MOLECULAR AND CELLULAR BIOLOGY, Sept. 1991, p. 4804–4807
0270-7306/91/094804-04$02.00/0

Vol. 11, No. 9

NOTES

Ribonucleoprotein Particles with LINE-1 RNA in Mouse Embryonal Carcinoma Cells

SANDRA L. MARTIN

Department of Cellular and Structural Biology and Program in Molecular Biology, University of Colorado Health Sciences Center, 4200 East Ninth Avenue, Denver, Colorado 80262

Received 15 January 1991/Accepted 4 June 1991

The LINE-1 repeat family is interspersed throughout mammalian genomes and is thought to be the result of duplicative transposition of LINE-1 sequences via an RNA intermediate. This report describes a ribonucleoprotein particle with LINE-1 RNA in the mouse embryonal carcinoma cell line F9. This ribonucleoprotein particle is a potential intermediate in the transposition of LINE-1 in the mouse genome.

Eucaryotic genomes are littered with repetitive DNA sequences of unknown function. One family of repetitive DNA, LINE-1 (L1), is interspersed throughout all mammalian genomes that have been examined (1). Genomic copies of L1 are characterized by several structural features that suggest that individual members of the family are the consequence of transposition events involving an RNA intermediate (13, 20, 21). In mouse DNA, the longest known elements of L1 are about 7 kb, end with a polyadenylation signal followed by a run of A-rich sequences, and are bounded by short direct repeats. There are two open reading frames (ORFs) that overlap by 14 bp; the sequence of the second ORF is related to those of known reverse transcriptases (11). Although some of these features are reminiscent of retroviruses (4), L1 differs in a major attribute that is critical for retrovirus replication, namely, no long terminal repeat-like structure has ever been identified. Elements similar to L1 are found in a wide variety of species, ranging from animals to plants to protists (22), and are known as the non-long terminal repeat, or class II, retrotransposons (17).

Most (>90%) of the ~100,000 copies of L1 DNA in the mouse genome are truncated and therefore do not contain both ORFs. The structure of L1, coupled with the observation that the truncated copies of L1 accumulate substitutions that lead to loss of the ORFs, suggests that the L1 family in mice is composed of a small number of functional transposable elements that are continuously giving rise to a large number of truncated pseudogenes via the process of retrotransposition (6). Evidence that L1 transposition is a continuous genetic process, and not merely an evolutionary relic, is provided most convincingly by independent germ line insertions of L1 into the human factor VIII gene, leading to hemophilia (9).

A model for L1 transposition, based on the DNA sequences of isolated genomic copies, proposes that the first step in mobilization of L1 is specific transcription of full-length functional elements. The L1 ORFs are then translated into proteins, which in turn bind the RNA (12), leading ultimately to the integration of a new L1 cDNA into the genome. Such a model predicts that a ribonucleoprotein particle (RNP) containing full-length, sense strand L1 RNA and L1-encoded protein will be an intermediate in the transposition of L1 (3, 8). This report describes the association of a 7.5-kb sense strand L1 RNA with L1-encoded protein. This RNP is a potential intermediate in the transposition of L1 in mammalian genomes.

In order to identify possible biochemical intermediates in the transposition of L1, extracts from 10^7 F9 cells (mouse embryonal carcinoma cells) were fractionated by sucrose density gradient centrifugation in the presence and absence of EDTA. F9 cells were obtained from B. Pierce (Pathology Department, University of Colorado Health Sciences Center) and maintained in Dulbecco's modified Eagle medium (GIBCO) with 10% fetal bovine serum (HyClone). Cells were harvested by scraping, washed in phosphate-buffered saline, and then lysed in 5 volumes of 140 mM NaCl–200 mM Tris-HCl (pH 8.5)–2 mM $MgCl_2$–0.25% Nonidet P-40 for 5 min on ice. After a 10-min centrifugation at 200 × *g*, the supernatant was recovered and adjusted to 10 μg of heparin per ml and 10 mM EDTA (as appropriate). This extract was overlaid onto 11 ml of 10 to 50% linear sucrose gradients in 25 mM Tris-HCl (pH 7.6)–40 mM KCl, with either 10 mM EDTA (TKE) or 7.5 mM $MgCl_2$ (TKM), on top of a 1-ml 60% sucrose pad in SW41 (Beckman) tubes. Centrifugation was for 2.5 h at 178,224 × *g* at 4°C. Fractions (1 ml) were collected from the top with an ISCO tube piercer and UA5 flow cell. For Northern (RNA) blots (Fig. 1a and b), a 400-μl aliquot from each fraction was extracted with phenol and then precipitated with ethanol. The RNA pellet was resuspended in denaturing sample buffer and resolved on formaldehyde-agarose (1%) gels (15). rRNA was detected after staining with ethidium bromide or by UV shadowing, and then RNA was blotted to nitrocellulose. Actin and L1 RNAs were detected after hybridization (15) to ^{32}P-labelled probes and autoradiography.

