Genetica **107:** 39–51, 1999.

Human L1 retrotransposition: insights and peculiarities learned from a cultured cell retrotransposition assay

John V. Moran
Departments of Human Genetics and Internal Medicine, The University of Michigan Medical School, Ann Arbor, MI 48109-0618, USA (E-mail: moranj@umich.edu)

Accepted 11 February 2000

Key words: cultured cells, LINE-1, mutagen, retrotransposon

Abstract

Long Interspersed Nuclear Elements (L1s or LINEs) are the most abundant retrotransposons in the human genome, and they comprise approximately 17% of DNA. L1 retrotransposition can be mutagenic, and deleterious insertions both in the germ-line and in somatic cells have resulted in disease. Recently, an assay was developed to monitor L1 retrotransposition in cultured human cells. This assay, for the first time, now allows for a systematic study of L1 retrotransposition at the molecular level. Here, I will review progress made in L1 biology during the past three years. In general, I will limit the discussion to studies conducted on human L1s. However, interesting parallels to rodent L1s and other non-LTR retrotransposons also will be discussed.

Introduction

Transposable elements are DNA sequences that can move (i.e., transpose) from one genomic location to another, and they are present in the genomes of virtually all eukaryotes. There are two general classes of transposable elements: transposons and retrotransposons. Transposons mobilize via a DNA intermediate and can transpose by a cut and paste, conservative, or replicative mechanisms (Bender & Kleckner, 1986; Kleckner, 1990). Retrotransposons mobilize via an RNA intermediate and transpose by a replicative mechanism termed retrotransposition (Boeke et al., 1985).

Retrotransposons can be subdivided into two general classes: Long-terminal repeat (LTR) and non-LTR retrotransposons. Autonomous LTR-retrotransposons, typified by the yeast Ty1 element, possess LTRs and resemble retroviruses in both their structure and replication mechanism (Boeke & Corces, 1989). However, LTR retrotransposons do not have an extracellular phase in their replication cycle because their envelope gene (*Env*) is either non-functional or absent.

Non-LTR retrotransposons (or poly A retrotransposons) lack LTRs and usually terminate in a polyadenylic acid (poly A) or short A-rich tail (Hutchison et al., 1989). They are a diverse group that includes the abundant mammalian long interspersed nuclear elements (LINEs or L1s), the CR1-like elements of birds and reptiles, the R1 and R2 elements of insects, the jockey, I, F, and G elements from *Drosophila*, and the Cin4 element of *Arabidopsis thaliana* (Xiong & Eickbush, 1988a). The reverse transcriptases (RTs) encoded by non-LTR retrotransposons also are related to the RTs encoded by retroplasmids from *Neurospora crassa* (Kuiper & Lambowitz, 1988), bacterial retrons (Lampson, Inouye & Inouye, 1991), mobile group II introns from yeast mitochondria (Kennell et al., 1993; Moran et al., 1995) and bacteria (Matsuura et al., 1997), and the catalytic subunit of telomerase (Nakamura & Cech, 1998). Thus, although quite different in structure, it appears that non-LTR retrotransposons share a common, ancient evolutionary ancestor.

The human genome also contains numerous copies of short interspersed nuclear elements (SINEs) and processed pseudogenes (Okada, 1991; Smit, 1996;

Deininger & Batzer, 1999). Many SINEs (e.g., Alu elements and tRNA-derived SINEs) and all processed pseudogenes structurally resemble non-LTR retrotransposons because they lack LTRs and introns, terminate in a polyA tail, and are flanked by variable-length target site duplications. However, SINEs and processed pseudogenes are non-autonomous retrotransposons because neither class of element encodes proteins that are required for their mobility.

We recently developed an assay to monitor L1 retrotransposition in cultured mammalian cells (Moran et al., 1996). The advent of this assay has allowed for the identification of retrotransposition-competent L1s (RC-L1s) in both the human and mouse genomes (Sassaman et al., 1997; DeBerardinis et al., 1998; Naas et al., 1998), and has yielded a number of unexpected findings about the impact of L1 retrotransposition on mammalian genomes (Moran et al., 1999a). Here, I will review progress made in L1 biology during the past three years.

Human L1s: Structure and abundance

L1s comprise approximately 17% of human DNA and are the most abundant retrotransposons in the genome (Smit, 1996). The overwhelming majority of L1s are unable to retrotranspose because they are 5′ truncated, rearranged or mutated (Grimaldi, Skowronski & Singer, 1984; Hutchison et al., 1989). However, an estimated 30–60 human L1s remain retrotransposition-competent (RC-L1s) (Sassaman et al., 1997). Human RC-L1s are 6.0 kb and contain a 5′ untranslated region (UTR) that harbors an internal promoter (Kurose et al., 1995; Minakami et al., 1992; Swergold, 1990), two non-overlapping open reading frames (ORF1 and ORF2) (Scott et al., 1987), and a 3′ UTR that ends in a poly A tail (Figure 1). In addition, L1s usually are flanked by variable-length target site duplications (Hutchison et al., 1989).

In vitro biochemical studies revealed that human ORF1 encodes a 40 kDa site-specific RNA binding protein (p40) that co-localizes with L1 RNA in cytoplasmic ribonucleoprotein particles (RNPs) (Hohjoh & Singer 1996, 1997; Holmes et al., 1992). RNAse T1 mapping experiments demonstrated that p40 binds two related AT-rich sites in L1 RNA. The first binding site is located just downstream of the putative ORF2 initiation codon between positions 1991–2039 in L1 RNA, while the second binding site is located between the RT and C domains at positions 4839–4875 of L1 RNA (Hohjoh & Singer, 1997). Notably, both of these sites are contained within the largest RNA fragments generated in the RNAse T1 digest of L1 RNA. Thus, it remains possible that the p40 binding specificity is an artifact of the *in vitro* assay.

The mouse L1 ORF1-encoded protein also is localized with L1 RNA in cytoplasmic RNPs (Martin, 1991). In contrast to the situation with human p40, *in vitro* experiments show that the mouse ORF1-encoded protein binds both RNA and DNA in a non-sequence specific manner (Kolosha & Martin, 1997). Although it remains unclear whether the mouse ORF1-encoded protein binds L1 RNA at specific sequences, the above studies have led to the hypothesis that cytoplasmic RNPs are intermediates in L1 retrotransposition (see Figure 2).

L1 ORF2 encodes a multifunctional protein containing at least four conserved domains, which are, in order from N to C terminus, an endonuclease (EN) domain (Feng et al., 1996), a domain of unknown function (Z) (Clements & Singer 1998), a reverse transcriptase (RT) domain (Mathias et al., 1991), and a carboxyl-terminal cysteine rich domain (C) (Fanning & Singer, 1987). Genetic and biochemical analyses revealed that the L1 ORF2-encoded protein contains both RT and endonuclease activities (Feng et al., 1996; Mathias et al., 1991). However, how ORF2 is translated, and whether the primary translation product undergoes post-translational processing remains unknown.

