

## Two additional potential retrotransposons isolated from a human L1 subfamily that contains an active retrotransposable element

BETH A. DOMBROSKI, ALAN F. SCOTT, AND HAIG H. KAZAZIAN, JR.

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

Communicated by Maxine F. Singer, March 22, 1993 (received for review January 13, 1993)

**ABSTRACT** We have previously reported the isolation of a human retrotransposable L1 element. This element, allele L1.2B at the *LRE-1* locus of chromosome 22, was shown by nucleotide sequence identity to be the direct precursor of a *de novo* retrotransposition event into the factor VIII gene on the X chromosome, resulting in hemophilia A in patient JH-27. We now report the isolation of the two remaining full-length members of the subfamily of L1 elements closely related to L1.2B present in the genome of the mother of JH-27. Since these elements, L1.3 and L1.4, are very similar in sequence to L1.2B and contain both open reading frames 1 and 2 intact, they are also likely to be active retrotransposable elements. This suggests that certain L1 subfamilies may contain multiple active elements.

L1 elements are highly repetitive sequences which have been found in all mammals and make up about 5% of the human genome (1). Of the 50,000–100,000 copies of L1 elements, greater than 90% are truncated at their 5' ends or rearranged (2). How many of the approximately 3500 full-length, 6-kb, elements (3) are active is unknown, but the number is thought to be small. By definition, an active transposable element is one that can be transcribed into RNA and reverse transcribed into cDNA, and its cDNA can be reinserted into the genome at a new location. Depending upon the genomic location into which the new element is reinserted, it could disrupt the expression of a gene and result in human disease. Previously, we reported two retrotranspositions of truncated L1 elements into an exon of the factor VIII gene on the X chromosome, resulting in hemophilia A (4). One of these insertions, in patient JH-27, was 3784 bp in length and corresponded to the 3' end of an L1 element. Subsequently, we reported the isolation of the full-length element, L1.2B, whose nucleotide sequence identity to the insertion indicated that it is the direct precursor of the JH-27 insertion (5). This element is one of several alleles at the *LRE-1* locus on chromosome 22 q11.1–21.1, is fixed in the human genome, and has been at the same chromosomal location for at least 6 million years. The element has both of the open reading frames, ORF1 and ORF2, predicted from the human L1 consensus sequence (6), and both ORFs are expressed in *in vitro* systems. ORF1 encodes a 40-kDa protein of unknown function (7), while ORF2 encodes a reverse transcriptase (RT) (8).

Recently, a number of disease-producing L1 insertions have been observed: two separate insertions into the dystrophin gene producing Duchenne muscular dystrophy (ref. 9; E. Bakker and G. J. B. van Omenn, personal communication) and a somatic insertion into the *APC* gene, a colorectal cancer suppressor gene (10). These insertions suggest that a large number of sites receptive to L1 integration exist in the human genome. Although the sequences of these 500- to 1000-bp insertions are similar to the sequence of the JH-27

insertion, they are not identical, suggesting that the human genome contains a number of active progenitor elements.

Since there is growing evidence for the existence of both a small number of master L1 genes and a larger number of active L1 elements (11), Deininger *et al.* hypothesized that each of a small number of master genes has given rise to an L1 subfamily, some members of which are active while others are inactive (11). We therefore decided to determine whether any of the other members of the JH-27 subfamily of L1 elements are likely to be active. We now report the isolation of two additional full-length elements from this subfamily, both of which have sequence characteristics of active retrotransposable elements.\*

### MATERIALS AND METHODS

**Hybridization of Genomic Digests.** Hybridization of the JH-27 oligomer to dried gels of genomic digests with various enzymes was carried out as described previously (5). This same technique was used to determine the gene frequency of L1.3 and L1.4 in various human populations. In the case of L1.3, the oligonucleotide probe was located from +18 to +37 nucleotides 3' to the element, while for L1.4, a 20-bp probe was located approximately 650 bp 5' to the element.

