5. J. M. Wilson *et al.*, *Clin. Biotechnol.* **3**, 21 (1991); J. C. Wiederkehr, G. T. Kondos, R. Pollak, *Transplantation* **50**, 466 (1990).
6. M. Grossman and J. M. Wilson, in preparation.
7. J. M. Wilson, D. E. Johnston, D. M. Jefferson, R. C. Mulligan, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 4421 (1988).
8. J. Price, P. Turner, C. Cepko, *ibid.* **84**, 156 (1987).
9. Animals were anesthetized in preparation for lobectomy by intravenous administration of 10 mg of pentobarbital sodium per kilogram of body mass and were maintained by incremental doses of 2 mg/kg. Amount of anesthesia was monitored by corneal reflex. After depilation and aseptic preparation of the abdomen, celiotomy was performed through a midline incision. We freed the left lateral lobe of the liver by resecting ligamentous attachments; vessels feeding this lobe were tied with a 0 silk ligature close to the hilus, and the lobe (approximately 30% of the liver mass) was removed. The abdomen was closed in three layers, including a layer of subcuticular stitches. Hepatocytes were isolated by immediate collagenase perfusion through the opening of the major vessels (*4*). Hepatocytes were plated in primary culture, infected with retroviruses, and harvested as described (*4*). The cells were transplanted into the rabbits by means of the spleen with the following procedure. The spleen was exposed through a left subchondral incision. The left lower pole was secured by a suture through the splenic mesentery, and the suspension of hepatocytes (~15 ml) was injected into the splenic pulp through a 27-gauge needle over a period of 8 min. During the injection, we saw the fluid flow through the splenic vein. A knot was tightened around the lower pole of the spleen, the needle was withdrawn, and the abdomen was closed in two layers. Approximately 65% of the animals survived the combination of partial hepatectomy and hepatocyte transplantation; the major cause of mortality occurred during the second procedure. This has been corrected through the use of improved anesthesia and careful perioperative monitoring.
10. The following tissues were analyzed by in situ hybridization to the rat albumin probe: W98 (4 sections from one tissue block), W204 (44 sections from four tissue blocks), and W185 (40 sections from three tissue blocks). The following tissues were analyzed by in situ hybridization to the consensus albumin probe: W204 (28 sections from three tissue blocks) and W185 (23 sections from three tissue blocks). There was substantial variation in the frequency of albumin-expressing cells. Review of 114 microscopic fields (×20 magnification) from the spleen of animal W185 revealed the following distribution of albumin-expressing cells with the rat albumin probe: zero cells per field, 25.8%; one to ten cells per field, 69.7%; and greater than ten cells per field, 4.5%. The same analysis with the same tissue revealed the following distribution of albumin-expressing cells with the consensus albumin probe: zero cells per field, 12.5%; one to ten cells per field, 68.7%; and greater than ten cells per field, 18.8%.
11. W. G. Cochran and G. M. Cox, *Experimental Designs* (Wiley, New York, ed. 2, 1957), pp. 127–131.
12. This hypothesis is based on the observation that receptor-negative individuals (<2% residual LDLR have more severe and mortal coronary heart disease than do receptor-defective individuals (>2% residual LDLR activity) (*1*). Reconstitution of 2 to 4% of residual LDLR activity, an amount achieved in this study, may convert a receptor-negative individual to a receptor-defective status, with the associated clinical advantages.
13. K. K. Kozarsky, H. A. Brush, M. Kreiger, *J. Cell Biol.* **102**, 1567 (1986).
14. M. G. Grossman, J. R. Baker, J. M. Wilson, unpublished data.
15. J. E. Pinter and D. I. Lugo, in *In Situ Hybridization: Applications to Neurobiology*, K. L. Valentino, J. H. Eberwine, J. D. Barchas, Eds. (Oxford Univ. Press, New York, 1987), pp. 179–196.
16. A. Pandura *et al.*, *Biochemistry* **25**, 1414 (1986); A. Dugaiczyk, S. W. Law, O. E. Dennison, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 71 (1982).
17. P. Trinder, *Ann. Clin. Biochem.* **12**, 226 (1974).
18. We thank L. L. Austin, D. McKinley, and P. Arscott for technical help with in situ hybridization and protein immunoblot analysis; K. Kozarsky and M. Kreiger for rabbit antiserum to LDLR; P. C. Yerneni and T. S. Huang for performing phlebotomies; S. Schmaltz for statistical analysis of data; D. Shafritz for the pGEM4Z plasmid with rat albumin cDNA; and D. Russell for the plasmid with partial WHHL LDLR cDNA. Supported by a grant from the Howard Hughes Medical Institute (to J.M.W.) and NIH grants RO1-DK42193-01 (to J.M.W.), P01-DK-42718 (to J.M.W., Program Project "Experimental Models of Gene Therapy"), R01-DK 34357 (to J.R.C.), R01-DK-39137 (to N.R.C.), and DK-201928 (to the Liver Research Center).