Human L1s and disease

L1 retrotranspositions first were recognized as causal agents of disease in 1988 when two separate insertions of 5′ truncated L1s were found to disrupt the factor VIII gene, resulting in hemophilia A (Kazazian et al., 1988). Since then, deleterious L1 insertions in the human germ-line or very early in development (i.e., before germ cell migration) were found to result in a variety of genetic disorders (Table 1; reviewed in Kazazian, 1998, 1999; Kazazian & Moran, 1998). For example, four separate truncated L1 insertions into the *dystrophin* gene were shown to result in muscular dystrophy (Bakker & van Ommen, personal communication; Narita et al., 1993; Holmes et al., 1994; Yoshida et al., 1998); whereas the insertion of different full-length L1s into the β-globin or RP2 gene were found to result in β-thalessemia and X-linked retinitis pigmentosa, respectively (Divoky et al., 1996;

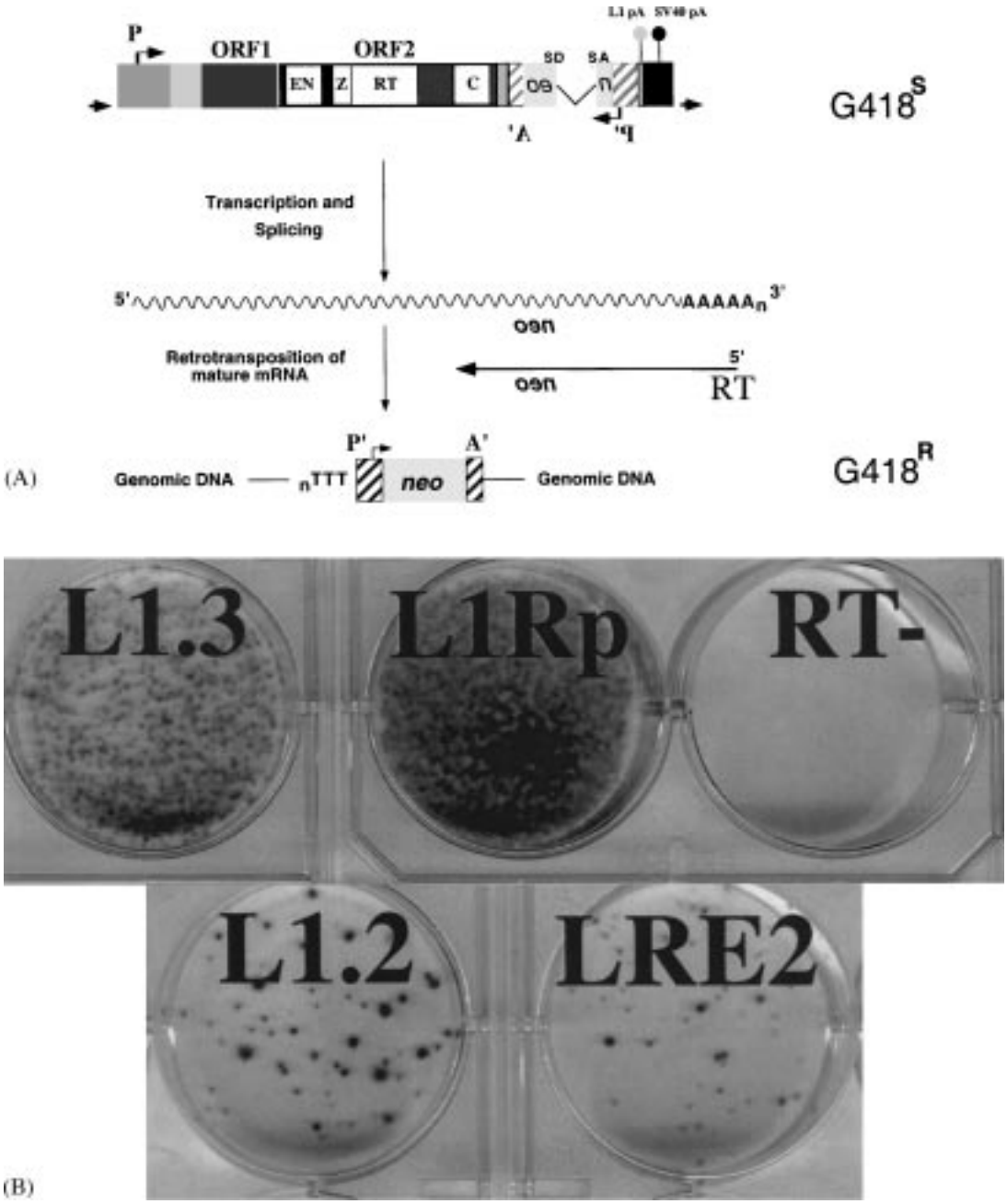

Figure 1. An assay to detect L1 retrotransposition in cultured mammalian cells. (A) The rationale of the assay. Candidate L1s are tagged with an indicator cassette (*mneoI*) designed to detect retrotransposition events. The cassette consists of a backward copy of the neomycin phosphotransferase gene (*neo*), which contains its own promoter (P′) and polyadenylation signal (A′). The backward *neo* gene is interrupted by an intron in the same transcriptional orientation of the L1 and the splice donor (SD) and splice acceptor sites (SA) are indicated. Transcription of L1 RNA from either its own promoter in the 5′UTR (light gray box) or a heterologous promoter (P) and subsequent RNA splicing results in the production of a polyadenylated mRNA. The ORF1- and ORF2-encoded proteins can be translated from the mRNA, but the spliced *neo* gene cannot be translated because it is backward. G418-resistant ($G418^R$) colonies only will arise if the mature L1 mRNA is reverse transcribed (RT) and integrated at a new genomic location. The retrotransposed indicator gene then can be expressed from its own promoter (P′) to produce a transcript, which can be translated to generate a functional neomycin phosphotransferase protein. Conserved domains in the L1-encoded proteins are described in the text. The variable length target site duplications flanking an L1 are denoted with arrows. (B) Results of a typical retrotransposition assay. Results from a typical assay with four different RC-L1s (L1.3, $L1_{Rp}$ L1.2, and LRE2) are shown. A negative control construct containing a point mutation in the L1 RT domain also is depicted.

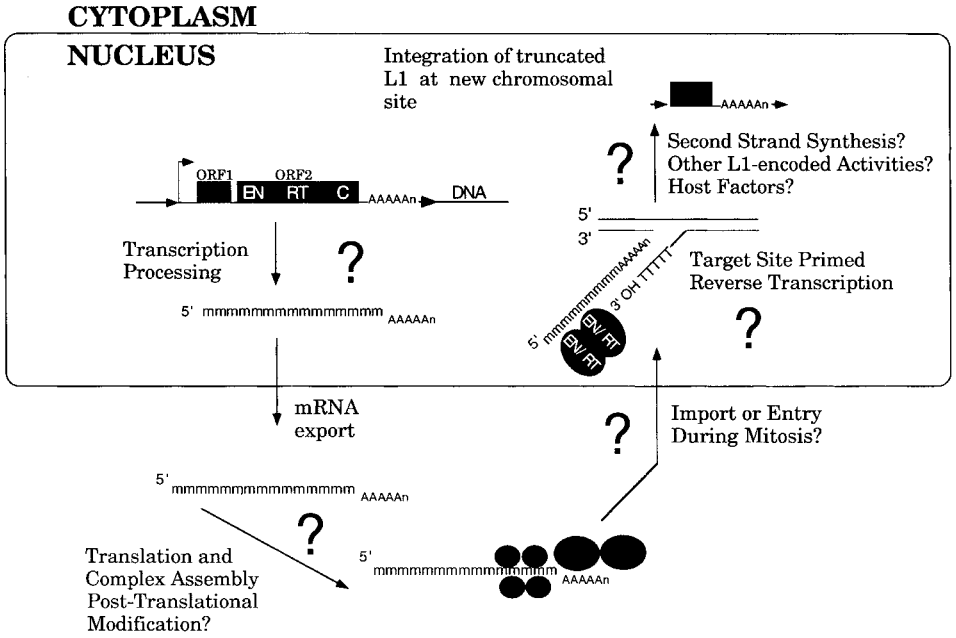

Figure 2. A working model for L1 retrotransposition. This model is adapted from Kazazian and Moran (1998). Question marks indicate steps in the pathway that remain unknown. Specific details are provided in the text.