**Construction and Screening of the Genomic JH-27M  $\lambda$  Library.** The construction of the genomic library from JH-27M, as well as the screening of the library with the L1 probes, has been reported (5). Portions of the L1.3 and L1.4 clones were subcloned in pGEM-7Zf(+) (Promega), and both strands were sequenced on a DNA sequencer (Applied Biosystems) using cycle sequencing with *Taq* DNA polymerase and dye-labeled dideoxynucleotide terminators.

**Somatic Cell Hybrid Mapping.** Genomic DNA (100 ng) from human-hamster cell hybrids (Bios, New Haven, CT) was amplified, using PCR primers for chromosomal mapping of L1.3 located 3' to the element from nucleotides +18 to +37 and from +522 to +541. These primers yielded a single band of 524 bp from the hybrid cells that contained a human chromosome 14. The PCR buffer for mapping L1.3 was 10 mM Tris-HCl/50 mM KCl/1.5 mM MgCl<sub>2</sub>, pH 8.3, and the PCR program was 95°C for 30 s, 55°C for 30 s, and 72°C for 30 s for 35 cycles. PCR primers for chromosomal mapping of L1.4 were located 3' to the element from nucleotides +99 to +118 and from approximately +1111 to +1130. These primers yielded a single band of approximately 1 kb from hybrid cells that contained a human chromosome 9. The PCR buffer for mapping L1.4 was 10 mM Tris-HCl/50 mM KCl/20 mM MgCl<sub>2</sub>, pH 8.3, and the PCR program was 95°C for 30 s, 55°C for 30 s, 72°C for 60 s for 40 cycles.

Abbreviations: ORF, open reading frame; RT, reverse transcriptase. \*The sequences discussed in this paper have been deposited in the GenBank data base (accession nos. as follows: L1.3 3' flank, L19086; L1.3 5' flank, L19087; L1.3, L19088; L1.4 3' flank-downstream, L19089; L1.4 3' flank-immediate, L19090; L1.4 5' flank, L19091; and L1.4, L19092).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

## RESULTS

To identify a candidate precursor for the JH-27 insertion among the roughly 3500 full-length L1 elements in the genome (3), we had previously chosen an oligomer from the JH-27 insertion sequence that differed at 3 of 20 bases from the L1 genomic consensus sequence (nucleotides 4250–4269) (5). When this JH-27 oligomer was used to probe genomic digests from the population, 6–10 hybridizing fragments were seen in each individual at single-copy intensity. Most fragments were polymorphic as to presence. These bands identified a subfamily of L1 elements specifically related to the

JH-27 insertion sequence. Fig. 1 shows the results of the JH-27 oligomer hybridization to genomic digests from patient JH-27 and his parents. In the *Bam*HI and *Eco*RI digests (Fig. 1 *B* and *C*), the new insertion in patient JH-27 is indicated by an arrow. All other bands can be explained by inheritance from his mother (JH-27M) or his father (JH-27F).

The number of bands in the *Bam*HI and *Eco*RI digests indicate that the genome of JH-27M contains seven (*Bam*HI) or eight (*Eco*RI) elements from this subfamily. The JH-27 oligomer hybridizes to the 5' end of a *Bam*HI-digested L1 element and to the 3' end of an *Eco*RI-digested element (Fig. 1A). *Acc* I recognition sites occur within the L1 element at

