Corso di laurea in Scienze Biologiche
Corso di laurea magistrale in Scienze Biomolecolari e dell'Evoluzione

Materiale didattico di supporto

Tutto il materiale fornito a supporto delle lezioni e reperibile nel minisito dell'insegnamento o sulla piattaforma online UniFE deve essere inteso come traccia degli argomenti svolti e non sostituisce il libro di testo.

Mobile Elements: Drivers of Genome Evolution

Haig H. Kazazian Jr.*

Mobile elements within genomes have driven genome evolution in diverse ways. Particularly in plants and mammals, retrotransposons have accumulated to constitute a large fraction of the genome and have shaped both genes and the entire genome. Although the host can often control their numbers, massive expansions of retrotransposons have been tolerated during evolution. Now mobile elements are becoming useful tools for learning more about genome evolution and gene function.

residues, then a glutamate) and a handlike three-dimensional structure (*6*, *8*).

Although these elements generally transpose to genomic sites less than 100 kb from their original site (e.g., the *Drosophila* P element), some are able to make distant "hops" (e.g., the fish Tc1/mariner element; see below).

LTR Retrotransposons

Watson BM Gene Capitolo 11

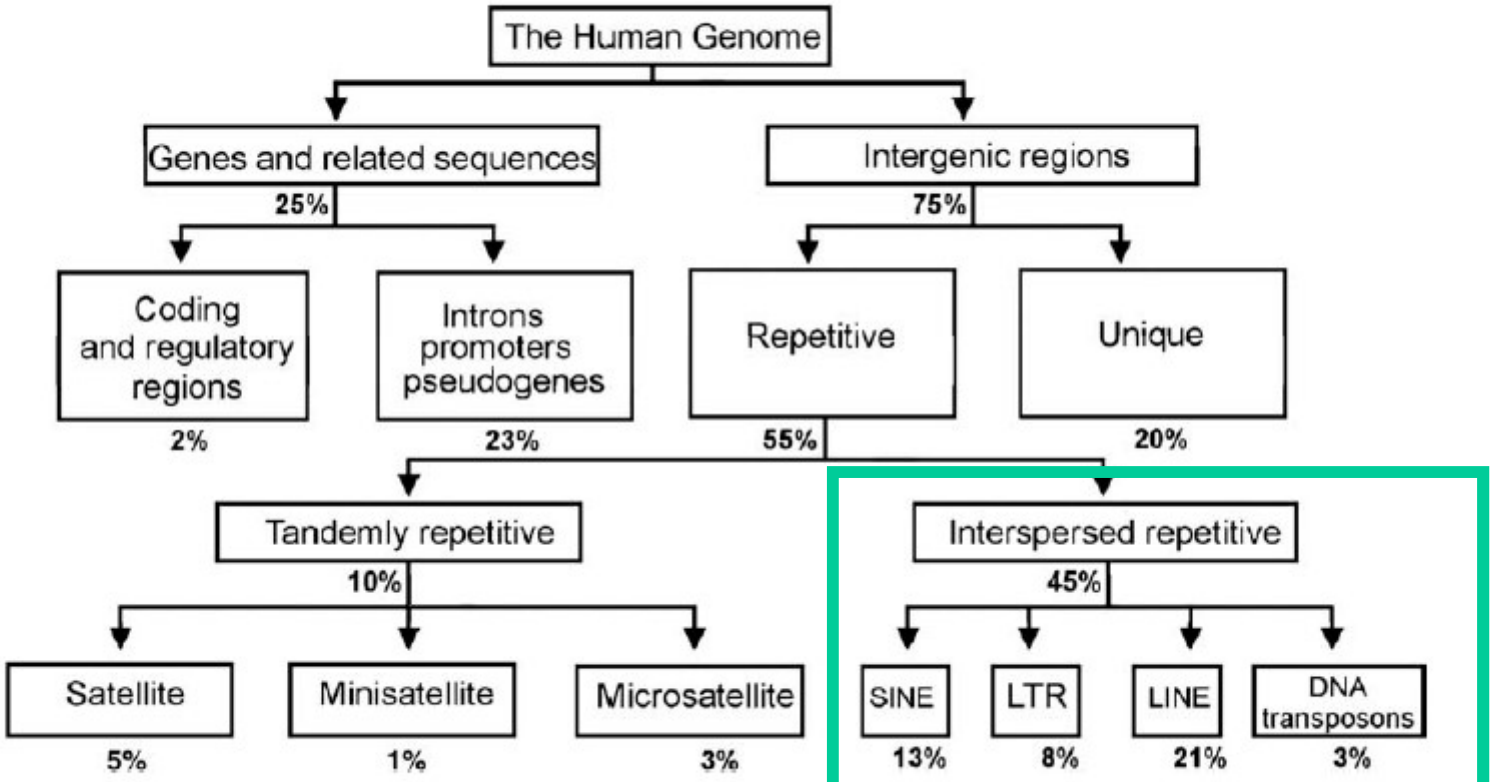
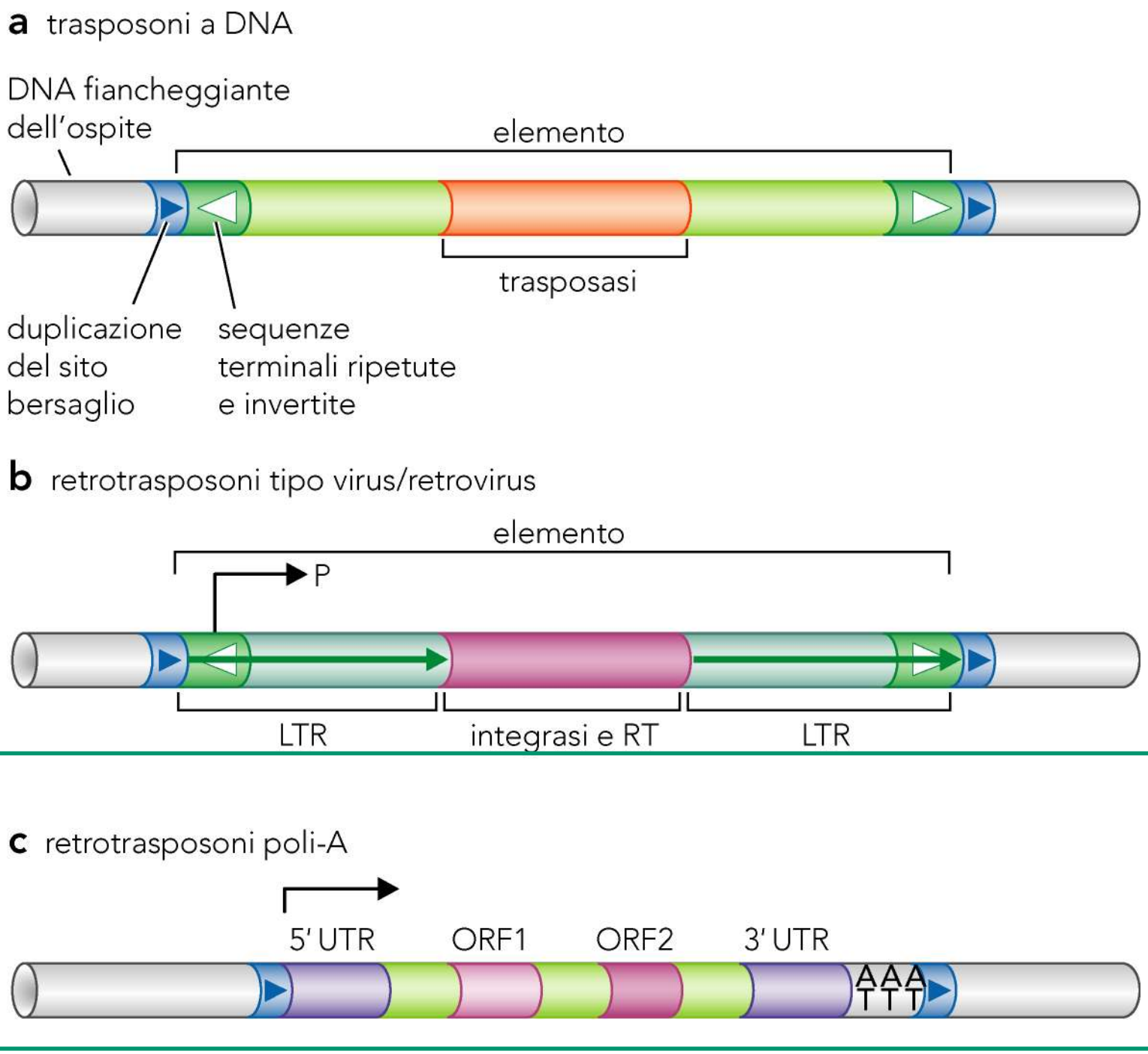
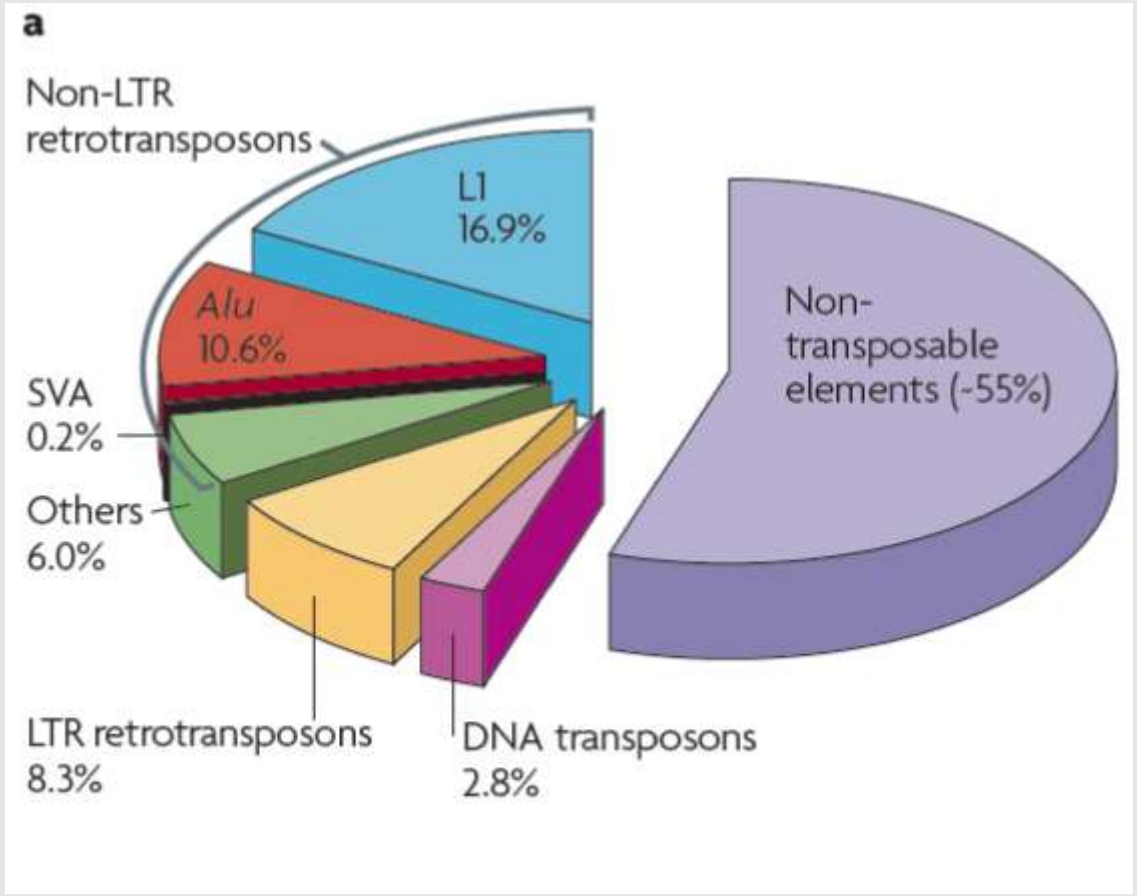
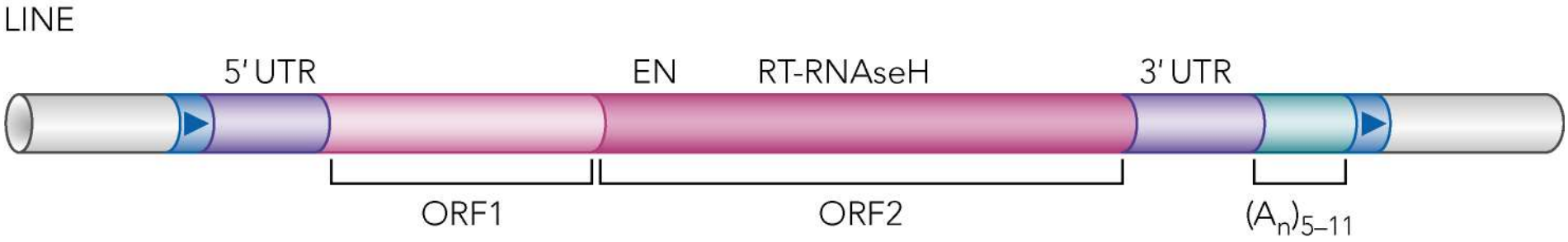


