Gene, 135 (1993) 183–188

GENE 07420

LINE-1: a human transposable element*

(Retrotransposition; non-LTR-retrotransposon; leucine zipper; transcription; reverse transcriptase; teratocarcinoma cells)

Maxine F. Singer, Veronica Krek**, Julie P. McMillan, Gary D. Swergold** and Ronald E. Thayer**

Laboratory of Biochemistry, National Cancer Institute, Building 37, Room 4A-01, National Institutes of Health, Bethesda, MD 20892, USA

Received by G. Bernardi: 19 May 1993; Accepted: 4 June 1993; Received at publishers: 16 July 1993

SUMMARY

Among the 10^5 *LINE-1* sequences (*L1Hs*) in the human genome are one or more 6-kb segments that are active retrotransposons. Expression of these retrotransposons appears to be favored in cells of germ line origin, as well as in some other tumor cells of epithelial origin. In such cells, the product of the first *L1Hs* open reading frame (ORF), a protein called p40, is detectable; p40 has no apparent similarity to gag proteins, but contains a leucine zipper region which may be responsible for the occurrence of p40 multimers. Transcription of *L1Hs* initiates at residue 1 although the transcriptional regulatory regions are downstream in the first 670 bp of the 5′ untranslated region; deletion of a YY1-binding site in the first 20 bp reduces transcription by fivefold. Translation of the second ORF, which encodes reverse transcriptase, is independent of the translation of the frame encoding p40.

INTRODUCTION

The human genome, like others, is subject to random alterations by the movement of transposable elements residing at certain genomic loci into new positions. Thus far, one such element has been shown to be actively transposable in humans, the retrotransposon *LINE-1* (or *L1Hs*). Another type of DNA segment, the so-called *Alu* sequence, also inserts in new positions in human chromosomes, but *Alu* segments do not appear to encode enzymes or other proteins that are expected to be required for their own transposition.

Several mutant alleles of human genes that owe their loss of function to the insertion of an *L1Hs* element have been reported. Among these are a few which represent new insertions, that is, they were not present in the parental chromosomes (Table I). Several of these insertions are likely to have occurred during meiosis or in early embryonic development as they appear in most if not all cells (Kazazian et al., 1988). In each such case so far identified, the insertion is in an X-chromosome gene, as might be expected. New transpositions are also possible in adult somatic cells as indicated by an insertion into an exon of the *APC* gene associated with familial adenomatous polyposis coli in the cells of a colon tumor but not in the corresponding gene in surrounding normal cells of the

Correspondence to: Dr. M.F. Singer, Laboratory of Biochemistry, National Cancer Institute, Building 37, Room 4A-01, National Institutes of Health, Bethesda, MD 20892, USA. Tel. (1-301) 496-1581, Fax (1-301) 402-3095.

*Presented at the COGENE Symposium, 'From the Double Helix to the Human Genome: 40 Years of Molecular Genetics', UNESCO, Paris, 21–23 April 1993.

**Present addresses: (V.K.) 5405 Wilson Lane, Bethesda, MD 20814, USA; (G.D.S.) Food and Drug Administration, Division of Cell and Gene Therapy Regulation, 1401 Rockville Pike, Rockville, MD 20852-1445, USA; (R.E.T.) Bldg. 10, Room 11B13, NIH, Bethesda, MD 20892, USA.

Abbreviations: aa, amino acid(s); bp, base pair(s); *Hs*, *Homo sapiens*; kb, kilobase(s) or 1000 bp; *L1Hs*, *L1Md*, *L1Rn*, *LINE-1* DNA sequences in the human, mouse and rat genomes, respectively; *LRE*, L1 retrotransposable element; *Md*, *Mus domestica*; nt, nucleotide(s); oligo, oligodeoxyribonucleotide; ORF, open reading frame; p40, protein encoded by ORF1 in *L1Hs*; PAGE, polyacrylamide-gel electrophoresis; *Rn*, *Rattus norvegicus*; SDS, sodium dodecyl sulfate; *tsp*, transcription start point; *UTR*, untranslated region; X, any aa.