18 June 1991; accepted 1 October 1991

# Isolation of an Active Human Transposable Element

BETH A. DOMBROSKI, STEPHEN L. MATHIAS, ELIZABETH NANTHAKUMAR, ALAN F. SCOTT, HAIG H. KAZAZIAN, JR.*

**Two de novo insertions of truncated L1 elements into the factor VIII gene on the X chromosome have been identified that produced hemophilia A. A full-length L1 element that is the likely progenitor of one of these insertions was isolated by its sequence identity to the factor VIII insertion. This L1 element contains two open-reading frames and is one of at least four alleles of a locus on chromosome 22 that has been occupied by an L1 element for at least 6 million years.**

TRANSPOSABLE ELEMENTS ARE A common feature of the genomes of eukaryotes (*1*). Retrotransposons (class I and class II) are a specific group of transposable elements that can be transcribed into RNA, reverse transcribed into cDNA, and then reintegrated as cDNAs into the genome at a new location. Examples of class II retrotransposons are the F, G, and I factors of *Drosophila melanogaster* and the long interspersed elements (LINE-1 or L1) of mammals (*2–4*). Class II retrotransposons do not contain long terminal repeats, have at least one open-reading frame (ORF), and have an A-rich tract at their 3′ ends. These retrotransposons are flanked by duplications of the insertion site that vary in length. The amino acid sequence derived from one of the reading frames contains a region similar to reverse transcriptases of retroviruses and class I retrotransposable elements (*5, 6*). Although retrotransposition by way of an RNA intermediate has been documented only for the I factor of *D. melanogaster* (*7*), the structural similarity of class II elements implies that they are all retrotransposons.

The L1 element is common in all mammalian genomes (*3*). In humans, L1 elements (full-length and truncated) compose about 5% of the genome and number 50,000 to 100,000 copies (*8*), most of which are truncated at the 5′ end (*9*). About 3500 copies are full-length (*10*). It has been proposed that rare L1 elements are capable of retrotransposition; however, no active individual sequences have been isolated.

A consensus human L1 element constructed from genomic L1 sequences predicted two ORFs that were not seen in any of the individual sequences used to construct the consensus (Fig. 1A) (*9*). On average, the nucleotide sequence of a genomic L1 element differs from the consensus sequence by 5% (*9*). In contrast, L1 cDNAs from a human embryonal tumor cell line differ in nucleotide sequence from the L1 consensus sequence by <1% (*11*). This suggests that the consensus sequence reflects the sequence of actively transcribed and potentially retrotransposable L1 elements.

Evidence that some L1 elements in the human genome are retrotransposable was provided by de novo transposition of truncated L1 DNA into exon 14 of the factor VIII gene producing hemophilia A in two individuals (*12*). Each new insertion contained the 3′ portion of an L1 element, including much of the second ORF in the consensus sequence (ORF2) (Fig. 1A) and a polyadenylic acid tail, and was flanked by a perfect target-site duplication (*13*). We report here the isolation of the L1 element that is the likely progenitor of the insertion from patient JH-27. It represents the first-identified active transposable element in humans.

From a region of the factor VIII insertion in patient JH-27 [nucleotides (nts) 4250 to 4269], we made a 20-nt oligomer (JH-27 oligomer) that differed at three positions from the consensus L1 sequence (Fig. 1A). (Nucleotide positions are numbered such that 1 represents the first nucleotide in the progenitor allele L1.2B.) When the JH-27 oligomer was hybridized to Bam HI-digested genomic DNA from 106 individuals, four to ten fragments of single-copy intensity were observed in the DNA from each individual (Fig. 1B) (*14*). Analysis of ge-

Center for Medical Genetics, Departments of Pediatrics and Medicine, The Johns Hopkins University School of Medicine, Baltimore, MD 21205.

