Since the early 1980's we have known about the major role that transposable elements play in the biology of a wide variety of organisms. In *Drosophila,* most spontaneous mutations result from new insertions of transposable elements. In yeast, about 1% of all mutations are due to insertions of mobile elements. In humans, the effects of transposition events are harder to estimate, but the best data indicate that they account for between 1 in 500 and 1 in 1,000 mutations.

In prokaryotes, most of transposition occurs through DNA itself. Transposon DNA is copied and reinserted into the genome at a new location. In eukaryotes, the process is one of retrotransposition in which the element is transcribed into RNA, reverse transcribed into cDNA, and the double-stranded cDNA is reinserted into the genome. In computer lingo the process is "copy and paste," as the original sequence remains intact. When a retrotransposon inserts into the genome, it is flanked by short direct repeats (target site duplications) of the endogenous sequence. These have been seen in nearly all retrotransposition events.

There are two general structures of eukaryotic retrotransposons, Class I and Class II elements (1). Class I elements have long terminal repeats at their two ends. They also have 1–3 protein-coding regions, one of which makes a reverse transcriptase used by the element for retrotransposition. Examples of these elements which are 4–6 kb in length are copia of *Drosophila,* Ty of yeast, and intracisternal A particle of mice. Retroviruses look very much like Class I elements, but they contain an envelope protein gene to help them move from cell to cell while only an exceptional Class I element contains an envelope gene. Evolutionary biologists have developed a "tree" for transposable elements based on reverse transcriptase sequences (2). From this "tree," retroviruses appear to have had a very recent origin, while Class II L1 elements (discussed below) appear near the root and are very old. Full-length Class II elements are 5–7 kb in size; they lack long terminal repeats, have 3′ poly A tails, and have one to two open reading frames (ORFs), one of which encodes a reverse transcriptase. Examples of these elements are the I factor of *Drosophila* and L1 elements of mammals.

L1 elements are present in the genome of all mammals: in humans, they are present in about 100,000 copies and account for about 5% of genomic DNA. Most of the L1 copies are 5′-truncated so that only about 3,500 are full-length, 6-kb elements. A number of facts suggested that some L1 elements were retrotransposons. They were present in many copies. They had a 3′ poly A tail and two ORFs in their consensus sequence. They had an internal promoter for transcription, and they also had a region of ORF2 that could potentially encode a polypeptide with reverse transcriptase homology. Five years ago, our group found two patients with hemophilia A in whom de novo insertion of a truncated L1 element into the Factor VIII gene on the X chromosome produced the disease (3). In 1991, the full-length precursor to one of those insertions was isolated (4). It contains the two ORFs predicted by the L1 consensus sequence. ORF1 makes a protein of unknown function in a transient expression assay (5) while ORF2 encodes a reverse transcriptase activity in a yeast assay system (6). This L1 element is present in two copies on chromosome 22 in all humans studied to date and it is flanked by a target site duplication, indicating that it originated as the product of a retrotransposition event.

Over the past two years, three separate insertions of Alu elements have been reported, all of which caused genetic disease (7–9). Alu elements presumably expanded to their present number of about 300,000 in the human genome by reinsertion of reverse transcripts. A similar process is thought to account for processed pseudogenes, i.e., the dispersion of mRNAs. The cognoscenti are betting that the reverse transcriptase for these events is provided by active L1 elements. Evidence suggests that L1 transcripts are sequestered in cytoplasmic particles within which L1 cDNA is produced (10). Alu RNAs and mRNAs could be trapped in L1 particles, reverse transcribed by L1 reverse transcriptase, and then be escorted back to the genome using L1 machinery for reinsertion. Although this scenario is speculative at present, it is attractive because it provides these wandering RNAs with the "goods" of a proven retrotransposon for reentering the genome.

Recently, two new twists to the L1 retrotransposition story have emerged. First, a somatic insertion was found in an adenomatous polyposis coli (APC) tumor suppressor gene (11). This insertion of roughly 500 bp of the 3′ sequence of an L1 element disrupted the coding region of an APC gene in the tumor, but was not present in the normal colonic epithelium of the patient. Thus, in this instance the retrotransposition was a somatic event which occurred in a dedifferentiating cell. Second, the paper by Narita et al. in this issue of *The Journal* describes an inherited L1 insertion into the dystrophin gene causing Duchenne muscular dystrophy (12). This insertion is of special interest because it lacks a target site duplication; instead, two nucleotides of the dystrophin sequence are deleted and one nucleotide is added at the point of insertion. This observation suggests that a minority of transposon insertions are processes whereby breaks in the DNA are repaired, in contrast to the active reintegration envisaged for most transposition events.

