EVENTI PATOLOGICI RARI

Direct insertional mutagenesis by L1 resulted in diseases including muscular dystrophy, hemophilia, and breast cancer

Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanism for mutation in man.
Kazazian HH Jr, Wong C, Youssoufian H, Scott AF, Phillips DG, Antonarakis SE.
We now report 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.

- **Characterization of a nondeleterious L1 insertion in an intron of the human factor VIII gene** Woods-Samuels P, Wong C, Mathias SL, Scott AF, Kazazian HH Jr, Antonarakis SE.

    - **A 20.7 kb deletion within the factor VIII gene associated with LINE-1 element insertion.**

      Van de Water N, Williams R, Ockelford P, Browett P.

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					
APC	5q	Colon cancer	L1 Ta	520	
CHM	Xq	Choroideremia	L1 Ta	6,017	
CYBB	Xp	CGD	L1 Ta	836	
CYBB	Xp	CGD	L1 Ta	1,722	
DMD	Xp	DMD	L1 Ta	1,400	
DMD	Xp	XLDCM	L1 Ta	530	
F8	Xq	Haemophilia A	L1 Ta	3,800	
F8	Xq	Haemophilia A	L1 preTa	2,300	
F9	Xq	Haemophilia B	L1 Ta	463	
F9	Xq	Haemophilia B	L1 Ta	163	
HBB	11p	β-Thalassemia	L1 Ta	6,000	
RP2	Xp	XLRP	L1 Ta	6,000	
RPS6KA3	Xp	CLS	L1 HS	2,800	
APC	5q	Desmoid tumor	*Alu*Yb8	278	
BCHE	3q	Acholinesterasemia	*Alu*Yb9	289	
BRCA2	13q	Breast cancer	*Alu*Yc1	281	
BTK	Xq	XLA	*Alu*Y	–[e]	

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”

Extensive transduction of nonrepetitive DNA mediated by L1 retrotransposition in cancer genomes Science 2014 Tubio et al

L1 elements propagate through RNA intermediates.

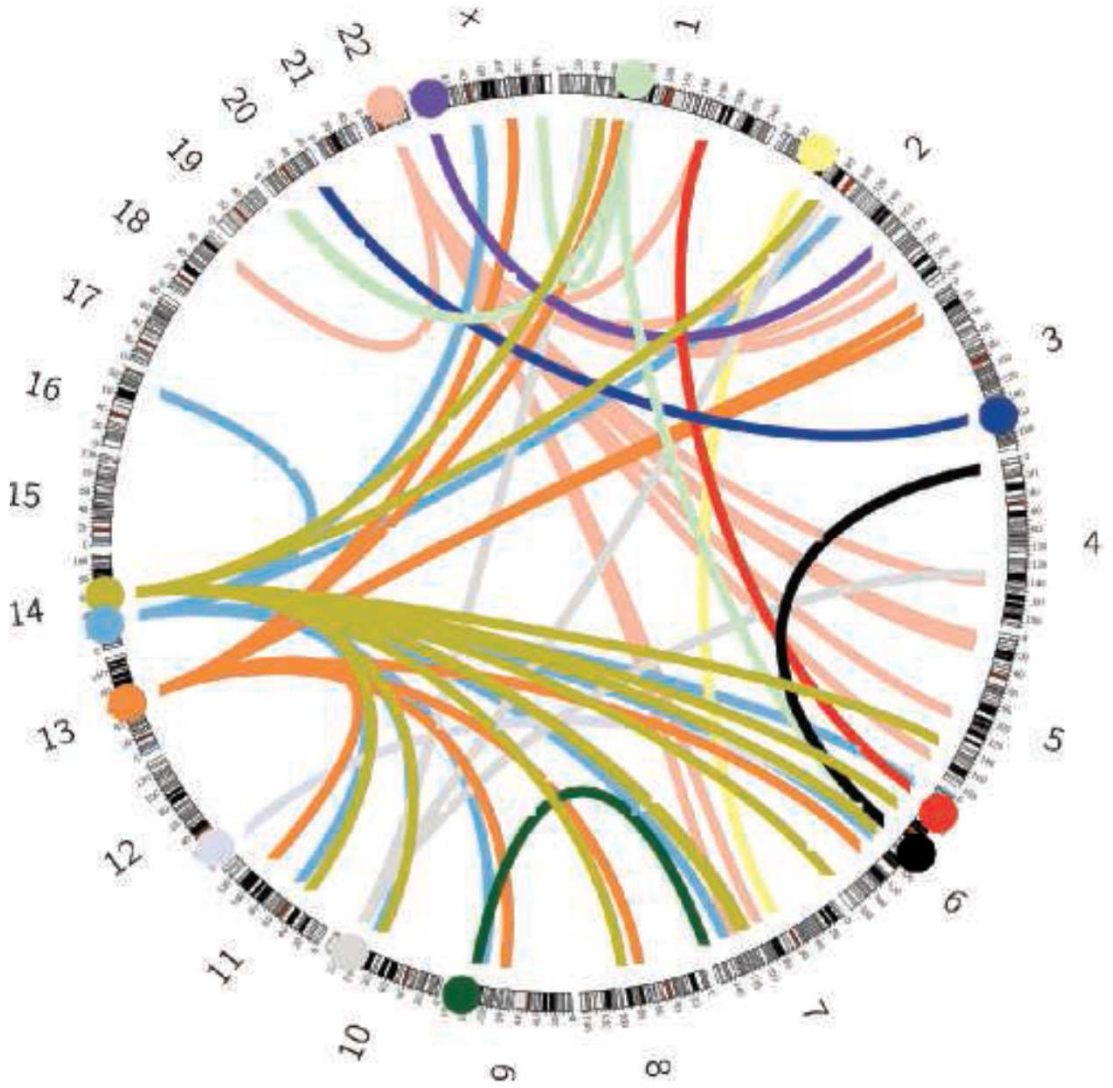
Studying cancer genomes from 244 patients, we found that tumors from 53% of the patients had somatic retrotranspositions ……..