*To whom correspondence should be addressed.

nomic DNA from these unrelated individuals showed that most hybridizing fragments were highly polymorphic as to their presence. However, 11-kb, 7-kb, and 5.5-kb fragments were most often present. All of the fragments in patient JH-27 could be explained by inheritance from his parents except for a band at approximately 4 kb (Fig. 1B). Subsequent hybridization with a probe from exon 14 of the factor VIII gene confirmed that this 4-kb band in patient JH-27 corresponded to the new L1 insertion.

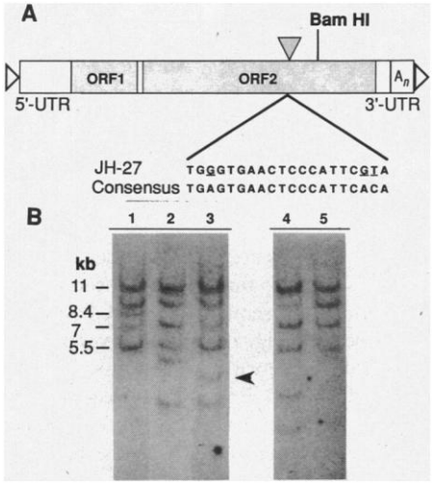


**Fig. 1.** (**A**) Structure of consensus 6.1-kb human L1 element. ORF1, first ORF; $A_n$, A-rich tract; open triangles flanking the consensus element, target-site duplication; shaded triangle above ORF2, position of the JH-27 oligomer; Bam HI, site at nt 4836; JH-27, the JH-27 oligomer (underlines indicate differences from consensus). (**B**) Hybridization pattern of the JH-27 oligomer to a dried agarose gel containing Bam HI–digested genomic DNA. Lanes 1, 2, and 3, family of patient JH-27: mother, father, and patient; lanes 4 and 5, DNA from two unrelated individuals. Arrow indicates the de novo insertion of L1 sequence into exon 14 of the factor VIII gene. The 11-kb Bam HI fragment corresponds to the L1.1 element, the 5.5-kb fragment to the L1.2 element, and the 8.4-kb fragment to the L1.3 element.

Nucleotide sequence identity between an individual L1 element and a new insertion identifies that element as a putative precursor of the insertion. Therefore, we searched for a full-length L1 element containing sequence identical to the 3784 bp of the insertion found in patient JH-27. We screened an anonymous human genomic library in bacteriophage λ with the JH-27 oligomer and with probes specific for the 5′ and 3′ ends of the consensus L1 element to detect full-length elements that included homology to the insertion found in patient JH-27 (*15*). Two recombinants, L1.1 and L1.2A, were isolated that hybridized to all three probes.

The high-frequency 11-kb Bam HI fragment that hybridized to the JH-27 oligomer is represented in L1.1 (Fig. 2). When Bam HI digests of genomic DNA from 62 unrelated individuals were hybridized with a portion of the 11-kb Bam HI fragment that lies 5′ to L1.1, only one band at 11 kb was seen in each individual (*16*). Nucleotide sequence analysis revealed that L1.1 is full-length, ends in a polyadenylic acid tail of 54 residues, and is flanked by a perfect 14-bp target-site duplication. There is a 1–base pair deletion at the beginning of the first predicted reading frame (nucleotide position 915), resulting in a frame-shift mutation and a premature termination codon (*17*). The second reading frame is open. The 3′ portion of L1.1 (3784 bp) differs from the insertion sequence found in patient JH-27 by six nucleotides, four of which change the encoded amino acid (Fig. 3). With the use of somatic cell hybrid DNAs, L1.1 has been mapped to chromosome 14 (*18*).

From the restriction map of L1.2A (Fig. 2), it was evident that the JH-27 oligomer fell within a 5.5-kb Bam HI fragment (Fig. 1B) that was observed in about 70% of human genomic DNAs studied. However, when a single-copy probe from the 5′ flanking sequence of L1.2A was hybridized to genomic DNA digested with Pst I, a fragment of the appropriate size (6.4 kb) was detected in all 40 individuals examined, and no variant fragment was seen in any individual (*16*). These data indicate that L1.2 defines a locus common to all individuals studied and that a variant allele contains changes in the JH-27 oligomer sequence that prevent the oligomer from hybridizing. Sequence analysis of L1.2A determined that it is full-length, ends in a polyadenylic acid tail of 27 residues, and is flanked by a perfect 15-bp target-site duplication (Fig. 3). Because both reading frames of L1.2A are open, this is the first human L1 element isolated that contains the two predicted ORFs. Functional studies with experimental constructs of L1.2A indicate that both ORFs can produce protein products (*19, 20*). Comparison of the nucleotide sequence in common between the insertion found in patient JH-27 and L1.2A revealed two differences at the 3′ end of ORF2, both of which result in an amino acid substitution (Fig. 3). The first is a T to G substitution at nt 5649 (Ile to Met), and the second is a C to T substitution at nt 5765 (Ser to Leu).