As shown in Fig. 1, the sedimentation behavior of L1 RNA in these sucrose gradients is unaffected by EDTA. In contrast, rRNA and polyribosome-associated mRNAs (such as actin) shifted to lighter fractions when polyribosomes were disrupted by the addition of 10 mM EDTA. This result indicates that the majority of L1 RNA in F9 cells is present in complexes distinct from polyribosomes.

The nature of this L1-containing complex was explored by treating a portion of material from fraction 8 of the EDTA-containing gradient with proteinase K and then analyzing its behavior in a second sucrose gradient. A 400-μl volume from fraction 8 of the EDTA-containing gradient (Fig. 1b) was diluted to 1 ml with ice-cold 1 mM EDTA, and 470 μl was

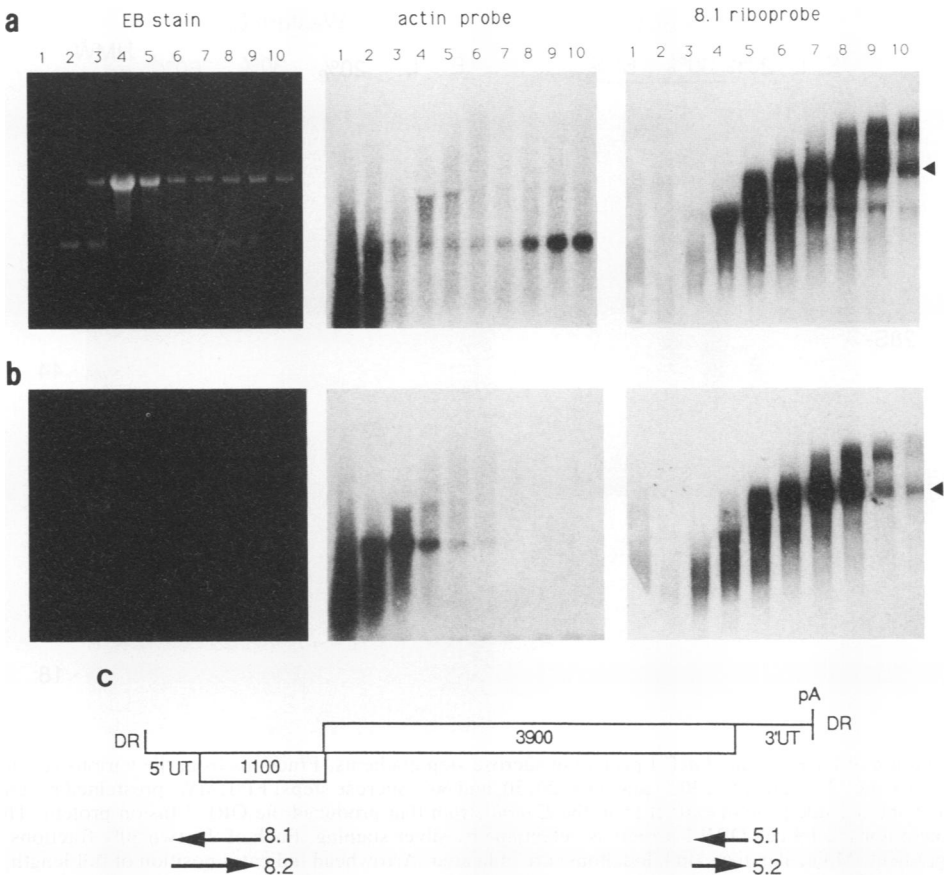

FIG. 1. Sucrose gradient sedimentation of F9 cell extracts. Parallel sucrose gradients were run in the absence (a) or presence (b) of 10 mM EDTA. RNA prepared from an equal volume from each fraction across the gradient was loaded on the gel (fraction 1 is the top). The 18 and 28S rRNAs were visualized by staining with ethidium bromide (EB), the 2-kb actin mRNA was visualized by hybridization to a random-primed (Boehringer Mannheim) ^{32}P-labelled actin fragment (5), and L1 RNA was visualized by hybridization to a ^{32}P-labelled 8.1 riboprobe, followed by autoradiography. The position of the 7.5-kb L1 RNA is indicated (◀). In this exposure, other species of RNA that hybridize to the 8.1 riboprobe are apparent. These are probably due, at least in part, to fortuitous transcription events, as described previously (3, 17, 18). Smaller species could also arise by degradation. (c) Structure of a typical long L1 in the mouse genome (11). The two ORFs are indicated by their size in nucleotides. UT, untranslated region; pA, the polyadenylation signal followed by an A-rich stretch; DR, target site duplication. Fragments used for riboprobe synthesis (by SP6 in vitro transcription; Boehringer Mannheim) in this study are indicated; the arrows show the direction of transcription. 