene			5?- flanking region (10 kb)		
	Quantity of the elements in the analyzed region	The length of the nucleotide sequence, which is occupied by the current type elements (b.p.)	Percentage of total length of the analyzed region (%)	Quantity of the elements in the analyzed area	The length of the nucleotide sequence, which is occupied by the current type elements, (b.p.)	Percentage of total length of the analyzed region, (%)
SINE:	33	8686	44,63	18	4719	47,19
Alu	30	8433	43,33	17	4575	45,75
MIR	3	253	1,30	1	144	1,44
LINE:	4	2372	12,19	6	1825	18,25
LINE 1	2	1995	10,25	0	0	0
LINE 2	2	377	1,94	6	1825	18,25
L3/CR 1	0	0	0	0	0	0
LTR elements	0	0	0	0	0	0
ДНК-transpos ons:	2	611	3,14	0	0	0
MER 1	0	0	0	0	0	0
MER 2	2	611	3,14	0	0	0
Total	39	11669	59,96	24	6544	65,44

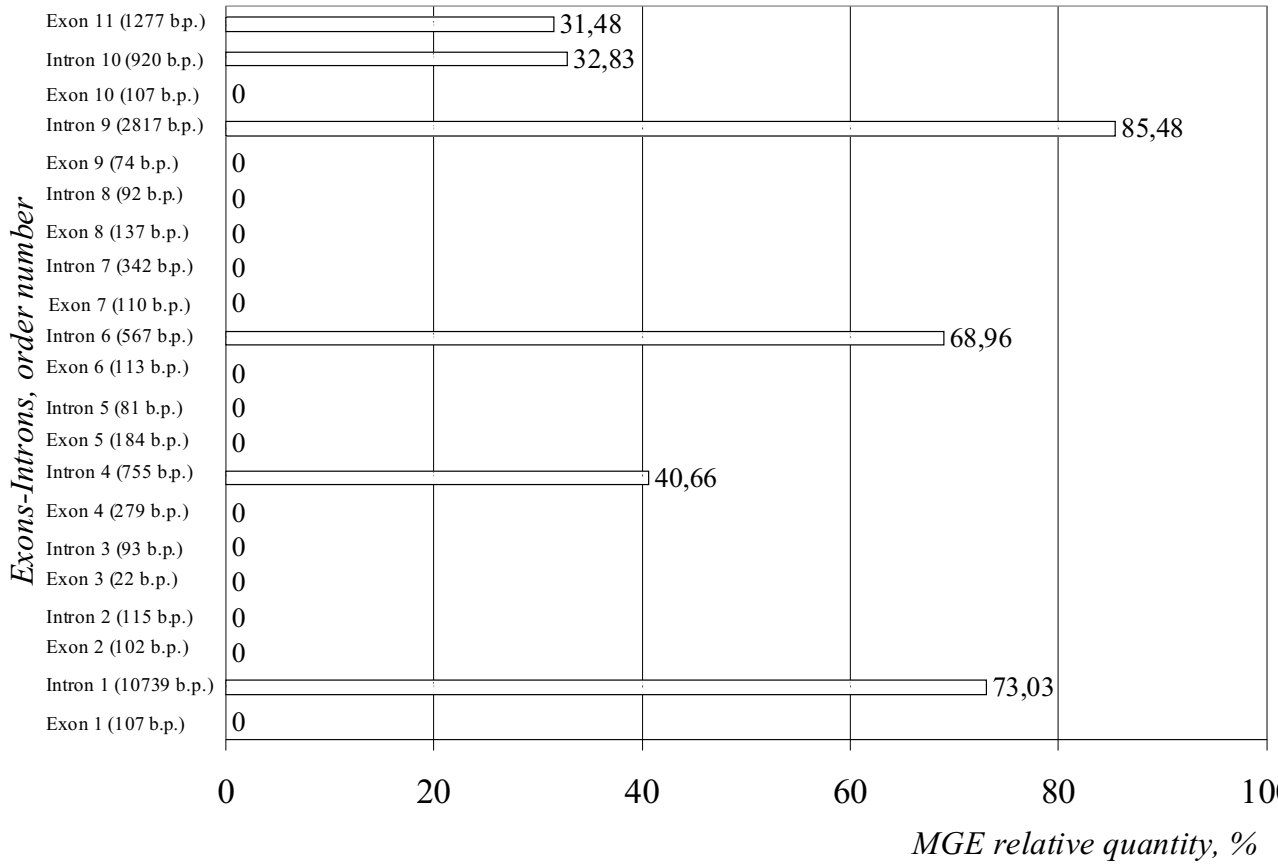

Figure 1. The distribution of MGE in human TP53 gene

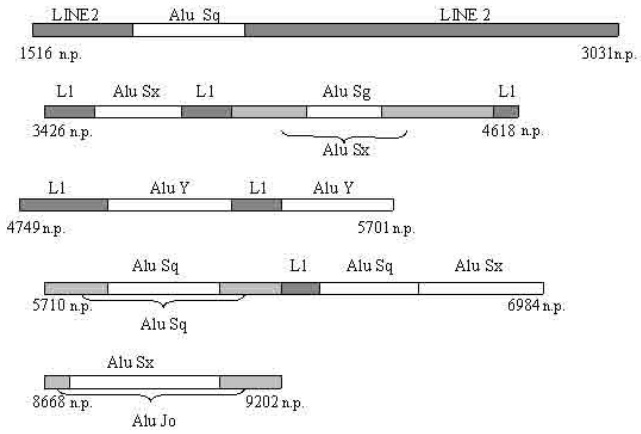

Fig 2. Compositional structures of different families/subfamilies of MGE within clusters of Intron 1 of human TP53 gene (numbers are the coordinates within present gene)