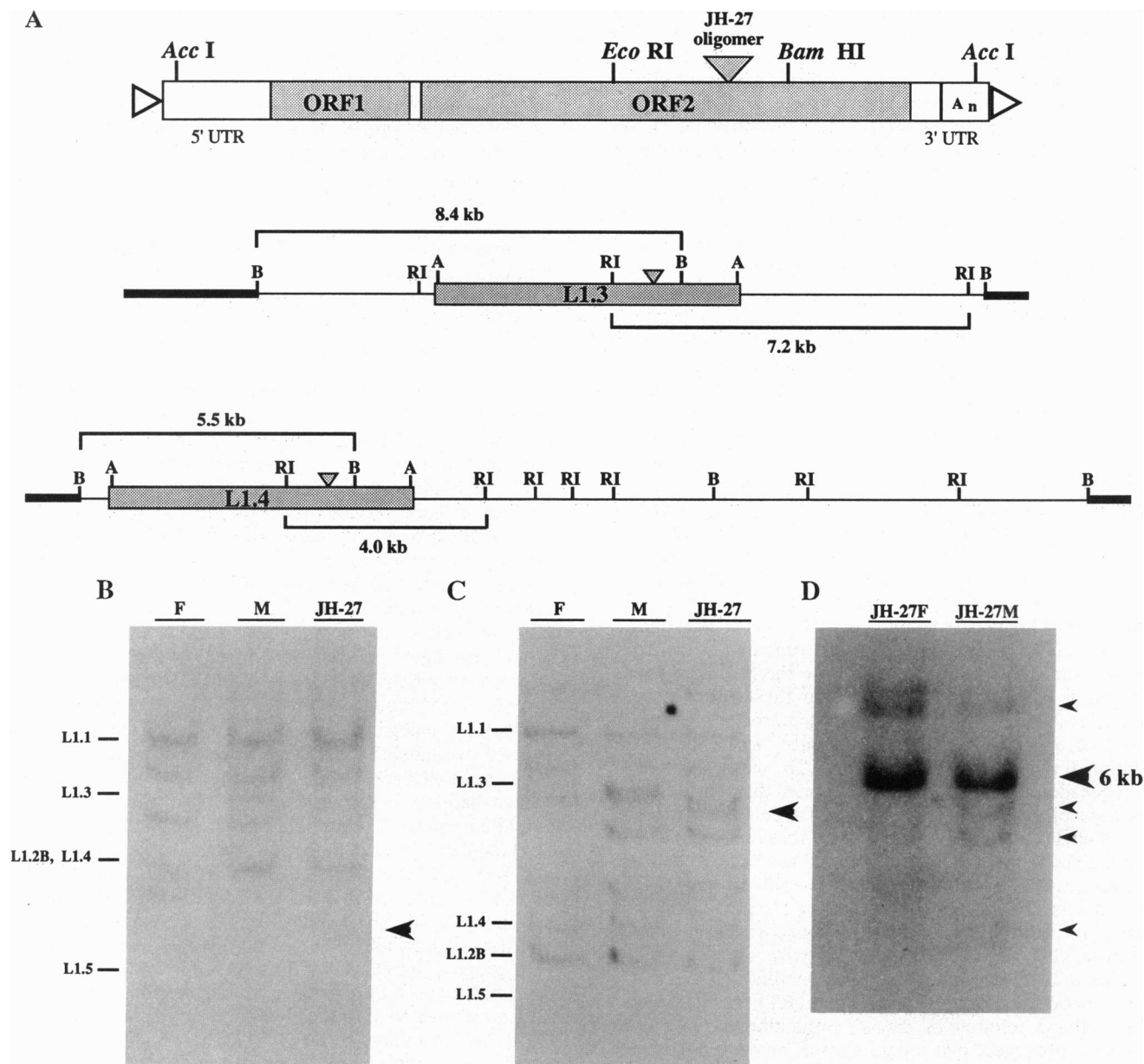


FIG. 1. (A) Structures of the consensus L1 element and restriction maps of  $\lambda$  phage inserts containing L1.3 and L1.4. For the L1 consensus element the 5' and 3' untranslated region (5' UTR and 3' UTR), the first and second ORFs (ORF1 and ORF2), and the poly(A) tail (A<sub>n</sub>) are indicated. The triangle between the *Eco*RI and *Bam*HI sites shows the location of the JH-27 oligomer sequence. Triangles flanking the element represent the target site duplication. For the restriction map of the phage containing L1.3, the JH-27 oligomer lies within an 8.4-kb *Bam*HI fragment and a 7.2-kb *Eco*RI fragment. For the restriction map of the phage containing the L1.4 element the JH-27 oligomer recognizes a 5.5-kb *Bam*HI fragment and a 4.0-kb *Eco*RI fragment. The thick lines represent the phage arms. Restriction sites: A, *Acc* I; B, *Bam*HI; RI, *Eco*RI. (B–D) Digests of genomic DNA hybridized with the JH-27 oligomer. Lanes containing DNA from the father of JH-27 (F or JH-27F), the mother of JH-27 (M or JH-27M), and JH-27 are shown. (B) *Bam*HI digests. (C) *Eco*RI digests. In B and C the arrow indicates the new insertion in JH-27. The fragments are identified by the L1 element of JH-27M from which they are derived. (D) *Acc* I digests. The large arrow indicates full-length 6-kb L1 elements. The small arrows indicate non-full-length elements in the genome of JH-27M. The 2.7-kb fragment is poorly seen in this *Acc* I digest of JH-27M DNA, but it was readily apparent in other *Acc* I digests of JH-27M DNA.