Schwahn et al., 1998). Moreover, somatic insertions of truncated L1s into the APC tumor suppressor gene and the *c-myc* proto-oncogene are implicated in colon and breast cancer (Morse et al., 1988; Miki et al., 1992). Interestingly, the overwhelming majority of disease producing L1 insertions are derived from a minor subfamily of expressed L1s denoted the Ta subset (Skowronski, Fanning & Singer, 1988). Thus, the Ta subset L1s likely account for the bulk of RC- L1s in the human genome.

The finding that L1 can induce mutations in somatic cells is significant because it indicates that L1 retrotransposition can lead to tumorigenesis (Miki et al., 1992). Intriguingly, biochemical and immunohistological studies have shown that L1 RNA and p40 are elevated in certain epithelial-derived tumors (Bratthauer & Fanning, 1992; Bratthauer & Fanning, 1993). Moreover, the L1 5′UTR is hypomethylated in certain p40 producing cell lines and malignant tumors (Alves, Tatro & Fanning, 1996; Thayer, Singer & Fanning, 1993). Whether L1 retrotransposition actually is increased in tumors is unknown. However, the above data suggest that unregulated L1 retrotransposition may accelerate tumor progression.

In addition to being autonomously mobile mutagens, it is hypothesized that the proteins encoded by RC-L1s mobilize certain SINEs (e.g., Alu elements and tRNA derived SINEs) and processed pseudogenes, which comprise another 10% of human DNA (Boeke, 1997; Jurka, 1997; Okada et al., 1997; Kazazian & Moran, 1998; Gilbert & Labuda, 1999). Alu retrotransposition can be mutagenic and 16 *de novo* Alu retrotranspositions either in the germ line or during early development have been found to cause human disease (reviewed in Deininger & Batzer 1999; Kazazian, 1998; Kazazian & Moran, 1998) (Table 1). Thus, elucidation of the mechanism of L1 retrotransposition likely will lead to an understanding of the retrotransposition mechanisms utilized by other nonautonomous retrotransposons.

It remains difficult to estimate the fraction of human disease-associated mutations due to the insertion of autonomous or non-autonomous non-LTR retrotransposons because most mutation-detection strategies utilize PCR and fail to detect large insertions. Moreover, retrotransposition events into autosomes may remain unnoticed because of the hemizygous nature of the mutation. However, as of February

Table 1. Disease-producing Alu and L1 retrotransposon insertions

Disease-producing retrotransposon insertions							
L1s					Alu		
Disrupted gene	Disease	Insertion size	Ta subset	L1-mediated transduction	Disrupted gene	Disease	Subfamily
Factor VIII	Hemophilia A	3.8 kb	Yes	No	Factor IX	Hemophilia B	Ya5
		2.2 kb	No	No	NF1	Neurofibromatosis	Ya5
		681 bp	Yes	No	FGFR2 (Two insertions)	Apert Syndrome	Ya5, Yb8
APC	Colon cancer	538 bp	Yes	Yes	APC	Hereditary colon cancer	Yb8
Dystrophin	Duchenne	608 bp	Yes	No	XLA	X-linked	Y
	muscular	878 bp	Yes	Yes		immunodeficiency	
	dystrophy	2.0 kb	Yes	No		syndromes	
		524 bp	Yes	No	XSCID	X-linked	Ya5
β-globin	β-Thalessemia	6.0 kb	Yes	No		immunodeficiency	
RP2	X-linked	6.0 kb	Yes	No		syndromes	
	retinitis-pigmentosa				BRCA2	Breast cancer	Y
CYBB	Chronic	1.7 kb	Yes	Yes	ChE	Acholinesteremia	Yb8
	ganulomatous	1.0 kb	Yes	No	C1 inhibitor	Hypocalciuric	Y
FCMD	Fukuyama-type	1.2 kb	Yes	No		hypercalcemia	
	congenital muscular				Huntington	Huntington	–
	dystrophy				disease gene	disease	
					CaR	Angioedema	Ya4
					EYA1	Branchio-oto-renal	Ya5
						syndrome	
					GK	Gycerol kinase	Y
						deficiency	
					PBGD	Acute intermittent	Ya5
						porphyria	
					MLVI-2	Leukemia	Ya5

Dr. Haig H. Kazazian maintains a complete database of new disease-producing retrotransposon insertions at the following address: www.med.upenn.edu/genetics/faculty/primary/hkazazian 2.html. The disrupted genes and resulting diseases are noted. The subfamily that gave rise to the retrotransposon insertion in noted. Two additional SINE disease-producing insertions into the FCMB (Fukuyama muscular dystrophy gene) and the BTK (Bruton X-linked a gammaglobulinemia gene) are not included in the table, but are available at the above internet address.

2000, 31 of the 19,687 spontaneous mutations (1 in 635) characterized in humans are due to the insertion of non-LTR retrotransposons (Table 1 and legend).

An assay to detect L1 retrotransposition in cultured cells

In 1996, we developed a genetic assay to study L1 retrotransposition in cultured human cells (Moran et al., 1996). First, we tested whether the likely progenitors of disease producing insertions in the factor VIII and dystrophin genes (L1.2 and LRE2) were indeed autonomous retrotransposons (Dombroski et al., 1991; Holmes et al., 1994). To accomplish this, a reporter cassette (*mneoI*) designed to detect rare retrotransposition events was subcloned into their 3′ UTR to create *L1.2mneoI* and *LRE2mneoI*, respectively (Freeman, Goodchild & Mager, 1994) (Figure 1A). The *mneoI* reporter cassette consists of an antisense copy of a selectable marker (*neo*), the SV40 promoter (P′), and a polyadenylation signal (A′). In addition, the *neo* gene is disrupted by an intron (IVS 2 of the γ-globin gene) in the opposite transcriptional orientation. This arrangement ensures that G418-resistant cells ($G418^R$) will arise only when a transcript initi-

ated from the promoter driving *L1.2mneoI* expression (P) is spliced, reverse transcribed, reintegrated into chromosomal DNA, and expressed from promoter P′. Transcripts that originate from P′ cannot be spliced, the neo gene product cannot be synthesized, and the cells will remain sensitive to G418 ($G418^S$).

Using the above assay, we demonstrated that both L1.2 and LRE2 could retrotranspose from an extrachromosomal episome into HeLa genomic DNA at a high frequency (Moran et al., 1996). Point mutations in conserved domains of the L1.2 ORF1 and L1.2 ORF2 encoded-proteins reduced retrotransposition by 2–3 orders of magnitude (Figure 1B), and the characterization of four *L1.2mneoI* integration events revealed the hallmarks of retrotransposition (Moran et al., 1996). However, the target site duplications flanking two of these newly inserted elements were somewhat unusual. They were longer (34 bp and 214 bp, respectively) than the target site duplications flanking genomic L1s. Thus, it remains formally possible that the placement of the *mneoI* indicator cassette in the L1 3′UTR adversely affects the integration process. However, our data demonstrate that we can, for the most part, faithfully reconstitute L1 retrotransposition in cultured human cells.

Applications of the cultured cell retrotransposition assay

Identification of other L1s in the human and mouse genomes

An immediate use of the cultured cell retrotransposition was to identify other RC-L1s from both the human and mouse genomes. During the characterization of L1.2, 15 other full-length L1s belonging to the Ta subset also were identified (Dombroski et al., 1991; Dombroski, Scott & Kazazian, 1993; Sassaman et al., 1997). Nine of those elements encoded a reverse transcriptase activity, as monitored in a heterologous yeast expression system (Dombroski et al., 1994; Sassaman et al., 1997), and nine contained intact open reading frames (Sassaman et al., 1997). Subsequent analysis revealed that five of those L1s were retrotransposition-competent, and those data were used to estimate that there are 30–60 RC-L1s present in the human genome.

