HUMAN MUTATION 24:9–13 (2004)

RAPID COMMUNICATION

Characterization of a Mutagenic B1 Retrotransposon Insertion in the Jittery Mouse

Nicolas Gilbert,[1]* Jamee M. Bomar,[2] Margit Burmeister,[1,2] and John V. Moran[1,3]

[1]*Department of Human Genetics, University of Michigan Medical School, Ann Arbor, Michigan;* [2]*Mental Health Research Institute, University of Michigan Medical School, Ann Arbor, Michigan;* [3]*Department of Internal Medicine, University of Michigan Medical School, Ann Arbor, Michigan*

Communicated by Haig H. Kazazian

B1 elements are an abundant class of short interspersed elements (SINEs) in the mouse genome and mobilize by a process known as retrotransposition. Here, we report the characterization of a mutagenic B1 insertion into exon 4 of the *Atcay* gene, which was previously shown to be responsible for the jittery mouse. Mutations in the human ortholog of this gene, *ATCAY*, are responsible for Cayman ataxia. The B1 insertion is ~150-bp long, ends in a 45–50-bp polyadenylic acid (poly A) tail, is flanked by a perfect 13-bp target-site duplication, and is inserted into a sequence that resembles a LINE-1 endonuclease consensus cleavage site. Computational analysis indicates that the mutagenic insertion is most closely related to elements of the B1-C subfamily, and we have identified two possible progenitor B1 sequences on mouse chromosome 19. Together, these data demonstrate that B1 retrotransposition is ongoing in the mouse genome and is consistent with the hypothesis that the reverse transcriptase and endonuclease encoded by LINE-1 elements mediate B1 mobility. Hum Mutat 24:9–13, 2004. © **2004 Wiley-Liss, Inc.**

KEY WORDS: retrotransposition; SINE; B1; LINE-1; mutagenic insertion; ATCAY; caytaxin

DATABASES:
ATCAY – OMIM: 601238, 608179; GenBank: AY349151, NT_039361.2, NT_039687.2

INTRODUCTION

Retrotransposons comprise approximately 37.5% of mouse DNA and their mobility continues to impact genome evolution [Waterston et al., 2002]. They can be subdivided into two general classes based on either the presence or absence of long terminal repeats (LTRs). LTR-retrotransposons (e.g., intracisternal A particles (IAPs), Etn elements, and MaLR elements) are similar to retroviruses except that they lack a functional envelope gene. By comparison, non-LTR retrotransposons (e.g., LINE-1 elements, B1, and B2 elements) lack LTRs, end in a polyadenylic acid (poly A) tail, and likely mobilize via a mechanism termed target site primed reverse transcription [Luan et al., 1993]. Representative members of both mobile element classes are active in the mouse genome, as deleterious IAP, Etn, MaLR, and LINE-1 insertions have resulted in a variety of mutations [reviewed in Deininger et al., 2003].

B1s are an abundant class of short interspersed elements (SINEs) that comprise ~2.7% of the mouse DNA and are present at ~550,000 copies per haploid genome [Waterston et al., 2002]. Individual B1 elements are ~150 bp and are similar in sequence to full-length Alu monomers (FLAMs), which are the likely precursors of primate Alu elements [Zietkiewicz and Labuda, 1996]. Like Alu elements, B1 elements are derived ancestrally from the 7SL RNA gene [Ullu and Tschudi, 1984], and probably have amplified by co-opting components of the LINE-1 retrotransposition machinery [Dewannieux et al., 2003].

Here, we report the characterization of a mutagenic B1 insertion into exon 4 of the mouse *Atcay* gene, which was previously shown to be responsible for the jittery mouse [Bomar et al., 2003]. Mutations in the human ortholog of this gene, *ATCAY* (MIM# 601238), are responsible for Cayman ataxia (MIM# 608179) [Bomar et al., 2003]. The B1 insertion contains the hallmarks of being dispersed by retrotransposition and inserted into a sequence that resembles a L1NE-1 endonuclease cleavage site. Further computational analysis showed that the mutagenic insertion is most closely related to elements of the B1-C subfamily, and we have identified two possible progenitors of the insertion on mouse chromosome 19. Thus, these data demonstrate that B1 retrotransposition is ongoing in the mouse genome.

Received 3 February 2004; accepted revised manuscript 19 March 2004.