8.1 and 8.2 are pSP64 clones of an 800-bp *Bam*HI fragment from the 5′ end of L1Md-A9 (16), and 5.1 and 5.2 are pSP64 clones of a 500-bp *Bam*HI fragment from the 3′ end of L1Md-A4 (12).

removed and treated for 30 min on ice with 100 μg of proteinase K (Boehringer Mannheim). The remainder was held in ice without proteinase K. Treated and untreated material was fractionated through a linear sucrose gradient, and 1-ml fractions were collected as described above. L1 RNA in a 200-μl aliquot from each fraction was detected after application to nitrocellulose by using a slot blot apparatus (Bio-Rad), hybridization to the L1 riboprobe (8.1 [Fig. 1c]), and autoradiography; direct quantification of L1 hybridization was done with a Bioscan 200.

As shown in Fig. 2, L1 RNA was shifted towards the top of the gradients after treatment with proteinase K. This result demonstrates that the faster-sedimenting complex is formed by the association of L1 RNA with protein. A similar result is obtained when the F9 cell extracts are extracted with phenol prior to fractionation in the sucrose gradients (data not shown).

To further characterize the composition of these L1 RNPs and their structure by methods that are less sensitive than Northern blot analysis, a step gradient fractionation procedure was developed. Step gradients were made in SW41 tubes with 5 ml of 20% sucrose, 1.2 ml of 30% sucrose, and

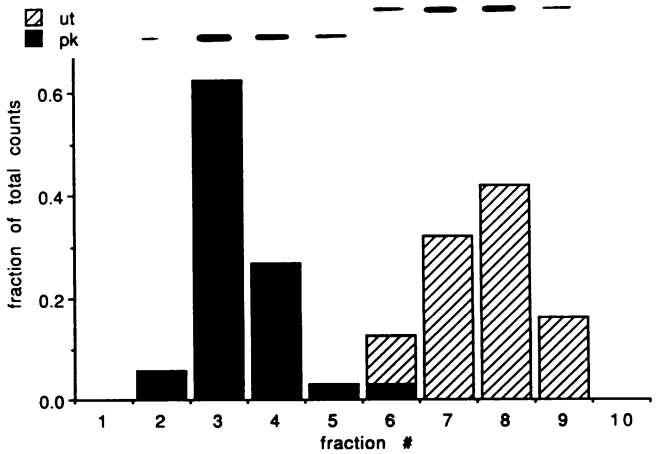

FIG. 2. Sucrose gradient sedimentation of the L1 particle after treatment with proteinase K. The histogram shows the relative number of counts (8.1 riboprobe) hybridizing to the RNA in each fraction from the two gradients. ut, untreated; pk, proteinase K. The corresponding autoradiogram is shown at the top.