We then used the polymerase chain reaction (PCR) to amplify the 3′ end of L1.2 containing nts 5649 and 5765 from genomic DNA of the parents of patient JH-27 (Fig. 2) (*21*). Sequence of this portion of L1.2 was homozygous in both parents of patient JH-27 and identical to the insertion sequence found in their son. In genomic DNA of five nuclear families, the changes at nts 5649 and 5765 always occurred together as a common allele of the L1.2 locus.

With the use of a bacteriophage λ library constructed from genomic DNA of the mother of patient JH-27, we isolated an allele (L1.2B) from the L1.2 locus that contains a sequence identical to the 3784-bp insertion found in patient JH-27 (*22*). The restriction map of the phage insert containing L1.2B is identical to the map of the insert containing L1.2A except that the L1.2B insert is 3 kb larger. Sequence analysis of this allele revealed that it is also full-length, ends in a polyadenylic acid tail of 27 residues, contains two ORFs, and is flanked by the same perfect 15-bp target-site duplication as allele L1.2A. The sequence of the 3′ portion (3784 bp) is identical to the entire insertion from patient JH-27 (*23*). Upstream of sequence in common with this insertion L1.2B differs from L1.2A by only a single silent change in ORF2 at nt 2109 (T to C) (Fig. 3).

The L1.2 locus is on chromosome 22 at 22q11.1-q11.2 (*18, 24*). Because we believe that the L1.2B allele at the L1.2 locus is responsible for the transposition event in patient JH-27, we designate this locus

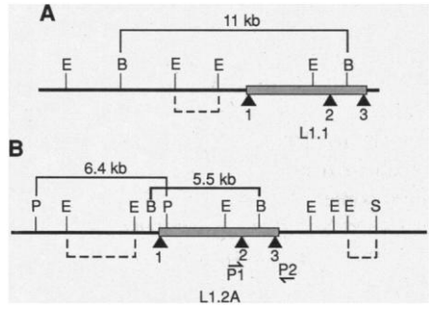


**Fig. 2.** Partial restriction maps of the phage inserts containing L1.1 and L1.2A. E, Eco RI; B, Bam HI; P, Pst I; S, Sph I. (**A**) Restriction map of the phage containing L1.1. Shaded box, L1.1; bold line, human genomic flanking sequences. Triangles indicate positions of probes we used to screen the bacteriophage libraries: 1, 5′ L1 oligomer (nts 61 to 80) derived from the L1 consensus; 2, the JH-27 oligomer; and 3, 3′ L1 oligomer (nts 5919 to 5938) derived from the cDNA consensus (*30*). The dashed bracket indicates the Eco RI fragment used as a single-copy flanking probe for Southern blot analysis. The 11-kb Bam HI fragment is indicated. (**B**) Restriction map of the phage containing L1.2A. Shaded box, L1.2A; bold line, human flanking sequences. Dashed brackets indicate 5′ and 3′ single-copy flanking probes used for Southern blot analysis. The bracket labeled 6.4 kb denotes the Pst I fragment detected by the 5′ flanking Eco RI probe. The bracket labeled 5.5 kb indicates the Bam HI fragment detected by the JH-27 oligomer. P1 and P2 indicate primers we used with the PCR to amplify the 3′ end of L1.2 from genomic DNA (*21*). Numbered triangles indicate the same probes as in (A).