*Correspondence to: Nicolas Gilbert, Institut de Génétique Humaine, CNRS, UPR 1142, 141 rue de la Cardonille, 34396 Montpellier cedex 5, France. E-mail: Nicolas.Gilbert@igh.cnrs.fr

Grant sponsor: National Institute of Health; Grant number: GM60518, NS32130; Grant Sponsor: March of Dimes.

DOI 10.1002/humu.20060
Published online in Wiley InterScience (www.interscience.wiley.com).

MATERIALS AND METHODS

Sequence Analysis

Isolation of the mutagenic B1 insertion in the *Atcay* gene was described previously [Bomar et al., 2003]. Sequence analyses to identify the possible progenitor of the mutagenic B1 insertion were performed using the nucleotide–nucleotide Blast interface (Blastn; www.ncbi.nlm.nih.gov/BLAST) [Altschul et al., 1997] or the February 2003 version of BLAT (http://genome.ucsc.edu/cgi-bin/hgBlat) [Kent, 2002]. Sequence alignments between the mutagenic insertion and the B1-C consensus sequence were generated using the Blast 2 sequences option (www.ncbi.nlm.nih.gov/blast/bl2seq/bl2.html) [Tatusova and Madden, 1999]. RNA secondary structure analyses were conducted using *mfold* (www.bioinfo.rpi.edu/applications/mfold/) [Zuker, 2003].

RESULTS AND DISCUSSION

Jittery is a spontaneous autosomal recessive mouse mutation that arose around 1935 during the process of maintaining the mouse strain Bagg albino [DeOme, 1945], which later became BALB/c [Beck et al., 2000]. Mice homozygous for the jittery mutation exhibit severe trunk and limb ataxia and die of dehydration and starvation by 3–4 weeks of age [DeOme, 1945]. It recently was determined that jittery is caused by a mutagenic B1 insertion into exon 4 of the *Atcay* gene [Bomar et al., 2003]. Sequence analysis reveals that the mutagenic B1 insertion ends in a 45–50-bp poly (A) tail and is flanked by a 13-bp target site duplication that was generated upon retrotransposition (Fig. 1A). Comparison of the pre-integration (i.e., wild-type) and post-integration (i.e., mutant) sites revealed that the B1 inserted into the sequence 5′-TTTC/G, where (/) indicates the phosphodiester backbone cleavage site. This insertion sequence resembles the consensus LINE-1 endonuclease cleavage site (5′-TTTT/A) [Feng et al., 1996; Gilbert et al., 2002; Morrish et al., 2002; Symer et al., 2002], supporting the hypothesis that B1 mobilizes by hijacking components of the L1 retrotransposition machinery [Jurka and Klonowski, 1996].

Further analysis indicates that the mutagenic B1 insertion has a single guanosine insertion between nucleotides 8 and 9, which is diagnostic for the B1-C subfamily [Quentin, 1989]. All CpG residues in the promoter region are conserved between the mutagenic B1 insertion and the B1-C consensus sequence, suggesting that both the mutagenic insertion and its potential progenitor may be transcriptionally active (Fig. 1B). However, the mutagenic insertion differs from the B1-C consensus sequence because it has an extra adenosine residue at the +1 position of the element and harbors guanosine to adenine transitions at positions 58 and 136, respectively (Fig. 1B). The extra adenosine residue could represent an untemplated nucleotide that was added during retrotransposition [Morrish et al., 2002; Symer et al., 2002] (Gilbert et al., unpublished observations). Alternatively, it could be present in the progenitor B1 that gave rise to the mutagenic insertion.

To identify the possible progenitor of the mutagenic B1 insertion, we used BLAT to screen the mouse genome draft sequence. We identified three sequences with 100% identity to the mutagenic insertion (i.e., excluding the extra adenosine at the +1 position). One is located on chromosome 6 (contig NT_039361.2, position 148178170 to 148178316), and two are located on chromosome 19 (contig NT_039687.2, position 4492805 to 4492952 and 4898913 to 4899059).

The B1 element on chromosome 6 ends in a 13-bp poly (A) tail, is flanked by a 14-bp target site duplication, integrated in a sequence that resembles a consensus L1 endonuclease cleavage site (TCTT/A), and inserted 500 bp upstream of an annotated expressed sequence tag (EST) (AK015591). However, it lacks the extra adenosine residue at the 5′ end of the element, and also lacks a potential RNA polymerase III terminator (5′-TTTTT) within 1 kb of its 3′ end, making it an unlikely progenitor sequence.