Fig. 1. Composition of the human genome. The percentage shares of various functional and non-functional sequences are shown.


a trasposoni a DNA
DNA fiancheggiante dell'ospite
elemento
trasposasi
duplicazione del sito bersaglio
sequenze terminali ripetute e invertite
b retrotrasposoni tipo virus/retrovirus
elemento
P
LTR
integrasi e RT
LTR
c retrotrasposoni poli-A
5' UTR
ORF1
ORF2
3' UTR
AAA
TTT


a
Non-LTR retrotransposons
L1 16.9%
Alu 10.6%
SVA 0.2%
Others 6.0%
LTR retrotransposons 8.3%
DNA transposons 2.8%
Non-transposable elements (~55%)


LINE
5′ UTR
EN
RT-RNAseH
3′ UTR
ORF1
ORF2
$(A_n)_{5–11}$

SVA

SINE-VNTR-Alu (SVA) elements are nonautonomous, hominid-specific non-LTR retrotransposons

composite mobile elements.

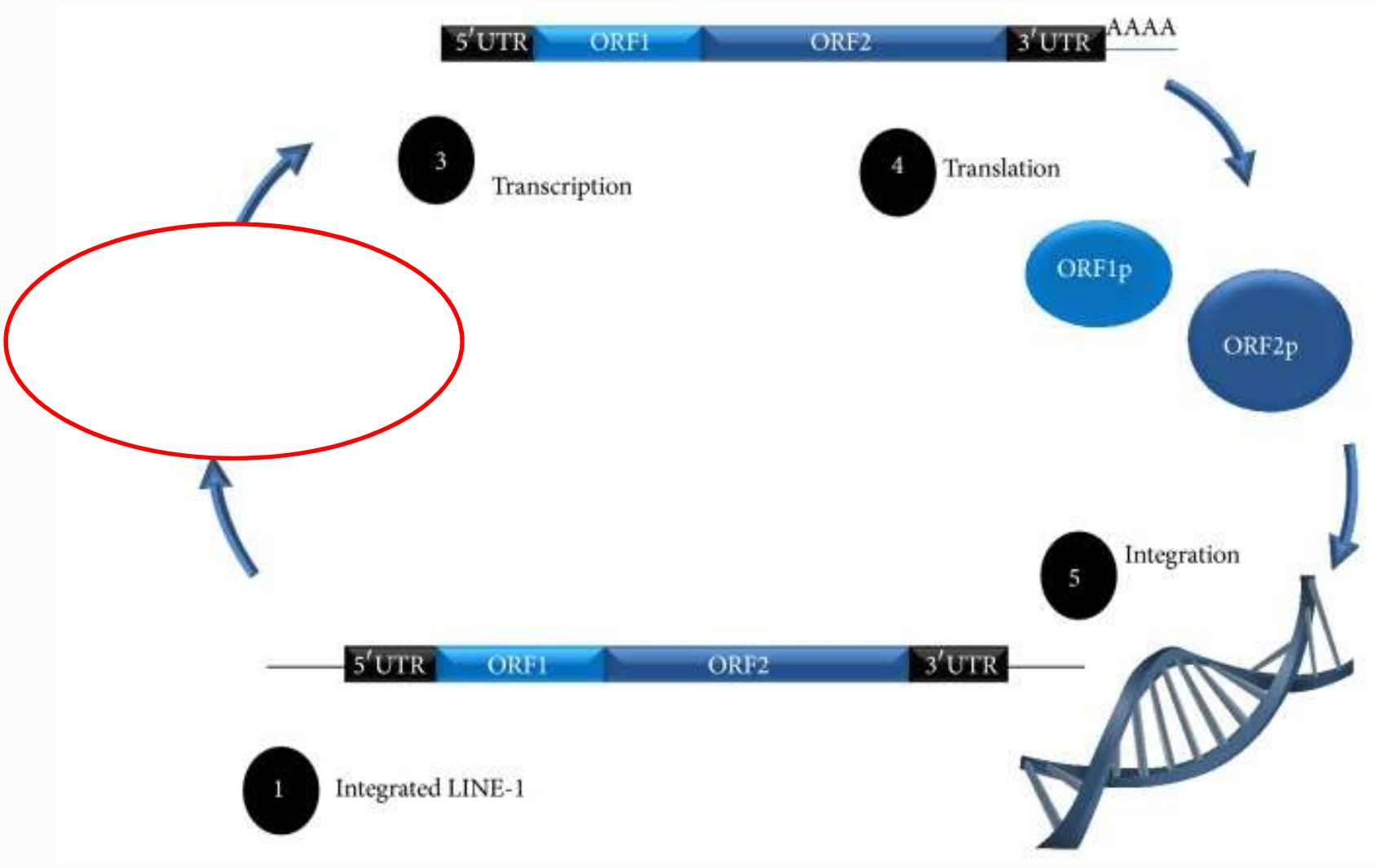
They represent the evolutionarily youngest, currently active family of human non-LTR retrotransposons