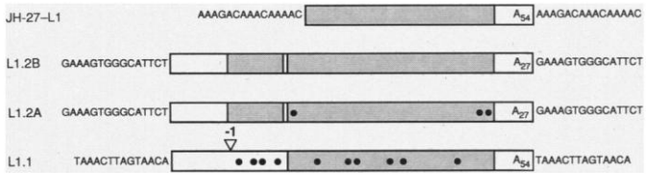


**Fig. 3.** Comparison of L1 elements (*31*) with sequence homology to the L1 fragment (JH-27–L1) inserted into the factor VIII gene of patient JH-27. Shaded boxes, ORFs; nucleotide sequence, target-site duplications; $A_n$ at 3′ end of each L1, polyadenylic acid tract of length *n*. Dots indicate single-nucleotide differences from L1.2B. In L1.2A from left to right, first dot is nt 2109 (T to C), second dot is nt 5649 (T to G, Ile to Met), and third dot is nt 5765 (C to T, Ser to Leu). In L1.1 from left to right, at nt 915 G is deleted (arrow), first dot is nt 1150 (A to T, Thr to Ser), second dot is nt 1641 (A to G), third dot is nt 1661 (T to C, Val to Ala), fourth dot is nt 1869 (A to G), fifth dot is nt 2565 (C to A), sixth dot is nt 3076 (G to A, Gly to Arg); seventh dot is nt 3121 (G to T, Ala to Ser), eighth dot is nt 3759 (A to T, Lys to Asn); ninth dot is nt 4055 (G to A, Arg to Gln), and tenth dot is nt 5622 (C to G).

LRE-1 (L1 retrotransposable element–1). We have identified four alleles of the LRE-1 locus: (i) L1.2B, (ii) L1.2A, (iii) an allele that is not recognized by the JH-27 oligomer, and (iv) an allele that at a minimum differs from L1.2B by a G to C transverison at position 5766. L1.2B has a gene frequency of roughly 0.2 in the 40 individuals we studied.

Southern (DNA) blot analysis of primate DNA with single-copy probes both 5′ and 3′ to L1.2 and 5′ to L1.1 (Fig. 2) showed that both L1.2 and L1.1 are present in chimpanzee and gorilla at the same genomic locations as in man. Digestion of genomic DNA with enzymes that have a recognition site in L1.2, followed by hybridization with a flanking probe, identified a fragment identical in size to the human fragment in two chimpanzees (six of six digests) and two gorillas (four of six digests) (*16*). Similar results were obtained for the L1.1 locus. These data indicate that L1.2 and L1.1 have been located at the same chromosomal sites at least since the evolutionary divergence of man, chimpanzee, and gorilla, approximately 6 million years ago (*25*).

We isolated another JH-27–specific full-length element, L1.3, from the bacteriophage λ library of genomic DNA from the mother of patient JH-27 that corresponds to the 8.4-kb Bam HI fragment (Fig. 1B). L1.3 contains at least two nucleotide changes from the insertion sequence found in patient JH-27. There are only six candidate progenitor elements in the parents of patient JH-27 (Fig. 1B) (those Bam HI fragments greater than 4836 bp). Three of these elements have now been characterized.

It is unlikely that an as yet unidentified element gave rise to the insertion found in patient JH-27. L1.2 has been at the same chromosomal location for at least 6 million years, and the insertion found in patient JH-27 occurred within the last 20 years. Given estimates of substitution rates in primate DNA, differences of 5 to 15 nts would be expected between L1.2 and the insertion found in patient JH-27 if these elements were diverging independently of each other (*26*). Instead, the sequence of the L1.2B allele is completely identical to the insertion found in patient JH-27, whereas the LRE-1 locus itself has diverged (four known alleles). Therefore, if the insertion arose from transposition of a different element, that element would probably have been derived from a recent transposition of L1.2B itself for there to be no sequence differences. In that case, L1.2B would still be the oldest precursor sequence and an active transposable element.

Two hypotheses of L1 expansion in the human genome have been proposed, a master gene model and a self-propagation model. The master gene produces only inactive copies (*27*), and a precursor retrotransposon would not itself be the product of a retrotransposition event. However, L1.2B, the putative progenitor of the insertion found in patient JH-27, is itself flanked by a target-site duplication and is therefore the product of a retrotransposition event. These data therefore favor the model of self-propagation, in which a number of active elements can produce both active and inactive progeny. If the self-propagation model is correct, then retrotransposable elements should contain their own promoter because an upstream promoter would be lost with each new secondary retrotransposition event. The 5′ untranslated regions (UTRs) of some human and rat L1 elements contain substantial promoter activity (*28*), and an internal RNA polymerase II promoter was reported for the L1-like element *jockey* from *D. melanogaster (29)*.