Many interesting questions about L1 retrotransposition are yet to be answered. For instance, we wonder whether these sequences are merely parasitic DNAs that we (as mammals) have been able to tolerate for a hundred million years or are they, like the mitochondria, symbiotes that provide some necessary biological function? If so, what might that function be? At present, the only known biochemical property of L1 is a reverse transcriptase activity. Does this reverse transcriptase play some role in the ova or sperm? In fact, is L1 normally expressed in germline cells and at what level? Does L1 really form a particle like retroviruses and can the particle trap other RNAs and reverse transcribe them?

Work cited above along with that of others suggests that L1 transcription and expression may be induced in some cancers. Is this unusual expression a consequence of general dedifferentiation? Could abnormal L1 expression be used as a tumor cell

J. Clin. Invest.

0021-9738/93/05/1859/02 $2.00
Volume 91, May 1993, 1859–1860

marker? Why does it happen and how common is it? Is it a consequence of methylation changes in the DNA or expression of transcription factors needed for L1 transcription?

What role might L1 insertions play in evolution? How frequently does insertion of potentially transcribed sequence near a gene alter the temporal and tissue-specific expression of that gene? Presumably, small effects in the timing of gene expression at hundreds or thousands of locations around the genome could have significant consequences on an organism's developmental program. The presence of large blocks of homologous sequence in chromosomes is likely to increase the opportunity for unequal crossing over. Mammalian evolution has been strongly influenced by gene duplication events. Changes can occur in duplicate genes to evolve new functions without the loss of the gene product from the primary gene. Perhaps the history of our genome is as much a consequence of L1-mediated genomic rearrangements as of cumulative point mutations.

What we know about L1 biology is much less than what remains to be known. The paper by Narita et al. is another important step in what should be an exciting and productive future for studies of this most interesting class of DNA.

Haig H. Kazazian, Jr.
Alan F. Scott
The Johns Hopkins University School of Medicine

References

1. Fanning, T. G., and M. F. Singer. 1987. LINE-1: a mammalian transposable element. *Biochem. Biophys. Acta.* 910:203–212.

2. Xiong, Y., and T. H. Eickbush. 1990. Origin and evolution of retroelements based upon their reverse transcriptase sequences. *EMBO (Eur. Mol. Biol. Organ.) J.* 9:3353–3363.

3. Kazazian, H. H., Jr., C. Wong, H. Youssoufian, A. F. Scott, D. G. Phillips, and S. E. Antonarakis. 1988. Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanism for mutation in man. *Nature (Lond.).* 332:164–166.

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

5. Holmes, S. E., M. F. Singer, and 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 element. *J. Biol. Chem.* 267:19765–19768.

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

7. Wallace, M. R., L. B. Andersen, A. M. Saulino, T. W. Gregory, and F. S. Collins. 1991. A de novo Alu insertion results in neurofibromatosis type 1. *Nature (Lond.).* 353:864–866.

8. Muratani, K., T. Hada, Y. Yamamoto, T. Kaneko, Y. Shigeto, T. Ohue, J. Furuyama, and K. Higashino. 1991. Inactivation of the cholinesterase gene by Alu insertion: possible mechanism for human gene transposition. *Proc. Natl. Acad. Sci. USA.* 88:11315–11319.

9. Vidaud, D., M. Vidaud, B. R. Bahnak, V. Siguret, S. G. Sanchez, Y. Laurian, D. Meyer, M. Goossens, and J. M. Lavergne. 1993. Haemophilia B due to a de novo insertion of a human-specific Alu subfamily member within the coding region of the factor IX gene. *Eur. J. Hum. Genet.* 1:30–36.

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

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

12. Narita, N., H. Nishio, Y. Kitoh, Y. Ishikawa, Y. Ishikawa, R. Minami, H. Nakamura, and M. Matsuo. 1992. 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.