nucleotides 38 and 5962 of the 6022-nucleotide full-length L1 element (Fig. 1A). Therefore, digestion of genomic DNA with *Acc* I and subsequent hybridization with the JH-27 oligomer reveals an intense band at 6 kb which corresponds to full-length elements and single-copy bands of differing size which represent truncated copies of L1 elements within the JH-27 subfamily. When this analysis was applied to DNA from the parents of JH-27, we found that the genome of JH-27M contains four JH-27-specific L1 elements that are non-full-length—i.e., *Acc* I fragments differing in size from 6 kb (Fig. 1D). Since the JH-27M genome has eight JH-27-specific elements seen on the *Eco*RI digest and four of these elements are non-full-length on the *Acc* I digest, the remaining four elements must account for the band at 6 kb in the *Acc* I digest and be full length.

We have previously reported the isolation of two of these four full-length elements, L1.1 and L1.2, from the JH-27 subfamily (5). These elements were identified through screening of genomic libraries with three oligomers, the JH-27 oligomer, an oligomer specific for the 5' end of the consensus L1 element, and a 3' oligomer derived from a subset (subset Ta) of L1 cDNA sequences (12). L1.1 corresponds to the 11-kb *Bam*HI fragment and the 10-kb *Eco*RI fragment (Fig. 1 B and C, respectively). This element is located on chromosome 14, ends in a stretch of 54 adenines, and is flanked by a perfect 14-bp target site duplication. Analysis of the nucleotide sequence revealed that L1.1 had a 1-bp deletion at the beginning of the first ORF predicted by the genomic L1 consensus sequence (6), resulting in a frameshift mutation and a premature termination codon. [*In vitro* translation studies using an RNA transcript from L1.1 yielded no detectable ORF1 product (13).] Although the second reading frame of L1.1 is open, its 3' 3784 bp in common with the JH-27 insertion sequence differ from that sequence at six nucleotides. From these results we concluded that L1.1 is not the precursor of the JH-27 insertion and is not an active element.

L1.2 corresponds to the 5.5-kb *Bam*HI and 3.5-kb *Eco*RI fragments (Fig. 1 B and C). L1.2 is full-length, ends in 27 adenines, is flanked by a perfect 15-bp target site duplication, and was the first isolated L1 element found to contain the two ORFs predicted from the consensus sequence. Two common alleles of L1.2 (L1.2A and L1.2B) were characterized. L1.2A, isolated from a commercial genomic library, contained two nucleotide differences from the JH-27 insertion sequence. However, L1.2B, which was isolated from the partial *Bam*HI library constructed from JH-27M genomic DNA, was identical in its 3' 3784 bp to the JH-27 insertion sequence, indicating that it was the direct precursor of the JH-27 insertion.

With one active L1 element identified from the JH-27 subfamily, we sought to determine whether this subfamily contains other potentially active elements. We screened the partial *Bam*HI genomic library from JH-27M with the three oligomers mentioned above, and the two remaining full-length clones, L1.3 and L1.4, were isolated. L1.3 is represented by the 8.4-kb *Bam*HI fragment and the 7.2-kb *Eco*RI fragment (Fig. 1 B and C). Nucleotide sequence analysis revealed that L1.3 ends in a 37-nt poly(A) tail and is flanked by a 17- to 19-bp perfect target site duplication. Both ORF1 and ORF2 are intact, and there are five nucleotide differences between L1.3 and L1.2B, three of which fall within the 3' 3784 bp in common with the insertion in patient JH-27 (Fig. 2). These changes are A to G at positions 999 and 1869 in ORF1, G to A at positions 3076 (Gly to Arg) and 4055 (Arg to Gln) in ORF2, and C to T at position 5160 in ORF2. PCR amplification of a 3' region flanking L1.3 in somatic cell hybrid DNAs mapped L1.3 to chromosome 14 (see *Materials and Methods*).