la trasposizione

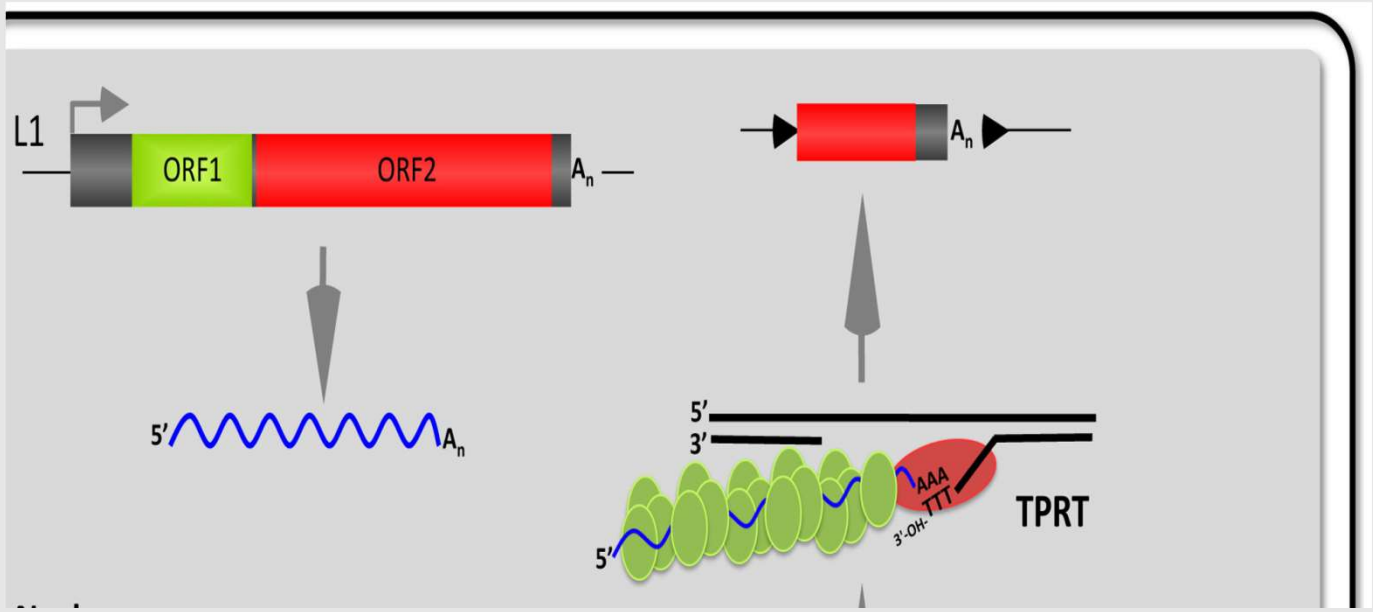
LINE-1

The non-LTR retrotransposon Long INterspersed Element-1 (or L1) is the only active autonomous TE.

In addition to mobilizing its own RNA to new genomic locations via a “copy-and-paste” mechanism, LINE-1 is able to retrotranspose other RNAs including Alu, SVA, and occasionally cellular RNAs.


5′UTR
ORF1
ORF2
3′UTR
AAAA
3
Transcription
4
Translation
ORF1p
ORF2p
5
Integration
5′UTR
ORF1
ORF2
3′UTR
1
Integrated LINE-1