Whether retrotransposition of L1 elements requires protein products translated from their ORFs remains an open question. L1.2A, expressed in transfected Ntera2D1 human teratocarcinoma cells, produced a 40-kD ORF1 (first ORF) protein, as predicted by the consensus sequence (*19*). In addition, Mathias and co-workers (*20*) have detected reverse transcriptase activity encoded by the ORF2 of L1.2A expressed in yeast.

The insertion found in patient JH-27 is an exact but truncated copy of its progenitor element L1.2B. If the translated ORF1 and ORF2 products of an L1 element function in trans (act on other L1 elements), an L1 insertion could be the product of a transcribed L1 element incapable of producing protein. That L1.2 alleles are the only human L1 elements known to date that contain two ORFs suggests that use of protein products in cis is preferred in L1 retrotransposition, as demonstrated for *Drosophila* I factor retrotransposition (*7*).

## REFERENCES AND NOTES


1. D. Berg and M. Howe, Eds. *Mobile DNA* (American Society of Microbiology, Washington, DC, 1989).
2. D. H. Fawcett, C. K. Lister, E. Kellett, D. J. Finnegan, *Cell* **47**, 1007 (1986).
3. C. A. Hutchinson III, S. C. Hardies, D. D. Loeb, W. R. Shehee, M. H. Edgell, in (*1*), pp. 593–617.
4. T. G. Fanning and M. F. Singer, *Biochim. Biophys. Acta* **910**, 203 (1987).
5. Y. Xiong and T. H. Eickbush, *EMBO J.* **9**, 3353 (1990).
6. R. F. Doolittle, D.-F. Feng, M. S. Johnson, M. A. McClure, *Q. Rev. Biol.* **64**, 1 (1989).
7. A. Pelisson, D. J. Finnegan, A. Bucheton, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 4907 (1991); S. Jensen and T. Heidmann, *EMBO J.* **10**, 1927 (1991).
8. R. K. Moyzis *et al.*, *Genomics* **4**, 273 (1989).
9. A. F. Scott *et al.*, *ibid.* **1**, 113 (1987).
10. J. W. Adams *et al.*, *Nucleic Acids Res.* **8**, 6113 (1980).
11. M. Singer, personal communication.
12. H. H. Kazazian *et al.*, *Nature* **332**, 164 (1988).
13. P. Woods-Samuels *et al.*, *Genomics* **4**, 290 (1989).
14. Genomic DNA (10 μg) was digested with Bam HI (Gibco-BRL) under appropriate conditions. The digested DNA was analyzed by electrophoresis in a 1% agarose gel. After we denatured and neutralized the gel, it was dried for 2 hours [with heat applied for the first hour (60°C)], hybridized with $^{32}$P-labeled JH-27 oligomer ($3 \times 10^6$ cpm/ml), and washed under stringent conditions as described [A. B. Studencki and R. B. Wallace, *DNA (New York)* **3**, 7 (1984); P. G. Waber *et al.*, *Blood* **67**, 551 (1986)]. The gel was autoradiographed with the use of an intensifying screen at −70°C for 1 to 3 days.
15. A human genomic library in bacteriophage λ (Clontech, Palo Alto, CA) was screened with the L1 probes (Fig. 2). Filters were washed in 2 × SSPE [0.30 M NaCl, 0.02 M $NaH_2PO_4 \cdot H_2O$, 0.002 M EDTA (pH 7.4)] and 0.1% SDS for 30 min at 4°C below the 50% denaturation temperature for each oligomer. Clones positive with all three probes were purified. We isolated phage DNA using Qiagen-tips (Qiagen) and constructed a restriction map by digesting 1 μg of phage DNA under appropriate conditions with each enzyme listed in Fig. 2 (Gibco-BRL), followed by analysis by electrophoresis in a 0.8% agarose gel. Various fragments were subcloned and sequenced.
16. B. A. Dombroski, unpublished data.
17. D. M. Leibold *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 6990 (1990).
18. PCR primers for chromosomal mapping of L1.1 were located 5′ to the element from nts −183 to −164 and from nts −18 to +2 and yielded a single band of 185 bp from human-hamster cell hybrids (BIOS) that contained a chromosome 14. PCR primers for chromosomal mapping of L1.2 were located 2.1 and 3.1 kb 3′ to the L1 sequence and yielded a single band of 1 kb from human-rodent cell hybrids (National Institute of General Medical Sciences Cell Repository and BIOS) that contained a chromosome 22.
19. S. Holmes and G. Swergold, personal communication.
20. S. L. Mathias, A. F. Scott, H. H. Kazazian, Jr., J. Boeke, A. Gabriel, *Science* **254**, 1808 (1991).
21. The 3′ end of L1.2A was amplified from 1 μg of