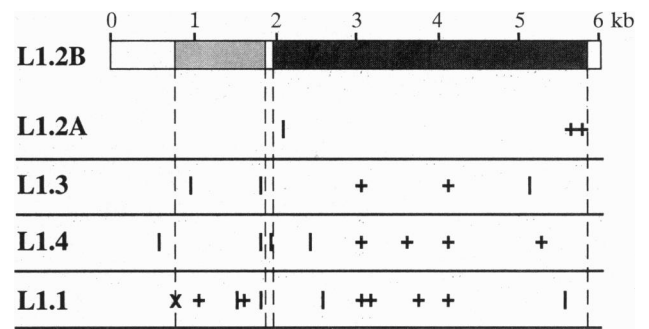


FIG. 2. Comparison of the nucleotide sequence of the full-length members of the JH-27 subfamily of L1 elements in JH-27M. The sequences are compared to the sequence of the active element, L1.2B. The length of L1.2B is 6022 bp. The broken lines separate the ORFs from untranslated regions. Vertical lines (|) indicate nucleotide differences, plus signs (+) represent amino acid substitutions, and an X indicates a nucleotide deletion. In L1.4, two nucleotide substitutions, at positions 3646 and 3647, result in a single amino acid substitution (Gln to Leu) and are represented by a single plus sign.

The restriction map of the fourth element isolated from the JH-27 subfamily, L1.4, indicated that it is represented by a 5.5-kb *Bam*HI fragment and a 4.0-kb *Eco*RI fragment (Fig. 1 B and C). Since the 5' *Bam*HI fragment of L1.2B is also 5.5 kb, the 5.5-kb band observed in the *Bam*HI digest is derived from two separate L1 elements. This explains the discrepancy between the number of bands in the *Bam*HI and *Eco*RI digests of JH-27M DNA identified by the JH-27 oligomer. L1.4 ends in a 31-nt poly(A) tail and is flanked by an 8- to 14-bp perfect target site duplication. ORF1 and ORF2 of this element are also intact. When the nucleotide sequence of L1.4 was compared to that of L1.2B, nine differences were found, and six of these fall within the 3784 bp of the JH-27 insertion sequence (Fig. 2). The differences 5' to the JH-27 insertion sequence include A to G at position 580 in the 5' untranslated region and at position 1869 in ORF1, and C to T at position 1934 in the inter-ORF region. The changes in ORF2 within the region of the JH-27 insertion sequence are C to G at position 2376, G to A at position 3076 (Gly to Arg), C to T at position 3646 and A to T at position 3647 within the same codon (Gln to Leu), G to A at position 4055 (Arg to Gln), and C to T at position 5207 (Ala to Val). We used PCR amplification of somatic cell hybrid DNAs to map L1.4 to chromosome 9 (see *Materials and Methods*).

A fifth L1 element (L1.5), which was grossly rearranged, was isolated from the JH-27M genomic library by using the three oligomers. It corresponds to the 2.6-kb *Bam*HI fragment, the 2.8-kb *Eco*RI fragment, and the 2.7-kb *Acc* I fragment (Fig. 1 B, C, and D).

We then digested genomic DNA from unrelated individuals with *Eco*RI and independently probed dried gels with oligomers flanking the 5' end of L1.3 (Fig. 3) and L1.4 (data not shown). We found that L1.3 is present in 7 of 50 chromosome 14s (gene frequency = 0.14) while L1.4 is present in 23 of 74 chromosome 9s (gene frequency = 0.3). Both L1.3 and L1.4 are found in all racial groups; however, the populations in which L1.3 and L1.4 attain the highest frequency (0.4) differ. The existence of these elements in all human populations with gene frequencies as high as 0.4 in one or more populations suggests that both of these elements have been present in the genome for 1–2 million years (14). As for the evolutionary relationship of L1.3 and L1.4 to the other full-length elements of the JH-27 subfamily (Fig. 2), one nucleotide change in ORF1 (position 1869) and two nucleotide changes in ORF2 (positions 3076 and 4055) are in common among L1.1, L1.3, and L1.4, suggesting that these elements may have a more recent common ancestor than L1.2B.

