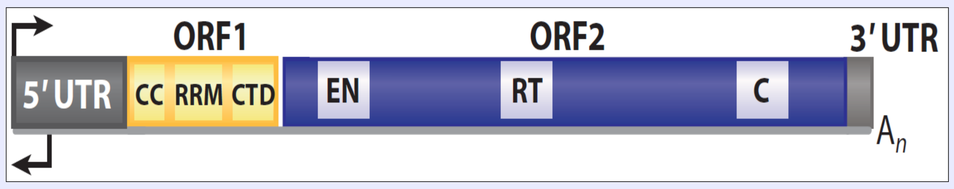
LINE-1 retrotransposons and their impact on the human genome and in disease-causing mutations

S. Corrales Lorite, Genetics Degree, Facultat de Biociències, Universitat Autònoma de Barcelona, Spain.

UAB Universitat Autònoma de Barcelona

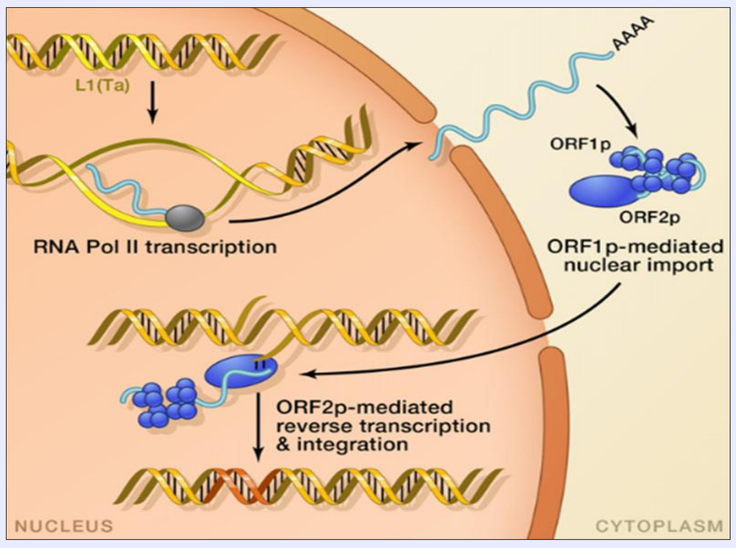
INTRODUCTION

Transposable elements (TEs) are DNA sequences capable of mobilizing and multiply within the genome of a host organism in both somatic and germ found in many species. In humans, comprise approximately 45% of the genetic material and, therefore, is not an insignificant fact, since although not part of the coding DNA (1.5%) are the largest source of variability among humans. TEs are divided into class I elements or retrotransposons (LTR as retroviruses and non-LTR as LINE, SINE and SVA) and class II or DNA transposons. This work is a review of TEs in humans, particularly LINE-1 retrotransposons (L1) and their impact on the human genome as well as the most characteristic disease-causing mutations caused directly by these, not in *trans* by other elements as *Alu* or SVA. The reason to focus only on the L1 retrotransposon mutagenesis is because they are the most abundant elements in whole human genome (16.9%), in addition to being the only currently active and autonomous elements, unlike *Alu* or SVA, which are active but non-autonomous.



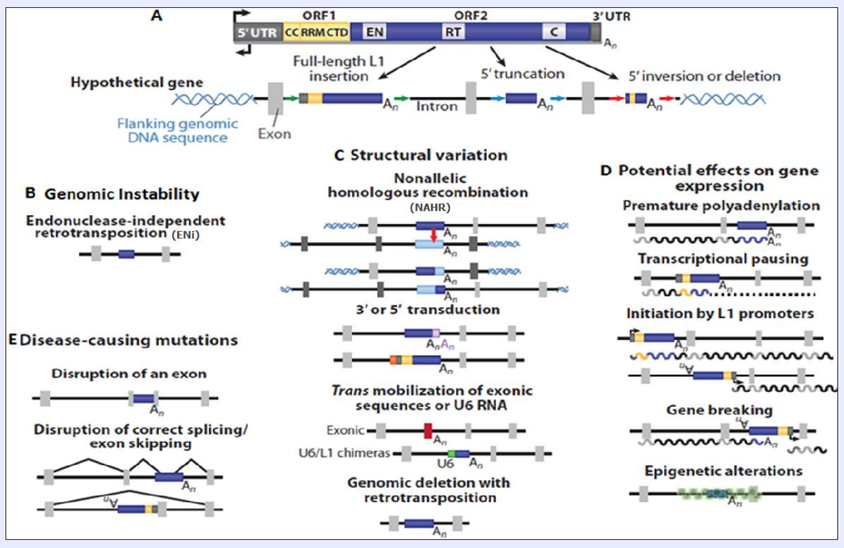
Schematic representation of the structural organization of a human full-length LINE-1 (6 kb) showing 5′UTR, ORF1, ORF2 and 3′UTR with poli(A) tail.