FIGURE 1. **Sequence analysis of the mutagenic B1 insertion and its possible progenitors. A: Sequence of the mutant jittery allele. Uppercase lettering indicates exon 4 sequences in the *Atcay* gene (Accession # AY349151). Lowercase lettering indicates the partial sequence of the upstream and downstream introns in the *Atcay* gene. Underlined bold lettering indicates the B1 insertion. The length of the poly (A) tail is estimated to be 45–50 adenosine residues; the size ambiguity likely reflects template slippage that occurred during the PCR amplification and/or DNA sequencing procedure. Blue lettering indicates a 13-bp target site duplication, which was generated upon B1 retrotransposition. Orange lettering indicates a potential untemplated nucleotide at the 5′ end of the B1 sequence. B: Alignment of the mutagenic B1 sequence with the B1-C consensus. Vertical bars indicate nucleotides shared by both the mutagenic B1 insertion and the B1-C consensus sequence. Only two mismatches (highlighted in green) and the untemplated adenosine residue (in orange) are noted. CpG dinucleotides are highlighted in gray. C: Alignment of the two potential B1 progenitors from chromosome 19. The sequences of the two B1 progenitors are shown. The first is located on contig NT_039687 at position 4492754 to 4493031. The second is located on contig NT_039687, at position 4898862 to 4899260. The green box indicates a potential TATA box upstream of the second B1 insertion. The red lettering indicates potential RNA polymerase III terminator sequences downstream of each of the elements.**

FIGURE 2. **Predicted RNA secondary structures of the candidate B1 progenitor transcripts. A: The predicted secondary structure of the chromosome B1 element located on contig NT_039687 at position 4492754 to 4493031. B: The predicted secondary structure of the chromosome B1 element located on contig NT_039687, at position 4898862 to 4899260. C: The predicted secondary structure of a BC1 "master" gene. The blue underlining indicates the region of the B1 RNA sequence that is conserved with the B1-C consensus sequence. The red circles indicate the predicted secondary structure of the RNA polymerase III terminator sequence that flanks the 3′ end of each element. The free energy of the structure (dG) is noted in kcal/mol. Structures were obtained using default parameters on the *mfold* server [Zuker, 2003]. In A and B, the structure that most closely resembles the secondary structure of BC1 RNA is indicated.**

The first B1 element on chromosome 19 (located at position 4492805 to 4492952) ends in a 16-bp poly (A) tail, integrated into the poly (A) of a previously inserted B1 element, and contains a potential RNA polymerase III terminator sequence 69 bp downstream of its poly (A) tail (Fig. 1C). This B1 contains