Two L1s, L1.3 and L1.4, yielded $G418^R$ foci at frequencies ~10-fold greater than L1.2. Interestingly, L1.3 and L1.2 only differ at four amino acid positions (Dombroski, Scott & Kazazian, 1993; Sassaman et al., 1997). Thus, it is likely that all or a subset of these amino acid substitutions account for the dramatic difference in their retrotransposition. One intriguing possibility is that these amino acid substitutions (or a subset of substitutions) alter the processivity of the ORF2-encoded RT.

Recently, two other full-length human L1s ($L1_{\beta thal}$ and $L1_{Rp}$) were identified as mutagenic insertions into the β-globin and retinitis pigmentosa-2 genes, respectively (Divoky et al., 1996; Schwahn et al., 1998). Subsequent studies revealed that both $L1_{\beta thal}$ and $L1_{Rp}$ could retrotranspose in cultured cells (Kimberland et al., 1999). Thus, to date, nine active human L1s have been identified. Though it remains possible that some of these L1s will be immobile because they reside in non-transcribed chromatin, it now is apparent that the human genome contains numerous RC-L1s. Clearly, the completion of the human genome project will lead to the discovery of even more RC-L1s.

The identification of RC-L1s in the mouse genome was due, in a large part, to serendipity. In 1994, the spastic mouse was found to result from a full-length L1 insertion ($L1_{spa}$) into the glycine receptor β-subunit gene (Kingsmore et al., 1994; Mulhardt et al., 1994); whereas in 1996, the Orleans reeler mouse was found to result from a full-length L1 insertion ($L1_{orl}$) into the reeler gene (Takahara et al., 1996). Since both of those L1s contained open reading frames, they became immediate candidates for retrotransposition-competent elements. Both $L1_{spa}$ and $L1_{orl}$ readily retrotransposed in the cultured cells, and DNA sequence analysis revealed that those elements belonged to a new subfamily (T_f) of mouse L1s (Naas et al., 1998; Saxton & Martin, 1998). Interestingly, subsequent studies revealed that seven of 11 (63%) full-length T_f elements randomly isolated from a mouse genomic DNA library were able to retrotranspose in cultured cells (DeBerardinis et al., 1998). Since there are approximately 2400 full-length T_f elements in the genome of *mus domesticus*, these data suggest that up to 1500 mouse T_f elements remain retrotransposition-competent.

We also demonstrated that human L1s can retrotranspose in cultured mouse L cells (Moran et al., 1996) and that mouse L1s readily retrotranspose in human Hela cells (Naas et al., 1998). Although it is possible that L1 encodes all of the functions required for its retrotransposition, it is likely that retrotransposition is facilitated by host-encoded cellular factors. If

so, our data would suggest that some of these factors are conserved between rodents and humans.

Identification of other essential and non-essential sequences within L1

The cultured cell retrotransposition assay also has been instrumental in identifying conserved functional domains within the L1-encoded proteins required for retrotransposition. In collaboration with the Boeke laboratory, we demonstrated that the amino terminus of L1 ORF2 encodes a novel endonuclease (EN) required for retrotransposition (Feng et al., 1996). The endonuclease motif resembles the apurinic/apyrimidinic (AP) endonucleases, but shows no preference for abasic sites. Instead, L1 EN preferentially cleaves the sequence 5′PO_4-TTT/AA-3′OH (where / denotes the insertion site), which resembles many *in vivo* L1 integration sites (Feng et al., 1996; Moran et al., 1996). The purified protein can generate site-specific 5′ PO_4 and 3′OH nicks either in supercoiled plasmids or in oligonucleotide substrates containing the recognition site (Cost & Boeke, 1998; Feng et al., 1996). Moreover, site-directed point mutations in conserved residues of L1 EN abolish this nicking activity *in vitro* and reduce L1 retrotransposition in cultured cells by 2–3 orders of magnitude (Feng et al., 1996). Thus, we propose that L1 is a site-specific retrotransposon that integrates preferentially at DNA sequences cleaved by L1 EN by a template DNA-primed reverse transcription mechanism (TPRT) (Luan et al., 1993; Figure 2; see below).

A parallel analysis further revealed that the L1 endonuclease-like domain is conserved among non-LTR retrotransposons from a wide variety of organisms, suggesting that those elements also retrotranspose by a similar TPRT mechanism (Martin et al., 1995; Feng et al., 1996). Indeed, the EN domain of the R1Bm element from *Bombyx mori* encodes a site-specific endonuclease, which likely is needed for its retrotransposition (Feng, Schumann & Boeke, 1998). Notably, though our study shows that L1 encodes an endonuclease likely required for its retrotransposition, it remains possible that L1 sometimes can utilize pre-existing nicks in chromosomal DNA to mediate its integration through an endonuclease independent pathway (Branciforte & Martin, 1994).

Comparative biological approaches also have been used to identify both conserved regions in L1 RNA and conserved amino acid residues in the L1-encoded proteins, which were hypothesized to be important for retrotransposition (Fanning & Singer, 1987; Xiong & Eickbush, 1990). The cultured cell assay now allows a rapid means to test many of those hypotheses experimentally. For example, we showed that conserved amino acids in the carboxyl terminus of the L1 ORF1-encoded protein are required for retrotransposition. Since ORF1 encodes an RNA binding protein, it is tempting to speculate that these amino acids play a crucial role in binding p40 to its nucleic acid template (Moran et al., 1996). We also demonstrated that the L1-ORF2 cysteine-rich domain provides a function required for retrotransposition, which is distinct from the L1 RT or L1 EN activity (Moran et al., 1999b). Again, it is tempting to speculate that these mutations affect binding of ORF2 to either L1 RNA or genomic DNA (Moran et al., 1996).

Interestingly, our studies also uncovered regions of L1, which apparently are dispensable for retrotransposition. For example, deletion of a conserved polypurine tract within the L1 3′UTR did not effect L1 retrotransposition (Moran et al., 1996). This finding remains somewhat enigmatic because: 1) the 3′UTRs of rat and mouse L1s also contain a conserved polypurine tract (Usdin & Furano, 1989); and (2) the 3′UTR of other non-LTR retrotransposons (notably R2Bm from *Bombyx mori*) functions as a *cis*-acting docking site for binding the element encoded RT *in vitro* (Luan et al., 1993). By the strictest definition, our data suggest that the 3′UTR neither is required for human L1 retrotransposition in cultured cells nor serves as a required binding site for the L1 RT. However, it remains formally possible that the L1 3′UTR functions *in vivo* to: (1) bind factors to help define the L1 pA site; (2) function in L1 RNA localization or translation; and (3) aid in the L1 integration process. Indeed, the unorthodox target site duplications observed in two of four characterized integration events (see previous section) may occur because the L1 3′UTR is interrupted by the *mneoI* indicator gene.

We also demonstrated that the L1 5′UTR could be replaced with a heterologous viral promoter (either the cytomegalovirus immediate early promoter (CMV) or the Rous sarcoma virus LTR (RSV)) and that the resultant constructs could retrotranspose efficiently (Moran et al., 1996). Therefore, we conclude that an RNA polymerase II transcript can serve as an intermediate in L1 retrotransposition. Moreover, our data show that L1 can retrotranspose from a heterologous transcript whose 5′ end is derived from non-L1 DNA sequences (Moran et al., 1996). Thus, it is possible that L1 can retrotranspose non-L1 sequences derived

from its 5′ flank to new genomic locations. Together, the above data are in accord with previous studies and suggest that the major function of the L1 5′ UTR is to serve as a promoter for L1 transcription (Swergold, 1990; Minakami et al., 1992; Kurose et al., 1995).

