ISSN 2079-0597, Russian Journal of Genetics: Applied Research, 2011, Vol. 1, No. 6, pp. 524–531. © Pleiades Publishing, Ltd., 2011.
Original Russian Text © K.K. Baskaev, A.A. Buzdin, 2011, published in Vavilovskii Zhurnal Genetiki i Selektsii, 2011, Vol. 15, No. 2, pp. 313–321.

Evolutionarily Recent Groups of Transposable Elements in the Human Genome

K. K. Baskaev and A. A. Buzdin

Shemyakin and Ovchinnikov Institute of Bioorganic Chemistry, Moscow, Russia
e-mail: bu3din@mail.ru

Abstract—Transposable elements (TEs) are fragments of DNA capable of self-reproduction in the genome of the host organism. They constitute ~40–50% of mammalian genomes. We have identified two TE families that formed relatively recently in the course of human evolution. Members of the first family are present only in the human genome but are absent from other primate DNAs. This family represented by ~80 members was formed by fusion of a portion of the CpG-island in the human *MAST2* gene with the 3'-terminal fragment of the *SVA* retrotransposon. According to our estimates, this hybrid family, termed *CpG-SVA*, is significantly more active than the ancestor *SVA* family. The regulatory region of *MAST2* allows copies of the new family to be transcribed in sperm precursor cells. The second family, called the family of chimera retrotranscripts, is older, but is still active today. Its representatives were formed by a rather unusual RNA recombination mechanism, which mediated the formation of fused DNA copies for diverse cellular transcripts. We showed that similar mechanisms operate in the genomes of other mammals and even fungi.

Keywords: human evolution, genetic instability, transposable elements, regulation of gene expression, genetic chimeras, hybrid family of retrotransposons.

DOI: 10.1134/S2079059711060049

INTRODUCTION

Repeating sequences occupy a vast part of every eukaryotic genome. They account for about half of the genomic DNA of mammals, whereas the exons of genes and protein-encoding sequences occupy only ~3 and 1%, respectively. A significant number of genome repeats contain genes that, as a rule, encode "egoistic" proteins necessary for the replication of transposable elements (TEs) in the host genome.

TEs amplify themselves either directly using their DNA-copies (DNA-transposons) or via RNA intermediates (retroelements). In the latter case, they use the mechanism of "reverse transcription" and the enzyme RNA-dependent DNA-polymerase called reverse transcriptase (RT). The newly formed copy of the element is integrated to the genome using a combination of proteins of the host cell and enzymes encoded by its own sequence of TEs (Temin, 1993; Jurka, 1997; Dewannieux et al., 2003; Buzdin, 2006).

Retroelements constitute more than 42% of human DNA and are at present the only class of transposable elements in mammals capable of transposition (Sverdlov, 2000; Buzdin, 2004). They also play an active role in a multitude of processes of functioning of the human genome (Wessler, 1998; Sverdlov, 1998, 2000; Deininger et al., 2003; Buzdin, 2004). TEs are "hot points of recombination" (for instance, recombinations between the *Alu* retroelements led to the deletion of about 400 kbp in human DNA (Sen et al., 2006)). It is known that retroelements can modify the activity of preexisting human genes (Brosius, 1999a; Deininger et al., 2003; Buzdin, 2004; van de Lagemaat et al., 2006). A third of all TEs are mapped in the gene composition (Mills et al., 2006).

Retroelements are divided into autonomous and nonautonomous ones, i.e., those that have a gene of reverse transcriptase themselves and those that borrow its protein product from others.

The main group of active retroelements of human beings is the family of *L1* autonomous retrotransposons. In primates, full-length representatives of *L1* occupy around 6 kbp and encode two open reading frames: one for the RT/integrase and another for the RNA-binding protein. The *L1* insertions are mostly considerably shortened from the 5'-end, apparently due to the abortive reverse transcription that often takes place (Boissinot et al., 2000). The next two active groups, the *Alu* retrotransposons (~300 bp long) and SVA (~1.5 kbp), are nonautonomous transposable elements that use the "alien" RT of the *L1* retrotransposons for replicating their copies (Wang et al., 2005). These two groups do not code for proteins and can be considered as parasites of the system of *L1* retrotransposition (Buzdin, 2004). Finally, autonomous endogenic retroviruses of the HERV-K (HML-2) group have the most complex organization among human transposable elements. They bear three typical

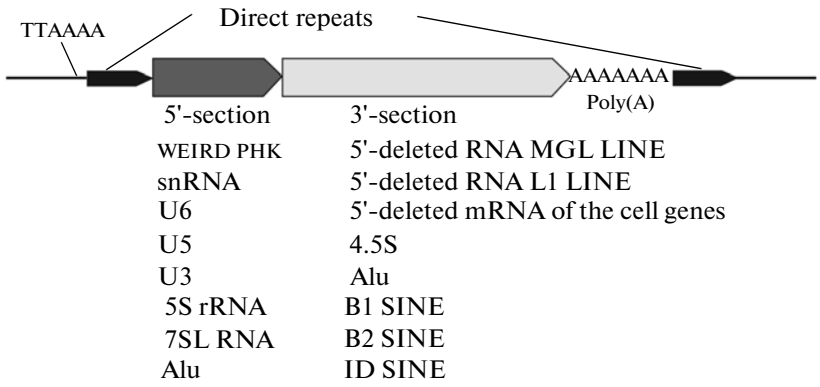

Fig. 1. A scheme of the double chimera TE. The TE insertion is flanked by direct repeats 10–20 bp long.

retrovirus functional genes and one additional gene that encodes a small regulator protein.