RETROTRANSPOSITION CYCLE OF LINE-1



L1 is transcribed by RNA Pol II, then the mRNA is exported to the cytoplasm, and translation of ORF1p and ORF2p leads to ribonucleoprotein (RNP) formation. Components of the L1 RNP are transported to the nucleus, and retrotransposition occurs by target-site primed reverse transcription (TPRT). During TPRT, the ORF2p endonuclease activity (EN) nicks genomic DNA, exposing a free 3′-OH that can serve as a primer for RT of the L1 mRNA. The process continues with a second-strand cleavage, second-strand cDNA synthesis, and completion of L1 integration. Resolution of the structure results in TSD.

DIFFERENT WAYS TO AFFECT THE HUMAN GENOME



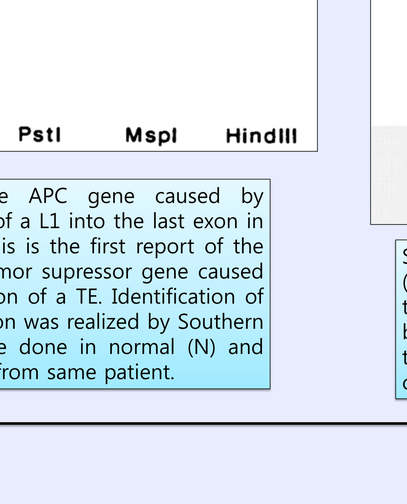
Schematic representation of the various ways that L1 retrotransposition events can affecte the human genome. **A** Hypothetical wild-type gene locus. **B** Genomic instability caused by L1 EN activity from ORF2p generating dsDNA breaks corrected by Eni retrotransposition. **C** Structural variation caused by NAHR, transductions or genomic deletions. **D** Potential effects on gene expression by initiation by L1 promoter, premature polyadenylation or epigenetic alterations. **E** Disease-causing mutations by disruption of an exon or disruption of correct splicing pattern.

DISEASE-CAUSING MUTATIONS CAUSED BY LINE-1 RETROTRANSPOSITION

Insertion	Gene	CHR	Reference	Disease	Subfamily	Size (nt)	PolyA tail length (nt)	Truncation	Transduction (nt)	Strand	Exon/intron/ mechanism
L1	*DMD*	X	Yoshida *et al.* 1998	XLDCM	L1 Ta	530	73	Y/5′TR	N	AS	5′-UTR/Loss of mRNA
L1	*DMD*	X	Narita *et al.* 1993	DMD	L1 Ta	608	16	Y/5′TR	N	AS	E
L1	*FVIII*	X	Kazazian *et al.* 1988	Hemophilia A	L1 Ta	3800	54	Y/5′TR	N	S	E
L1	*FVIII*	X	Kazazian *et al.* 1988	Hemophilia A	L1 preTa	2300	77	Y/5′TR/INV	N	AS	E
L1	*APC*	5	Miki *et al.* 1992	Colon cancer	L1Ta	520	222	Y/5′TR/INV	N	S	E
L1	*PDHX*	11	Mine *et al.* 2007	PDHc deficiency	L1 Hs	6086	67	FL	N	S	46 kb Deletion
L1	*CYBB*	X	Meischl *et al.* 2000	CGD	L1 Ta	836	69	Y/5′TR/INV	N	S	I/Splicing

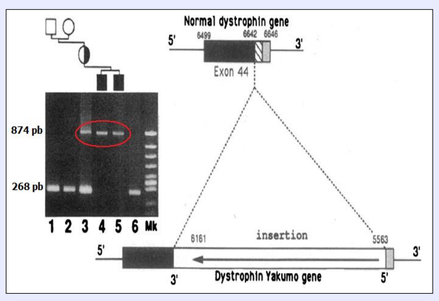
Most carachteristic disease-causing mutations by L1 described until 2012.