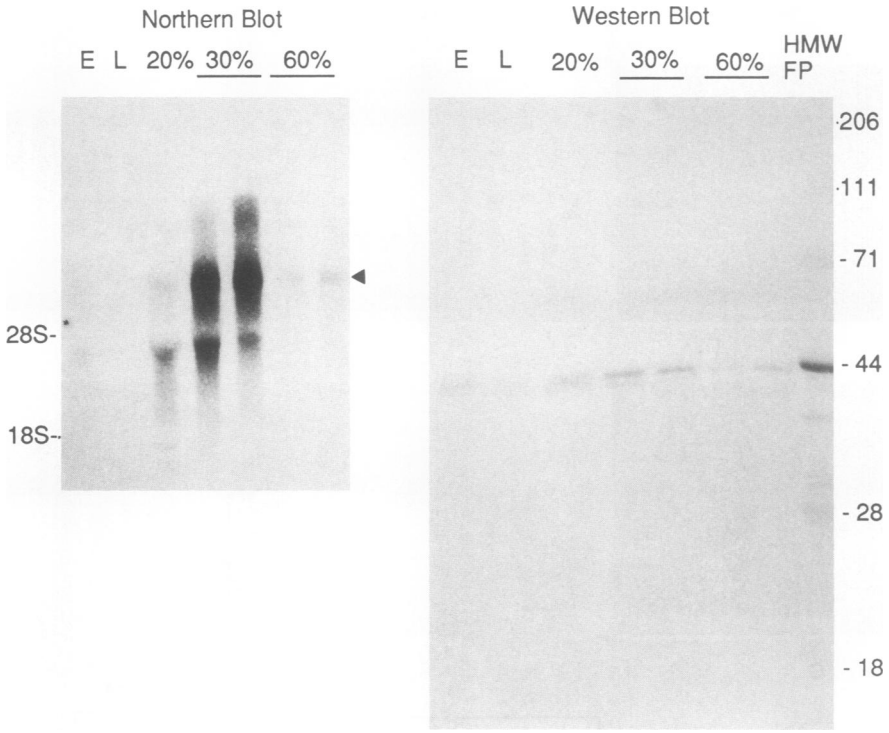

FIG. 3. Cosedimentation of L1 RNA and ORF 1 protein in sucrose step gradients. Fractions from the various regions of the gradient are indicated. Lanes: E, F9 extract; L, load; 20%, 30%, and 60%, 20, 30, and 60% sucrose steps; FP/HMW, prestained protein markers (Bethesda Research Laboratories) and a crude protein extract from the *E. coli* strain that produced the ORF 1 fusion protein. This amount of neither the ORF 1 fusion protein nor the F9 cell ORF 1 protein is detectable by silver staining. Each of the two 30% fractions contains about 1.5% of the protein in the gradient. Molecular sizes (in kilodaltons) are indicated. Arrowhead indicates position of full-length L1 RNA as detected by hybridization to the 8.1 riboprobe.

1.5 ml of 60% sucrose in TKE. These gradients were overlaid with extract from 8.3×10^7 cells. Centrifugation was as described above, but gradient collection was adjusted to recover the load material (~4 ml) and the 20% sucrose region, each in a single tube, and then the remainder as 0.5-ml fractions. The protein concentration of each fraction was determined (modified Bradford assay; Bio-Rad). A volume containing 5 μg of protein was either analyzed by Northern blotting with the 8.1 riboprobe as described above or fractionated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis in 10% acrylamide. The proteins were transferred to nitrocellulose by electrophoresis (Novablot; LKB), and L1 ORF 1 protein was detected by using an affinity-purified anti-ORF 1 antibody, followed by an alkaline phosphatase-conjugated goat anti-rabbit immunoglobulin G (ProtoBlot system; Promega). ORF 1 antiserum was prepared by immunization of rabbits with gel-purified (7) ORF 1 fusion protein (364 amino acids from the *Nhe*I site at position 1720 in L1Md-A2 [11] behind 13 amino acids from pET3b and 12 amino acids from a multiple-cloning-site polylinker) made in *Escherichia coli* under the control of a T7 promoter (19). The anti-ORF 1 antibody was affinity purified on an ORF 1 column (AminoLink; Pierce). Preimmune serum was purified by Affi-Gel Blue (Bio-Rad) chromatography.

In step gradients, cofractionation of L1 RNA with L1-encoded ORF 1 protein is observed (Fig. 3). The main protein detected with the anti-ORF 1 antibody on Western blots (immunoblots) is enriched in the 30% sucrose step, as is L1 RNA. This immunoreactive protein has an apparent molecular size of 42 kDa, in good agreement with the predicted molecular weight of 41,226 for the ORF 1 protein. No proteins were detected in F9 extracts by using the purified immunoglobulin fraction from preimmune serum. In addition to the major component, two to three proteins with a similar molecular size were detected by the antibody but were more evenly distributed across the gradient. These variants of the ORF 1 polypeptide may be the result of protein degradation in the extracts, translation of multiple polymorphic alleles of L1, or posttranslational processing of the ORF 1 polypeptide. It is possible that only one of the many subfamilies of mouse L1 encodes an ORF 1 protein that associates with L1 RNA or that some type of posttranslational modification is required before the ORF 1 protein can bind L1 RNA to form the RNP. Any of these possibilities could explain the observation that only one species of ORF 1 polypeptide appears to coenrich with L1 RNA.