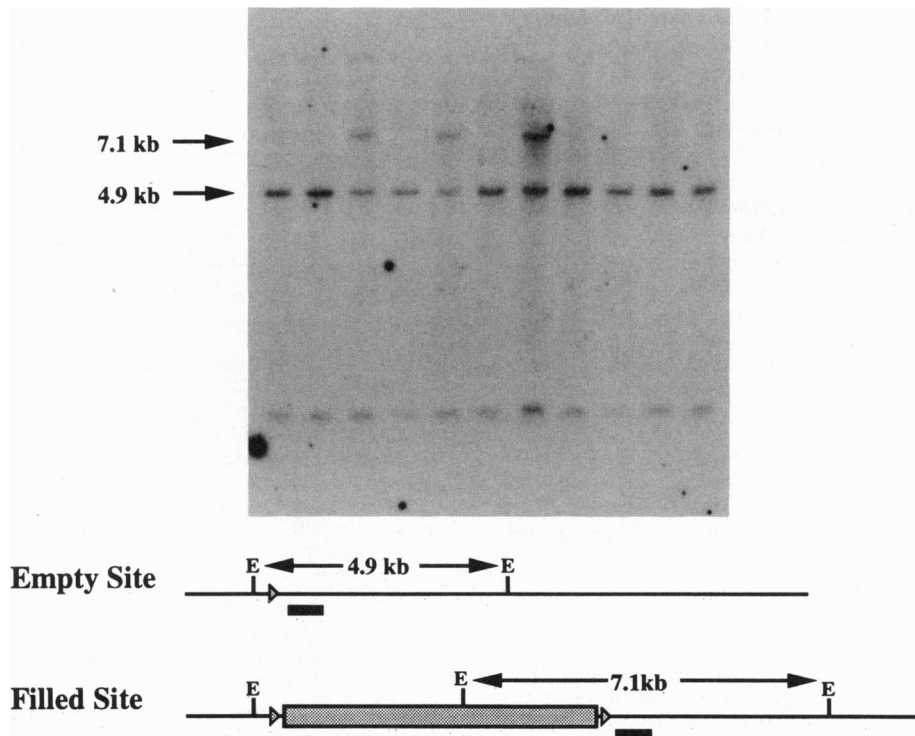


FIG. 3. The presence of L1.3 in the genome is polymorphic. Genomic DNA from unrelated individuals was digested with *EcoRI* and hybridized with an oligomer (depicted as a small black rectangle in the diagrams at the bottom) flanking the 3' end of L1.3. Empty sites (L1.3 absent) are represented by a 4.9-kb *EcoRI* fragment, while filled sites (L1.3 present) are indicated by a 7.1-kb *EcoRI* fragment. The triangle represents the target site of L1.3 integration. E, *EcoRI* restriction site.

## DISCUSSION

There are eight JH-27 subfamily elements in the genome of JH-27M, and four are known by digestion with *Acc I* to be truncated. Therefore, we have now isolated all four full-length, JH-27-related L1 elements from the genome of JH-27M. One of these elements, L1.2B, is known to be active. Two others, L1.3 and L1.4, have two ORFs and are potentially active. On the other hand, it is possible that a full-length L1 element whose nucleotide sequence encodes all of the activities necessary for retrotransposition is inactive due to its location in a heterochromatic, inactive region of the genome. The remaining element, L1.1, is believed to be inactive due to a 1-bp deletion resulting in a frameshift in ORF1. Although other full-length, JH-27-related elements do not exist in the genome of JH-27M, it is still possible that polymorphic, JH-27-related, full-length elements in addition to L1.1–L1.4 exist in the genomes of other individuals.