TABLE I

Recent *L1Hs* transpositions

L1Hs	Gene encoding	Size (kb)	Change vs. *L1.2B*[a] (bp)	Type
JH-27[b]	Factor VIII	3.8	0	germline
JH-28[b]	Factor VIII	2.2	16	germline
	APC[c]	0.538	4	somatic

[a] *L1.2B* is the source of the JH-27 *L1Hs* insertion (Dombroski et al., 1991). The number of bp differences between *L1.2B* and the indicated *L1Hs* is stated.

[b] Kazazian et al. (1988); Dombroski et al. (1991).

[c] Adenomatous polyposis coli (Miki et al., 1992).

affected individual (Miki et al., 1992). We do not now have any information on the frequency of *L1Hs* transposition in any human cell type, nor indeed in that of any mammal, all of which carry species-specific versions of *LINE-1* elements. On evolutionary time scales, it has been estimated that in *Mus domesticus*, half of the approximately 10^5 *L1Md* sequences currently fixed (found) in the genome were placed there within the last 3 million years (Hutchison et al., 1989).

LINE-1 elements fall into the class of non-LTR retrotransposons (also termed poly(A)$^+$ retrotransposons) (Boeke and Corces, 1989). Thus, the 5′ and 3′ ends of the elements carry no repeats, direct or indirect, although elements are generally surrounded by duplications of the target sites into which they transposed (Fig. 1). The consensus *L1Hs* sequence has two ORFs on one strand, ORF1 and ORF2, separated by a short inter-ORF region containing multiple stop codons (Scott et al., 1987; Skowronski et al., 1988). ORF1 and ORF2 are in the same frame. A 5′ *UTR* of about 900 bp precedes ORF1 and a 200-bp 3′ *UTR* follows ORF2. A variable length, A-rich stretch follows the 3′ *UTR* on the coding strand. Only about 4000 full length *L1Hs* elements occur in the human genome, the other approximately 10^5 being truncated and/or rearranged, to varying extents, usually at

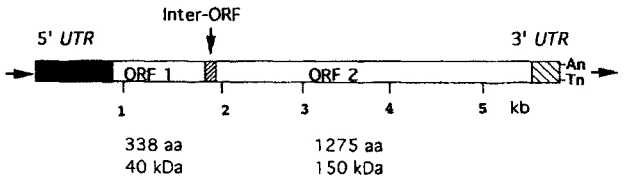

Fig. 1. Schematic diagram of a full length (6-kb) *L1Hs* element. Shown with different markings are the 5′ untranslated region (5′ *UTR*), the two ORFs, ORF1 and ORF2, the inter-ORF region that contains multiple, in-frame stop codons, the 3′ untranslated region (3′ *UTR*), and the A-rich stretch on the coding strand (upper strand). The numbers along the bottom indicate kb. The small arrows at either end indicate the duplications of the target site into which the element transposed. Beneath is shown the size of the polypeptide predicted by each ORF in number of aa and the corresponding calculated M_r.

the 5′ end (Adams et al., 1980; Scott et al., 1987). Of the 4000, many have ORFs that are closed by single bp changes.

Current models propose that a subset of the full length *L1Hs* elements with open ORFs are actively transposable elements, capable of being transcribed, translated to produce proteins including reverse transcriptase, and reverse transcribed to provide the DNA found in new target sites. It is possible that, like the LTR retrotransposons such as the Ty elements of yeast, the products of *L1* transcription and translation are associated in intracellular particles. Deragon et al. (1990) presented evidence for a high molecular weight complex containing reverse transcriptase activity and *L1Hs* RNA in human teratocarcinoma cells. Martin (1991) described ribonucleoprotein particles containing *L1Md* RNA and the protein encoded by the first open reading frame of *L1Md* in mouse embryonal carcinoma cells.