However, retroelements often function as regulators of transcription of host RNA (Gogvadze, Buzdin, 2009). The repeating transcripts can also serve as multifunctional RNA, take part in antisense regulation of genes, and compete with the host transcripts over binding with cell proteins (Zabolotneva et al., 2010).

It was shown in experiments on cell cultures that polymorph copies *L1*- and *Alu*-elements, which are situated in the introns of human genes, decrease the transcription of alleles that contain these retroelements as compared with the alleles that do not contain them (Ustyugova et al., 2006; Lebedev et al., 2007). Genome repeats also include regulator sequences, such as promoters, enhancers, splicing sites, polyadenylation signals, and insulators, which influence the genome and transcriptome actively (Schumann et al., 2010).

THE FAMILY OF HUMAN TRANSPOSABLE ELEMENTS FORMED USING THE MECHANISM OF RNS-RECOMBINATION

A typical *L1* element codes for two proteins: the RNA-binding protein ORF1p, which promotes reverse transcription as a chaperone of nucleic acids (Martin, 2006), and ORF2p, which is the reverse transcriptase and endonuclease (Kazazian, 2004).

In addition to its copies of itself, *LINE* elements are capable of transposition of other sequences, mainly nonautonomous TEs that belong to the *SINE* class, as well as copies of various cellular RNA, which leads to the formation of pseudogenes (Dewannieux et al., 2003). We have recently shown that the *L1* elements are included in the formation of double and triple chimera TEs, which are formed in the process of reverse transcription in many genomes, including all the genomes of mammals and some fungi genomes (Buzdin et al., 2002, 2003, 2005; Fudal et al., 2005; Gogvadze et al., 2005) (Fig. 1).

All in all, 82, 116, 66, and 31 such chimera elements were found in human, mouse, rat, and *Magnaporthe grisea* DNA, respectively (Buzdin et al., 2002, 2003; Fudal et al., 2005; Gogvadze et al., 2005). These elements consist of DNA copies of cellular transcripts both directly linked with each other and, more often, joined with the 3'-end part of the *L1* element. Apart from the TE copies, the chimeras included copies of various cellular transcripts corresponding to mRNA of protein-encoding genes, ribosome RNA, various small nuclear RNA, and 7SL RNA. Such chimeras met the following criteria: 1) their 5'-end parts were full-length copies of cellular RNA, 2) their 3'-end parts were shortened from the 5'-end by copies of other RNA (most often, TEs), 3) interruption sites were located in random places of the corresponding RNA, 4) both parts were united in one and the same direction, 5) the chimeras had poly(A)-tail at their 3'-end, and 6) the chimeras were flanked by short straight repeats.

This last structural characteristic proves that these chimeras integrated to the genome in the form of a single DNA copy. As a matter of fact, the chimeras in the mammalian genome bore at their 5'-ends T_2A_4-hexanucleotides or their variants (Buzdin et al., 2002; Gogvadze et al., 2005), which respond to the consensus sequence T_2A_4, which is used by the *L1* elements for the initiation of reverse transcription (Jurka, 1997). Synchronous integration of both parts of the chimeras was also confirmed by the data from the locus-specific PCR with the genome DNA of various primate species (Buzdin et al., 2002, 2003).

Thus, chimeras were formed with the use of a special mechanism. This mechanism often unites cellular transcripts that do not have anything in common with transposable elements (Buzdin et al., 2007). Many chimeras can be considered new genes; thus, their transcription was shown, as well as the fact that some of them are transcribed tissue-specifically (Buzdin et al., 2003; Gogvadze et al., 2005, 2007; Gogvadze, Buzdin, 2009). We conjectured that these chimera retrogenes were formed according to a mechanism that