target-site primed reverse transcription



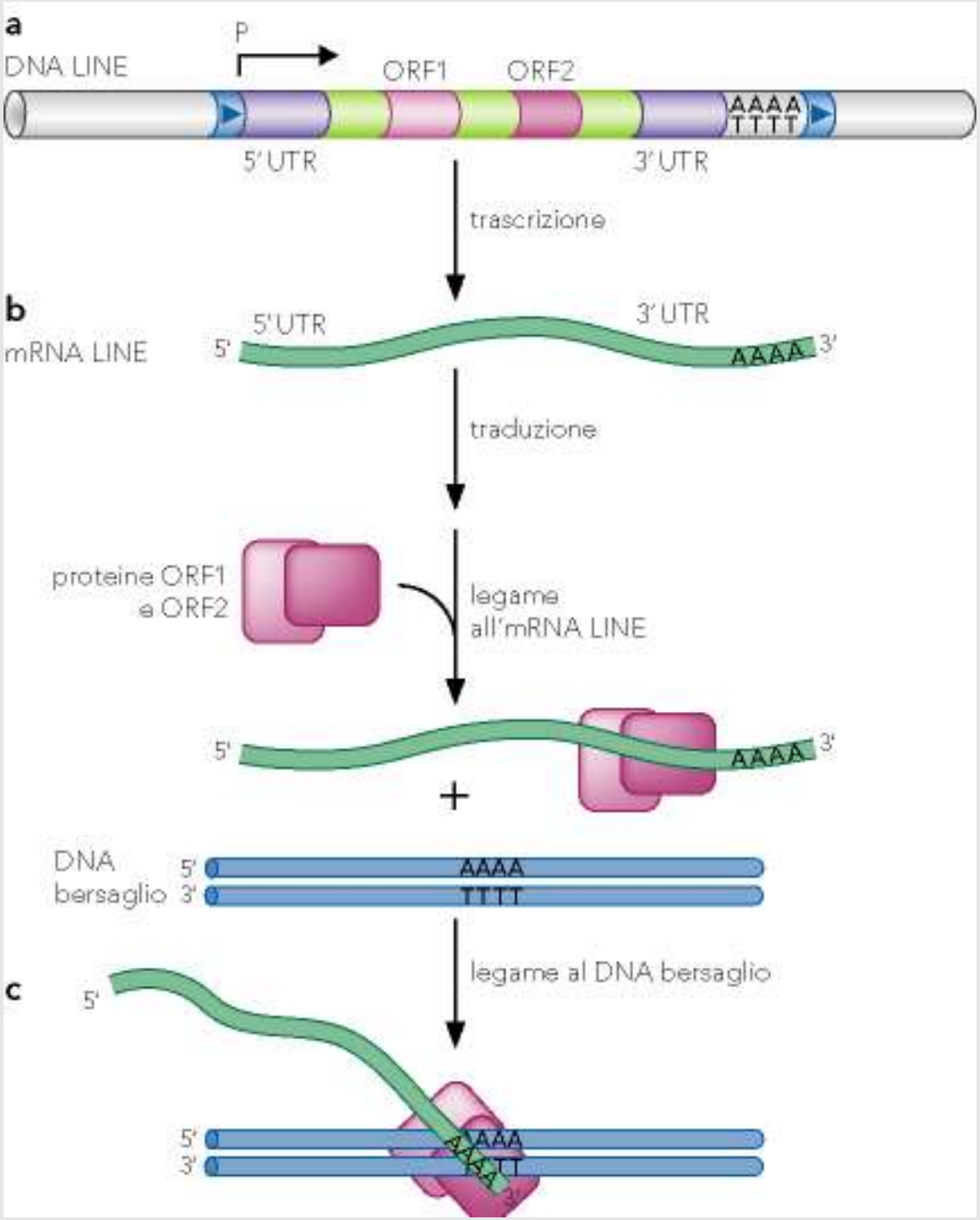
Trascrizione L1

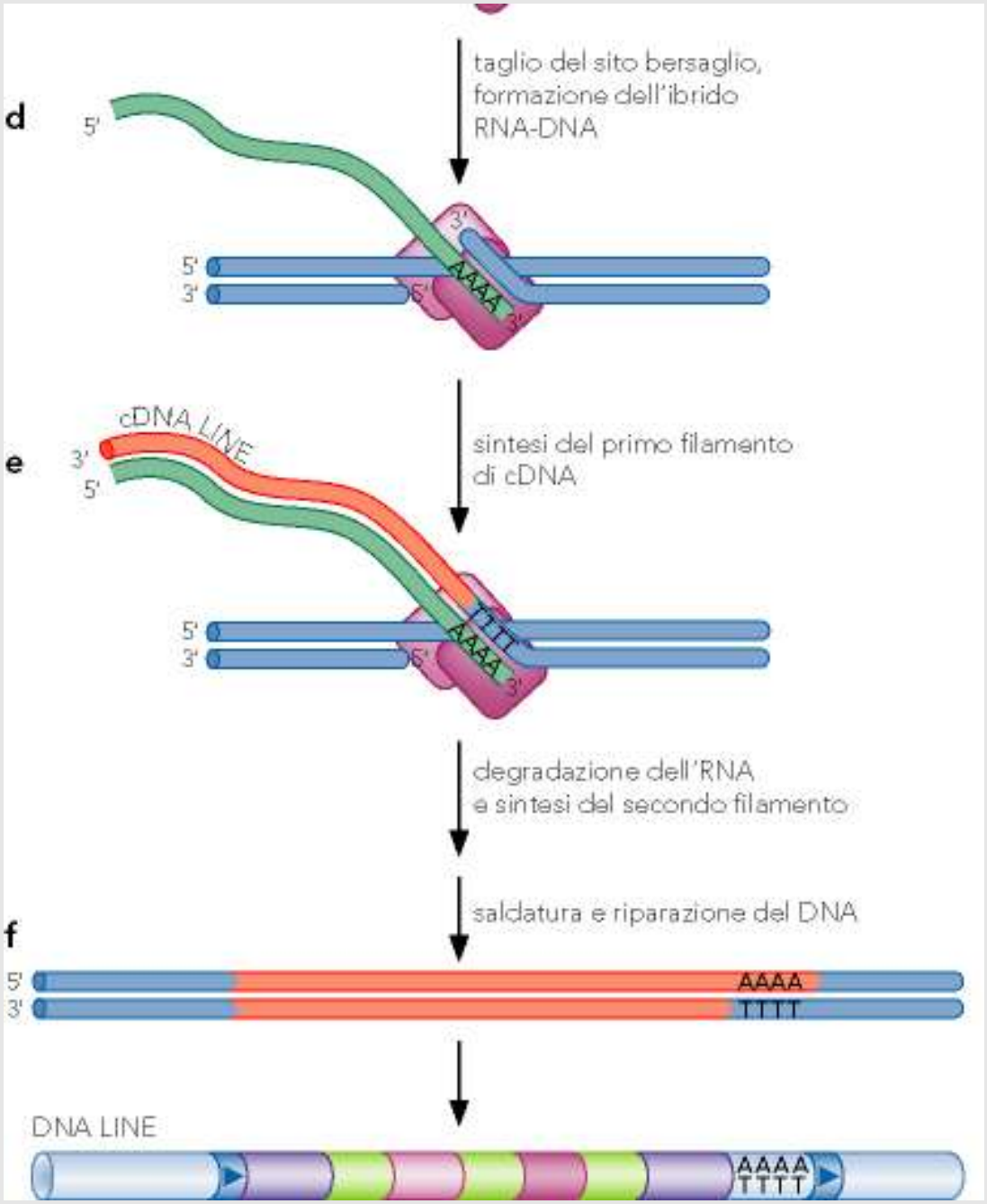
- RNApolymeraseII
- Transcription factors

SOX11

YY1

RUNX3

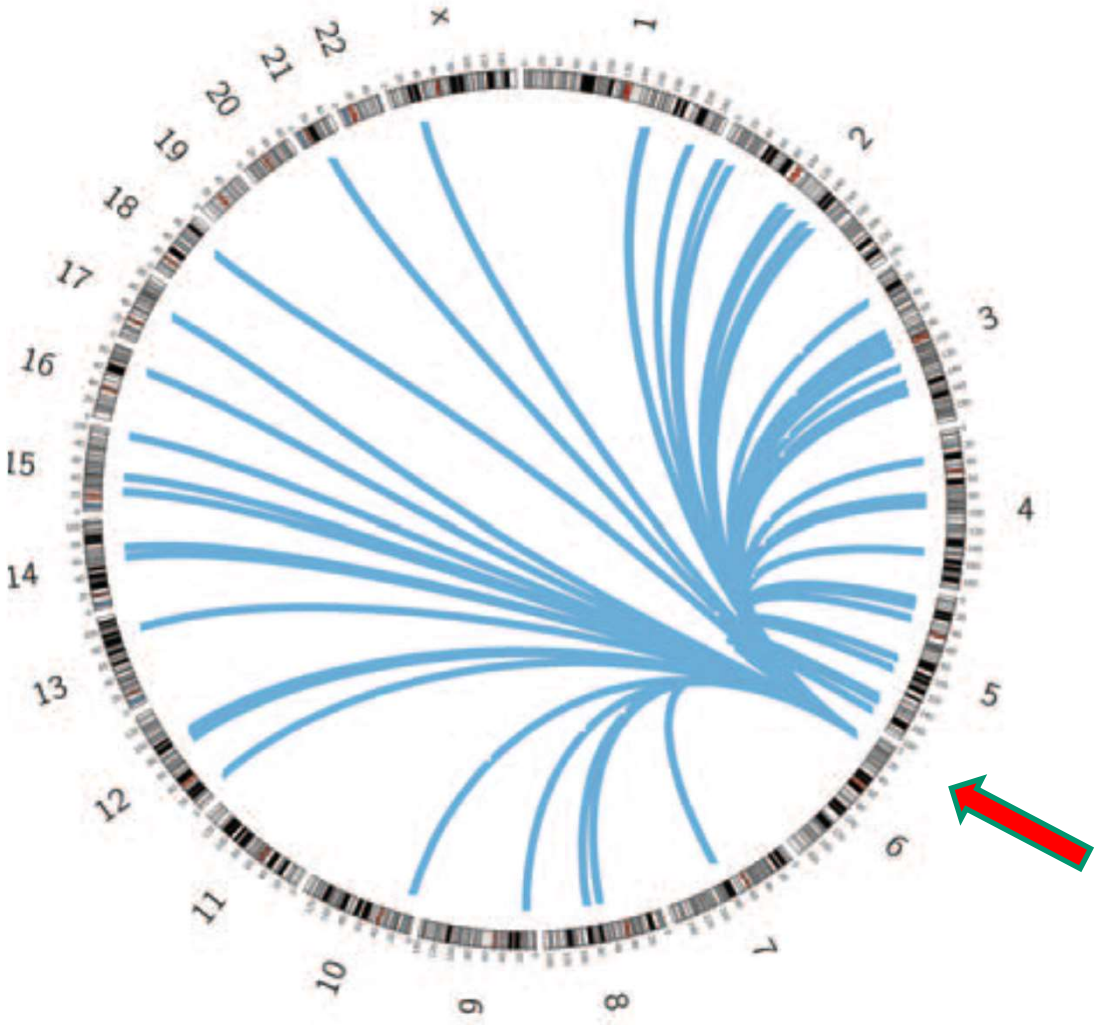

a
DNA LINE
P
ORF1
ORF2
AAAA
TTTT
5' UTR
3' UTR
trascrizione
b
mRNA LINE
5' UTR
3' UTR
5'
AAAA
3'
traduzione
proteine ORF1
e ORF2
legame
all'mRNA LINE
5'
AAAA
3'
+
DNA
bersaglio
5'
3'
AAAA
TTTT
legame al DNA bersaglio
c
5'
5'
3'
AAA
TT
3'