We have been interested in studying how and under what conditions *L1Hs* elements are transcribed and translated and the nature of the encoded proteins as a means of understanding the transposition mechanism and its control. In recent years, this work has been aided by the isolation of *L1Hs* elements that appear to be active. Dombroski and her colleagues (1991; 1993), using a distinctive oligo segment in the ORF2 region of an *L1Hs* newly inserted into the factor VIII-encoding gene in a hemophiliac boy, detected a subset of between four and 10 *L1Hs* elements that might have been a source for the transposition. Among the genomic members of this subset are a group of alleles at the *LRE-1* locus on chromosome 22q that contain full length *L1Hs* elements. An *LRE-1* allele cloned from the genome of the mother of the patient (JH27 in Table I) was identical in sequence to the newly transposed element, indicating that it was the likely origin of the transposed element; randomly selected *L1Hs* elements differ from a consensus sequence by as much as 5% or more in sequence (Skowronski and Singer, 1986). The availability of the cloned *LRE-1* alleles has solved one of the primary challenges to the study of *L1Hs* transposition, namely, the identification and isolation of an active element from among the many related genomic sequences. We are grateful to Kazazian and his colleagues for making these clones available.

Besides the sequence identity between an *L1Hs* allele at *LRE-1* and the newly transposed element, several other observations confirm the characterization of the segments at *LRE-1* as active elements. First, the sequence of the *LRE-1* alleles falls in a subfamily of genomic *L1Hs* elements, the Ta subset, previously identified as being transcribed to yield cytoplasmic, polyadenylated RNA (Skowronski et al., 1988) and which are transcribed upon transfection into human teratocarcinoma cells (Holmes

et al., 1992). Second, the *LRE-1* allele *L1.2A* has been shown to encode an active reverse transcriptase (Mathias et al., 1991), as predicted some time ago from the *L1Hs* consensus nt sequence (Hattori et al., 1986; Skowronski and Singer, 1986). Finally, the p40 polypeptide encoded by ORF1 of *L1.2* elements was shown, in transfection experiments, to have an electrophoretic mobility (under denaturing conditions) identical to that of intracellular p40 (Holmes et al., 1992); this is significant because most of the p40s translated from isolated Ta subset clones have different mobilities and thus do not appear to contribute to the intracellular form (Leibold et al., 1990; Holmes et al., 1992).

RESULTS AND DISCUSSION

(a) Cell type specificity of *L1Hs* expression

Full length, cytoplasmic, polyadenylated *L1Hs* transcripts have been detected and characterized in a human teratocarcinoma cell line, NTera2D1 (Skowronski and Singer, 1985; Skowronski et al., 1988). Such RNAs were also detected in JEG3 choriocarcinoma cell lines but not in HeLa cells. Additional evidence for preferential expression in certain cell types comes from the use of antiserum prepared against p40 synthesized in *E. coli* (Leibold et al., 1990). Western blots prepared with cell extracts (Leibold et al., 1990) or immunocytochemical staining of whole cell preparations (Bratthauer and Fanning, 1992; 1993) indicated the presence of p40 in the cell types already mentioned as well as in cell lines 2102EP (human teratocarcinoma) and A431 and in the cells of human germ cell tumors. Only very low amounts of p40, if any, were detected in HeLa, HL60, and 293 cells. Finally, it is apparent that the transcriptional regulatory region in the *L1Hs* 5′ *UTR* functions most efficiently in teratocarcinoma cells (Swergold, 1990). Thus, expression appears to be favored in cells of germ line origin as well as in some other tumor cells of epithelial origin.

(b) Synthesis and characterization of p40 in human teratocarcinoma cells

Both in situ immunocytochemical analysis (Bratthauer and Fanning, 1992; 1993) and cell fractionation studies combined with SDS-PAGE and Western blotting (R.E.T., V.K., J.P.McM. and M.F.S., in preparation) indicate that the bulk and perhaps all of the p40 in teratocarcinoma cells is in the cytoplasm. The protein is approximately 40 kDa, as predicted from ORF1, and is phosphorylated, as indicated by the effect of phosphatase treatment on electrophoretic mobility and the incorporation of ^{32}P from [γ-^{32}P]ATP supplied to teratocarcinoma cells (R.E.T., V.K., J.P.McM. and M.F.S., in preparation). A variety of sites appropriate for phosphorylation by known protein kinases exist in p40 (Fig. 2).

An interesting feature within the central region of p40 is a potential leucine zipper structure very similar, in important residues, to the well-characterized GCN4 leucine zipper (Fig. 2) (Holmes et al., 1992). A basic region follows the zipper segment in p40. Experiments designed to test whether p40 is a DNA-binding protein yielded negative results. Cross-linking by glutaraldehyde of p40 present in teratocarcinoma cell extracts (R.E.T., V.K., J.P.McM. and M.F.S., in preparation) and of p40 synthesized in *E. coli* (H. Hohjoh, unpublished experiments) indicates that the polypeptide forms homomultimeric complexes, possibly through leucine zipper interactions.