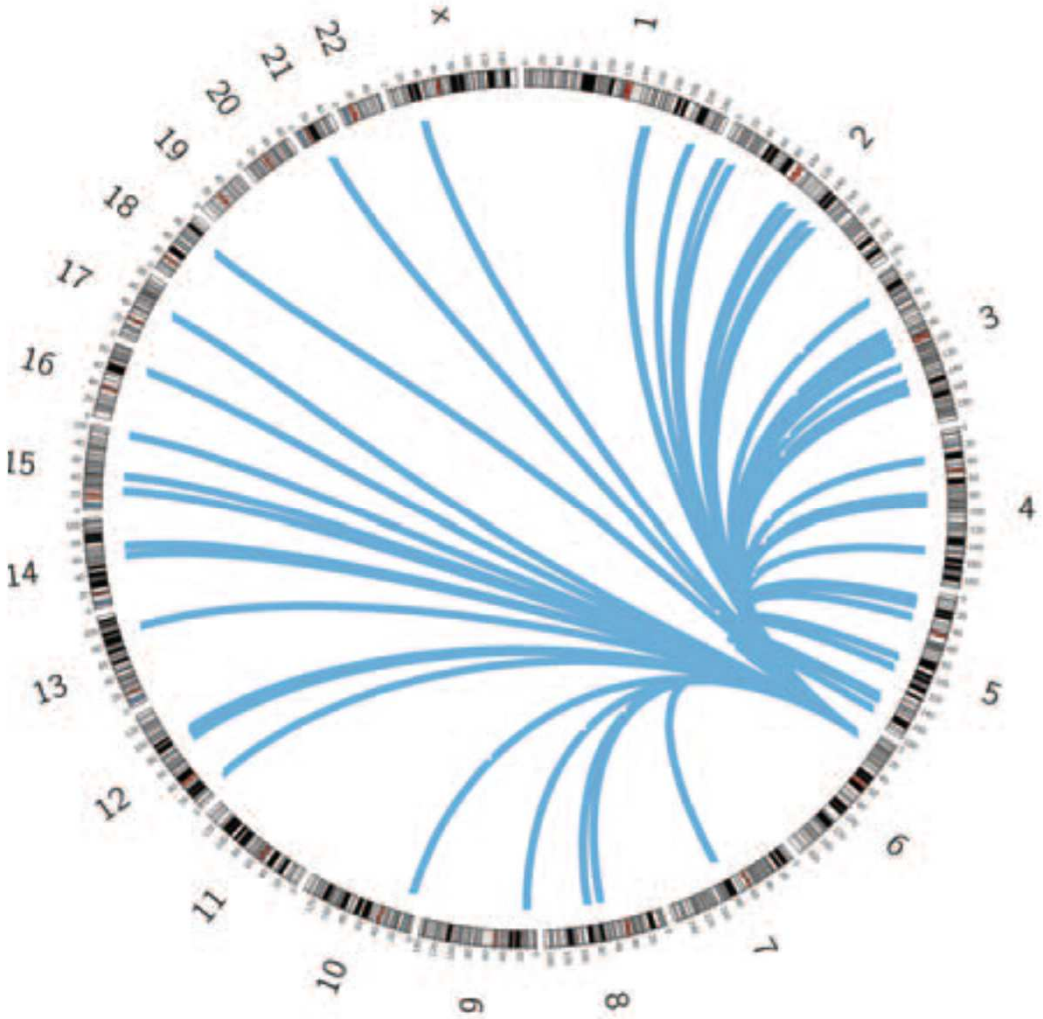
Fingerprinting of donor L1s revealed that a handful of source L1 elements in a tumor can spawn from tens to hundreds of 3′ transductions, which can themselves seed further retrotranspositions.