L1 RNP from a sucrose step gradient was run a second time through a linear sucrose gradient. L1-containing fractions were identified by Northern blot analysis, and then RNA from the peak fraction was further characterized. Figure 4 shows the results of hybridization of this RNP RNA with sense and antisense probes prepared from 5′ and 3′ fragments of L1. These results demonstrate that the L1 RNA contained in the RNP is sense strand specific and contains 5′ and 3′ ORF regions as typified by previously characterized L1 (11).

This report describes the association of a 7.5-kb, sense strand L1 RNA with protein to form RNPs. Whether this L1 RNP is an intermediate in the transposition of L1 remains to be proven; however, this particle exhibits an essential characteristic of such an intermediate, i.e., the apparent association of full-length, intact, L1 sense strand RNA with ORF 1-encoded protein. In addition, the expression of this L1 RNP in undifferentiated F9 cells is consistent with the

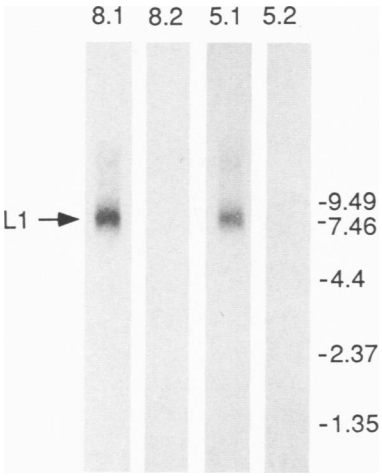

FIG. 4. Northern blot analysis of L1 RNP RNA. RNA was isolated from the peak fraction of a secondary linear gradient, fractionated in agarose-formaldehyde gels, and hybridized to single-stranded probes (as indicated above the lanes) from the 5′ and 3′ ends of L1 (Fig. 1c). The migration of RNA size markers (Bethesda Research Laboratories) is indicated. The presence of full-length, sense strand L1 RNA has been reported previously for poly(A)$^+$ fractions prepared from a human teratocarcinoma cell line (18) as well as a number of murine lymphoid cells (3).

evolutionary data that require L1 transposition to occur during gametogenesis or early embryogenesis (14, 18).

In addition to an association of full-length, sense strand L1 RNA with ORF 1 protein, it is possible that the L1 RNP also contains ORF 2-encoded polypeptides, including reverse transcriptase. Efforts to detect ORF 2 gene products in these L1 RNP preparations, either by Western analysis or by reverse transcriptase activity assays, have been unsuccessful to date, probably because of insufficient material. Deragon et al. (2) have reported fractionation of a macromolecular complex containing reverse transcriptase activity from a human embryonal carcinoma cell line, NTera2D1. Apparent cofractionation of L1 RNA is detected by polymerase chain reaction amplification of nucleic acids present in the peak fractions of reverse transcriptase. Interestingly, silver staining revealed a protein of 37 kDa, which is similar in size to the human L1 ORF 1 protein detected in the same cell line by Liebold et al. (10). Understanding of the relationship between the L1 RNP reported here, the reverse transcriptase-containing particle from NTera2D1 cells, and the intermediates involved in transposition of L1 awaits further study and will be greatly aided by the development of an in vitro system, or a heterologous in vivo system, that is capable of supporting the replication and integration steps of L1 retrotransposition.

I thank S. Nordeen for helpful discussion, A. Franzusoff for critical review of the manuscript, and D. Branciforte for affinity-purified anti-ORF 1 antibody.

This work was supported by Public Health Service grant GM-40367 from the National Institute of General Medical Sciences.