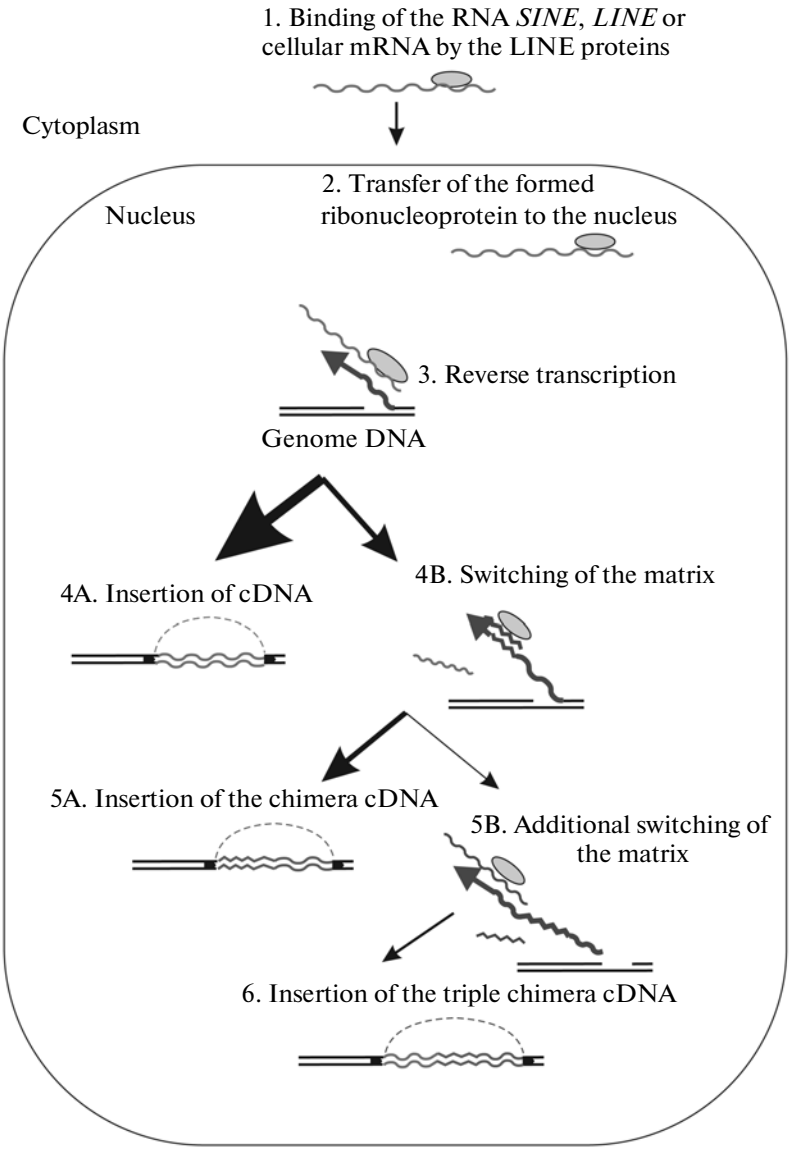

Fig. 2. The mechanism of formation of the chimera using the *L1* enzymatic apparatus. Step 1: the *L1* preintegration protein complex binds the RNA-copy *L1, SINE* or cellular RNA in the cytoplasm; step 2: the formed ribonucleoproteid is transferred to the nucleus; step 3: reverse transcription of the bound RNA primed by the single-chain rupture of the genome DNA; step 4A: successful integration of the transcribed cDNA-copy back to the genome DNA; step 4B: a "jump" to another RNA matrix in the process of reverse transcription; step 5A: integration of the chimera to the genome DNA; and step 5B: second "jump" of the matrix to another RNA with subsequent DNA-reparation, as a result of which insertions of the retrogen of triple chimera structure form. The normal way of the integration of the *L1* element corresponding to its "life cycle": steps (1)- (2)- (3)- (4A).

includes RNA-recombination in the process of reverse transcription of the cellular RNA (Fig. 2). This model implies a "jump" of the forming first cDNA chain from one RNA, which serves as a matrix for reverse transcription of the 3'-end part of the chimera, onto another RNA-matrix, with which the 5'-part is formed. The chimera is subsequently integrated to the genome (Buzdin, 2004).

Although the activity of the reverse transcriptase, which is considered the main one, is directed at the uninterrupted synthesis of cDNA on the RNA matrix, the enzyme is capable of switching to other matrices in the process of reverse transcription. For example, in retroviruses, reverse transcriptase "jumps" from one site of the RNA-matrix to another at least twice, which is necessary for the synthesis of full-length long terminal repeats (LTRs). In addition, retrovirus particles usually contain two matrix RNA molecules (Temin, 1993), along which "jumps" of the matrix occur with high frequency, which significantly increase the degree of changeability of the retrovirus through the recombination between the two matrices

(Kandel, Nudler, 2002). A similar type of recombination, but with foreign RNA-matrices, leads to the formation of a mosaic structure of the genome insertions of some retroviruses (Swanstrom et al., 1983; Jamain et al., 2001).

This model of the formation of chimeras was subsequently confirmed by the results obtained for the human element *L1* using an experimental system of retrotransposition in vitro (Gilbert et al., 2005). The authors characterized 100 retrotranspositions de novo in the HeLa cell line. One insertion (1%) represented an unknown chimera similar to those that we found in the human genome, consisting of a full-length U6 snRNA joined with the 5'-shortened element *L1*. Later, similar results were obtained in vivo using the model of retrotransposition of *L1* in transgenic mice: 33 new retrotransposition events were characterized, 13% of which occurred with the change of the matrix in the process of reverse transcription (Babushok et al., 2006).

The assumption has recently been made that the jumps of the reverse transcriptase from its own *L1* matrix onto the host DNA promote successful integration of the copy of the element into the genome and can thus represent, not the exception, but the rule, in the functioning of this enzyme (Bibillo, Eickbush, 2004; Babushok et al., 2006). Apart from the formation of chimera retrogenes, the jump of the matrix during the reverse transcription can lead to the formation of chimera nonautonomous *SINE* elements (Nishihara et al., 2006) and, apparently, to the mosaicism of the structure of the *L1*-elements themselves (Hayward et al., 1997; Brosius, 1999b). The evolution of separate *L1* families can also include RNA recombination, which leads to the unification of the 3'-end part of *L1* with the new sequence at the 5'-end. As a comparison of the initial structure, 5'-nontranslated regions of human, mouse, rat, and rabbit families are not homologous to each other (Furano, 2000). Interestingly, the reverse transcriptase encoded by another member of the *LINE* superfamily (to which *L1* belongs), R2 from the genome of arthropoda, jumps from one matrix onto another in model experiments in vitro with the formation of the corresponding chimeras (Bibillo, Eickbush, 2004).

