FEBS Letters 581 (2007) 2877–2882

Minireview

Chimeric retrogenes suggest a role for the nucleolus in LINE amplification

Anton Buzdin[a,*], Elena Gogvadze[a], Marc-Henri Lebrun[b]

[a] *Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Moscow 117871, Russia*
[b] *UMR2847 CNRS/Bayer CropScience, 14-20 rue Pierre Baizet, 69263 Lyon, Cedex 09, France*

Received 22 March 2007; revised 3 May 2007; accepted 16 May 2007

Available online 25 May 2007

Edited by Lev Kisselev

Abstract **Chimeric retrogenes, found in mammalian and fungal genomes, are bipartite elements composed of DNA copies of cellular transcripts either directly fused to each other or fused to the 3′ part of a LINE retrotransposon. These cellular transcripts correspond to messenger RNAs, ribosomal RNAs, small nuclear RNAs and 7SL RNA. The chimeras are likely formed by RNA template switches during reverse transcription of LINE elements by their retrotranspositional machinery. The 5′ part of chimeras are copies of nucleolar RNAs, suggesting that the nucleolus plays a significant role in LINE retrotransposition. RNAs from the nucleolus might have protective function against retroelement invasion or, alternatively, the nucleolus may be required for retrotranspositional complex assembly and maturation. These hypotheses will be discussed in this review.**

Keywords: L1 LINE; Retrotransposition; Chimeric retroelements; Reverse transcription; Nucleolus; U6 small nuclear RNA

1. Introduction

Reverse transcription is one of the key processes that shape eukaryotic genomes. At least 40% of the mammalian genome is resulting from reverse transcription events [1,2]. This phenomenon was discovered by Temin and Baltimore while they purified and characterized the first retroviral RNA-dependant DNA polymerase (reverse transcriptase, RT), which catalyzes the synthesis of complementary DNA using a RNA template [3]. Since, RT sequences were found in diverse retroviruses, mitochondrial group II introns, bacterial retrointrons, plasmids and in genetic elements termed retroelements (REs). REs are transposable elements that proliferate through RNA intermediates using self-encoded or exogenous RT and insert the newformed DNA copy of the element into the host genome.

Autonomous retroelements that carry their own RT genes can be subdivided into two major classes: long terminal repeat (LTR) containing elements, and non LTR retrotransposons [4]. Autonomous non LTR REs are mostly long interspersed nuclear elements (LINEs) found in essentially all eukaryotic genomes [5,6]. LINEs are 3.5–8 kb long, encode a RT and few other proteins necessary for their transposition [7]. LINEs also provide their RT enzyme for the proliferation of non-autonomous REs such as short interspersed nuclear elements (SINEs) [8]. Novel insertions of LINEs in a genome can be easily identified by the presence of 10–20 bp long target site duplications flanking the REs, that are formed during the integration process. LINEs contain canonical polyadenylation signals, oligo(A) tails or, sometimes, other A-rich sequences at their 3′-termini [9]. LINEs are transcribed by the cellular RNA polymerase II from an internal promoter located in their 5′-untranslated region [10]. Another LINE distinguishing feature is their frequent 5′-truncation that likely result from an interruption of LINE RNA reverse transcription, as RT could frequently dissociate from its RNA template before having completed a full cDNA synthesis [11].

Most LINEs found in eukaryotic genomes are inactive 5′-truncated copies that are transpositionally deficient while only a small number of actively transposing full-sized elements are present [12]. However, LINEs have frequently expanded during genome evolution as observed for the human genome that contains 5×10^5 L1 elements representing 17% of the total human genomic DNA [1,2]. The presence of such a number of LINEs in genomes affects many cellular processes. Highly repetitive LINE sequences may serve as recombination hot spots, causing frequent host DNA rearrangements [13]. Moreover, LINEs may disrupt preexisting gene exon-intronic structures [14] and in different ways interfere with host gene expression [15–17]. Another interesting property of LINEs is their ability to transfer their 3′-flanking DNA to new genomic loci, termed L1 transduction [18,19]. Taken together, this transduced DNA makes up ~0.6–1% of the human genome.

The full-sized LINE (+) RNA has a dual role as transpositional RNA intermediate and template for protein synthesis [20]. LINE transposition is known to proceed in several steps including RNA Pol II transcription of an active element, reverse transcription of the RNA formed with a LINE-encoded RT, and integration of the cDNA into a new position within the genome using an endonuclease [21] (Fig. 1A). A typical LINE element encodes two proteins: ORF1p that is a RNA binding protein which likely helps reverse transcription as a nucleic acid chaperone [22], and ORF2p, the reverse transcriptase and the endonuclease [21]. Due to a '*cis*-preference', the enzymatic machinery of a retrotransposition-competent LINE predominantly transposes its own copies [23] (Fig. 1A). However, LINEs are able to mediate the transposition of other sequences, mostly non-autonomous elements termed SINEs, but

*Corresponding author. Fax: +7 495 3306538.
E-mail address: anton@humgen.siobc.ras.ru (A. Buzdin).