REFERENCES

1. **Burton, F. H., D. D. Loeb, C. F. Voliva, S. L. Martin, M. H. Edgell, and C. A. Hutchison III.** 1986. Conservation throughout Mammalia and extensive protein-encoding capacity of the highly repeated DNA long interspersed sequence 1. J. Mol. Biol. **187**:291–304.
2. **Deragon, J. M., D. Sinnett, and D. LaBuda.** 1990. Reverse transcriptase activity from human embryonal carcinoma cells. EMBO J. **9**:3363–3368.
3. **Dudley, J. P.** 1987. Discrete high molecular weight RNA transcribed from the long interspersed repetitive element L1Md. Nucleic Acids Res. **15**:2581–2592.
4. **Fanning, T., and M. Singer.** 1987. The LINE-1 DNA sequences in four mammalian orders predict proteins that conserve homologies to retrovirus proteins. Nucleic Acids Res. **15**:2251–2260.
5. **Gunning, P., P. Ponte, H. Okayama, J. Engel, H. Blau, and L. Kedes.** 1983. Isolation and characterization of full-length cDNA clones for human α-, β-, and γ-actin mRNAs: skeletal but not cytoplasmic actins have an amino-terminal cysteine that is subsequently removed. Mol. Cell. Biol. **3**:787–795.
6. **Hardies, S. C., S. L. Martin, C. F. Voliva, C. A. Hutchison III, and M. H. Edgell.** 1986. An analysis of replacement and synonymous changes in the rodent L1 repeat family. Mol. Biol. Evol. **3**:109–125.
7. **Harlow, E., and D. Lane.** 1988. Antibodies. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
8. **Hutchison, C. A., III, S. C. Hardies, D. D. Loeb, W. R. Shehee, and M. H. Edgell.** 1989. LINEs and related retroposons: long interspersed repeated sequences in the eucaryotic genome, p. 593–617. *In* D. E. Berg and M. M. Howe (ed.), Mobile DNA. American Society for Microbiology, Washington, D.C.
9. **Kazazian, H. H., Jr., C. Wong, H. Youssoufian, A. F. Scott, D. G. Phillips, and S. E. Antonarakis.** 1988. Haemophilia A resulting from *de novo* insertion of L1 sequences represents a novel mechanism for mutation in man. Nature (London) **332**: 164–166.
10. **Liebold, D. M., G. D. Swergold, M. F. Singer, R. E. Thayer, B. A. Dombroski, and T. G. Fanning.** 1990. Translation of LINE-1 elements *in vitro* and in human cells. Proc. Natl. Acad. Sci. USA **87**:6990–6994.
11. **Loeb, D. D., R. W. Padgett, S. C. Hardies, W. R. Shehee, M. B. Comer, M. H. Edgell, and C. A. Hutchison III.** 1986. The sequence of a large L1Md element reveals a tandemly repeated 5′ end and several features found in retrotransposons. Mol. Cell. Biol. **6**:168–182.
12. **Martin, S. L., C. F. Voliva, F. H. Burton, M. H. Edgell, and C. A. Hutchison III.** 1984. A large interspersed repeat found in mouse DNA contains a long open reading frame that evolves as if it encodes a protein. Proc. Natl. Acad. Sci. USA **81**:2308–2312.
13. **Rogers, J.** 1983. A straight line story. Nature (London) **306**:113–114.
14. **Rogers, J.** 1986. The origin of retroposons. Nature (London) **319**:725.
15. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
16. **Shehee, W. R., S. F. Chao, D. D. Loeb, M. B. Comer, C. A. Hutchison III, and M. H. Edgell.** 1987. Determination of a functional ancestral sequence and definition of the 5′ end of A-type mouse L1 elements. J. Mol. Biol. **196**:757–767.
17. **Skowronski, J., T. G. Fanning, and M. F. Singer.** 1988. Unit-length Line-1 transcripts in human teratocarinoma cells. Mol. Cell. Biol. **8**:1385–1397.
18. **Skowronski, J., and M. F. Singer.** 1985. Expression of a cytoplasmic LINE-1 transcript is regulated in a human teratocarcinoma cell line. Proc. Natl. Acad. Sci. USA **82**:6050–6054.
19. **Studier, F. W., A. H. Rosenberg, J. J. Dunn, and J. W. Dubendorff.** 1990. Use of T7 RNA polymerase to direct expression of cloned genes. Methods Enzymol. **185**:60–89.
20. **Thayer, R. E., and M. F. Singer.** 1983. Interruption of an α-satellite array by a short member of the *Kpn*I family of interspersed, highly repeated monkey DNA sequences. Mol. Cell. Biol. **3**:967–973.
21. **Voliva, C. F., C. L. Jahn, M. B. Comer, C. A. Hutchison III, and M. H. Edgell.** 1983. The L1Md long interspersed repeat family in the mouse: almost all exmaples are truncated at one end. Nucleic Acids Res. **11**:8847–8859.
22. **Xiong, Y., and T. H. Eickbush.** 1990. Origin and evolution of retroelements based upon their reverse transcriptase sequences. EMBO J. **9**:3353–3362.