In addition, it was hypothesized that all LTR-containing TEs and many *SINE*-elements represent similarly formed chimeras (Ohshima et al., 1996; Kramerov, Vassetzky, 2001, 2005; Malik, Eickbush, 2001). Indeed, a phylogenetic analysis of the domain of ribonuclease H showed that LTR-containing elements could be formed as chimeras between the DNA-transposons and LTR-noncontaining retrotransposons, for example, belonging to *LINE* (Malik, Eickbush, 2001). *SINE*-elements that take root in tRNA probably evolved from the so-called "strong-stop DNA" retrovirus, which is a DNA-copy of the short first fragment of the retrovirus genome formed in the course of reverse transcription (Ohshima et al., 1996). Such *SINE* contain two areas: a conservative one, which includes a promoter of tRNA and a core domain, and a variable one, analogous to the 3'-end sequence of various *LINE* families. The core domain of tRNA-like *SINE* contains conservative sequences similar to the fragments of long end repeats of lysine-tRNA of primering retroviruses.

It was suggested that the *SINE*-elements originating from tRNA appeared as a result of the integration of the retrovirus strong-stop DNA next to the 3'-end part of *LINE*. The formed retroelements could be transcribed by the RNA polymerase III and could spread in the genome. This mechanism of *SINE* formation explains how they can spread in the genome. It is possible that for this they attract the fermentative system of *LINE*-elements, mimicking them using a very similar "tail" sequence (Ohshima et al., 1996).

THE HUMAN-SPECIFIC HYBRID FAMILY *CPG-SVA*

A detailed structural analysis of the insertions of retrotransposons *SVA* in the human genome allowed us to discover 76 elements of an unusual structure. The 5'-ends of such elements carried copies of the first exon of the *MAST2* gene, whereas the 3'-end carried the 3'-end sequence of the *SVA* retrotransposon.

The boundary between the exon area and *SVA* area was situated precisely between the canonic acceptor splice-site AG from the exon part and the noncanonic donor splice-site CC from the *SVA* part. The length of both parts of the chimera element could differ significantly: from 35 to 383 bp for the 5'-end area and from 662 to 4255 bp for the 3'-end area. The border point between the two areas was invariable for all chimeras (Fig. 3).

CpG-SVA elements were found only in the human DNA, but not in other primates, whereas the *SVA* retrotransposons and sequences of *MAST2* exons themselves exist in the genomes of all higher primates. Thus, *CpG-SVA* is the only known human-specific family of retrotransposons (Bantysh, Buzdin, 2009).

In other works simultaneously published that describe the same family of hybrid retrotransposons, this family was also called *MAST2-SVA* (Hancks et al., 2009) or *SVA-F1* (Damert et al., 2009).

Based on the structural specificities of the discovered members of the *CpG-SVA* family, we made a conjecture about the mechanism of their formation (Fig. 4).

At the first stage, the *SVA* retrotransposon apparently entered into the first intron of the *MAST2* gene in a direct orientation. During the transcription of the *MAST2* gene, an aberrant pre-mRNA terminally processed along the 3'-end signal of polyadenylation of the in-built *SVA* element was formed. Then, this RNA was spliced at the acceptor AG splice-site from the gene side and at the donor CC splice-site from the *SVA* side, owing to which a unification of the sequences of

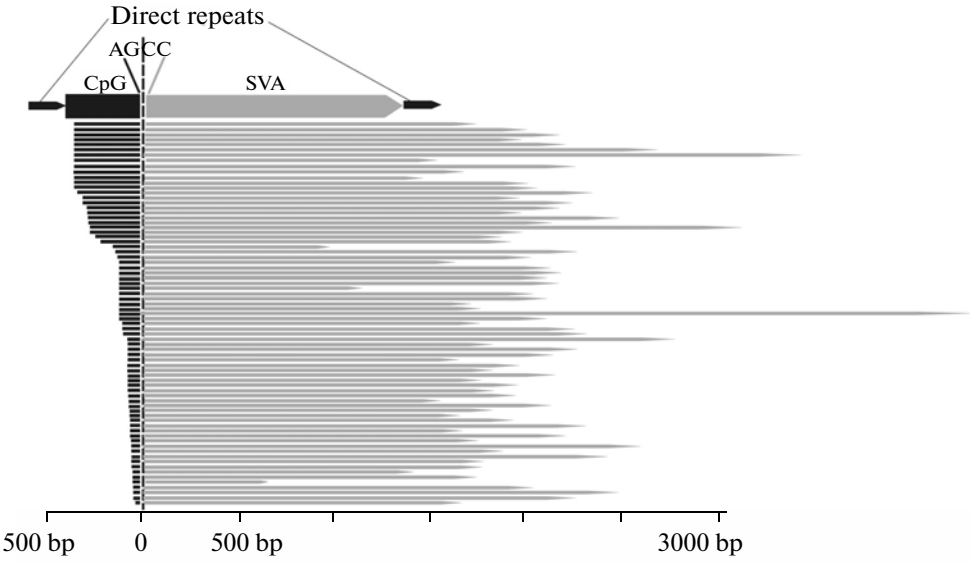

Fig. 3. Structures of the chimera retrotransposon *CpG-SVA*. The *CpG-SVA* insertions are flanked by direct repeats. The length of the 5'-end (exon) part is from 35 to 383 bp, while the length of the 3'-end (*SVA*-derivative) section ranges from 662 to 4255 bp. The 5'-end section is homologous to the first exon of the *MAST2* gene, and the 3'-end part is homologous to the *SVA* retrotransposon. The point of junction between the two parts is identical in all elements of *CpG-SVA* (canonic acceptor splice-site AG from the exon part and noncanonic donor splice-site CC from the *SVA*). All *SVA*-fragments begin from position 396 of the consensus *SVA* sequence.