DNA damage response

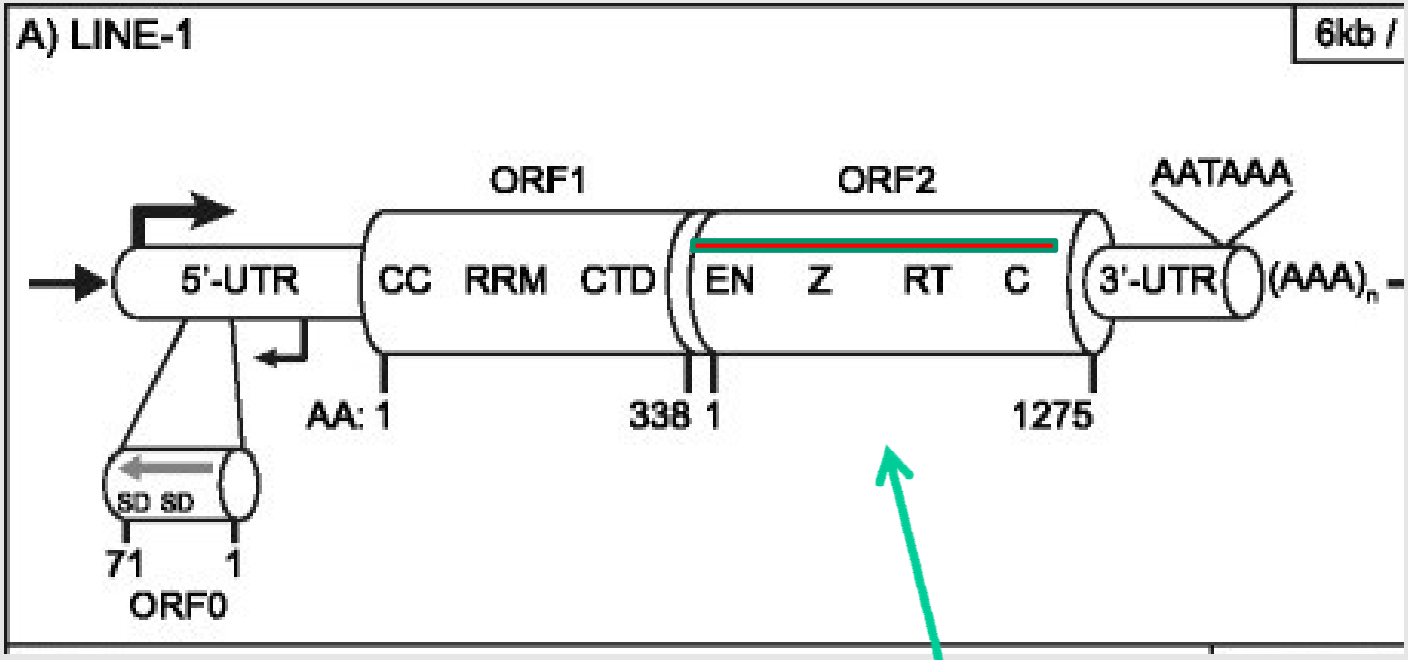
POCHISSIMI ELEMENTI L1 SONO ATTIVI

it is estimated, on the basis of full-length L1 elements with preserved open reading frames and activity in in vitro retrotransposition assays, that there are 50 to 120 currently active L1 repeats in the human genome, of which a small number are highly active -“hot-L1s”



Circos diagram mapping the distribution throughout the human genome of insertions of specific L1 elements

Le proteine/ENZIMI/complessi della trasposizione

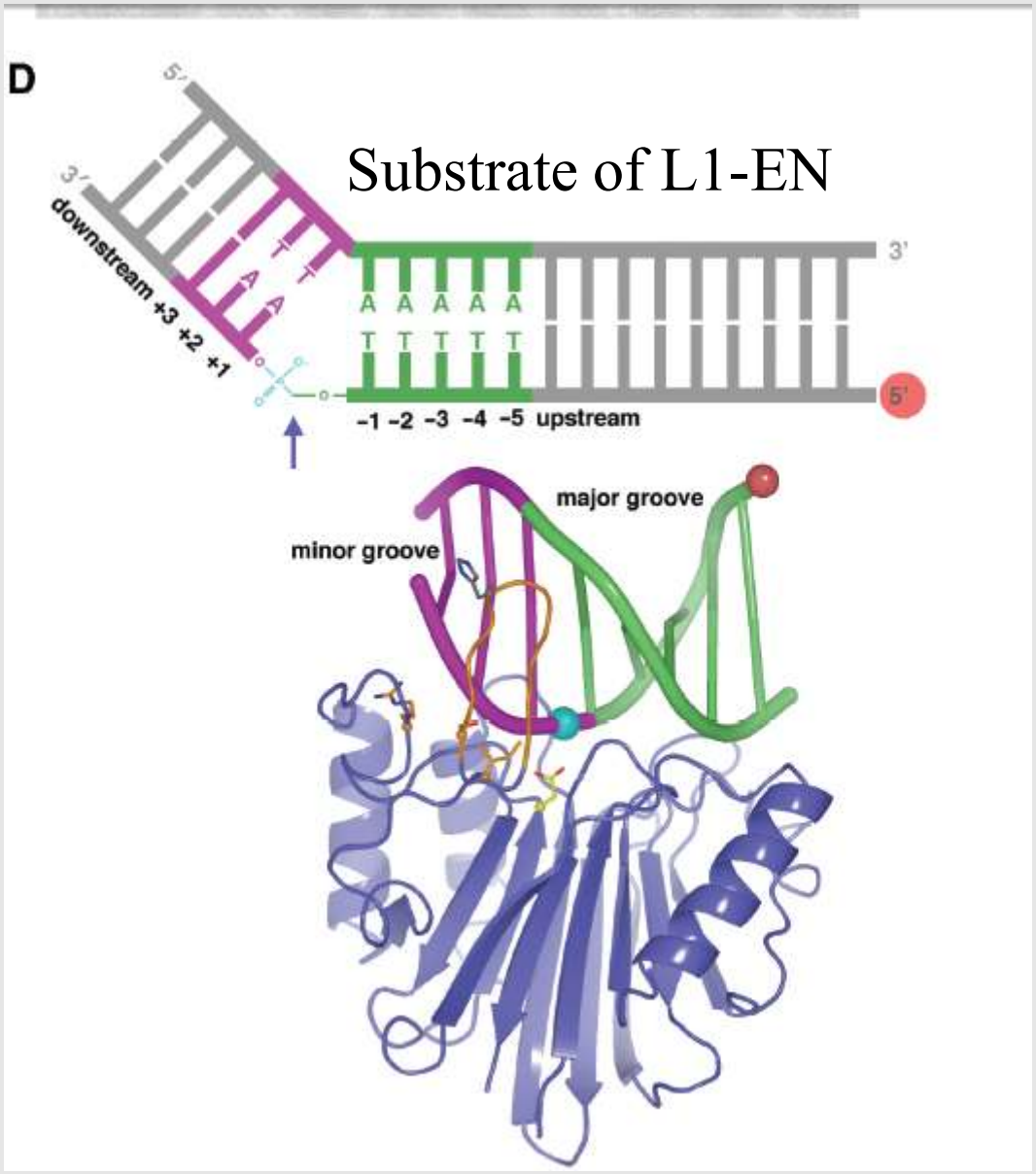


EN-endonuclease

Z domain

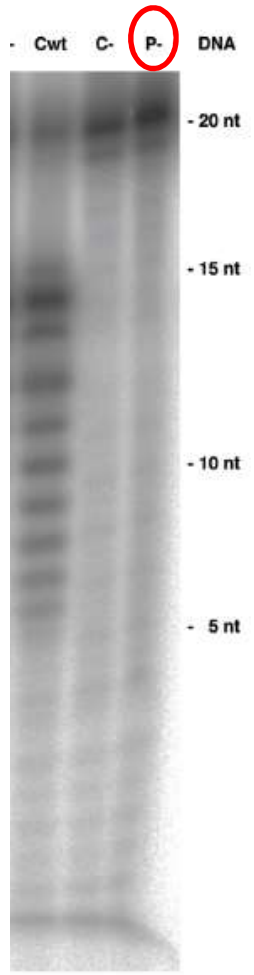
RT-reverse transcriptase

C-cysteine-rich.



Substrate of L1-EN

L1-EN cleavage



(P-): Cwt without L1-EN protein

Specificity of L1-EN

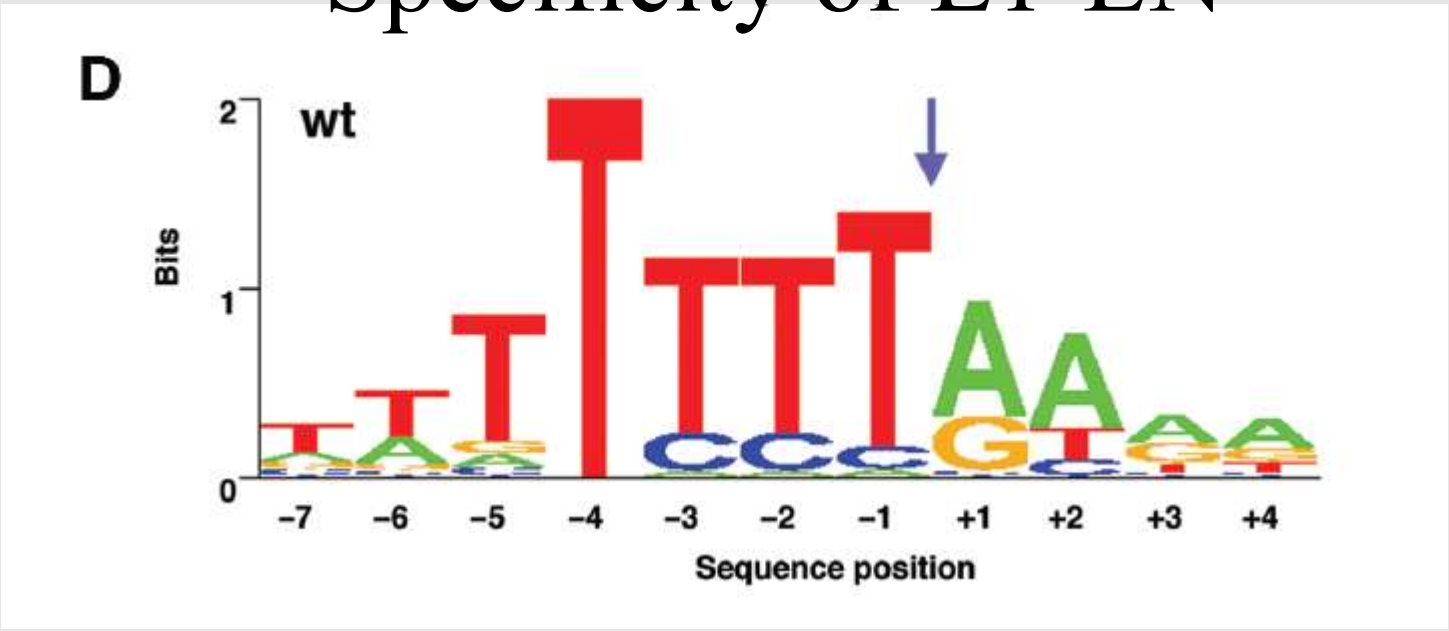


Table 1. Comparison of retrotransposition frequencies *in vivo* and plasmid nicking activities *in vitro*

L1-EN variant	Retrotransposition frequency[a], %	Plasmid nicking activity[b], %
wt	100 ± 17.1	100 ± 0.8
LTx	21 ± 2.4	29 ± 2.6
LR1	2 ± 2.3	6 ± 0.8
L3G	0 ± 2.2	10 ± 1.8
D145A	0[c]	3 ± 1.0
R155A	12 ± 3.3	19 ± 3.4
T192V	5 ± 3.0	–
S202A	32 ± 7.8	28 ± 2.2

EN mutants

Proporzionalità !!