0014-5793/$32.00
doi:10.1016/j.febslet.2007.05.034

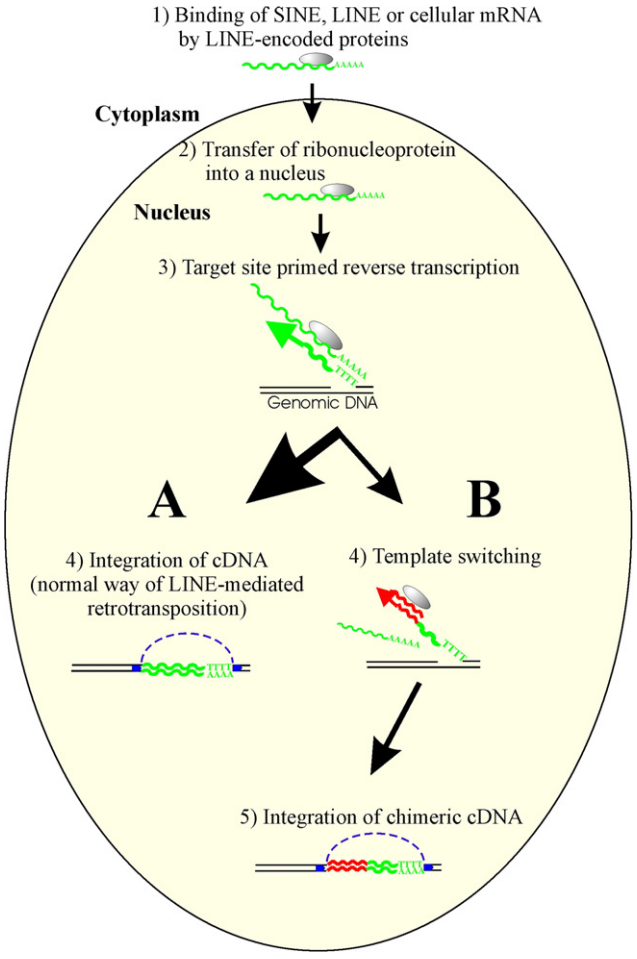

Fig. 1. Mechanism for the chimeras' formation by LINE enzymatic machinery. (Step 1) LINE pre-integration complex binds LINE, SINE or host mRNAs in the cytoplasm. (Step 2) The resulting ribonucleoprotein is transferred to the nucleus. (Step 3) Reverse transcription of the bound mRNA primed by a genomic DNA single-stranded break within the TTTTAA sequence (target site primed reverse transcription). (Step 4A) Successful integration of the reverse transcribed LINE cDNA copy into the genomic DNA. (Step 4B) Switch of templates to another RNA during the reverse transcription. (Step 5A) Integration of the chimera formed after RNA template switch into genomic DNA. This event leads to the formation of a bipartite chimeric retrogene carrying a poly(A) sequence at the 3′ terminus and flanked by short direct repeats. The normal LINE integration pathway is: Steps (1)–(2)–(3)–(4A).

also cDNAs originating from different cellular RNAs, leading to the formation of processed pseudogenes [24]. Recently, we have shown that LINEs are involved in the formation of bipartite chimeric retrogenes during reverse transcription in many genomes including human and fungi [25–29].

2. Template switching generates bipartite chimeric retrogenes

Bipartite chimeric retrogenes with an unusual structure were recently identified in three mammalian and in one fungal genomes (Fig. 2): a total of 82, 116, 66 and 31 elements were found in human, mouse, rat and rice blast fungus *Magnaporthe grisea* DNAs, respectively [25,27–29]. These elements are composed of DNA copies from cellular transcripts either directly fused

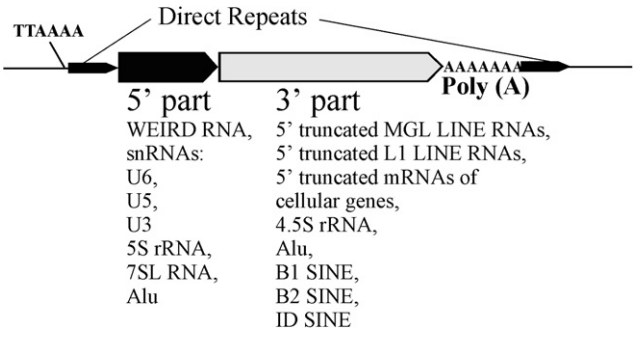

Fig. 2. Schematic representation of the bipartite chimeric retrogenes identified in eukaryotic genomes. The insertions in mammalian genomes are located downstream of the TTAAAA hexanucleotide motif. Insertions in mammalian and fungal genomes harbour poly(A) sequence and are flanked by 10–25 bp long genomic direct repeats.

to each other or more frequently fused to the 3′ part of a LINE retroposon. The various cellular transcripts found in these chimeras correspond to messenger RNAs, ribosomal RNAs, small nuclear RNAs, and 7SL RNA.