Although ORF1 occupies a position in *L1Hs* that is analogous to that of gag and gag-like polypeptides in LTR-containing retrotransposons and retroviruses, it has no homology to these proteins as determined by searches and alignment tests against GenBank sequences. Moreover, as indicated above, p40 does not appear to be subject to proteolytic maturation as are the primary translation products of gag coding regions. Thus, it is difficult to speculate on the significance, if any, of p40 to the transposition process at this time.

(c) Transcription of *L1Hs* elements

L1Hs elements of the Ta subset contain, within the 5′ *UTR*, *cis*-acting sequences sufficient to promote transcription in a cell-specific manner (Swergold, 1990; Minikami et al., 1992) and to specify the transcriptional start point (*tsp*) at nt 1 of the elements (Swergold, 1990). Thus, although each element is in a distinctive genomic environment, transcription is coordinated. Experiments utilizing *lacZ* as a reporter gene fused in frame after the first few ORF1 codons, have indicated that all the signals

```
MGKKQNRKTG NSKTQSASPP PKERSSSPAT EQSWMENDFD ELREEGFRRS 50
NYSELREDIQ TKGKEVENFE KNLEECITRI TNTEKCLKEL MELKTKAREL 100
REECRSLRSR CDQLEERVSA MEDEMNEMKR EGKFREKRIK RNEQSLQEIW 150
DYVKRPNLRL IGVPESDVEN GTKLENTLQD IIQENFPNLA RQANVQIQEI 200
QRTPQRYSSR RATPRHIIVR FTKVEMKEKM LRAAREKGRV TLKGKPIRLT 250
VDLSAETLQA RREWGPIFNI LKEKNFQPRI SYPAKLSFIS EGEIKYFIDK 300
QMLRDFVTTR PALKELLKEA LNMERNNRYQ PLQNHAKM 338
```

Fig. 2. The aa sequence of p40. The 338-aa sequence is deduced from the nt sequence of ORF1 in *L1.2A* (Dombroski et al., 1991; Mathias, 1992). Italics indicate aa that can form a leucine zipper. Bold type indicates a site that could be a target for phosphorylation by cAMP-dependent protein kinase. Underlined sequences are consensus sites for protein kinase C. In addition there are target sites for casein kinase II (S/TXXD/E).

required for initiation of transcription at residue 1 and contributing to the efficiency of transcription as well as to specific transcription in NTera2D1 cells reside in the first 670 bp of the 900-bp 5′ *UTR*; deletion of different portions of the 5′ *UTR* indicates that several important regulatory segments are spread throughout the region (Swergold, 1990).

The first 100 bp are especially important, as their deletion reduces transcription by 300-fold in transfected NTera2D1 cells while deletions in other regions have lesser effects on transcription. Deletion of the first 18 bp alone reduces transcription about fivefold and no further decrease was observed when the deletion was extended to bp 32 (Becker et al., 1993). Inspection of the sequences close to the *tsp* indicated the presence of the sequence 5′-GGCCATCTT-3′ (nt 21–13 on the bottom strand, Fig. 1), a binding site for the known transcription factor YY1 (Hariharan et al., 1991; Flanagan et al., 1992). Nuclear extracts of NTera2D1 cells as well as those from cells previously known to contain YY1 contain a protein that forms complexes with oligos representing the first 40 bp of *L1Hs* element *L1.2* and these complexes are ablated by antibody specific for YY1. Thus, YY1 appears to be important for *L1Hs* transcription. It is unlikely, however, that YY1 is responsible for the cell-type specificity of transcription because it is ubiquitous. It is interesting to point out that YY1 is important in the transcriptional regulation of other mammalian genes that have promoters downstream from the *tsp*.