DMD – X-Linked Dilated Cardiomyopathy

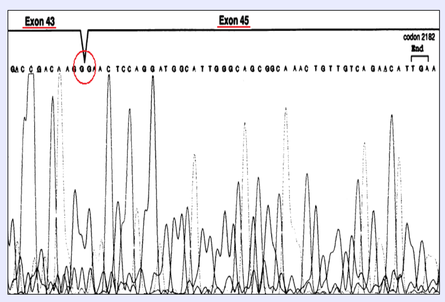


Identification by PCR of a unique *de novo* L1 5′-truncated insertion in the 5′UTR muscle exon 1 of DMD gene in three XLDCM patients from two unrelated Japanese families which affected the stability of the muscle form of dystrophin transcripts but not that of the brain or Purkinje cell form.

DMD – Duchenne Muscular Dystrophy

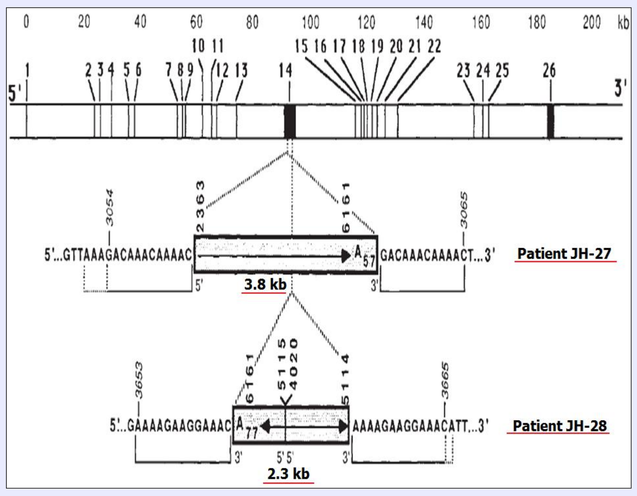


When the PCR was used to amplify a region of the DMD gene encompassing exon 44 from genomic DNA of two Japanese brothers with DMD, it was found to be 600 bp larger than expected.



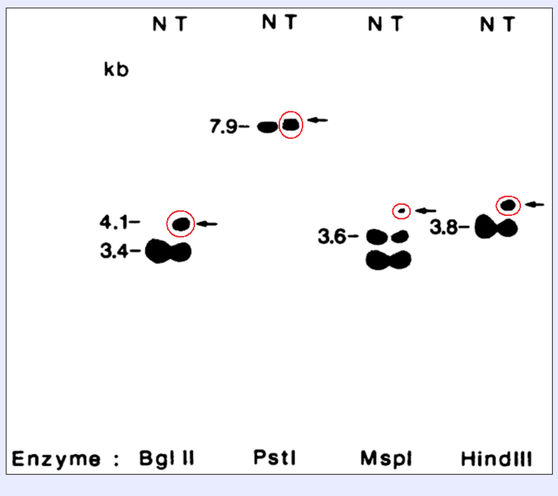
Sequencing of the mRNA disclosed that the sequence of exon 43 was directly joined to that of exon 45, suggesting an exon 44 skipping from the transcript.