their participation in the regulation of this gene. Therefore, in our further researches we plan on more detailed analysis of the representatives of Alu-repeats subfamily to determine various functional sites in order to define the MGE role, Alu-repeats in particular, in the expression of the human TP53 gene.

Acknowledgements.. We express sincere gratitude to S.M. Kvasha, PhD. (biology), for methodological and consulting help.

О. В. Пидпала, А. П. Яцышина, Л. Л. Лукаш

Анализ распределения мобильных генетических элементов в гене TP53 человека и его 5'-фланкирующем участке

Резюме

Проведен компьютерный анализ распространения мобильных генетических элементов (МГЭ) в генетте TP53 человека и его 5'-фланкирующем участке. Не выявлено существенной разницы для SINE-

и LINE-элементов, однако показано, что LINE1-элементы присутствуют исключительно в гене, тогда как LINE2 ? преимущественно в 5'-фланкирующем участке. Среди МГЭ чаще всего встречаются Alu-повторы. Совсем не обнаружено LTR-элементов, а ДНК-транспозоны в незначительном количестве представлены лишь в гене. Как в гена, так и его 5'-фланкирующем участке МГЭ преимущественно формируют кластерные мозаичные структуры, в состав которых входят элементы разных семейств и подсемейств.

Ключевые слова: ген TP53человека, мобильные генетические элементы, Alu-повторы, кластерные мозаичные структуры.