Unexpected findings uncovered from the cultured cell retrotransposition assay

The cultured cell assay also has yielded unexpected information about L1 retrotransposition. First, we found that 5–10% of new L1 retrotransposition events in cultured cells occurs into the introns of actively transcribed genes (Moran, DeBerardinis & Kazazian, 1999a). Data from the Washington University genome center estimate that 15% of human DNA consist of genes (exons plus introns; Ian Korf, personal communication). Thus, our findings suggest there is not an active mechanism to preclude L1 from retrotransposing into genes.

The finding that L1 could readily retrotranspose into genes was somewhat surprising based on earlier studies, which demonstrated that L1 predominates in AT rich non-transcribed heterochromatin (Korenberg & Rykowski, 1988). However, those studies are subject to selective pressures that have operated on the accumulation of L1s during human genome evolution. Clearly, L1 can be a mutagen. Thus, over evolutionary time, it is likely that L1 retrotransposition events into intergenic regions would be tolerated more easily than L1 insertions into genes. This effective 'neutrality' could, in principle, lead to the preponderance of L1 sequences in heterochromatin. Alternatively, there may exist yet undiscovered recombination-based mechanisms to remove L1s (as well as other DNA sequences) from the human genome.

Since our study detects only new L1 retrotransposition events, it likely reflects L1 integration preferences quite accurately. The completion of the human genome project undoubtedly will enable a thorough analysis of the genomic distribution of L1 sequences. In fact, it will be interesting to determine whether recent L1 insertions (such as those derived from the Ta subset) are dispersed randomly throughout the genome or are clustered in specific chromosomal regions.

We also demonstrated that L1 could efficiently retrotranspose non-L1 DNA derived from its 3′ flank to new genomic locations (Figure 3; Moran, DeBerardinis & Kazazian, 1999a). In principle, this process (termed L1-mediated transduction) provides a simple mechanism to introduce non-L1 sequences (such as exons or promoters) into existing genes. L1-mediated transduction is unique because:

(1) it does not depend on homologous DNA sequences;
(2) the relative genomic positions of the 'shuffled' sequences essentially are irrelevant;
(3) it occurs via an RNA intermediate; thus, the original 'donor sequence' will not be changed by the process; and
(4) it allows a mechanism to exchange delimited amounts of genetic information between non-homologous chromosomes, providing a powerful means to generate diversity in randomly mating sexual populations.

Furthermore, since most L1s are 5′ truncated, it is likely that some L1-mediated transduction events leave no trace of L1 sequence (see middle panel in Figure 3). Indeed, our data predict that some processed pseudogenes and microsatellite markers may have arisen or been amplified by L1-mediated transduction.

Initially the finding that L1 could readily retrotranspose DNA derived from 3′ flanking sequences was a surprise. However, upon reflection, our data is explained rather easily. First, L1 undergoes unorthodox polyadenylation because the polyA tail is added directly to the 5′- AAUAAA-3′ CPSF (cleavage and polyadenylation specificity factor) binding site (Colgan & Manley, 1997). In addition, conserved elements that reside downstream of standard RNA polymerase II cleavage and polyA addition sites clearly are absent in L1. Thus, we propose that the L1 polyadenylation (L1 pA) site is 'weak', and can be bypassed if 'stronger' pA sites are present in 3′ flanking genomic sequences. Indeed, perhaps the presence of a 'weak' pA site allows L1s to reside within introns and not wreak havoc on gene expression (Moran, DeBerardinis & Kazazian, 1999a). Notably, the ability of L1 readthrough transcripts to retrotranspose further supports our hypothesis that the L1 poly A tail, and not RNA sequences in the 3′ end of L1 RNA, is critical for retrotransposition (Moran et al., 1996).

Retrotranspositions derived from readthrough L1 transcripts also have been identified *in vivo* (Table 1). The first occurrence was published in 1992, when a mutagenic insertion into the APC gene was accompanied by 23 bp of non-L1 sequence (Miki et al., 1992). However, these data were clearly overlooked. The phenomenon of L1-mediated transduction became apparent in 1994, when a mutagenic insertion into the

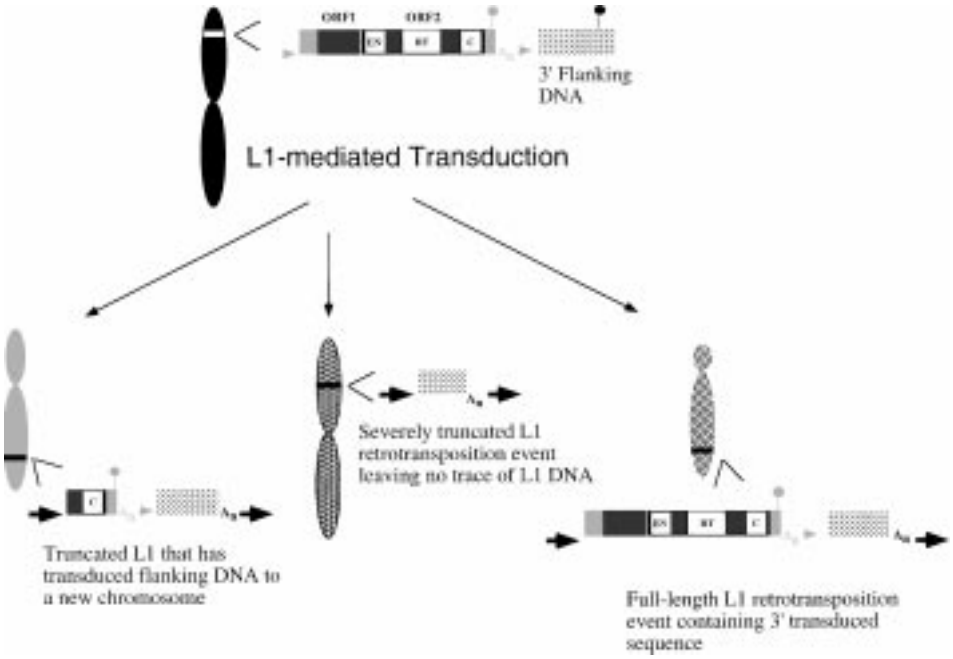

Figure 3. L1-mediated transduction (adapted form Moran et al., 1999). A retrotransposition-competent L1 resides at a chromosomal location denoted by the white bar at the top of the figure. The native L1 pA site and a fortuitous pA site in 3′ flanking DNA are denoted by the gray and black lolipops, respectively. In principle, three types of L1-mediated transduction events can occur if the L1pA site is bypassed and the pA site in flanking DNA is utilized. Each type of event is described in the text and is noted in the figure. The gray arrows flanking L1 (top) represent the original target site duplications flanking the element. New target site duplications generated by the retrotransposition of a readthrough L1 transcript are denoted with the black arrows (bottom figures). Non-homologous chromosomes are indicated by the different shading patterns.

dystrophin gene was accompanied by 489 bp of non-L1 sequence (Holmes et al., 1994). More recently, a mutagenic L1 insertion causing chronic granulomatosis disease also likely arose from the retrotransposition of a readthrough transcript (Meischl, Boer & Roos, 1998). Since three of 13 (23%) of L1 insertions in human disease genes are derived from readthrough transcripts, we predict that a detailed analysis of mammalian genomes should reveal many more examples of L1-mediated transduction events. Indeed, a preliminary examination of human DNA sequences deposited in public databases suggests that up to 15% of L1s arose via the retrotransposition of a readthrough transcript (Goodier, Ostertag & Kazazian, 2000).