The chimeras have the following common features: (i) 5′-parts are full-length copies of cellular RNAs; (ii) 3′-parts are 5′-truncated copies of the corresponding RNAs (mostly LINEs); (iii) sites of these truncations occur at random in the corresponding RNA; (iv) both parts are directly joined with the same transcriptional orientation; (v) chimeras have a poly(A) tail at their 3′ end, and (vi) chimeras are flanked by short direct repeats.

The last structural feature demonstrate that these elements were transposed as bipartite DNA copies. Indeed, mammalian chimeras carried at their 5′ ends a T_2A_4 hexanucleotide or its variants [25,27,28] that correspond to the T_2A_4 genomic site used by LINEs to initiate reverse transcription on oligo(A) motifs and separate newly inserted DNA by short tandem repeats [30]. The simultaneous integration of both parts of these chimeras was further supported by the data came from PCR-based evolutionary insertion polymorphism assay [25,27].

The number of mouse and rat chimeras is likely underestimated as their 3′-terminal parts are often missing because they correspond to gaps in the genome sequence (23 and 33 cases for mouse and rat genomes, respectively). This significant sampling suggests that these bipartite elements are generated by a specific active mechanism. This mechanism frequently combines functional cellular transcripts that have nothing in common with transposable elements. Many of the chimeras can be considered as new genes, as they were shown to be transcribed, some of them in a tissue-specific manner [25,28,31] Gogvadze, 2007, unpublished for *M. grisea*.

We further hypothesised that these chimeric retrogenes were generated through a mechanism involving RNA recombination during the reverse transcription of cellular RNAs (Fig. 1B). This model includes a switch from the nascent cDNA serving as template for the reverse transcription of the 3′ part of the chimera to another RNA template corresponding to the 5′ part, followed by the chimera integration into the host genome [6]. Although RT main enzymatic activity is the continuous synthesis of the cDNA on RNA template, RT is able to switch templates during reverse transcription. For example, in retroviruses, RT jumps from one site of the RNA template to another site, are necessary for the synthesis

of LTRs. Moreover, as retroviral particles usually contain two independent RNA molecules [32], the high template switch frequency significantly increases the retroviral diversity through recombination between these RNAs [33]. These recombination events most probably account for the mosaic structure of most retroviruses [34,35].

Besides generating chimeric retrogenes, template switching events during LINE reverse transcription could give rise to chimeric SINE elements [36] and to mosaic rodent L1 structures, likely resulting from RNA recombination between L1 templates [37,38]. Evolution of certain LINE families might also involve RNA–RNA recombination, resulting in the fusion of the 3′ part of a LINE to a new sequence at their 5′ end, as suggested by the observation that the 5′-untranslated regions of human, murine, rat and rabbit L1 families are not homologous to each other [11]. Interestingly, RT encoded by another member of LINE superfamily – R2 from arthropods, was documented to jump from one template to another *in vitro*, with R2–R2 chimeras being formed [39].

This model for the chimera formation was further supported by results obtained with human L1 LINE element using an elegant experimental system of retrotransposition *in vitro* [40]. The authors managed to characterize 100 de novo retrotransposition events in HeLa cells. Importantly, one insert (1%) represented a newly formed chimera similar to those we identified in human genome, consisting of a full length U6 snRNA fused to a 5′ truncated L1. Similar results were obtained *in vivo* with a transgenic mouse model for L1 retrotransposition by Babushok and coauthors that characterized 33 novel retrotransposition events. Thirteen percent of these events likely result from template switching during reverse transcription [41]. Interestingly, it has been recently postulated that RT template jumps from LINE RNA to host genomic DNA might facilitate integration and, thus, could be normally required for successful LINE retrotransposition [39,41].

Overall, these genome analyses, evolutionary studies and experimental evidence strongly suggest that RNA template switching occurs during reverse transcription of LINE retroelements. This property of LINE RTs is likely responsible for the generation of mosaic retroelements and chimeric retrogenes. We were able to find such chimeras in the available mammalian genomes, but not in those from amphibian, fishes or invertebrates. However, such chimeras were identified in the genome of rice blast fungus *Magnaporthe grisea* [29] suggesting a conservation of this mechanism across animal and fungal kingdoms.