L1-associated genomic regions are deleted in somatic cells of the healthy human brain.

Erwin JA Nat Neurosci. 2016 Sep 12

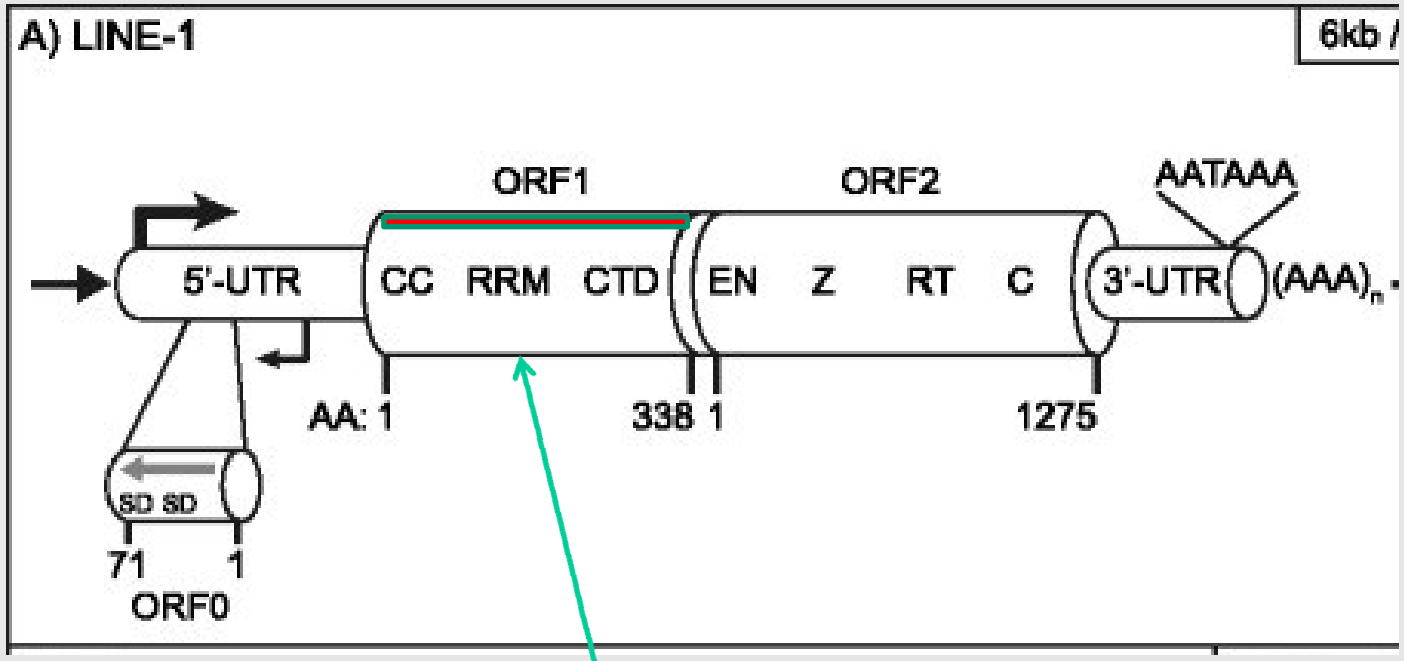
….somatic deletions generated by L1 endonuclease cutting activity.

L1-associated genomic regions …..contributor to somatic mosaicism.

…. present in **crucial neural genes** ….affect 44-63% of cells of the cells in the healthy brain.

PSD93
postsynaptic density proteins

- ORF1 Protein

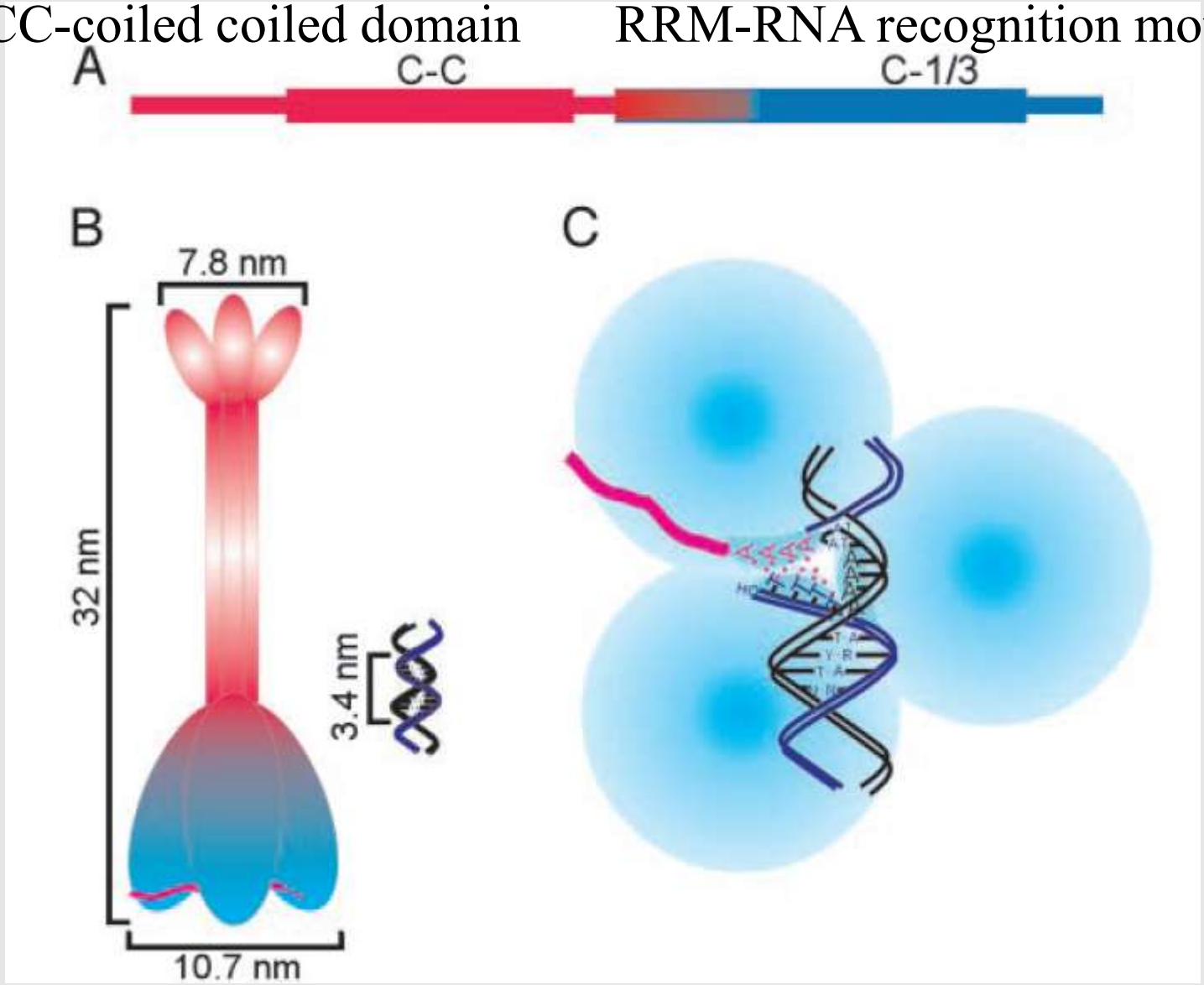


CC-coiled coiled domain

RRM-RNA recognition motif .