Interestingly, L1-mediated transduction does not seem specific for members of the Ta subset. Two ancient human L1 insertions likely resulted from the retrotransposition of a readthrough transcript (McNaughton et al., 1997; Rozmahel et al., 1997). Moreover, a full-length A-type L1 transcript from RNPs from mouse F9 cells was shown to contain about 1 kb of 3′ flanking sequence (Martin, 1994). Subsequent analyses demonstrated that this transcript was derived from a contiguous sequence in mouse genomic DNA, indicating that it likely represents an authentic example of L1-mediated transduction. Thus, we conclude that L1-mediated transduction occurs in mouse and man, and it is likely that the process has an ancient origin. However, the extent to which this process impacted mammalian genome evolution, and whether L1-mediated transduction actually can lead to the evolution of new genes *in vivo* remains to be determined.

A working model for L1 retrotransposition

The similarity of the RTs encoded by non-LTR retrotransposons from different organisms suggests that these elements retrotranspose by similar mechanisms. The best model for non-LTR retrotransposition comes from studies of a site-specific non-LTR retrotransposon (R2Bm) found in the ribosomal DNA of the silkworm *Bombyx mori* (Luan et al., 1993). *In vitro* biochemical studies conducted in the Eickbush laboratory showed that R2Bm encodes a single protein

with both sequence specific endonuclease and reverse transcriptase activities (Luan et al., 1993; Xiong & Eickbush, 1988b). The R2Bm protein can generate a site-specific endonucleolytic nick in target DNA to liberate a 3′ hydroxyl (OH) that subsequently is used by the R2Bm protein as a primer to reverse transcribe the 3′ end of R2Bm RNA (Luan et al., 1993). This mode of priming is termed target site DNA-primed reverse transcription (TPRT).

Studies of mobile group II introns support the general TPRT retrotransposition model for non-LTR retrotransposon. Genetic experiments demonstrate that certain group II introns are site-specific retrotransposons, which can move from intron-containing alleles to intronless alleles by a process termed homing (Lazowska, Meunier & Macarde, 1994; Moran et al., 1995). Mobile group II introns encode a multifunctional protein that promotes their splicing (maturase activity) (Carignani et al., 1983; Moran et al., 1994) and mobility (reverse transcriptase and site-specific endonuclease activities) (Kennell et al., 1993; Moran et al., 1995; Zimmerly et al., 1995a, b). Mobility is initiated by an endonuclease containing the intron-encoded protein and the excised intron RNA and occurs by a two step process (Zimmerly et al., 1995a, b). First, the sense strand of recipient DNA is cleaved precisely at the exon junction by a partial or a complete reverse splicing of the intron RNA. Second, the antisense strand is cleaved 10 nucleotides downstream by the intron-encoded endonuclease to liberate a 3′ hydroxyl residue, which is used as a primer for the reverse transcription of the intron RNA by the intron-encoded RT.

Our work supports the notion that L1 retrotransposition also occurs by TPRT (Kazazian & Moran, 1998) (see Figure 2). First, RC-L1s are transcribed from an internal promoter located in the 5′UTR (Kurose et al., 1995; Minakami et al., 1992; Swergold, 1990). Next, the bicistronic L1 RNA is transported to the cytoplasm, where it is translated (Leibold et al., 1990; McMillan & Singer 1993). It is hypothesized that the L1 encoded proteins bind to the RNA from which they were translated (Moran et al., unpublished data) to form a cytoplasmic RNP – a proposed intermediate in the retrotransposition process (Martin, 1991; Hohjoh & Singer, 1996, 1997). However, how RNPs are formed and whether the particles contain the ORF2-encoded protein remains uncertain.

After its formation, the L1 RNP presumably is transported to the nucleus to carry out retrotransposition. It is not known whether p40 and the ORF2-encoded protein are critical to the transport of L1 RNA into the nucleus, or whether L1 RNA bound to ORF2 protein reaches chromatin passively after nuclear breakdown in mitosis. However, it is likely that L1 retrotransposition occurs by TPRT. We hypothesize that the L1 endonuclease nicks chromosomal DNA at the sequence 5′PO_4-TTT/AA-3′OH (where / denotes the cleavage site). The liberated 3′ hydroxyl residue then is thought to act as a primer for reverse transcription of the L1 RNA by the L1 RT (Figure 2; Feng et al., 1996). Subsequent steps in the retrotransposition pathway such as second strand cleavage, second strand DNA synthesis, and ligation of the resultant cDNA to genomic DNA all remain to be explained.

Potential limitations of the cultured cell retrotransposition assay

The cultured cell L1 retrotransposition assay has enabled the identification of RC-L1s in both the human and mouse genomes, and has revealed functional domains within the L1-encoded proteins. Moreover, we also have uncovered a novel mechanism by which L1 retrotransposition can contribute to genome diversity. However, the cultured cell assay has some potential limitations.

First, activation of the *mneoI* gene requires RNA splicing. Since L1s normally lack introns, it remains possible that the presence of the indicator cassette diverts L1 RNA from a splicesome-independent to a splicesome-dependent RNA export pathway (Stutz & Rosbash, 1998). Thus, the presence and function of RNA export signals in L1 RNA may be compensated for by the artificial inclusion of an intron into L1 RNA.

Next, in every instance, we have assayed for L1 retrotransposition in transformed somatic mammalian cultured cells, which have an abnormal karyotype. It will be interesting to determine whether L1 actually can retrotranspose in non-transformed somatic cells or embryonic stem cells. Indeed, the somatic insertion into the APC gene suggests that a subset of L1s can retrotranspose in somatic cells (Miki et al., 1992).

Finally, in every instance we have assayed for L1s ability to retrotranspose from an extrachromosomal episome to genomic DNA. Thus, all of these L1s have been placed in a privileged expression context. It remains possible that some of those RC-L1s are integrated in transcriptionally-repressed regions of chromatin, which would not allow for their expression and retrotransposition *in vivo*.

Closing remarks

We just are beginning to realize the consequences of L1 retrotransposition on the human genome. Clearly, L1 is a mutagen. Moreover, because of the abundance of L1s, it is likely that they provide scaffolds for illegitimate recombination, which may contribute to the genome instability seen in many tumors. Thus, with a fundamental mechanistic knowledge of L1 biology, we will gain a further understanding of both the molecular mechanisms that underlie human disease and the molecular processes that helped sculpt mammalian genomes.

Acknowledgements

I thank Haig Kazazian, Jef Boeke, John Goodier, and Eric Ostertag for helpful discussions. I thank Reid Alisch, Nicolas Gilbert, and Jyoti Athanikar for critical reading of the manuscript. J. V. M. was supported by a grant (GM60518) from the National Institutes of Health.

References

Alves, G., A. Tatro & T. Fanning, 1996. Differential methylation of human LINE-1 retrotransposons in malignant cells. Gene 17: 39–44.

Bender, J. & N. Kleckner, 1986. Genetic evidence that Tn10 transposes by a nonreplicative mechanism. Cell 45: 801–815.

Boeke, J.D., 1997. LINEs & Alus – the polyA connection [news; comment]. Nat Gene 16: 6–7.

Boeke, J.D. & V.G. Corces, 1989. Transcription and reverse transcription of retrotransposons. [Review]. Annu Rev Microbiol 43: 403–434.

Boeke, J.D., D.J. Garfinkel, C.A. Styles & G.R. Fink, 1985 Ty elements transpose through an RNA intermediate. Cell 40: 491–500.