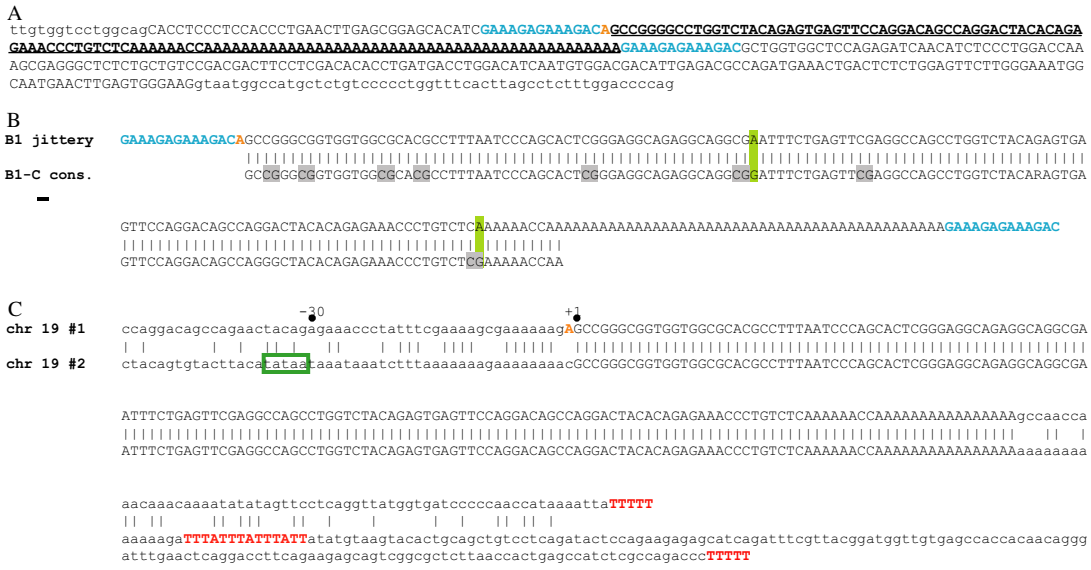

Fig. 1

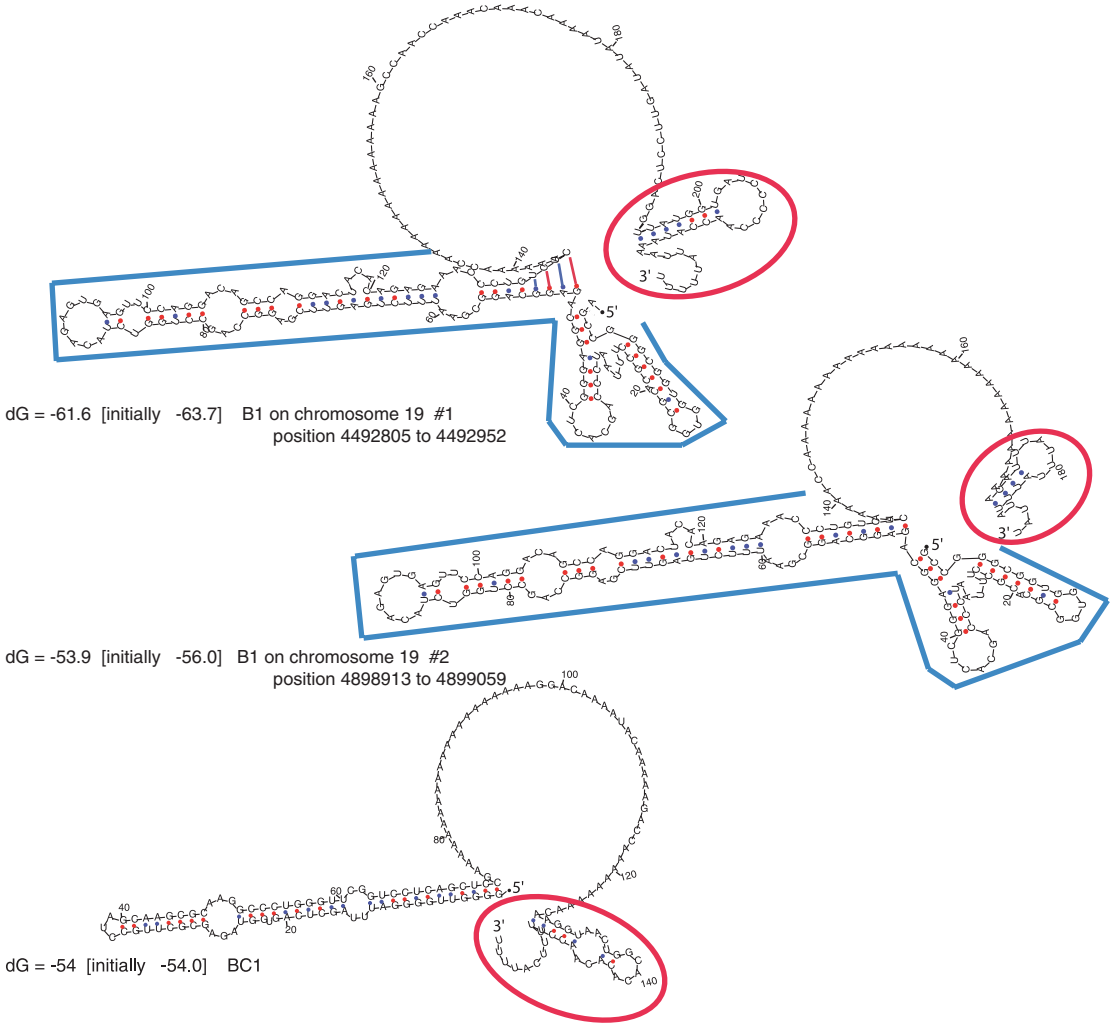

Fig. 2

the adenosine residue at the +1 position and is located in a large intron of the mouse Neosin gene (Accession # X94344).