Among the full-length members of the JH-27 subfamily of L1 elements, the nucleotide sequences are very similar. In the 909 bp of the 5' untranslated region there is only one nucleotide difference that is observed in L1.4 (Fig. 2). It has been demonstrated that the 5' end of the L1 element contains a promoter region (15) with binding sites for a number of cellular transcription factors (16, 17). The base change in L1.4 occurs 3' to the important promoter sequences. The nucleotide sequences of L1.2A, L1.2B, L1.3, and L1.4 indicate that the ORF1 proteins encoded by all of these elements are identical. From *in vivo* studies in which ORF1 from L1.2A was transfected into Ntera2D1 human teratocarcinoma cells, it was determined that this ORF1 encodes a 40-kDa protein with the same electrophoretic mobility as the ORF1 protein encoded by endogenous L1 elements (7). Of the amino acid differences seen in ORF2 among L1.2B, L1.3, and L1.4 (Fig. 2) only one falls in an RT conserved sequence domain (18). The nucleotide changes in L1.4 within the same codon at positions 3646 and 3647 (CA to TT, Gln to Leu) fall

within domain 2 of RT sequence homology. It is unclear whether this conservative amino acid substitution, which is also present in RTs of other transposable elements and retroviruses, has any effect on RT activity (18).

Sequence comparison of the four full-length JH-27 subfamily elements over the 3784 bp of the JH-27 insertion confirms that L1.2B is the progenitor of the JH-27 insertion. The fact that the JH-27 insertion is identical in sequence to L1.2B indicates that no transcription or reverse transcription errors occurred during the process of retrotransposition. Therefore, the genomic L1.2B sequence represents the actual sequence of an active element. This would suggest, but not prove, that other recent L1 insertions which differ in sequence from the JH-27 insertion [the JH-25 (19) and JH-28 (4) insertions into the factor VIII gene, an insertion into the dystrophin gene (9), and the somatic insertion into the *APC* gene (10)] arose from other active progenitor elements, one or more of which could be an allele of L1.2.

Since most L1 elements in the genome are truncated at the 5' end, it was unexpected that three elements from the same subfamily would be full-length and potentially active. This result supports the hypothesis that a larger number of active L1 elements exist in a limited number of subfamilies, each of which is derived from a master L1 element (11). In fact, the reported sequences of five recently inserted L1 elements differ by less than 0.7%, while on average human L1 elements differ from each other by 5% (6), again suggesting that the presently active L1 elements are derived from retrotranspositions of a small number of master L1 elements. The situation for L1 elements would then be similar to that proposed for *Alu* sequences, which are thought to be propagated by a few master genes (11, 20, 21). Moreover, since L1.2B, L1.3, and L1.4 are flanked by target site duplications and therefore appear to be the product of retrotransposition events, second-generation retrotransposition is possible. That is, an active element can give rise to active and inactive

progeny and the active progeny, in turn, can give rise to new insertions.

Active L1 elements in the human genome may also participate in the retrotransposition of *Alu* sequences (22–24). Since *Alu* elements do not encode an RT activity and retrotransposition requires reverse transcription, it has been postulated that the RT encoded by active L1 elements is utilized in the retrotransposition of *Alu* sequences (8).

Retrotransposition has been successfully demonstrated in culture for the intracisternal A particle of mouse (25) and in whole organisms for the I factor of *Drosophila* (26, 27). Through use of an assay system to detect retrotransposition events, one could demonstrate directly whether or not a potentially active L1 element, such as L1.3 or L1.4, has the capacity to retrotranspose.

We thank S. E. Holmes and Dr. C. W. Fearon for technical assistance. This work was supported by a National Institutes of Health grant to H.H.K. and A.F.S.