(d) Translation of *L1Hs* RNA

Several features of *L1Hs* suggest special questions regarding translation of ORF1 and ORF2. First, there is the very long and G+C-rich 5′ *UTR*. Computer analysis indicates that the 900-bp segment has the potential to form stable secondary structures. Moreover, each of the *L1.2* alleles at locus *LRE-1*, as well as other characterized members of the Ta subset, has at least one AUG codon in the 5-*UTR*, upstream from the AUG codon that initiates translation of ORF1. The upstream AUGs could initiate short ORFs of from three to 20 codons. These structural considerations suggest that translation of ORF1 might be impeded if a scanning 40S ribosome, starting at the 5′ end, had to traverse the whole 5′ *UTR*. As expected from these considerations, in vitro translation of ORF1 from an mRNA with a very short leader sequence is appreciably more efficient than from *L1Hs* RNA (J.P.McM. and M.F.S., in preparation). Nevertheless, as the experiments summarized above indicate, ORF1 is translated in vitro and in cells.

Additional questions arise about the translation of ORF2. Unlike p40, no products of ORF2 translation have been detected in human teratocarcinoma cells. ORF1 and ORF2 are in the same frame, but they are separated by an inter-ORF region of 33 bp bracketed by two conserved in-frame stop codons; some *L1Hs* elements contain additional in-frame stop codons in the inter-ORF. How is ORF2 translated? A number of mechanisms are known to account for the translation of bicistronic mRNAs in eukaryotic cells. These include ribosomal frameshifts at the overlap region between overlapping ORFs to produce a fusion protein (Hatfield et al., 1992), suppression of a single termination codon and readthrough which again produces a fusion protein (Hatfield et al., 1992), reinitiation by attached ribosomes following termination of translation at the end of the first ORF, or independent internal initiation by newly attached ribosomes (Chang et al., 1990; Schultze et al., 1990).

None of the translation experiments with *L1Hs* to date, including both in vitro and in vivo, give any evidence for formation of an ORF1/ORF2 fusion protein, apparently eliminating suppression of the multiple stop codons as a mechanism whereby ORF2 could be translated. These experiments have utilized reporter gene constructs in which the *Escherichia coli lacZ* gene is fused, in frame, after the first 15 codons of the *L1Hs* ORF2 and β-galactosidase production have been assayed either as an immunologically reactive protein of expected mobility (in vitro) or by enzymatic assay of cell extracts (after transfection of teratocarcinoma cells) or by histochemical staining of cells (Swergold, 1990; J.P.McM. and M.F.S., in preparation). After in vitro translation of in vitro synthesized mRNA with a rabbit reticulocyte lysate, the full length product expected from the ORF2/*lacZ* construct as well as p40 are readily detectable by SDS-PAGE (J.P.McM. and M.F.S., in preparation). Moreover, no obvious qualitative effects were seen on ORF2 products when constructs containing deletions or other debilitating modifications of ORF1 were translated in vitro. Similar results were obtained when the mRNAs contained the full length *L1Hs* ORF2; ORF2 polypeptide was synthesized efficiently but no ORF1/ORF2 fusion protein was detectable. These experiments indicated that ORF1 translation is neither essential for, nor inhibitory to the translation of ORF2 in vitro and suggest that either reinitiation by scanning ribosomes or independent initiation by newly attached ribosomes is involved in ORF2 translation. Recent experiments on *L1Rn*, in which the two ORFs overlap and are in different reading frames reached a similar conclusion (Ilves et al., 1992).

The translation of the *lacZ* reporter gene fused into the beginning of ORF2 was also examined after transfection of plasmids into NTera2D1 cells. In contrast to the in vitro translations, neither enzyme activity nor immunologically cross-reacting material (to β-galactosidase)

was reproducibly detectable in cell extracts prepared from transfected cells although a small number of cells always were positive in in situ tests for enzyme activity. Thus, some factor(s) in the intracellular environment appear to suppress translation of ORF2.

A series of recent experiments give some hint of the mechanism whereby ORF2 translation is inhibited in the cells and also permit distinguishing between initiation of ORF2 translation by scanning ribosomes or reinitiation by newly attached ribosomes (J.P.McM. and M.F.S., in preparation). A stable hairpin structure was introduced into the region of the 5′ *UTR* between that known to be important for transcription (nt 1–660) and the beginning of ORF1. When mRNA synthesized in vitro from such constructs was translated in vitro, translation of ORF1 (either p40 or *lacZ* fused in frame within ORF1) was decreased approximately fivefold. Similarly, the stable hairpin structure decreased the translation of ORF1 when such constructs were transfected into teratocarcinoma cells. These results indicate that the translation of ORF1 may initiate following the loading of 40S ribosomal subunits at some point 5′ of the site of the hairpin insertion followed by scanning.