3. Why are there so many nucleolar RNAs in bipartite LINE chimeras?

Essentially all RNAs participating in the chimera formation are abundant cellular transcripts. This observation is in line with the observed prevalence of abundant RNA copies among the usual (singular) retropseudogenes [42]. The 5′-parts of the bipartite chimeric retrogenes correspond to DNA copies of nucleolar RNAs, while the 3′-parts are copies of cytoplasmic RNAs including LINE transcripts. The 3′ parts of the chimeras are mostly (>80%) composed of 5′-truncated copies of a LINE. However, we identified in 12–20% of the chimeras a 3′ part corresponding either to pseudogenes of various protein coding mRNAs or to SINE retroposons. In mammalian genomes, the 5′ parts of these chimeras correspond mostly to pseudogenes encoding small nuclear RNAs (snRNAs) including U6 (82–93%), U3 (1–10%) and U5 (0–1%). Other pseudogenes identified at low frequency at their 5′-termini include Alu SINE (0–2%), 7SL RNA (0–1%) and 5S ribosomal RNA (3–6%). All these RNAs have known nucleolar localization. U3 snRNA is involved in nucleolar ribosomal RNA processing [43]. Although U5 and U6 normally reside outside of nucleolus in speckles (spliceosomes), their maturation is performed in the nucleolus (U6–2′-*O*-methylation and pseudouridylation of nucleotides) [44]. 7SL and Alu RNAs are processed in the nucleolus [45] as well as 5S rRNA (pre-rRNA processing). Unfortunately, in the fungus *M. grisea*, we were not able to identify cellular localization of the RNA corresponding to the 5′ part of the fungal chimeric retrogene family termed *MINE*.

The 5′ parts of *MINE* correspond to a full length copy of a 1.1 kb long non-coding transcript of unknown function called *WEIRD* that is constitutively expressed at high level [29]. U6 encoding gene was identified in *M. grisea* genome and it is not associated with the chimeras [Gogvadze, 2007 unpublished]. The 3′ parts of *MINE* correspond to 5′ truncated copies of the *MGL* LINE element [29]. As the fungal *MINE* element is similar in structure to mammalian chimeras, we hypothesize that the RNA corresponding to its 5′ part (WEIRD), although it differs from U6 snRNA, is likely located in the nucleolus or associated with *MGL* LINE RNA or particles.

The prevalence of DNA copies corresponding to nucleolar RNAs at the 3′ terminus of mammalian chimeric retrogenes suggests that some phases of the LINE life cycle are associated with the nucleolus. In agreement with this hypothesis, Goodier et al. [46] have shown that both human LINE-encoded proteins (ORF1p and ORF2p) may enter the nucleolus. The authors identified a functional nucleolar localization signal in ORF2p (RT/integrase) and found that the C-terminus truncated ORF2 protein was cytoplasmic, nuclear and nucleolar in the cultured cells. ORF1p was mainly localized in the cytoplasm with a speckled pattern although it also colocalized with ORF2p in the nucleolus of a subset of cells. Therefore, the nucleolus is a cell compartment in which mammalian LINE proteins and possibly RNAs are occurring. This finding is in line with the recent identification of numerous mammalian nucleolar RNA-derived retropseudogenes most probably created by LINE enzymatic machinery [47,48]. One such pseudogene was even a bipartite chimera consisting of a full-length housekeeping gene *RPS3A* processed pseudogene followed by a copy of H/ACA box U70 snoRNA [48].

However, it is clear that LINE reverse transcription is not limited to the nucleolus. Otherwise, a majority of LINEs would be inserted into ribosomal RNA genes, which are the nucleoli-forming centers. Obviously, this is not the case [49].

A second hypothesis postulate that RNAs from the nucleolus have a role in protecting the genome from LINE invasions. As the 5′-terminal LINE components always correspond to 5′ truncated copies, these chimeras could result from unsuccessful LINE retrotransposition events. According to this hypothesis, 5′ truncations of newly inserted LINEs could reflect RNA attacks of retrotranspositional complexes. This hypothesis is in agreement with the fact that U6 snRNA is the key spliceosomal attacking component that may act on other RNAs. A survey of the three mammalian genomes studied for 5′-truncated

LINE inserts showed that their truncation sites are quite similar to those observed in bipartite chimeric retrogenes (Fig. 3). Interestingly, L1 LINE RT has at least five pausing sites during reverse-transcription of L1 RNA [50]. Surprisingly, these sites coincide with some of the hot spots for LINE 5′ truncation and chimerization (Fig. 3). In *in vitro* experiments with R2 LINE RT, it was shown that RNA template switches occur only when the RT has reached the 5′ end of the RNA template [39]. If this is true for L1 RT as well, template switch and chimeric retrogene production could only occur at the 5′ end of a template RNA. Therefore, the template corresponding to the 3′-terminus of the chimera can only be a nicked or damaged (LINE) RNA. 5′ truncated LINE RNAs have been already reported by Belancio et al. [51] in human and mouse as a result of *in vivo* splicing.