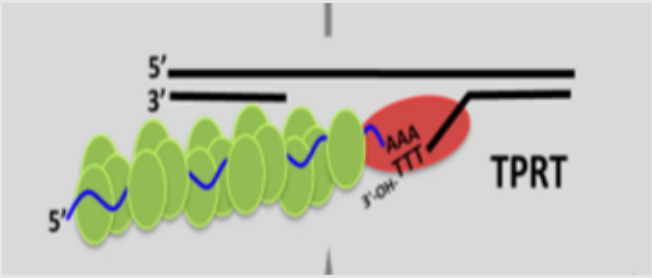
acidic(red) basic (blue)

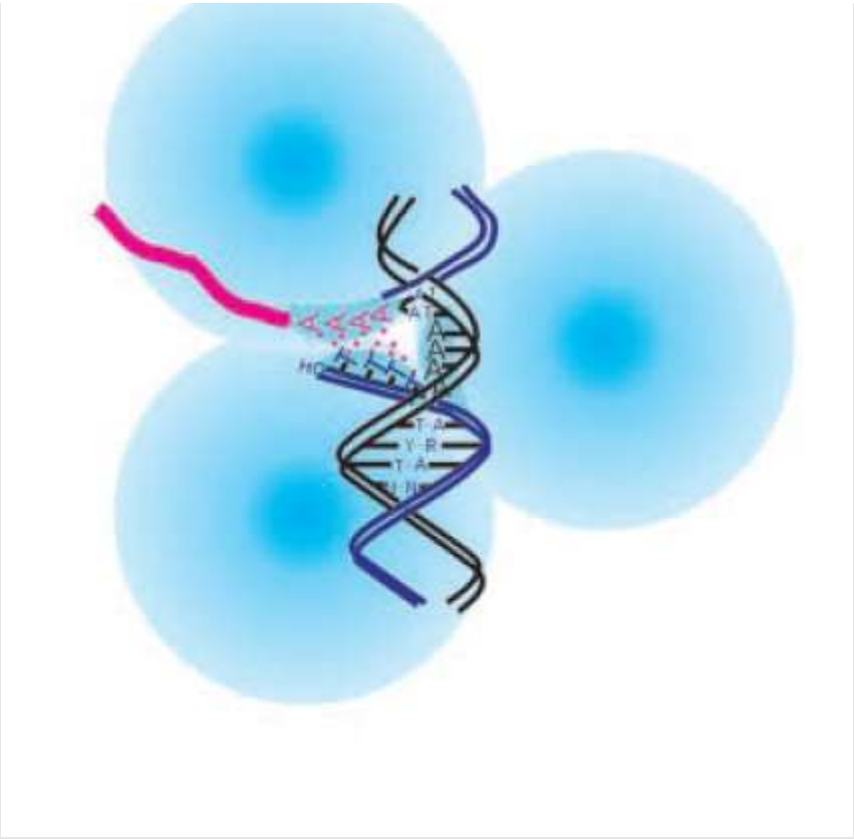
CC-coiled coiled domain RRM-RNA recognition motif



ORF1p binds single-stranded nucleic acids (L1 RNA and DNA) And functions as a nucleic acid chaperone.

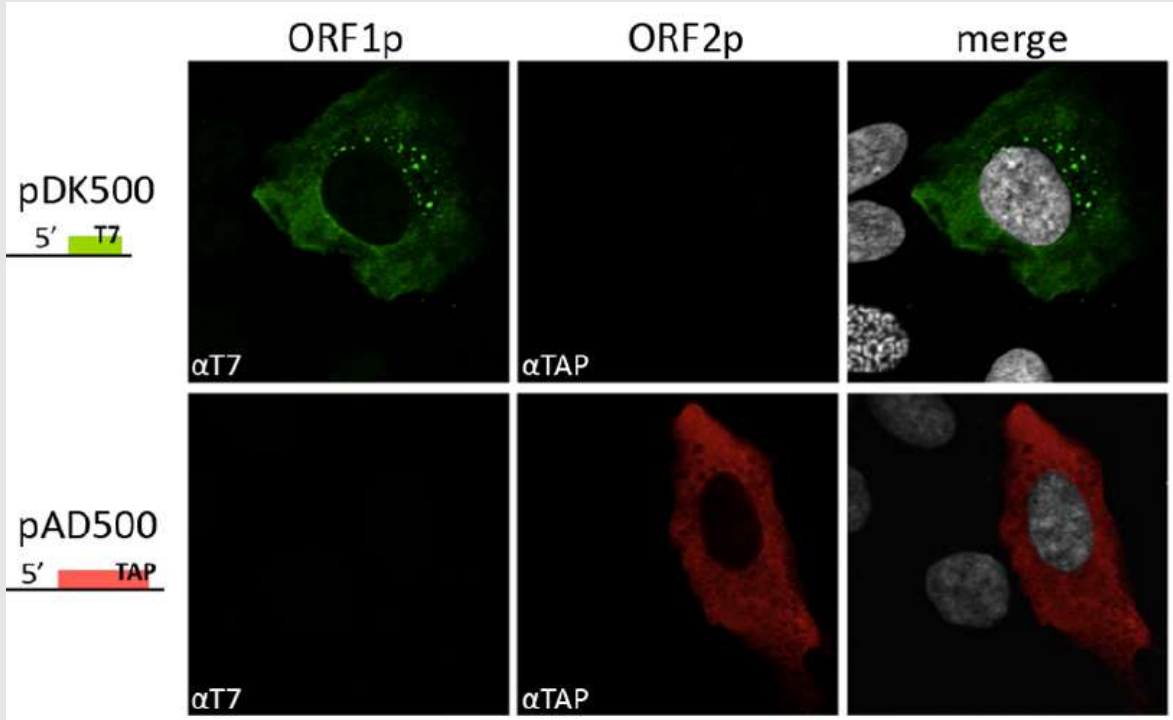
1 Each subunit of the trimer contains one single-stranded nucleic acid binding interface which is bound with one of the DNA target strands or the polyA tail of the L1 RNA (red). The double-stranded regions of the target are not bound