We then investigated the influence on ORF2 translation of the decrease in ORF1 translation consequent to the introduction of the stable hairpin in the 5′ *UTR*. These experiments utilized the constructs in which *lacZ* is fused, in frame, after the first 15 codons of ORF2. The diminished translation of ORF1 had little or no effect on the translation of ORF2 in vitro, consistent with the translation of ORF2 being independent of that of ORF1. Moreover, the presence of the hairpin and the resulting suppression of p40 translation in transfected teratocarcinoma cells was accompanied by an increase in the number of cells producing β-galactosidase (detected by in situ, histochemical staining). Thus, when the translation of p40 is decreased, the translation of ORF2 is enhanced, consistent with an internal initiation by newly attached ribosomes as the mechanism of translation of ORF2.

(e) Conclusions

The expression of *L1Hs* elements appears to involve a series of known, but uncommon mechanisms including internal transcriptional regulatory signals, a long and complex 5′ *UTR*, normally suppressed translation of ORF1, and highly suppressed translation of ORF2. Moreover, most of even the full length *L1Hs* elements in the human genome, including those that are specifically transcribed in NTera2D1 cells have bp sequences that result in closure of ORF1 or ORF2 or both (Skowronski et al., 1988). The evolution of *L1Hs* has resulted in a very large family of nonfunctional and minimally functional elements. This can be seen as a 'stand-off' between the *L1Hs* family evolving to sustain its existence and the rest of the genome, which might be more stable in the absence of these integral insertional mutagens.

REFERENCES

Adams, J.W., Kaufman, R.E., Kretschmer, P.J., Harrison, M. and Nienhuis, A.W.: A family of long reiterated DNA sequences, one copy of which is next to the human beta globin gene. Nucleic Acids Res. 8 (1980) 6113–6128.

Becker, K.G., Swergold, G.D., Ozato, K. and Thayer, R.E.: Binding of the ubiquitous nuclear transcription factor YY1 to a *cis* regulatory sequence in the human LINE-1 transposable element. Hum. Mol. Genet. (1993) in press.

Boeke, J.D. and Corces, V.G.: Transcription and reverse transcription of retrotransposons. Annu. Rev. Microbiol. 43 (1989) 403–434.

Bratthauer, G.L. and Fanning, T.G.: Active LINE-1 retrotranposons in human testicular cancer. Oncogene 7 (1992) 507–510.

Bratthauer, G.L. and Fanning, T.G.: LINE-1 retrotransposon expression in pediatric germ cell tumors. Cancer 71 (1993) 2383–2386.

Chang, L.-J., Ganem, D. and Varmus, H.E.: Mechanism of translation of the hepadnaviral polymerase (*P*) gene. Proc. Natl. Acad. Sci. USA 87 (1990) 5158–5162.

Deragon, J.-M., Sinnett, D. and Labuda, D.: Reverse transcriptase activity from human embryonal carcinoma cells NTera2D1. EMBO J. 9 (1990) 3363–3368.

Dombroski, B.A., Mathias, S.L., Nanthakumar, E., Scott, A.F. and Kazazian Jr., H.H.: Isolation of an active human transposable element. Science 254 (1991) 1805–1808.

Dombroski, B.A., Scott, A.F. and Kazazian Jr., H.H.: Two new potential retrotransposons isolated from a human L1 subfamily which contains an active retrotransposable element. Proc. Natl. Acad. Sci. USA 90 (1993) 6513–6517.

Flanagan, J.R., Becker, K.G., Ennist, D.L., Gleason, S.L., Driggers, P.H., Levi, B.-Z., Appela, E. and Ozato, K.: Cloning of a negative transcription factor that binds to the upstream conserved region of Maloney murine leukemia virus. Mol. Cell. Biol. 12 (1992) 38–44.