FVIII – Hemophilia A



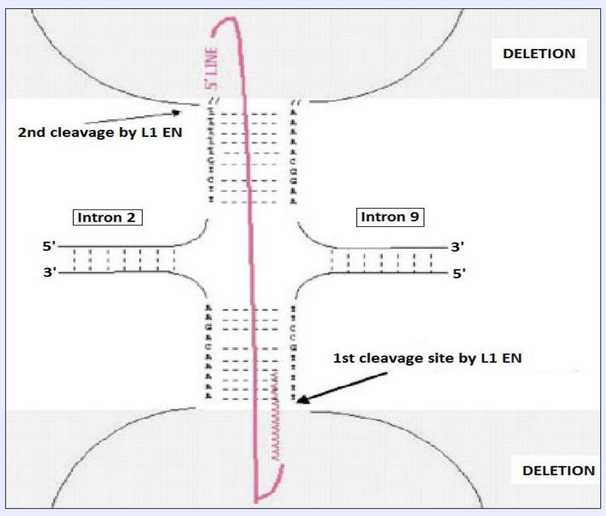
Schematic representation of *de novo* insertions of L1 elements into exon 14 of the FVIII gene in two of 240 unrelated patients with hemophilia A (3.8 kb insertion in family JH-27 and 2.3 kb in family JH-28).

APC – Colon Cancer



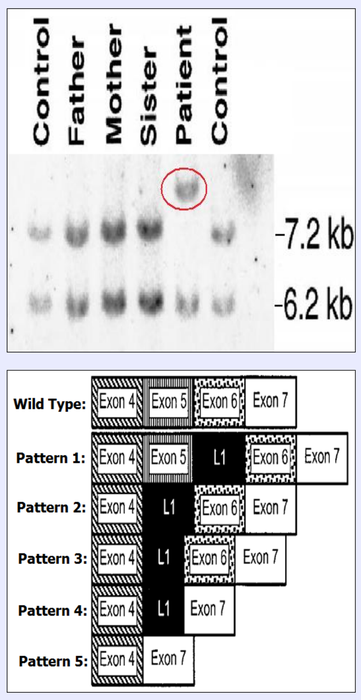
Disruption of the APC gene caused by somatic insertion of a L1 into the last exon in a colon cancer. This is the first report of the disruption of a tumor supressor gene caused by somatic insertion of a TE. Identification of the 750 pb insertion was realized by Southern Blot. Analysis were done in normal (N) and tumoral (T) tissue from same patient.

PDHX– Pyruvate Dehydrogenase complex deficiency



Schematic representation of the two hairpin structure (formed by inverted repeats present immediately 5′ to the upstream breakpoint and 3′ to the downstream breakpoint) responsible of a deletion of 46 kb by template jumping model between intron 2 and 9 caused by a L1 full-length L1 insertion.

CYBB – Chronic Granulomatous Disease



In this patient, CGD is caused by the insertion of an 1 kb L1 into intron 5 of the CYBB gene.

The L1 insertion into the intron sequence introduce new splice sites resulting in a highly heterogeneous splicing pattern. No wild-type cDNA was found in the patient.

CONCLUSIONS

- ✓ Transposable elements comprise nearly 50% of the genome, and within these, LINE-1 elements are the most abundant group (16.9% in the entire human genome) and the only active besides autonomous elements.
- ✓ LINE-1 retrotransposons can affect the genome causing genomic instability, structural variations, potential effects on gene expression or epigenetics and disease-causing mutations.
- ✓ Disease-causing mutations comprise various mechanisms, such as exonic disruption (colon cancer and hemophilia A), altered splicing (CGD), exon skipping (DMD), large deletion (PDHc deficiency) or transcript instability (XLDCM).
- ✓ Transposable elements also provide beneficial aspects on the host genome such as gene duplication, genome rearrangements, gene and epigenetic regulation or providing variability that plays a key role throughout the evolution.
- ✓ With the improvement of technology, it presents a great opportunity to expand knowledge about transposable elements.

REFERENCES

Yoshida K et *al.* **Insertional mutation by transposable element, L1, in the DMD gene results in X-linked dilated cardiomyopathy**. Hum Mol Genet. 1998.

Narita N et *al.* **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. 1993.

Kazazian HH Jr et *al.* **Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanism for mutation in man**. Nature. 1988.

Miki Y et *al.* **Disruption of the APC gene by a retrotransposal insertion of L1 sequence in a colon cancer**. Cancer Res. 1992.

Miné M et *al.* **A large genomic deletion in the PDHX gene caused by the retrotranspositional insertion of a full-length LINE-1 element**. Hum Mutat. 2007.

Meischl C et *al.* **A new exon created by intronic insertion of a rearranged LINE-1 element as the cause of chronic granulomatous disease.** Eur J Hum Genet. 2000.