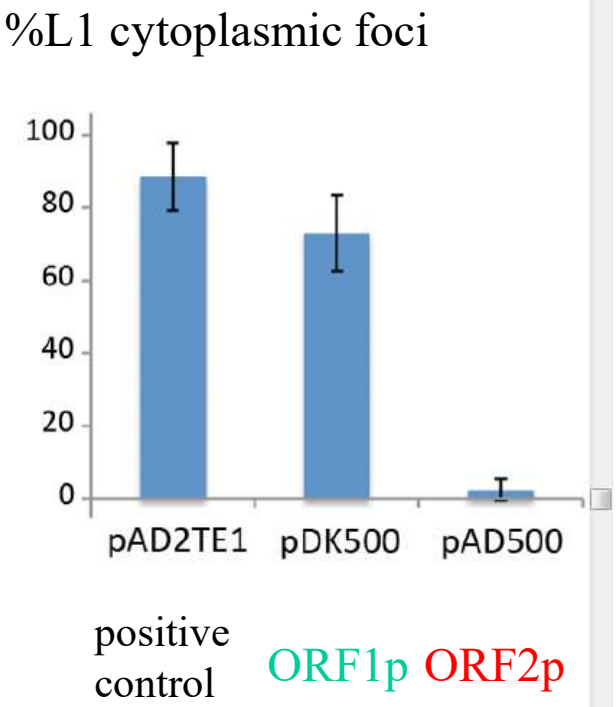




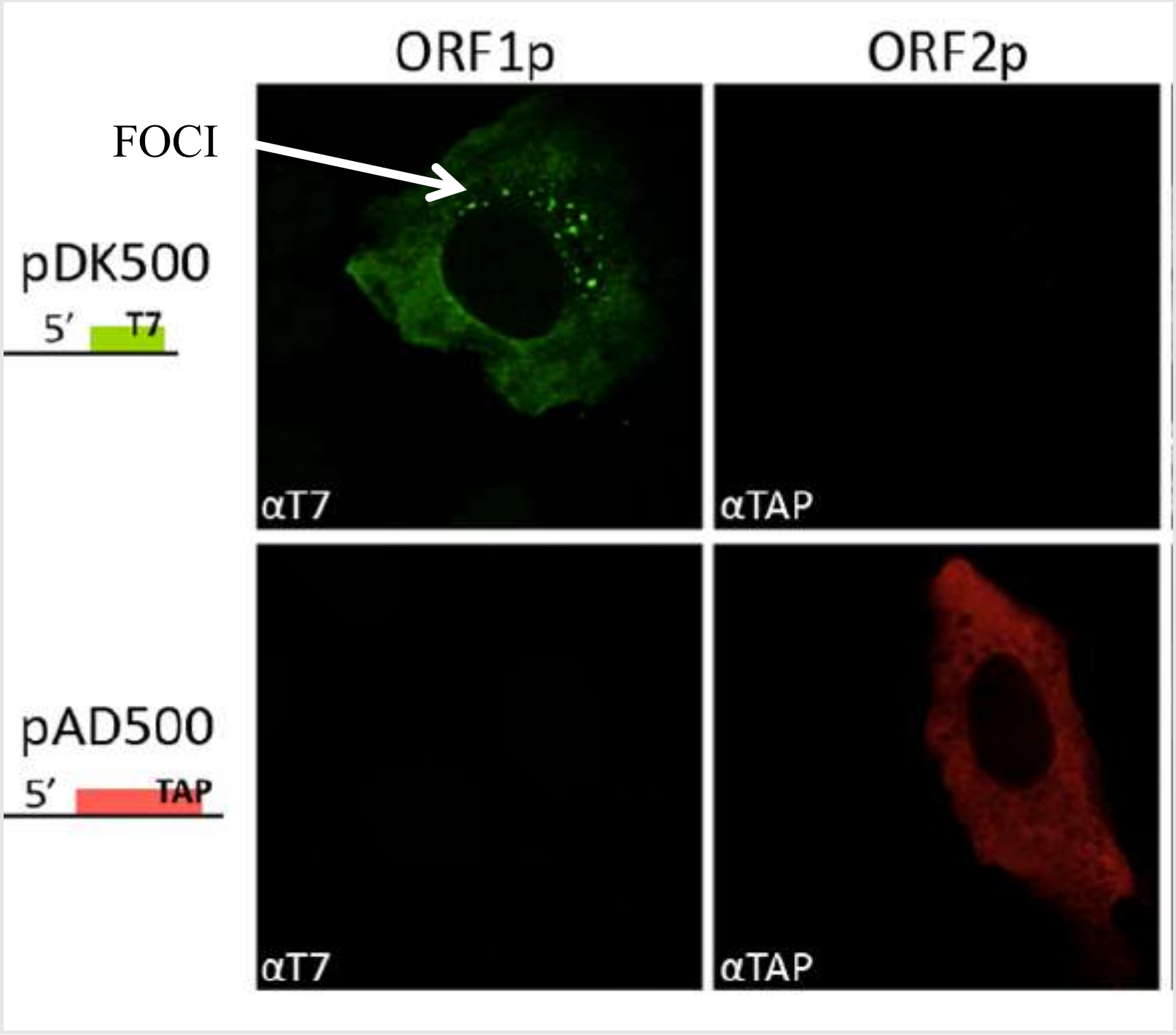
2 ORF1p coats the entire 7-kb L1 RNA to form a ribonucleoprotein particle
The nucleic acid chaperone activity of ORF1p melts the DNA and then facilitates formation of the RNA:DNA hybrid

ORF1p is necessary and sufficient for L1 cytoplasmic foci formation





T7-tagged ORF1p green
TAP-tagged ORF2p red;


ORF1p
ORF2p
FOCI
pDK500
5′
T7
αT7
αTAP
pAD500
5′
TAP
αT7
αTAP

Roles for retrotransposon insertions in human disease

Over evolutionary time, the dynamic nature of a genome is driven, in part, by the activity of transposable elements (TE) such as retrotransposons.

On a shorter time scale it has been established that new TE insertions can result in single-gene disease in an individual.

MUTAGENESI INSERZIONALE

To date in humans, > 1000 LINE-1-mediated insertions which result in genetic diseases have been reported.

Haemophilia A

Haemophilia A resulting from de novo insertion ofL1 sequences represents a novel mechanism for mutation in man.

Kazazian Antonarakis SE.

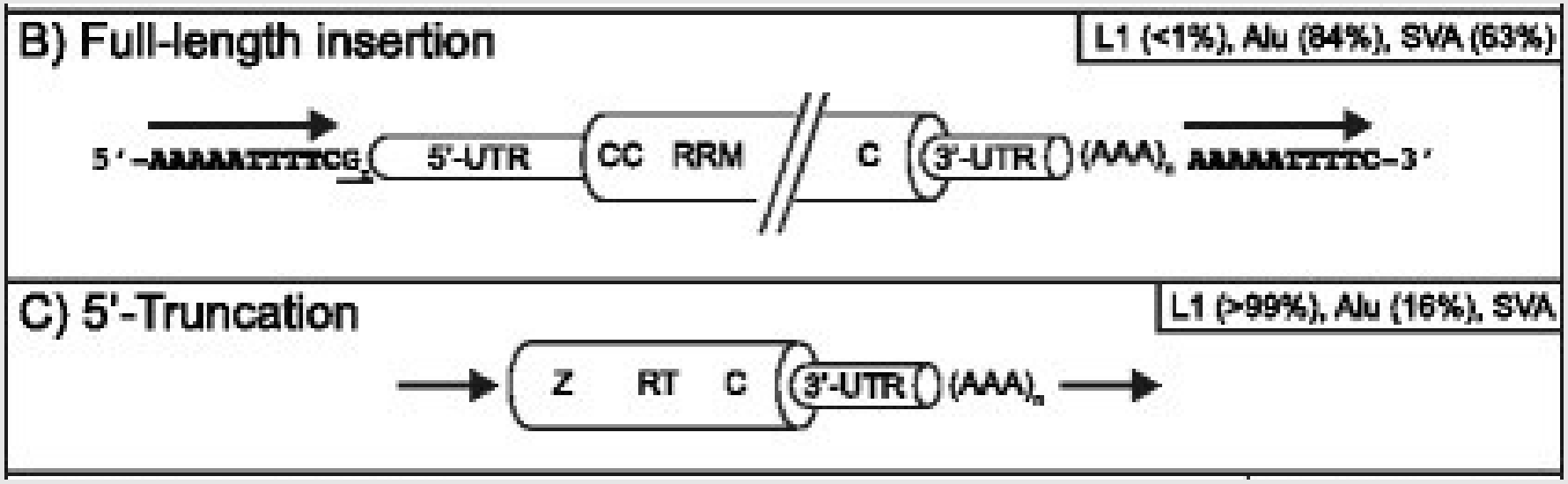
Insertions of L1 elements into exon 14 of the factor VIII gene in two of 240 unrelated patients with haemophilia A.

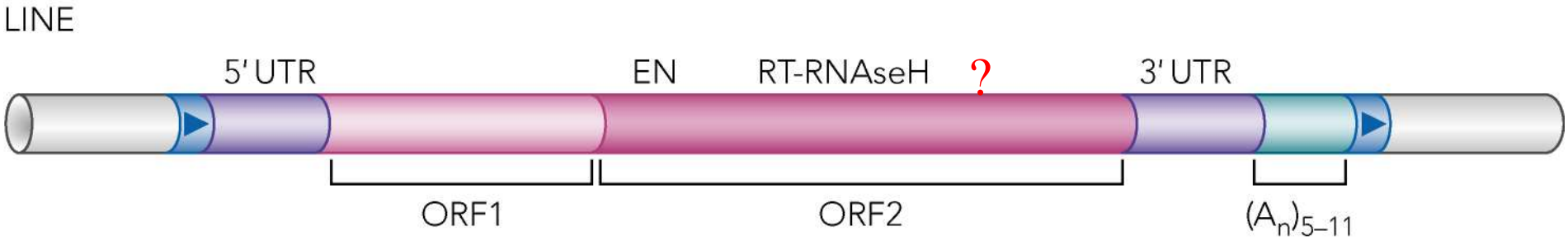
Both of these insertions (3.8 and 2.3 kilobases respectively) contain 3' portions of the L1 sequence, including the poly (A) tract, and create target site duplications of at least 12 and 13 nucleotides of the factor VIII gene.

Table 2 L1 EN-mediated retrotranspositions associated with human genetic diseas

Disrupted gene[a]	Chromosomal location	Disorder[b]	Inserted element	Insertion size (bp)	
Simple insertions					
	5q	Colon cancer	L1 Ta	520	
	Xq	Choroideremia	L1 Ta	6,017	
	Xp	CGD	L1 Ta	836	
	Xp	CGD	L1 Ta	1,722	
	Xp	DMD	L1 Ta	1,400	
	Xp	XLDCM	L1 Ta	530	
	Xq	Haemophilia A	L1 Ta	3,800	
	Xq	Haemophilia A	L1 preTa	2,300	
	Xq	Haemophilia B	L1 Ta	463	
	Xq	Haemophilia B	L1 Ta	163	
	11p	β-Thalassemia	L1 Ta	6,000	
	Xp	XLRP	L1 Ta	6,000	
	Xp	CLS	L1 HS	2,800	
	5q	Desmoid tumor	*Alu*Yb8	278	
	3q	Acholinesterasemia	*Alu*Yb9	289	
	13q	Breast cancer	*Alu*Yc1	281	
	Xq	XLA	*Alu*Y	–[e]	

retrotransposon insertions





516000 copies 100 copies/genome ACTIVE
0,9Kb mean length
–Truncated! Full-length 6 KB

20% genome

Half of the genome has stemmed from L1-mediated mobilization
¼ of promoters contain repetitive sequences