the first exon *MAST2* with the 3'-end *SVA* fragment (that starts from position 393 of the consensus sequence *SVA*) took place. The donor site of the CC splicing inside the *SVA* element is noncanonic, which is explained by the specificities of the exon–intron structure of *MAST2,* where noncanonic splicing sites exist as a majority. The spliced chimera RNA was then "fortunate" enough to be subjected to reverse transcription with the help of the *L1* fermentative system, which was aided by integration of the formed cDNA to the genome. This event led to the appearance of the so-called "master copy" of the *CpG-SVA* element in the human DNA. The master copy to be proved transcriptionally active, probably due to the insertion of the CpG island of the gene, which led to the formation of a new family of TEs (Fig. 4).

Interestingly, at present, no known inclusion of the *SVA* element into the intron of the *MAST2* gene in the human genome exists. Apparently, the ancestor allele, which contained the whole *SVA* element in the intron of the gene, was eliminated through negative selection, since such a gene could not form a functional RNA due to the disturbance in the splicing of the pre-mRNA and/or premature polyadenylation of the transcript along the *SVA* sequence.

Among the *CpG-SVA* elements, we found cases of 5'- and 3'-end transduction flanking the unique genome DNA. As in the classical mechanism of 3'-transduction, the downstream genome fragments were, apparently, caught due to the skip of a faint signal of the polyadenylation of *SVA* with subsequent termination on the unique downstream genome sequence by the RNA complex of polymerase II. In the case of 5'-end transduction, the transcription of the *CpG-SVA* element was, apparently, initiated from the upstream genome promoters. Altogether, we identified 18 and 11 cases for the 5'- and 3'-end *CpG-SVA* transduction, respectively. The size of the transposed genome sequence varied from 8 to 854 bp for 5' and from 141 to 734 bp for the 3'-end transduction. It is remarkable that four elements of *CpG-SVA* contain both 5'- and 3'-transduced sequences. These four elements possess a high degree of identity and consist of 364 bp of an exon fragment of *MAST2* and an SVA sequence of varying length: 2143–3361 bp. Variations of the length of the *SVA* fragment can be explained by the instability of its inner microsatellite repeating modules. Elements of *CpG-SVA* with double transduction were flanked by the *Alu* sequences (members of the evolutionarily ancient family *AluSc*) at the 5'-end and by a 400-bp-long sequence including an evolutionarily ancient *AluSp* element at the 3'-end. This similarity of the structure proves that all four elements originate from one ancestor *CpG-SVA* element.

The exon part of chimeras varies in length, but not in structure, whereas the part that originated from *SVA* varies considerably both in length and initial structure. Various genetic changes are present in the *SVA* sections, such as insertions, deletions, duplications, changes in the number of tandem repeats, and even insertions of other TEs.

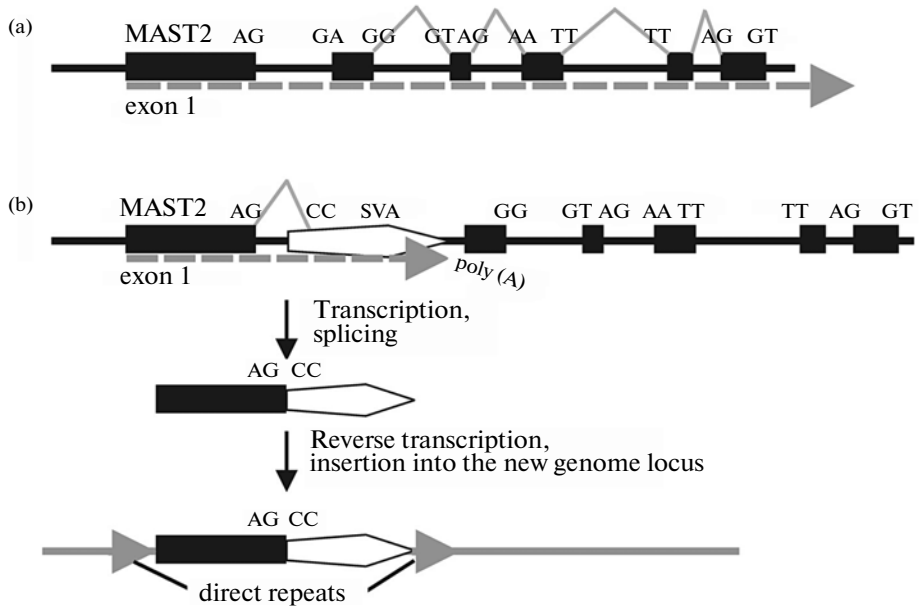

Fig. 4. Assumed mechanism of formation of the *CpG-SVA* family. (a) Schematic image of the genome locus, which contains the human *MAST2* gene. The dotted arrow designates the direction of the transcription, exons, and splicing sites. (b) The insertion of the *SVA* retrotransposon in direct orientation changed the exon–intron structure and led to an incorrectly spliced mRNA polyadenylated along the *SVA* region. A copy of this mRNA was inserted into the new locus of the human genome. This led to formation of the *CpG-SVA* family, which continues its amplification in the human DNA. However, the ancestor allele of the *MAST2* gene that contained the *SVA* insertion was lost due to negative selection.