Branciforte, D. & S.L. Martin, 1994. Developmental and cell type specificity of LINE-1 expression in mouse testis: implications for transposition. Mol Cell Biol 14: 2584–2592.

Bratthauer, G.L. & T.G. Fanning, 1992. Active LINE-1 retrotransposons in human testicular cancer. Oncogene 7: 507–510.

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

Carignani, G., O. Groudinsky, D. Frezza, E. Schiavon, E. Bergantino et al., 1983. An mRNA maturase is encoded by the first intron of the mitochondrial gene for the subunit I of cytochrome oxidase in *S. cerevisiae*. Cell 35: 733–742.

Clements, A.P. & M.F. Singer, 1998. The human LINE-1 reverse transcriptase: effect of deletions outside the common reverse transcriptase domain. Nucl Acids Res 26: 3528–3535.

Colgan, D.F. & J.L. Manley, 1997. Mechanism and regulation of mRNA polyadenylation. Gen Dev 11: 2755–2766.

Cost, G.J. & J.D. Boeke, 1998. Targeting of human retrotransposon integration is directed by the specificity of L1 endonuclease for regions of unusual DNA structure. Biochemistry 37: 18081–18093.

DeBerardinis, R.J., J.L. Goodier, E.M. Ostertag & H.H. Kazazian, Jr., 1998. Rapid amplification of a retrotransposon subfamily is evolving the mouse genome. Nat Genet 20: 288–290.

Deininger, P.L. & M.A. Batzer, 1999. Alu repeats and human disease. Mol Genet Metab 67: 183–193.

Divoky, V., K. Indra, M. Mrug, V. Brabec, T.H.J. Huisman et al., 1996. A novel mechanism of β-thalessemia. The insertion of L1 retrotransposable element into β globin IVSII. Blood 88: 148a.

Dombroski, B.A., Q. Feng, S.L. Mathias, D.M. Sassaman, A.F. Scott et al., 1994. An *in vivo* assay for the reverse transcriptase of human retrotransposon L1 in *Saccharomyces cerevisiae*. Mol Cell Biol 14: 4485–4492.

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

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

Fanning, T. & M. Singer, 1987. The LINE-1 DNA sequences in four mammalian orders predict proteins that conserve homologies to retrovirus proteins. Nucl Acids Res 15: 2251–2260.

Feng, Q., J.V. Moran, H.H. Kazazian & J.D. Boeke, 1996. Human L1 retrotransposon encodes a conserved endonuclease required fo retrotransposition. Cell 87: 905–916.

Feng, Q., G. Schumann & J.D. Boeke, 1998. Retrotransposon R1Bm endonuclease cleaves the target sequence. Proc Natl Acad Sci USA 95: 2083–2088.

Freeman, J.D., N.L. Goodchild & D.L. Mager, 1994. A modified indicator gene for selection of retrotransposition events in mammalian cells. BioTechniques 17: 47–52.

Gilbert, N. & D. Labuda, 1999. CORE-SINEs: eukaryotic short interspersed retroposing elements with common sequence motifs. Proc Natl Acad Sci USA 96: 2869–2874.

Goodier, J.L., E.M. Ostertag & H.H. Kazazian, 2000. Transduction of 3′ flanking sequences is common in L1 retrotransposition. Hum Molec Genetics, 9: 653–657.

Grimaldi, G., J. Skowronski & M.F. Singer, 1984. Defining the beginning and end of *KpnI* family segments. EMBO J 3: 1753–1759.

Hohjoh, H. & M.F. Singer, 1996. Cytoplasmic ribonucleoprotein complexes containing human LINE-1 protein and RNA. EMBO J 15: 630–639.

Hohjoh, H. & M.F. Singer, 1997. Sequence-specific single-strand RNA binding protein encoded by the human LINE-1 retrotransposon. EMBO J 16: 6034–6043.

Holmes, S.E., B.A. Dombroski, C.M. Krebs, C.D. Boehm & H.H. Kazazian, Jr., 1994. A new retrotransposable human L1 element from the LRE2 locus on chromosome 1q produces a chimaeric insertion. Nat Genet 7: 143–148.

Holmes, S.E., M.F. Singer & G.D. Swergold, 1992. 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: 19765–19768.

Hutchison, C.A., S.C. Hardies, D.D. Loeb, W.R. Shehee & M.H. Edgell, 1989. LINES and related retroposons: long interspersed sequences in the eucaryotic genome, pp. 593–617 in Mobile DNA, edited by D.E. Berg & M.M. Howe. ASM Press, Washington, DC.

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

Kazazian, H.H., Jr., 1998. Mobile elements and disease. Curr Opin Genet Dev 8: 343–350.

Kazazian, H.H., 1999. An estimated frequency of endogenous insertional mutations in humans. Nat Genet 22: 130.

Kazazian, H.H. & J.V. Moran, 1998. The impact of L1 retrotransposons on the human genome. Nat Genet 19: 19–24.

Kazazian, H.H., Jr., C. Wong, H. Youssoufian, A.F. Scott, D.G. Phillips et al., 1988. Haemophilia A resulting from *de novo* insertion of L1 sequences represents a novel mechanism for mutation in man. Nature 332: 164–166.

Kennell, J.C., J.V. Moran, P.S. Perlman, R.A. Butow & A.M. Lambowitz, 1993. Reverse transcriptase activity associated with maturase-encoding group II introns in yeast mitochondria. Cell 73: 133–146.

Kimberland, M.L., V. Divoky, J. Prchal, U. Schwahn, W. Berger et al., 1999. Full-length human L1 insertions retain the capacity for high frequency retrotransposition in cultured cells. Hum Mol Genet 8: 1557–1560.

Kingsmore, S.F., B. Giros, D. Suh, M. Bieniarz, M.G. Caron et al., 1994. Glycine receptor b-subunit gene mutation in *spastic* mouse associated with LINE-1 element insertion. Nat Genet 7: 136–142.

Kleckner, N., 1990. Regulation of transposition in bacteria. Annu Rev Cell Biol 6: 297–327.

Kolosha, V.O. & S.L. Martin, 1997. *In vitro* properties of the first ORF protein from mouse LINE-1 support its role in ribonucleoprotein formation during retrotransposition. Proc Natl Acad Sci USA 94: 10155–10160.

Korenberg, J.R. & M.C. Rykowski, 1988. Human genome organization: Alu, lines, and the molecular structure of metaphase chromosome bands. Cell 53: 391–400.

Kuiper, M.T. & A.M. Lambowitz, 1988. A novel reverse transcriptase activity associated with mitochondrial plasmids of Neurospora [published erratum appears in Cell 1989 Jan 13;56(1): following 139]. Cell 55: 693–704.

Kurose, K., K. Hata, M. Hattori & Y. Sakaki, 1995. RNA polymerase III dependence of the human L1 promoter and possible participation of the RNA polymerase II factor YY1 in the RNA polymerase III transcription system. Nucl Acids Res 23: 3704–3709.

Lampson, B.C., S. Inouye & M. Inouye, 1991. msDNA of bacteria. [Review]. Prog Nucleic Acid Res Mol Biol 40: 1–24.

Lazowska, J., B. Meunier & C. Macarde, 1994. Homing of a group II intron in yeast mitochondrial DNA is accompanied by unidirectional co-conversion of upstream-located markers. EMBO J. 13: 4963–4972.

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

Luan, D.D., M.H. Korman, J.L. Jakubczak & T.H. Eickbush, 1993. Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. Cell 72: 595–605.

Martin, F., C. Maranon, M. Olivares, C. Alonso & M.C. Lopez, 1995. Characterization of a non-long terminal repeat retrotransposon cDNA (L1Tc) from *Trypanosoma cruzi*: homology of the first ORF with the ape family of DNA repair enzymes. J Mol Biol 247: 49–59.