genomic DNA with P1 (150 nM) and P2 (300 nM) as primers for the PCR [K. B. Mullis and F. A. Faloona, *Methods Enzymol.* **155**, 335 (1987)]. P1 is an oligomer whose sequence is shifted 1 nt in the 5′ direction with respect to the JH-27 oligomer to place the two changes from consensus at its 3′ end; P2 is a single-copy oligomer located at 17 to 36 bp 3′ to L1.2A. The 1836-bp PCR fragment was sequenced with the use of an internal primer [R. K. Saiki *et al.*, *Science* **239**, 487 (1988); C. Wong *et al.*, *Nature* **330**, 384 (1987)].

22. Genomic DNA (100 μg) from the mother of patient JH-27 was partially digested with Bam HI (10 μg of DNA per digest). Digests were pooled and centrifuged in a 10 to 40% sucrose gradient at 60,000*g* for 22 hours. Aliquots containing fragments of 10 to 18 kb were pooled, dialyzed, ligated to digested EMBL3 bacteriophage arms, and packaged (Stratagene). Recombinants ($2.1 \times 10^6$) were screened with the three L1 probes (Fig. 2), and one recombinant containing L1.2B was purified, and phage DNA was isolated.
23. The L1.2B element was amplified by PCR, and we sequenced overlapping 1-kb products of both strands on a DNA sequencer (Applied Biosystems) using cycle sequencing with Taq polymerase and dye-labeled dideoxynucleotide terminators.
24. M. L. Budarf, M. Simon, B. S. Emanuel, *Cytogenet. Cell Genet.* (abstr.), in press.
25. R. Holmquist, M. M. Miyamoto, M. Goodman, *Mol. Biol. Evol.* **5**, 217 (1988).
26. M. Goodman *et al.*, *Genome* **31**, 316 (1989); R. J. Britten, *Science* **231**, 1393 (1986).
27. R. J. Britten, D. B. Stout, E. H. Davidson, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 3718 (1989).
28. G. D. Swergold, *Mol. Cell. Biol.* **10**, 6718 (1990); A. V. Furano, S. M. Robb, F. T. Robb, *Nucleic Acids Res.* **16**, 9215 (1988).
29. L. J. Mizrokhi *et al.*, *Cell* **54**, 685 (1988).
30. J. Skowronski, M. F. Singer, *Cold Spring Harbor Symp. Quant. Biol.* **51**, 457 (1986).
31. GenBank accession numbers are as follows: L1.1, M80340; L1.1 3′ flanking sequence, M80341; L1.1 5′ flanking sequence, M80342; L1.2 alleles, M80343; L1.2 3′ flanking sequence, M80344; and L1.2 5′ flanking sequence, M80345.
32. We thank M. Singer, G. Swergold, T. Fanning, R. Thayer, L. Shapiro, and S. Holmes for helpful discussions, and patient JH-27 and his parents for their contribution to this work. Supported in part by NIH grants (to H.H.K. and A.F.S.). B.A.D. is supported by a fellowship from the Sloan Foundation. S.L.M. is supported by an NIH predoctoral training grant in human genetics.

5 July 1991; accepted 3 October 1991

# Reverse Transcriptase Encoded by a Human Transposable Element

STEPHEN L. MATHIAS, ALAN F. SCOTT, HAIG H. KAZAZIAN, JR., JEF D. BOEKE, ABRAM GABRIEL*

**L1 elements are highly repeated mammalian DNA sequences whose structure suggests dispersal by retrotransposition. A consensus L1 element encodes a protein with sequence similarity to known reverse transcriptases. The second open reading frame from the human L1 element L1.2A was expressed as a fusion protein targeted to Ty1 virus-like particles in *Saccharomyces cerevisiae* and shown to have reverse transcriptase activity. This activity was eliminated by a missense mutation in the highly conserved amino acid motif Y/F-X-D-D. Thus, L1 represents a potential source of the reverse transcriptase activity necessary for dispersion of the many classes of mammalian retroelements.**

LONG INTERSPERSED ELEMENTS (LINE-1 or L1) are a large class of mammalian repeated sequences with structural similarities to retrotransposons (*1*). The predicted protein encoded by the second open reading frame (ORF2) of L1 elements contains domains similar to retroviral and other reverse transcriptases (RT) (*2, 3*). However, the classification of a sequence as a retrotransposon based solely on structural similarities or the presence of RT amino acid sequence motifs is tenuous; RT sequence similarity does not by itself prove RT activity. The availability of L1.2A, a putatively functional human L1 element (*4*), presented an opportunity to test directly for RT activity.