Together with the presence of transduced genome sequences, this gave us the possibility to build a phylogenetic tree for the members of the *CpG-SVA* family to determine the degree of their mutual kinship. The *CpG-SVA* elements proved to be grouped in three main clusters. Although correlations exist between the size of the exon section and the localization sequence on the tree, all branches of the tree contain elements with exon parts of varying lengths. In addition, no link exists between the position on the phylogenetic tree and the length of the *SVA* sections. In several cases, various tree branches include elements with identical lengths of exons sections. For instance, branch 2 contains one element of *CpG-SVA* with an exon section 364 bp long, whereas branch 1 contains five identical elements. The exon part of seven elements on branch 2 and one element on branch 3 are 148 bp long. The same coincidences were found for lengths of 64, 76, 88, and 361 bp, which indicates multiple independent events in the course of the formation of the *CpG-SVA* elements with identical exon parts.

The observed specificities of the distribution of lengths in the exon section of *CpG-SVA* could be explained by the following factors: 1) there could exist multiple sites of the beginning of transcription within *CpG-SVA* or 2) in some cases, the reverse transcription of RNA *CpG-SVA* could be terminated prematurely, before the copying of the matrix sequences was completed. The resulting shortened insertions of *CpG-SVA* could then give rise to others with even shorter exon parts, etc.

What is the function of the exon part of the *CpG-SVA* elements? Taking into account that 1) the first exon of the *MAST2* gene includes a CpG-island, 2) CpG-islands play an extremely important role in the regulation of the transcription of genes, and 3) the *MAST2* gene is expressed tissue-specifically in testicles, it can be assumed that the exon part can be useful to intensify the transcription of the representatives of *CpG-SVA* in testicles. This may be profitable for the *CpG-SVA* elements as it facilitates the fixation of new insertions in the genome. Indeed, to be transmitted to the offspring, the TE insertion has to be realized in the line of germ cells, which, for example, can be localized in the testicles.

From the point of view of the spread in the genome, the evolutionarily young *CpG-SVA* family discovered by us is very successful, since the offspring of only one of more than a thousand copies of *SVA* elements, which existed in the DNA of human ancestor at the time of divergence with the chimpanzee line (i.e., <0.1%), created 76 new insertions fixed in the genome (~9% of all 860 hominid-specific *SVA* insertions) (Bantysh, Buzdin, 2009).

Our further studies showed that the sequence of the CpG-island of the *MAST2* gene can indeed intensify the transcription of the 3'-end part of *SVA* demonstrating double the promoter activity as compared with the 5'-end part of *SVA* (Zabolotneva et al., unpublished

data). In addition, in mammals, the *MAST2* gene is characterized by the tissue-specific expression in the forming spermatogonial cells. Waves of total demethylation of the genome are observed there as well at certain stages of pubescence, which is accompanied by an increase in the expression of the retroelements, usually repressed with the help of methylation. Accordingly, the involvement of regulator sequences of *MAST2* could have provided the elements of *CpG-SVA* at the transcription activation "at the right place, at the right time."

CONCLUSIONS

Our studies show that transposable elements of the genome are intensely involved in the evolutionary process and are far from existing in "petrified" forms. Systematic and detailed analysis of the insertions of both the known transposable elements and various pseudogenes will probably allow researchers to discover a large number of groups unknown at present. Many of them will, apparently, include complex composite elements similar to the ones described in this article.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (grant no. 10-04-00593-a), the President of the Russian Federation (grant MD-480.2010.4), and the "Molecular and Cell Biology" program of the Presidium of the Russian Academy of Sciences.

REFERENCES

Buzdin, A., Transposable Elements and Their Use for Target Site Specific Gene Delivery, *Curr. Pharmacogenomics,* 2006, vol. 4, no. 1, pp. 1–8.

Buzdin, A.A., Retroelements and Formation of Chimeric Retrogenes, *Cell Mol. Life Sci.*, 2004, vol. 61, no. 16, pp. 2046–2059.

Brosius, J., Genomes Were Forged by Massive Bombardments with Retroelements and Retrosequences, *Genetics*, 1999b, vol. 107, nos. 1/3, pp. 209–238.

Buzdin, A., Gogvadze, E., and Lebrun, M.H., Chimeric Retrogenes Suggest a Role for the Nucleolus in LINE Amplification, *FEBS Lett.*, 2007, vol. 581, no. 16, pp. 2877–2882.

Buzdin, A., Gogvadze, E., Kovalskaya, E., et al., The Human Genome Contains Many Types of Chimeric Retrogenes Generated through *in vivo* RNA Recombination, *Nucleic Acids Res.*, 2003, vol. 31, no. 15, pp. 4385–4390.

Buzdin, A., Ustyugova, S., Gogvadze, E., et al., A New Family of Chimeric Retrotranscripts Formed by a Full Copy of U6 Small Nuclear RNA Fused to the 3 Terminus of 11, *Genomics*, 2002, vol. 80, no. 4, pp. 402–406.