Hariharan, N.D., Kelley, E. and Perry, R.P.: δ, a transcription factor that binds to downstream elements in several polymerase II promoters, is a functionally versatile zinc finger protein. Proc. Natl. Acad. Sci. USA 88 (1991) 9799–9803.

Hatfield, D.L., Levin, J.G., Rein, A. and Oroszlan, S.: Translational suppression in retroviral gene expression. Adv. Vir. Res. 41 (1992) 193–239.

Hattori, M., Kuhara, S., Takenaka, O. and Sakaki, Y.: The L1 family of long interspersed repetitive DNA sequences in primates may be derived from a sequence encoding a reverse transcriptase related protein. Nature 321 (1986) 625–628.

Holmes, S.E., Singer, M.F. and Swergold, G.D.: Studies on p40, the leucine zipper motif-containing protein encoded by the first open reading frame of an active human LINE-1 transposable element. J. Biol. Chem. 267 (1992) 19765–19768.

Hutchison III, C.A., Hardies, S.C., Loeb, D.D., Shehee, W.R. and Edgell, M.H.: LINEs and related retrotransposons: long interspersed repeated sequences in the eucaryotic genome. In: Berg, D.E. and Howe, M.M. (Eds.), Mobile DNA. American Society for Microbiology, Washington, DC, 1989, pp. 593–617.

Ilves, H., Kahre, O. and Speek, M.: Translation of the rat LINE bicistronic RNAs in vitro involves ribosomal reinitiation instead of frameshifting. Mol. Cell. Biol. 12 (1992) 4242–4248.

Kazazian Jr., H.H., Wong, J.C., Youssoufian, H., Scott, A.F., Phillips,

D.G. and Antonarakis, S.: Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanism for mutation in man. Nature 332 (1988) 164–166.

Leibold, D.M., Swergold, G.D., Singer, M.F., Thayer, R.E., Dombroski, B.A. and Fanning, T.G.: Translation of LINE-1 DNA elements in vitro and in human cells. Proc. Natl. Acad. Sci. USA 87 (1990) 6990–6994.

Martin, S.L.: Ribonucleoprotein particles with LINE-1 RNA in mouse embryonal carcinoma cells. Mol. Cell. Biol. 11 (1991) 4804–4807.

Mathias, S.L.: Functional Studies of a Human Transposable Element. Ph.D. Thesis. The Johns Hopkins University, Baltimore, MD, 1992.

Mathias, S.L., Scott, A.F., Kazazian Jr., H.H., Boeke J.D. and Gabriel, A.: Reverse transcriptase encoded by a human transposable element. Science 254 (1991) 1808–1810.

Miki, Y., Nishisho, I., Horii, A., Miyoshi, Y., Utsunomia, J., Kinzler, K.W., Vogelstein, B. and Nakamura, Y.: Disruption of the APC gene by a retrotransposable insertion of L1 sequence in a colon cancer. Cancer Res. 52 (1992) 643–645.

Schultze, M., Hohn, T. and Jiricny, J.: The reverse transcriptase gene of cauliflower mosaic virus is translated separately from the capsid gene. EMBO J. 9 (1990) 1177–1185.

Scott, A.F., Schmeckpeper, B.J., Abdelrazik, M., Comey, C.T., O'Hara, B., Rossiter, J.P., Cooley, T., Heath, P., Smith, K.D. and Margolet, L.: Origin of the human L1 elements: proposed progenitor genes deduced from a consensus sequence. Genomics 1 (1987) 113–125.

Skowronski, J. and Singer, M.F.: Expression of a cytoplasmic LINE-1 transcript is regulated in a human teratocarcinoma cell line. Proc. Natl. Acad. Sci. USA 82 (1985) 6050–6054.

Skowronski, J. and Singer, M.F.: The abundant LINE-1 family of repeated DNA sequences in mammals: genes and pseudogenes. Cold Spring Harbor Symp. Quant. Biol. 51 (1986) 457–464.

Skowronski, J., Fanning, T.G. and Singer, M.F.: Unit-length line-1 transcripts in human teratocarcinoma cells. Mol. Cell. Biol. 8 (1988) 1385–1397.

Swergold, G.D.: Identification, characterization, and cell specificity of a human LINE-1 promoter. Mol. Cell Biol. 10 (1990) 6718–6729.