REFERENCES

1. *International* Human Genome Sequencing Consortium: Initial sequencing and analysis of the human genome// Nature. – 2001. – Vol. 409, N 6822. – P. 860-921.
2. *Eukaryotic* transposable elements as mutagenic agent/ Eds Lambert M.E., McDonald J.F., Weinstein J.B. N.G.: Cold Spring Harbor Press, 1988. 345 p.
3. *Георгиев П.Г.* Роль мобильных элементов в мутагенезе, индуцированном химическими и физическими агентами: Автореф. дис. канд. биол. наук. М.: ИМБ АН СССР, 1991. 24 с.
4. *Kazazian H.H.J.* Mobile elements and disease// Curr. Opin. Genet. Dev. – 1998. – Vol. 8, N 3. – P. 343-350.
5. *Хесин Р.Б.* Непостоянство генома. М.: Наука, 1984. 472 с.
6. *Mobile* DNA/ Eds Berg D.E., Howe M.M. Washington, D.C. (USA); Amer.Soc. Microbiol., 1989. 972p.
7. *Ратнер В.А., Васильева Л.А.* Мобильные генетические элементы (МГЭ): "эгоистическая ДНК" или функциональная часть генома?// Современные концепции эволюционной генетики/ Под ред. Шумного Б.К., Маркеля А.Л. Новосибирск ИциГ СЩ РАН , 2000, с. 128-150.
8. *Shapiro J.A.* Repetitive DNA, genome system architecture and genome reorganization// Res. Microbiol. – 2002. – Vol. 153, N 7. – P. 447-453.
9. *McClintock B.* The significance of responses of the genome to challenge// Science. – 1984. –Vol. 226, N 4676. – P. 792-801.
10. *Fedoroff N., Botstein D.* The dynamic genome: Barbara McClintock's ideas in the century of genetics// Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press, 1992. 422 p.
11. *Capy P., Gasperi G., Biemont C., Bazin C.* Stress and transposable elements: co-evolution or useful parasites?// Heredity. – 2000. – Vol. 85, N 2. P. 101-106.
12. *Fedoroff N.V.* Transposable elements as a molecular evolutionary force// Ann. NYAcad. Sci. – 2002. – Vol. 981. – P. 154-188.
13. *Britten R.J.* DNA sequence insertion and evolutionary variation in gene regulation// Proc. Natl. Acad. Sci. USA. – 1996. – Vol. 93, N 18. – P. 9374-9377.
14. *Kidwell M.G., Lisch D.* Transposable elements as sources of variation in animals and plants// Proc.Natl. Acad. Sci. USA. – 1997. Vol. 94, N 15. - P. 7704-7711.
15. *Jordan I.K., Rogozin I.B., Glasko G.V., Koomin E.V.* Origin of a substantial fraction of human regulatory sequences from transposable elements// Trends Genet. – 2003. Vol. 19, N 2. – P. 68-72.
16. *Britten R.J.* Mobile elements inserted in the distant past have taken on important functions// Gene. – 1997. – Vol. 205, N 1-2. –P. 177-182.
17. *Brosius J.* RNAs from all categories generate retrosequences that may be exapted as novel genes or regulatory elements// Gene. – 1999. – Vol. 238, N 1. - P. 115-134.
18. *van de Lagemaat L.N., Landry J.R., Mager D.L., Medstrand P.* Transposable elements in mammals promote regulatory variation and diversification of genes with specialized functions// Trends Genet. – 2003. - Vol. 19, N 10. – 10. –P.530-536.
19. *Nekrutenko A., Li W-H.* Transposable elements are found in a large number of human protein-coding genes// Trends Genet. – 2001. – Vol. 17, N 11. – P.619-621.
20. *Hon L.S., Jain A.N.* Compositional structure of repetitive elements is quantitatively related to coexpression of gene pairs// J. Mol. Biol. – 2003. – Vol. 332, N 2. – P.305-310.
21. *Hollstein M., Sidransky D., Vogelstein B., Harris C.C.* P53 mutations in human cancer// Science. – 1991. – Vol. 253, N 5015. – P.49-53.
22. *Levine A.J., Momand J., Finlay C.A.* The p53 tumor suppressor gene// Nature. – 1991. – Vol. 351, N 6326. – P.453-456.
23. *Lane D.P.* P53, guardian of the genome// Nature. – 1992. – Vol. 358, N 6381. – P.15-16.
24. *Bargonetti J., Manfredi J.J.* Multiple roles of the tumor suppressor p53// Curr. Opin. Oncol. – 2002. – Vol. 14, N 1. – P.86-91.
25. *Fridman J.S., Lowe S.W.* Control of apoptosis by p53// Oncogene. – 2003. – Vol. 22, N 56. – P.9030-9040.
26. *Futreal P.A., Barrett J.C., Wiseman R.W.* An Alu polymorphism intragenic to the TP53 gene// Nuclear Acids Res. – 1991. – Vol. 19, N 24. – P.6977.
27. *Fu L., Ma W., Benchimol S.* A translation repressor element resides in the 3? untranslated region of human p53 mRNA // Oncogene.- 1999.-Vol.18, № 47. - P.6419-6424.
28. *Шахмурадов И.А., Колчанов Н.А., Капитонов В.В.* Распространение повторов Alu человека по геному: формирование кластеров и особенности участков встраивания// Мол. биол. – 1989. – Т. 23, вып. 2. – С.526-536.
29. *Pavlicek A., Jabbari K., Paces J., Paces V., HeJnar J.V., Bernardi G.* Similar integration but different stability of Alus and LINEs in the human genome// Gene. – 2001. – Vol. 276, N 1-2. –P.39-45.
30. *Jurka J., Krujajic M., Kapitonov V.V., Stenger J.E., KoHany O.* Active Alu elements are passed primarily through paternal germlines// Theor. Popul. Biol. – 2002. – Vol. 61, N 4. – P.519-530.
31. *Toda G., Tomita M.* Alu elements as an aid in deciphering genome rearrangements// Gene. – 1997. – Vol. 205, N 1-2. P.173-176.
32. *Kulski J.K., Gandierri S., Bellgard M., Balmer L., Giles K., Inoko H., Dawkins R.L.* The evolution of MNC diversity by segmental duplication and transposition of retroelements// J. Mol. Evol. – 1997. - Vol. 45, N 6. - P. 599-609.
33. *Jurka J., Kohany O., Pavlicek A., Kapitonov V.V., Jurka M.V.* Duplication, coclustering, and selection of human Alu retrotransposones// PNAS. - 2004. - Vol. 101, N 5. - P. 1268-1272.