A third hypothesis, likely more realistic, explains the presence of LINE ribonucleoproteins in the nucleolus as a need for their RNA modification(s) that would be required for the activation of retrotransposition. This hypothesis relies partly on the fact that numerous RNA modifications are performed in the nucleolus, many being uncharacterized. Additionally, the nucleolus might be the compartment in which final assembly of the LINE ribonucleoprotein complex is taking place, as observed for several retroviruses [52]. The recent identification of conserved functional nucleolar localization signals in mammalian LINE RTs supports this hypothesis [19]. It is believed that LINE ribonucleoprotein complex is normally pre-assembled in cytoplasm and enters later in the nucleus. According to this hypothesis, this ribonucleoprotein would be further transferred to nucleolus, where it would maturate and will be finally assembled. During this maturation process, a subset of abundant nucleolar RNAs could be occasionally captured by the LINE ribonucleoprotein, especially by its ORF1p RNA binding protein component. The final active complex could move away from the nucleolus and start the retrotransposition when in contact with chromosomal DNA. At this stage, reverse-transcription pause may occur at low frequency [50]. This pause could be followed by the dissociation of RT from the original RNA template creating new 5′-truncated LINE copy, or by a switch of the RT to a new RNA template likely captured by LINE ribonucleoprotein complex in the nucleolus, creating a novel bipartite chimeric retrogene. This hypothesis could also explain the high proportion of chimeras with a U6 copy as a consequence of its high affinity for the LINE RNA binding protein. Indeed, as shown indirectly by Garcia-Perez et al. [53], ORF1p was absolutely required for the successful U6 pseudogenization de novo, whereas other RNAs could be successfully mobilized by the RT alone. In addition, U6 was shown to display a high affinity for Gag proteins from retroviruses or LTR retrotransposons that are nucleic acid chaperones functionally similar to LINE ORF1p [54] and U6 molecules are frequently captured by retroviral particles [55,56].

Although these different hypotheses are still equally probable, they reflect an unexpected role of the nucleolus and its abundant RNAs in the reverse transcription and transposition of LINE elements. We believe that the cellular localization of