Buzdin, A., Vinogradova, T., Lebedev, Y., and Sverdlov, E., Genome-Wide Experimental Identification and Functional Analysis of Human Specific Retroelements, *Cytogenet. Genome Res.,* 2005, vol. 1101, nos. 1/4, pp. 468–474.

Brosius, J., RNAs from All Categories Generate Retrosequences That May Be Exapted as Novel Genes or Regulatory Elements, *Gene*, 1999a, vol. 238, no. 1, pp. 115–134.

Babushok, D.V., Ostertag, E.M., Courtney, C.E., et al., *L1* Integration in a Transgenic Mouse Model, *Genome Res.,* 2006, vol. 16, no. 2, pp. 240–250.

Bantysh, O.B. and Buzdin, A.A., Novel Family of Human Transposable Elements Formed Due to Fusion of the First Exon of Gene *MAST2* with Retrotransposon SVA, *Biochemistry (Moscow),* 2009, vol. 74, no. 12, pp. 1393–1399.

Boissinot, S., Chevret, P., and Furano, A.V., *L1* (LINE-1) Retrotransposon Evolution and Amplification in Recent Human History, *Mol. Biol. Evol.*, 2000, vol. 17, no. 6, pp. 915–928.

Bibillo, A. and Eickbush, T.H., End-to-End Template Jumping by the Reverse Transcriptase Encoded by the R2 Retrotransposon, *J. Biol. Chem.*, 2004, vol. 279, no. 15, pp. 14945–14953.

Damert, A., Raiz, J., Horn, A.V., et al., 5-Transducing SVA Retrotransposon Groups Spread Efficiently Throughout the Human Genome, *Genome Res.,* 2009, vol. 19, no. 11, pp. 1992–2008.

Dewannieux, M., Esnault, C., and Heidmann, T., *LINE*-Mediated Retrotransposition of Marked *Alu* Sequences, *Nat. Genet.*, 2003, vol. 35, no. 1, pp. 41–48.

Deininger, P.L., Moran, J.V., Batzer, M.A., and Kazazian, H.H., Mobile Elements and Mammalian Genome Evolution, *Curr. Opin. Genet. Dev.*, 2003, vol. 13, no. 6, pp. 651–658.

Fudal, I., Bohnert, H.U., Tharreau, D., and Lebrun, M.H., Transposition of *LINE*, a Composite Retrotransposon, in the Avirulence Gene *ACE1* of the Rice Blast Fungus *Magnaporthe grisea*, *Fungal. Genet. Biol.*, 2005, vol. 42, no. 9, pp. 761–772.

Furano, A.V., The Biological Properties and Evolutionary Dynamics of Mammalian *LINE-1* Retrotransposons, *Prog. Nucl. Acid Res. Mol. Biol.*, 2000, vol. 64, pp. 255–294.

Gilbert, N., Lutz, S., Morrish, T.A., and Moran, J.V., Multiple Fates of *L1* Retrotransposition Intermediates in Cultured Human Cells, *Mol. Cell Biol.*, 2005, vol. 25, no. 17, pp. 7780–7795.

Gogvadze, E. and Buzdin, A., Retroelements and Their Impact on Genome Evolution and Functioning, *Cell Mol. Life Sci.*, 2009, vol. 66, no. 23, pp. 3727–3742.

Gogvadze, E., Barbisan, C., Lebrun, M.H., and Buzdin, A., Tripartite Chimeric Pseudogene from the Genome of Rice Blast Fungus *Magnaporthe grisea* Suggests Double Template Jumps during Long Interspersed Nuclear Element (LINE) Reverse Transcription, *BMC Genomics,* 2007, vol. 8, p. 360.

Gogvadze, E.V., Buzdin, A.A., and Sverdlov, E.D., Multiple Template Switches on LINE-Directed Reverse Transcription: The Most Probable Formation Mechanism for the Double and Triple Chimeric Retroelements in Mammals, *Russ. J . Bioorg. Chem.*, 2005, vol. 31, no. 1, pp. 74–81.

Gogvadze, E.V. and Buzdin, A.A., A New Mechanism of Retrogene Formation in Mammalian Genomes: *In Vivo* Recombination during RNA Reverse Transcription, *Mol. Biol. (Moscow)*, 2005, vol. 39, no. 3, pp. 321–330.

Hayward, B.E., Zavanelli, M., and Furano, A.V., Recombination Creates Novel *L1* (*LINE-1*) Elements in *Rattus norvegicus*, *Genetics,* 1997, vol. 146, no. 2, pp. 641–654.

Hancks, D.C., Ewing, A.D., Chen, J.E., et al., Exon-Trapping Mediated by the Human Retrotransposon SVA, *Genome Res.*, 2009, vol. 19, no. 11, pp. 1983–1991.

Jurka, J., Sequence Patterns Indicate An Enzymatic Involvement in Integration of Mammalian Retroposons, *Proc. Natl. Acad. Sci. U.S.A.*, 1997, vol. 94, no. 5, pp. 1872–1877.

Jamain, S., Girondot, M., Leroy, P., et al., Transduction of the Human Gene *FAM8A1* by Endogenous Retrovirus During Primate Evolution, *Genomics*, 2001, vol. 78, nos. 1/2, pp. 38–45.