To assay the ORF2 protein encoded by L1.2A for RT activity, we used an expression system (*5*) based on retrotransposon Ty1 of *Saccharomyces cerevisiae* (*6*) (Fig. 1). Ty1 encodes two overlapping ORFs. The first, *TYA*, encodes a *gag*-like protein. The second, *TYB*, encodes a *pol*-like protein that is expressed as a fusion protein with TYA protein. Ty1 proteins and RNA are coassembled into virus-like particles (VLPs). We used a Ty1-L1 chimeric construct (pSM2) in which the L1.2A ORF2 was fused in-frame with *TYB* immediately 3′ to the protease domain, replacing the integrase, RT, and RNAseH domains of *TYB* (Fig. 1). As a negative control, the sequence between the fusion site and the downstream long terminal repeat (LTR) was deleted (pSM5). The parent plasmid, which contained a functional Ty1 element (pSM1), was used as a positive control. All constructs were regulated by the *GAL1* promoter. Upon galactose induction, yeast carrying these plasmids produced and accumulated large amounts of VLPs, which were then partially purified and assayed for RT activity. To reduce RT activity attributable to expression of endogenous Ty1 and Ty2 elements, an *spt3* host strain was used. Transcripts initiated from the intact endogenous Ty1 LTR promoter are reduced 10- to 20-fold in *spt3* strains, whereas transcription of *GAL1*-promoted elements is unaffected (*5, 7–9*).

Yeast cell extracts containing VLPs were fractionated on sucrose gradients and the fractions were assayed for RT activity (Fig. 2A). The L1-associated activity sedimented with a slightly slower velocity than the Ty1 RT. This may be due to an altered structure of the L1-Ty1 VLPs. No RT activity was detected in VLPs from the negative control plasmid, pSM5. The same fractions used for the RT profiles were immunoblotted with an antiserum raised against Ty1-VLPs (*10*) (Fig. 2B). The L1 RT peak cofractionated with TYA proteins, as did the Ty1 RT. VLPs were produced from the negative control construct (pSM5) (Fig. 2B), therefore absence of RT from pSM5 was not due to a lack of VLP production.

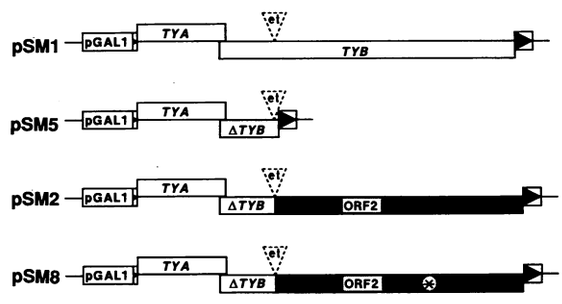


**Fig. 1.** Summary of plasmid inserts. Plasmids were constructed as described (*22*) and carry the yeast 2 μm plasmid origin of replication, the yeast selectable genes *URA3* and *TRP1*, and the bacterial selectable gene *bla*. pGAL1, *GAL1* promoter; *TYA*, first open reading frame of Ty1; *TYB*, second open reading frame of Ty1; Δ*TYB*, *TYB* truncated at base pair 2173; black triangles, Ty1 LTRs; ORF2, second open reading frame from L1.2A (*4*); dashed triangles, epitope tags (et) inserted in the parent plasmids to create pSM1et, pSM5et, pSM2et, and pSM8et (*16*). The asterisk indicates a missense mutation which produces a D to Y substitution in the RT domain of ORF2 (*22*).

S. L. Mathias and A. F. Scott, Center for Medical Genetics, Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, MD 21205.
H. H. Kazazian, Jr., Center for Medical Genetics, Departments of Medicine and Pediatrics, Johns Hopkins University School of Medicine, Baltimore, MD 21205.
J. D. Boeke and A. Gabriel, Department of Molecular Biology and Genetics, Johns Hopkins University School of Medicine, Baltimore, MD 21205.

*To whom correspondence should be addressed.