The second B1 element on chromosome 19 (located at position 4898913 to 4899059) ends in a 29-bp poly (A) tail, integrated into the poly (A) of a previously inserted B2 element, and contains two potential RNA polymerase III terminator sequences. The first is located immediately downstream of the poly (A) tail, whereas the second is located 185 bp downstream of the poly (A) tail (Fig. 1C). This B1 lacks the adenosine residue at the +1 position and is located in a repeat rich region that harbors the sequences of two ESTs in the opposite transcriptional orientation of the B1 (BI134002 and BY791586, respectively). Notably, since both of the B1s on chromosome 19 had inserted into the poly (A) tail of a preexisting SINE, we were unable to unambiguously determine whether they are flanked by target site duplications.

We next used *mfold* to predict the RNA secondary structure of the possible progenitors noted above. In each instance, we selected the structure with the lowest free energy that most closely resembled the predicted secondary structure of B1-C consensus RNA [Labuda et al., 1991] (Fig. 2). We then compared those B1 RNA structures to the structure of an RNA derived from a BC1 progenitor gene [Kim et al., 1994]. This comparison was done because: 1) it was previously demonstrated that this BC1 element is responsible for the expansion of rodent ID SINE elements [Shen et al., 1997]; 2) the unique 3′ flanking sequences of this BC1 element can be folded into a stem loop that may contribute to its expression efficiency and/or RNA stability [Rozhdestvensky et al., 2001; Shen et al., 1997; West et al., 2002]; and 3) BC1 and ID elements also likely exploit the LINE-1 retrotransposition machinery to mediate their mobility [Jurka, 1997]. Interestingly, RNAs derived from both of the possible B1 progenitors can be folded into a similar structure as the progenitor BC1 RNA, providing additional evidence that these B1s could be the actual progenitor of the mutagenic insertion (Fig. 2).

Previous studies have shown that 5′ flanking sequences can also influence the transcription of certain RNA polymerase III transcribed genes (e.g., 7SL, BC1, U6) [Chesnokov and Schmid, 1996; Kobayashi and Anzai, 1998; Ma and Hernandez, 2002; Roy et al., 2000; Ullu and Weiner, 1985]. Indeed, the presence of such sequences may be an important factor in establishing SINE "master" elements [Arnaud et al., 2001; Chesnokov and Schmid, 1996]. Thus, we looked for upstream sequences that could influence the expression of the possible B1 progenitors. Interestingly, we found a putative TATA box at a position −31 to −35 upstream of the second B1 on chromosome 19 (Fig. 1C), which could, in principle, influence the transcription of this B1.

In sum, we have characterized the first mutagenic B1 insertion in the mouse genome. These data indicate that some variant members of the B1-C subfamily remain retrotransposition competent, and suggest that B1 elements present on chromosome 19 may, in fact, represent master genes. Our data also are consistent with the hypothesis that components of the LINE-1 retrotransposition machinery mediate the mobility of B1 elements [Jurka, 1997]. The recent development of a system to analyze SINE retrotransposition in cultured cells [Dewannieux et al., 2003] coupled with the identification of these potential B1 master genes now allows an opportunity to test this hypothesis experimentally.

ACKNOWLEDGMENTS

We thank Dr. David Kass for helpful discussion and Dr. Mark Batzer for critically reading the manuscript. This work was supported by grants from the National Institutes of Health (GM60518 (J.V.M.) and NS32130 (M.B.)) and a March of Dimes grant to M.B.

REFERENCES

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–402.

Arnaud P, Yukawa Y, Lavie L, Pelissier T, Sugiura M, Deragon JM. 2001. Analysis of the SINE S1 Pol III promoter from Brassica; impact of methylation and influence of external sequences. Plant J 26:295–305.

Beck JA, Lloyd S, Hafezparast M, Lennon-Pierce M, Eppig JT, Festing MF, Fisher EM. 2000. Genealogies of mouse inbred strains. Nat Genet 24:23–25.

Bomar JM, Benke PJ, Slattery EL, Puttagunta R, Taylor LP, Seong E, Nystuen A, Chen W, Albin RL, Patel PD, Kittles RA, Sheffield VC, Burmeister M. 2003. Mutations in a novel gene encoding a CRAL-TRIO domain cause human Cayman ataxia and ataxia/dystonia in the jittery mouse. Nat Genet 35:264–269.

Chesnokov I, Schmid CW. 1996. Flanking sequences of an Alu source stimulate transcription in vitro by interacting with sequence-specific transcription factors. J Mol Evol 42: 30–36.