Kramerov, D.A. and Vassetzky, N.S., Structure and Origin of a Novel Dimeric Retroposon *B1-DiD*, *J. Mol. Evol.*, 2001, vol. 52, no. 2, pp. 137–143.

Kramerov, D.A. and Vassetzky, N.S., Short Retroposons in Eukaryotic Genomes, *Int. Rev. Cytol.*, 2005, vol. 247, pp. 165–221.

Kazazian, H.H., Jr. Mobile Elements: Drivers of Genome Evolution, *Science*, 2004, vol. 303, no. 5664, pp. 1626–1632.

Kandel, E.S. and Nudler, E., Template Switching by RNA Polymerase II *in vivo*. Evidence and Implications from a Retroviral System, *Mol. Cell*, 2002, vol. 10, no. 6, pp. 1495–1502.

Lebedev, Y.B., Amosova, A.L., Mamedov, I.Z., et al., Most Recent *AluY* Insertions in Human Gene Introns Reduce the Content of the Primary Transcripts in a Cell Type Specific Manner, *Gene*, 2007, vol. 390, nos. 1/2, pp. 122–129.

Martin, S.L., The ORF1 Protein Encoded by *LINE-1*: Structure and Function during *L1* Retrotransposition, *J. Biomed. Biotechnol.,* 2006, vol. 2006, pp. 1–6.

Mills, R.E., Bennett, E.A., Iskow, R.C., et al., Recently Mobilized Transposons in the Human and Chimpanzee Genomes, *Am. J. Hum. Genet.*, 2006, vol. 78, no. 4, pp. 671–679.

Malik, H.S. and Eickbush, T.H., Phylogenetic Analysis of Ribonuclease H Domains Suggests a Late, Chimeric Origin of LTR Retrotransposable Elements and Retroviruses, *Genome Res.,* 2001, vol. 11, no. 7, pp. 1187–1197.

Nishihara, H., Smit, A.F., and Okada, N., Functional Noncoding Sequences Derived from *SINEs* in the Mammalian Genome, *Genome Res.*, 2006, vol. 16, no. 7, pp. 864–874.

Ohshima, K., Hamada, M., Terai, Y., and Okada, N., The 3 Ends of TRNA-Derived Short Interspersed Repetitive Elements Are Derived from the 3 Ends of Long Interspersed Repetitive Elements, *Mol. Cell Biol.*, 1996, vol. 16, no. 7, pp. 3756–3764.

Sverdlov, E.D., Perpetually Mobile Footprints of Ancient Infections in Human Genome, *FEBS Lett.*, 1998, vol. 428, nos. 1/2, pp. 1–6.

Sverdlov, E.D., Retroviruses and Primate Evolution, *BioEssays*, 2000, vol. 22, no. 2, pp. 161–171.

Swanstrom, R., Parker, R.C., Varmus, H.E., and Bishop, J.M., Transduction of a Cellular Oncogene: The Genesis of Rous Sarcoma Virus, *Proc. Natl. Acad. Sci. U.S.A.*, 1983, vol. 80, no. 9, pp. 2519–2523.

Schumann, G.G., Gogvadze, E.V., Osanai-Futahashi, M., et al., Unique Functions of Repetitive Transcriptomes, *Int. Rev. Cell Mol. Biol.*, 2010, vol. 285, pp. 115–188.

Sen, S.K., Han, K., Wang, J., et al., Human Genomic Deletions Mediated by Recombination between *Alu* Elements, *Am. J. Hum. Genet.*, 2006, vol. 79, no. 1, pp. 41–53.

Temin, H.M., Retrovirus Variation and Reverse Transcription: Abnormal Strand Transfers Result in Retrovirus Genetic Variation, *Proc. Natl. Acad. Sci. U.S.A.*, 1993, vol. 90, no. 15, pp. 6900–6903.

Ustyugova, S.V., Lebedev, Y.B., and Sverdlov, E.D., Long *L1* Insertions in Human Gene Introns Specifically Reduce the Content of Corresponding Primary Transcripts, *Genetica*, 2006, vol. 128, no 1/3, pp. 261–272.

Ustyugova, S.V., Amosova, A.L., Lebedev, Yu.B., and Sverdlov, E.D., A Tissue-Specific Decrease in the Pre-mRNA Level of *L1*- and *Alu*-Containing Alleles

of Human Genes, *Bioorg. Khim.,* 2006, vol. 32, no. 1, pp. 83–95.

van de Lagemaat, L.N., Medstrand, P., and Mager, D.L., Multiple Effects Govern Endogenous Retrovirus Survival Patterns in Human Gene Introns, *Genome Biol.*, 2006, vol. 7, no. 9, p. R86.

Wessler, S.R., Transposable Elements and the Evolution of Gene Expression, *Symp. Soc. Exp. Biol.*, 1998, vol. 51, pp. 115–122.

Wang, H., Xing, J., Grover, D., et al., SVA Elements: a Hominid-Specific Retroposon Family, *J. Mol. Biol.*, 2005, vol. 354, no. 4, pp. 994–1007.

Zabolotneva, A., Tkachev, V., Filatov, F., and Buzdin, A., How Many Antiviral Small Interfering RNAs May Be Encoded by the Mammalian Genomes?, *Biol. Direct.,* 2010, vol. 5, p. 62.