34. Kolomietz E., Meyn M.S., Pandita A., Squire J.A. The role of Alu repeat cluster as mediators of recurrent chromosomal aberrations in tumors// Genes, Chromosomes, Cancer. - 2002. - Vol. 35, N 2. - P. 97-112.

35. Ory K., Legros Y., Auguin C., Soussi T. Analysis of the most representative tumor-derived p53 mutants reveals that changes in protein conformation are not correlated with loss of transactivation or inhibition of cell proliferation// EMBO J. - 1994. - Vol. 13, N 15. - P. 3496-3504.

36. Slebos R.S., Resnick M.A., Taylor J.A. Inactivation of the p53 tumor suppressor gene via a novel Alu rearrangement// Cancer Res.- 1998. - Vol. 58, N 23. - P. 5333-5336.

37. Bougeard G., Brugieres L., Chompret A., Gesta P., Charbonnier F., Valent A., Martin C., Raux G., Feunteun J., Paillerets B.B., Frebourg T. Screening for TP53 rearrangements in families with the Li-Fraumeni syndrome reveals a complete deletion of the TP53 gene// Oncogene. - 2003. - Vol.22, N 6. - P. 840-846.

38. Ратнер В.А., Васильева Л.А. Роль мобильных генетических элементов (МГЭ) в микроэволюции// Генетика. - 1992. - Т.28, N 12. - С. 5-15.

39. 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. - Vol.92, N 18. - P. 8229-8233.

40. Babich V., Aksenov N., Alexeenko V., Oei S.L., Buchlow G., Tomilin N. Association of some potential hormone response elements in human genes with the Alu family repeats// Gene. - 1999. - Vol. 239, N 2. - P. 341-349.

УДК575.113+577.2+599.89
Надійшла до редакції 07.12.04