Deininger PL, Moran JV, Batzer MA, Kazazian HH Jr. 2003. Mobile elements and mammalian genome evolution. Curr Opin Genet Dev 13:651–658.

DeOme KB. 1945. A new recessive lethal mutation in mice. Univ. Calif. Publ. Zool. 53:41–55.

Dewannieux M, Esnault C, Heidmann T. 2003. LINE-mediated retrotransposition of marked Alu sequences. Nat Genet 35: 41–48.

Feng Q, Moran JV, Kazazian HH Jr, Boeke JD. 1996. Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. Cell 87:905–916.

Gilbert N, Lutz-Prigge S, Moran JV. 2002. Genomic deletions created upon LINE-1 retrotransposition. Cell 110:315–325.

Jurka J, Klonowski P. 1996. Integration of retroposable elements in mammals: selection of target sites. J Mol Evol 43:685–689.

Jurka J. 1997. Sequence patterns indicate an enzymatic involvement in integration of mammalian retroposons. Proc Natl Acad Sci USA 94:1872–1877.

Kent WJ. 2002. BLAT—the BLAST-like alignment tool. Genome Res 12:656–664.

Kim J, Martignetti JA, Shen MR, Brosius J, Deininger P. 1994. Rodent BC1 RNA gene as a master gene for ID element amplification. Proc Natl Acad Sci USA 91:3607–3611.

Kobayashi S, Anzai K. 1998. An E-box sequence acts as a transcriptional activator for BC1 RNA expression by RNA polymerase III in the brain. Biochem Biophys Res Commun 245:59–63.

Labuda D, Sinnett D, Richer C, Deragon JM, Striker G. 1991. Evolution of mouse B1 repeats: 7SL RNA folding pattern conserved. J Mol Evol 32:405–414.

Luan DD, Korman MH, Jakubczak JL, Eickbush TH. 1993. Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. Cell 72:595–605.

Ma B, Hernandez N. 2002. Redundant cooperative interactions for assembly of a human U6 transcription initiation complex. Mol Cell Biol 22:8067–8078.

Morrish TA, Gilbert N, Myers JS, Vincent BJ, Stamato TD, Taccioli GE, Batzer MA, Moran JV. 2002. DNA repair mediated by endonuclease-independent LINE-1 retrotransposition. Nat Genet 31:159–165.

Quentin Y. 1989. Successive waves of fixation of B1 variants in rodent lineage history. J Mol Evol 28:299–305.

Roy AM, West NC, Rao A, Adhikari P, Aleman C, Barnes AP, Deininger PL. 2000. Upstream flanking sequences and transcription of SINEs. J Mol Biol 302:17–25.

Rozhdestvensky TS, Kopylov AM, Brosius J, Huttenhofer A. 2001. Neuronal BC1 RNA structure: evolutionary conversion of a tRNA(Ala) domain into an extended stem-loop structure. RNA 7:722–730.

Shen MR, Brosius J, Deininger PL. 1997. BC1 RNA, the transcript from a master gene for ID element amplification, is able to prime its own reverse transcription. Nucleic Acids Res 25: 1641–1648.

Symer DE, Connelly C, Szak ST, Caputo EM, Cost GJ, Parmigiani G, Boeke JD. 2002. Human l1 retrotransposition is associated with genetic instability in vivo. Cell 110:327–338.

Tatusova TA, Madden TL. 1999. BLAST 2 Sequences, a new tool for comparing protein and nucleotide sequences. FEMS Microbiol Lett 174:247–250.

Ullu E, Tschudi C. 1984. Alu sequences are processed 7SL RNA genes. Nature 312:171–172.

Ullu E, Weiner AM. 1985. Upstream sequences modulate the internal promoter of the human 7SL RNA gene. Nature 318:371–374.

Waterston RH, Lindblad-Toh K, Birney E, Rogers J, Abril JF, Agarwal P, Agarwala R, Ainscough R, Alexandersson M, An P, et al. 2002. Initial sequencing and comparative analysis of the mouse genome. Nature 420:520–562.

West N, Roy-Engel AM, Imataka H, Sonenberg N, Deininger P. 2002. Shared protein components of SINE RNPs. J Mol Biol 321:423–432.

Zietkiewicz E, Labuda D. 1996. Mosaic evolution of rodent B1 elements. J Mol Evol 42:66–72.

Zuker M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res 31:3406–3415.