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

Martin, S.L., 1994. Characterization of a LINE-1 cDNA that originated from RNA present in ribonucleoprotein particles: implications for the structure of an active mouse LINE-1. Gene 153: 261–266.

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

Matsuura, M., R. Saldanha, H. Ma, H. Wank, J. Yang et al., 1997. A bacterial group II intron encoding reverse transcriptase, maturase, and DNA endonuclease activities: biochemical demonstration of maturase activity and insertion of new genetic information within the intron. Genes Dev 11: 2910–2924.

McMillan, J.P. & M.F. Singer, 1993. Translation of the human LINE-1 element, L1Hs. Proc Natl Acad Sci. 90: 11533–11537.

McNaughton, J.C., G. Hughes, W.A. Jones, P.A. Stockwell, H.J. Klamut et al., 1997. The evolution of an intron: analysis of a long, deletion-prone intron in the human dystrophin gene. Genomics 40: 294–304.

Meischl, C., M. De Boer & D. Roos, 1998. Chronic granulomatous disease caused by a LINE-1 retrotransposon. Eur J Haem 60: 349–350.

Miki, Y., I. Nishisho, A. Horii, Y. Miyoshi, J. Utsunomiya et al., 1992. Disruption of the APC gene by a retrotransposal insertion of L1 sequence in a colon cancer. Cancer Res. 52: 643–645.

Minakami, R., K. Kurose, K. Etoh, Y. Furuhata, M. Hattori et al., 1992. Identification of an internal *cis*-element essential for the human L1 transcription and a nuclear factor(s) binding to the element. Nucl Acids Res 20: 3139–3145.

Moran, J.V., R.J. DeBerardinis & H.H. Kazazian, 1999a. Exon shuffling by L1 retrotransposition. Science 283: 1530–1534.

Moran, J.V., Q. Feng, B.A. Dombroski, T.P. Naas, R.J. DeBerardinis et al., 1999b. L1-mediated processed pseudogene formation in *Saccharomyces cerevisiae*. In preparation.

Moran, J.V., S.E. Holmes, T.P. Naas, R.J. DeBerardinis, J.D. Boeke et al., 1996. High frequency retrotransposition in cultured mammalian cells. Cell 87: 917–927.

Moran, J.V., K.L. Mecklenburg, P. Sass, S.M. Belcher, D. Mahnke et al., 1994. Splicing defective mutants of the *COX1* gene of yeast mitochondrial DNA: initial definition of the maturase domain of the group II intron aI2. Nucl Acids Res 22: 2057–2064.

Moran, J.V., S. Zimmerly, R. Eskes, J.C. Kennell, A.M. Lambowitz et al., 1995. Mobile group II introns of yeast mitochondrial DNA are novel site-specific retroelements. Mol Cell Biol 15: 2828–2838.

Morse, B., P.G. Rotherg, V.J. South, J.M. Spandorfer & S.M. Astrin, 1988. Insertional mutagenesis of the myc locus by a LINE-1 sequence in a human breast carcinoma. Nature 333: 87–90.

Mulhardt, C., M. Fischer, P. Gass, D. Simon-Chazottes, J.-L. Guénet et al., 1994. The spastic mouse: aberrant splicing of glycine receptor b subunit mRNA caused by intronic insertion of L1 element. Neuron 13: 1003–1015.

Naas, T.P., R.J. DeBerardinis, J.V. Moran, E.M. Ostertag, S.F. Kingsmore et al., 1998. An actively retrotransposing, novel subfamily of mouse L1 elements. EMBO J. 17: 590–597.

Nakamura, T.M. & T.R. Cech, 1998. Reversing time: origin of telomerase. Cell 92: 587–590.

Narita, N., H. Nishio, Y. Kitoh, Y. Ishikawa, Y. Ishikawa et al., 1993. Insertion of a 5′ truncated L1 element into the 3′ end of exon 44 of the dystrophin gene resulted in skipping of the exon during splicing in a case of Duchenne muscular dystrophy. J Clin Invest 91: 1862–1867.

Okada, N., 1991. SINEs. Curr Opin Genet Dev 1: 498–504.

Okada, N., M. Hamada, I. Ogiwara & K. Ohshima, 1997. SINEs and LINEs share common 3′ sequences: a review. Gene 205: 229–243.

Rozmahel, R., H.H. Heng, A.M. Duncan, X.M. Shi, J.M. Rommens et al., 1997. Amplification of CFTR exon 9 sequences to multiple locations in the human genome. Genomics 45: 554–561.

Sassaman, D.M., B.A. Dombroski, J.V. Moran, M.L. Kimberland, T.P. Naas et al., 1997. Many human L1 elements are capable of retrotransposition [see comments]. Nat Genet 16: 37–43.

Saxton, J.A. & S.L. Martin, 1998. Recombination between subtypes creates a mosaic lineage of LINE-1 that is expressed and actively retrotransposing in the mouse genome. J Mol Biol 280: 611–622.

Schwahn, U., S. Lenzner, J. Dong, S. Feil, B. Hinzmann et al., 1998. Positional cloning of the gene for X-linked retinitis pigmentosa 2. Nat Genet 19: 327–332.

Scott, A.F., B.J. Schmeckpeper, M. Abdelrazik, C.T. Comey, B. O'Hara et al., 1987. Origin of the human L1 elements: proposed progenitor genes deduced from a consensus DNA sequence. Genomics 1: 113–125.

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

Smit, A.F.A., 1996. The origin of interspersed repeats in the human genome. Curr Opin Genet Dev 6: 743–748.

Stutz, F. & M. Rosbash, 1998. Nuclear RNA export. Genes Dev 12: 3303–3319.

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

Takahara, T., T. Ohsumi, J. Kuromitsu, K. Shibata, N. Sasaki et al., 1996. Dysfunction of the *Orleans* reeler gene arising from exon skipping due to transposition of a full-length copy of an active L1 sequence into the skipped exon. Hum Mol Genet 5: 989–993.

Thayer, R.E., M.F. Singer & T.F. Fanning, 1993. Undermethylation of specific LINE-1 sequences in human cells producing a LINE-1 encoded protein. Gene 133: 273–277.

Usdin, K. & A.V. Furano, 1989. The structure of the guanine-rich polypurine:polypyrimidine sequence at the right end of the rat L1 (LINE) element. J Biol Chem 264: 15681–15687.

Xiong, Y. & T.H. Eickbush, 1988a. Similarity of reverse transcriptase-like sequences of viruses, transposable elements, and mitochondrial introns. Mol Biol Evol 5: 675–690.

Xiong, Y. & T.H. Eickbush, 1990. Origin and evolution of retroelements based upon their reverse transcriptase sequences. EMBO J 9: 3353–3362.

Xiong, Y.E. & T.H. Eickbush, 1988b. Functional expression of a sequence-specific endonuclease encoded by the retrotransposon R2Bm. Cell 55: 235–246.

Yoshida, K., A. Nakamura, M. Yazaki, S. Ikeda & S. Takeda, 1998. Insertional mutation by transposable element, L1, in the DMD gene results in X-linked dilated cardiomyopathy. Hum Mol Genet 7: 1129–1132.

Zimmerly, S., H. Guo, R. Eskes, J. Yang, P.S. Perlman et al., 1995a. A group II intron RNA is a catalytic component of a DNA endonuclease involved in intron mobility. Cell 83: 529–538.

Zimmerly, S., H. Guo, P.S. Perlman & A.M. Lambowitz, 1995b. Group II intron mobility occurs by target DNA-primed reverse transcription. Cell 82: 545–554.