- Hutchison, C. A., III, Hardies, S. C., Loeb, D. D., Shehee, W. R. & Edgell, M. H. (1989) in *Mobile DNA*, eds Berg, D. & Howe, M. (Am. Soc. Microbiol., Washington, DC), p. 593.
- Fanning, T. G. & Singer, M. F. (1987) *Biochim. Biophys. Acta* **910**, 203–212.
- Moyzis, R. K., Torney, D. C., Meyne, J., Buckingham, J. M., Wu, J.-R., Burks, C., Sirotkin, K. M. & Goad, W. B. (1989) *Genomics* **4**, 273–289.
- Kazazian, H. H., Jr., Wong, C., Youssoufian, H., Scott, A. F., Phillips, D. G. & Antonarakis, S. E. (1988) *Nature (London)* **332**, 164–166.
- Dombroski, B. A., Mathias, S. L., Nanthakumar, E., Scott, A. F. & Kazazian, H. H., Jr. (1991) *Science* **254**, 1805–1808.
- Scott, A. F., Schmeckpeper, B. J., Abdelrazik, M., Comey, C., O'Hara, B., Rossiter, J., Cooley, T., Heath, P., Smith, K. D. & Margolet, L. (1987) *Genomics* **1**, 113–125.
- Holmes, S. E., Singer, M. F. & Swergold, G. D. (1992) *J. Biol. Chem.* **267**, 19765–19768.
- Mathias, S. L., Scott, A. F., Kazazian, H. H., Jr., Boeke, J. & Gabriel, A. (1991) *Science* **254**, 1808–1810.
- Narita, N., Nishio, H., Kitoh, Y., Ishikawa, Y., Ishikawa, Y., Minami, R., Nakamura, H. & Matsuo, M. (1993) *J. Clin. Invest.* **91**, 1862–1867.
- Horii, A., Nakatsuru, S., Miyoshi, Y., Ichii, S., Nagase, H., Kato, Y., Yanagisawa, A. & Nakamura, Y. (1992) *Cancer Res.* **52**, 3231–3233.
- Deininger, P. L., Batzer, M. A., Hutchison, C. A., III, & Edgell, M. H. (1992) *Trends Genet.* **8**, 307–311.
- Skowronski, J. & Singer, M. F. (1986) *Cold Spring Harbor Symp. Quant. Biol.* **51**, 457–464.
- Liebold, D. M., Swergold, G. D., Singer, M. F., Thayer, R. E., Dombroski, B. A. & Fanning, T. G. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 6990–6994.
- Kimura, M. & Ohta, T. (1973) *Genetics* **75**, 199–212.
- Swergold, G. D. (1990) *Mol. Cell. Biol.* **10**, 6718–6729.
- Mathias, S. L. & Scott, A. F. (1993) *Biochem. Biophys. Res. Commun.* **191**, 625–632.
- Minakami, R., Kurose, K., Etoh, K., Furuhashi, Y., Hattori, M. & Sakaki, Y. (1992) *Nucleic Acids Res.* **20**, 3139–3145.
- Xiong, Y. & Eickbush, T. H. (1990) *EMBO J.* **9**, 3353–3363.
- Woods-Samuels, P., Wong, C., Mathias, S. L., Scott, A. F., Kazazian, H. H., Jr., & Antonarakis, S. E. (1989) *Genomics* **4**, 290–296.
- Deininger, P. L. & Slagel, V. K. (1988) *Mol. Cell. Biol.* **8**, 4566–4569.
- Matera, A. G., Hellmann, U. & Schmid, C. W. (1990) *Mol. Cell. Biol.* **10**, 5424–5432.
- Wallace, M. R., Andersen, L. B., Saulino, A. M., Gregory, T. W. & Collins, F. S. (1991) *Nature (London)* **353**, 864–866.
- Muratani, K., Hada, T., Yamamoto, Y., Kaneko, T., Shigeto, Y., Ohue, T., Furuyama, J. & Higashino, K. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 11315–11319.
- Vidaud, D., Vidaud, M., Bahnak, B. R., Siguret, V., Sanchez, S. G., Laurian, Y., Meyer, D., Goossens, M. & Lavergne, J. M. (1993) *Eur. J. Hum. Genet.* **1**, 30–36.
- Heidmann, O. & Heidmann, T. (1991) *Cell* **64**, 159–170.
- Jensen, S. & Heidmann, T. (1991) *EMBO J.* **10**, 1927–1937.
- Pelisson, A., Finnegan, D. J. & Bucheton, A. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 4907–4911.