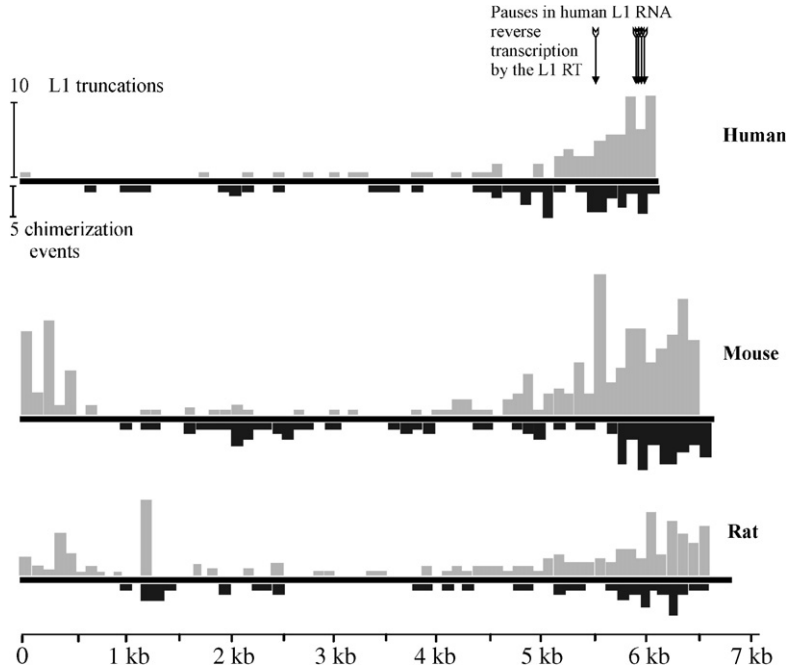

Fig. 3. Comparison of frequencies of 5′-truncation versus chimerization events along the mammalian L1 sequences. Upper columns (in gray) represent each the number of L1 5′-truncations per 100 nt, whereas chimerization events are shown in black (scale bars on the left). Co-clusterization of both truncation- and chimerization-rich regions in the 3′-terminal part of L1 sequence is seen for all three genomes investigated. For these plots, chimerization points for all mammalian L1-containing chimeras were compared with 5′-truncation sites from randomly chosen 250 human, 350 mouse and 210 rat L1s. Arrows represent transcriptional pause sites, observed for human L1 RT when reverse-transcribing L1 template *in vitro*. For human L1s, 50 L1PA1, 50 L1PA2, 50 L1PA3, 50 L1PA5 and 50 L1PA7 elements were taken. These and other mammalian LINE families were chosen because their representatives were found among the chimera 3′ parts. For mouse, 150 L1_MM, 150 L1Md_F and 50 L1Md_T elements were sampled, whereas for the rat we took 210 L1_RN elements. Full length elements were excluded from the sampling. All mammalian L1 elements were recovered with BLAT search engine at the UCSC web site (http://genome.ucsc.edu/cgi-bin/hgBlat?command=start) using as a query 50 3′-terminal nucleotides of the consensus sequence for each respective L1 family [57].

the WEIRD RNA, a major component of the fungal bipartite chimeras, will help understanding this challenging phenomenon.

Acknowledgements: This work was supported by the Molecular and Cellular Biology Program of the Presidium of the Russian Academy of Sciences and by the Grants 05-04-50770-a, 05-04-48682-a from the Russian Foundation for Basic Research, by the Grant MK-4227.2007.4 by the president of the Russian Federation and by the FEBS short-term fellowship.

References

[1] Mouse Genome Sequencing Consortium (2002) Initial sequencing and comparative analysis of the mouse genome. Nature 420, 520–562.
[2] International Human Genome Sequencing Consortium (2001) Initial sequencing and analysis of the human genome. Nature 409, 860–921.
[3] Temin, H.M. (1982) Viruses, protoviruses, development, and evolution. J. Cell Biochem. 19, 105–118.
[4] Leib-Mosch, C. and Seifarth, W. (1995) Evolution and biological significance of human retroelements. Virus Genes 11, 133–145.
[5] Volff, J.N., Korting, C. and Schartl, M. (2000) Multiple lineages of the non-LTR retrotransposon Rex1 with varying success in invading fish genomes. Mol. Biol. Evol. 17, 1673–1684.
[6] Buzdin, A.A. (2004) Retroelements and formation of chimeric retrogenes. Cell Mol. Life Sci. 61, 2046–2059.
[7] Finnegan, D.J. (1997) Transposable elements: how non-LTR retrotransposons do it. Curr. Biol. 7, R245–R248.
[8] Matsutani, S. (2006) Links between repeated sequences. J. Biomed. Biotechnol. 2006, 13569.
[9] Birnstiel, M.L., Busslinger, M. and Strub, K. (1985) Transcription termination and 3′ processing: the end is in site! Cell 41, 349–359.
[10] Mizrokhi, L.J., Georgieva, S.G. and Ilyin, Y.V. (1988) jockey, a mobile Drosophila element similar to mammalian LINEs, is transcribed from the internal promoter by RNA polymerase II. Cell 54, 685–691.
[11] Furano, A.V. (2000) The biological properties and evolutionary dynamics of mammalian LINE-1 retrotransposons. Prog. Nucleic Acid Res. Mol. Biol. 64, 255–294.
[12] Brouha, B., Schustak, J., Badge, R.M., Lutz-Prigge, S., Farley, A.H., Moran, J.V. and Kazazian Jr., H.H. (2003) Hot L1s account for the bulk of retrotransposition in the human population. Proc. Natl. Acad. Sci. USA 100, 5280–5825.
[13] Burwinkel, B. and Kilimann, M.W. (1998) Unequal homologous recombination between LINE-1 elements as a mutational mechanism in human genetic disease. J. Mol. Biol. 277, 513–517.
[14] Narita, N., Nishio, H., Kitoh, Y., Ishikawa, Y., Minami, R., Nakamura, H. and Matsuo, M. (1993) Insertion of a 5′ truncated L1 element into the 3′ end of exon 44 of the dystrophin gene resulted in skipping of the exon during splicing in a case of Duchenne muscular dystrophy. J. Clin. Invest. 91, 1862–1867.
[15] Matlik, K., Redik, K. and Speek, M. (2006) L1 antisense promoter drives tissue-specific transcription of human genes. J. Biomed. Biotechnol. 2006, 71753.
[16] Deininger, P.L., Moran, J.V., Batzer, M.A. and Kazazian Jr., H.H. (2003) Mobile elements and mammalian genome evolution. Curr. Opin. Genet. Dev. 13, 651–658.
[17] Roy-Engel, A.M., El-Sawy, M., Farooq, L., Odom, G.L., Perepelitsa-Belancio, V., Bruch, H., Oyeniran, O.O. and Deininger, P.L. (2005) Human retroelements may introduce intragenic polyadenylation signals. Cytogenet. Genome Res. 110, 365–371.
[18] Pickeral, O.K., Makalowski, W., Boguski, M.S. and Boeke, J.D. (2000) Frequent human genomic DNA transduction driven by LINE-1 retrotransposition. Genome Res. 10, 411–415.
[19] Goodier, J.L., Ostertag, E.M. and Kazazian Jr., H.H. (2000) Transduction of 3′-flanking sequences is common in L1 retrotransposition. Hum. Mol. Genet. 9, 653–657.
[20] Martin, S.L., Li, W.L., Furano, A.V. and Boissinot, S. (2005) The structuRes. of mouse and human L1 elements reflect their insertion mechanism. Cytogenet. Genome Res. 110, 223–228.
[21] Kazazian Jr., H.H. (2004) Mobile elements: drivers of genome evolution. Science 303, 1626–1632.
[22] Martin, S.L. (2006) The ORF1 protein encoded by LINE-1: structure and function during L1 retrotransposition. J. Biomed. Biotechnol. 2006, 45621.
[23] Wei, W., Gilbert, N., Ooi, S.L., Lawler, J.F., Ostertag, E.M., Kazazian, H.H., Boeke, J.D. and Moran, J.V. (2001) Human L1 retrotransposition: *cis* preference versus *trans* complementation. Mol. Cell Biol. 21, 1429–1439.
[24] Dewannieux, M., Esnault, C. and Heidmann, T. (2003) LINE-mediated retrotransposition of marked Alu sequences. Nat. Genet. 35, 41–48.
[25] Buzdin, A., Gogvadze, E., Kovalskaya, E., Volchkov, P., Ustyugova, S., Illarionova, A., Fushan, A., Vinogradova, T. and Sverdlov, E. (2003) The human genome contains many types of chimeric retrogenes generated through *in vivo* RNA recombination. Nucleic Acids Res. 31, 4385–4390.
[26] Buzdin, A., Vinogradova, T., Lebedev, Y. and Sverdlov, E. (2005) Genome-wide experimental identification and functional analysis of human specific retroelements. Cytogenet. Genome Res. 110, 468–474.
[27] Buzdin, A., Ustyugova, S., Gogvadze, E., Vinogradova, T., Lebedev, Y. and Sverdlov, E. (2002) A new family of chimeric retrotranscripts formed by a full copy of U6 small nuclear RNA fused to the 3′ terminus of L1. Genomics 80, 402–406.
[28] Gogvadze, E.V., Buzdin, A.A. and Sverdlov, E.D. (2005) Multiple template switches on LINE-directed reverse transcription: the most probable formation mechanism for the double and triple chimeric retroelements in mammals. Bioorg. Khim. 31, 82–89.
[29] Fudal, I., Bohnert, H.U., Tharreau, D. and Lebrun, M.H. (2005) Transposition of MINE, a composite retrotransposon, in the avirulence gene ACE1 of the rice blast fungus *Magnaporthe grisea*. Fungal Genet. Biol. 42, 761–772.
[30] Jurka, J. (1997) Sequence patterns indicate an enzymatic involvement in integration of mammalian retroposons. Proc. Natl. Acad. Sci. USA 94, 1872–1877.
[31] Gogvadze, E.V. and Buzdin, A.A. (2005) New mechanism of retrogene formation in mammalian genomes: *in vivo* recombination during RNA reverse transcription. Mol. Biol. (Mosk). 39, 364–373.
[32] Temin, H.M. (1993) Retrovirus variation and reverse transcription: abnormal strand transfers result in retrovirus genetic variation. Proc. Natl. Acad. Sci. USA 90, 6900–6903.
[33] Kandel, E.S. and Nudler, E. (2002) Template switching by RNA polymerase II *in vivo*. Evidence and implications from a retroviral system. Mol. Cell 10, 1495–1502.
[34] Swanstrom, R., Parker, R.C., Varmus, H.E. and Bishop, J.M. (1983) Transduction of a cellular oncogene: the genesis of Rous sarcoma virus. Proc. Natl. Acad. Sci. USA 80, 2519–2523.
[35] Jamain, S., Girondot, M., Leroy, P., Clergue, M., Quach, H., Fellous, M. and Bourgeron, T. (2001) Transduction of the human gene FAM8A1 by endogenous retrovirus during primate evolution. Genomics 78, 38–45.
[36] Nishihara, H., Smit, A.F. and Okada, N. (2006) Functional noncoding sequences derived from SINEs in the mammalian genome. Genome Res. 16, 864–874.
[37] Brosius, J. (1999) Genomes were forged by massive bombardments with retroelements and retrosequences. Genetica 107, 209–238.
[38] Hayward, B.E., Zavanelli, M. and Furano, A.V. (1997) Recombination creates novel L1 (LINE-1) elements in *Rattus norvegicus*. Genetics 146, 641–654.
[39] Bibillo, A. and Eickbush, T.H. (2004) End-to-end template jumping by the reverse transcriptase encoded by the R2 retrotransposon. J. Biol. Chem. 279, 14945–14953.
[40] Gilbert, N., Lutz, S., Morrish, T.A. and Moran, J.V. (2005) Multiple fates of L1 retrotransposition intermediates in cultured human cells. Mol. Cell Biol. 25, 7780–7795.
[41] Babushok, D.V., Ostertag, E.M., Courtney, C.E., Choi, J.M. and Kazazian Jr., H.H. (2006) L1 integration in a transgenic mouse model. Genome Res. 16, 240–250.
[42] Zhang, Z., Carriero, N. and Gerstein, M. (2004) Comparative analysis of processed pseudogenes in the mouse and human genomes. Trends Genet. 20, 62–67.

[43] Verheggen, C., Lafontaine, D.L., Samarsky, D., Mouaikel, J., Blanchard, J.M., Bordonne, R. and Bertrand, E. (2002) Mammalian and yeast U3 snoRNPs are matured in specific and related nuclear compartments. EMBO J. 21, 2736–2745.
[44] Ganot, P., Jady, B.E., Bortolin, M.L., Darzacq, X. and Kiss, T. (1999) Nucleolar factors direct the 2′-*O*-ribose methylation and pseudouridylation of U6 spliceosomal RNA. Mol. Cell Biol. 19, 6906–6917.
[45] Chen, Y., Sinha, K., Perumal, K., Gu, J. and Reddy, R. (1998) Accurate 3′ end processing and adenylation of human signal recognition particle RNA and alu RNA *in vitro*. J. Biol. Chem. 273, 35023–35031.
[46] Goodier, J.L., Ostertag, E.M., Engleka, K.A., Seleme, M.C. and Kazazian Jr., H.H. (2004) A potential role for the nucleolus in L1 retrotransposition. Hum. Mol. Genet. 13, 1041–1048.
[47] Luo, Y. and Li, S. (2007) Genome-wide analyses of retrogenes derived from the human box H/ACA snoRNAs. Nucleic Acids Res. 35, 559–571.
[48] Weber, M.J. (2006) Mammalian small nucleolar RNAs are mobile genetic elements. PLoS Genet. 2, e205.
[49] Boissinot, S., Entezam, A., Young, L., Munson, P.J. and Furano, A.V. (2004) The insertional history of an active family of L1 retrotransposons in humans. Genome Res. 14, 1221–1231.
[50] Piskareva, O. and Schmatchenko, V. (2006) DNA polymerization by the reverse transcriptase of the human L1 retrotransposon on its own template *in vitro*. FEBS Lett. 580, 661–668.
[51] Belancio, V.P., Hedges, D.J. and Deininger, P. (2006) LINE-1 RNA splicing and influences on mammalian gene expression. Nucleic Acids Res. 34, 1512–1521.
[52] Hiscox, J.A. (2002) The nucleolus-a gateway to viral infection? Arch Virol. 147, 1077–1089.
[53] Garcia-Perez, J.L., Doucet, A.J., Bucheton, A., Moran, J.V. and Gilbert, N. (2007) Distinct mechanisms for *trans*-mediated mobilization of cellular RNAs by the LINE-1 reverse transcriptase. Genome Res., Epub ahead of print.
[54] Gabus, C., Ivanyi-Nagy, R., Depollier, J., Bucheton, A., Pelisson, A. and Darlix, J.L. (2006) Characterization of a nucleocapsid-like region and of two distinct primer tRNALys,2 binding sites in the endogenous retrovirus Gypsy. Nucleic Acids Res. 34, 5764–5777.
[55] Onafuwa-Nuga, A.A., King, S.R. and Telesnitsky, A. (2005) Nonrandom packaging of host RNAs in moloney murine leukemia virus. J. Virol. 79, 13528–13537.
[56] Giles, K.E., Caputi, M. and Beemon, K.L. (2004) Packaging and reverse transcription of snRNAs by retroviruses may generate pseudogenes. RNA 10, 299–307.
[57] Jurka, J., Kapitonov, V.V., Pavlicek, A., Klonowski, P., Kohany, O. and Walichiewicz, J. (2005) Repbase update, a database of eukaryotic repetitive elements. Cytogenet. Genome Res. 110, 462–467.