The activity of individual L1 elements correlated with L1 promoter hypomethylation.

The transductions disseminated genes, exons, and regulatory elements to new locations, **most often to heterochromatic regions** of the genome

In a lung tumor, hundreds of 3′ transductions arose from a small number of active L1 source elements (colored circles)




1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
X

..Somatic transduction occurs frequently in human tumors which can scatter exons, genes, and regulatory elements
widely across the genome.
Dissemination of these sequences appears to be due to a small number of highly active L1 elements.

The majority of the retrotransposition events are likely to be harmless “passenger” mutations.

REGOLAZIONE DELLA TRASCRIZIONE DA TRASPOSONI NEL GENOMA

Transcriptional disruption by the L1 retrotransposon and implications for mammalian transcriptomes

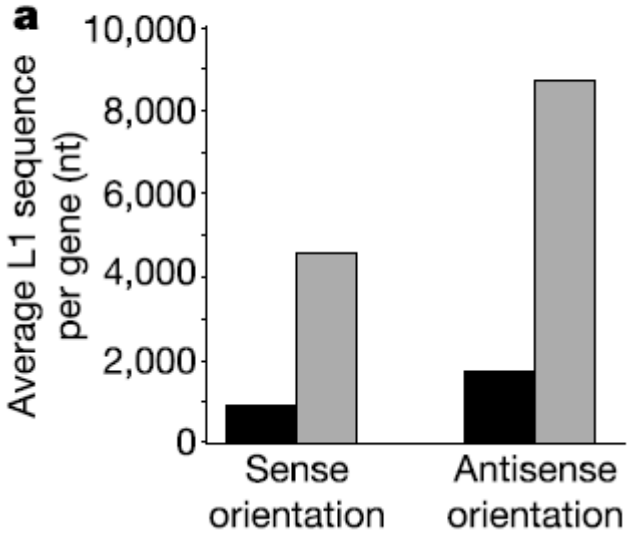
Jeffrey S. Han[1], Suzanne T. Szak[2] & Jef D. Boeke[1]

[1]*Department of Molecular Biology and Genetics and High Throughput Biology Center, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, USA*
[2]*Biogen, Inc., Cambridge, Massachusetts 02142, USA*

LINE-1 (L1) elements are the most abundant autonomous retrotransposons in the human genome, accounting for about 17% of human DNA. The L1 retrotransposon encodes two proteins, open reading frame (ORF)1 and the ORF2 endonuclease/reverse transcriptase. L1 RNA and ORF2 protein are difficult to detect in mammalian cells, even in the context of overexpression systems. Here we show that inserting L1 sequences on a transcript significantly decreases RNA expression and therefore protein expression. This decreased RNA concentration does not result from major effects on the transcription initiation rate or RNA stability. Rather, the poor L1 expression is primarily due to inadequate transcriptional elongation. Because L1 is an abundant and broadly distributed mobile element, the inhibition of transcriptional elongation by L1 might profoundly affect expression of endogenous human genes. We propose a model in which L1 affects gene expression genome-wide by acting as a 'molecular rheostat' of target genes. Bioinformatic data are consistent with the hypothesis that L1 can serve as an evolutionary fine-tuner of the human transcriptome.

L1 intragenica ed espressione genica

High Low



L1 intragenica ed espressione genica

High Low

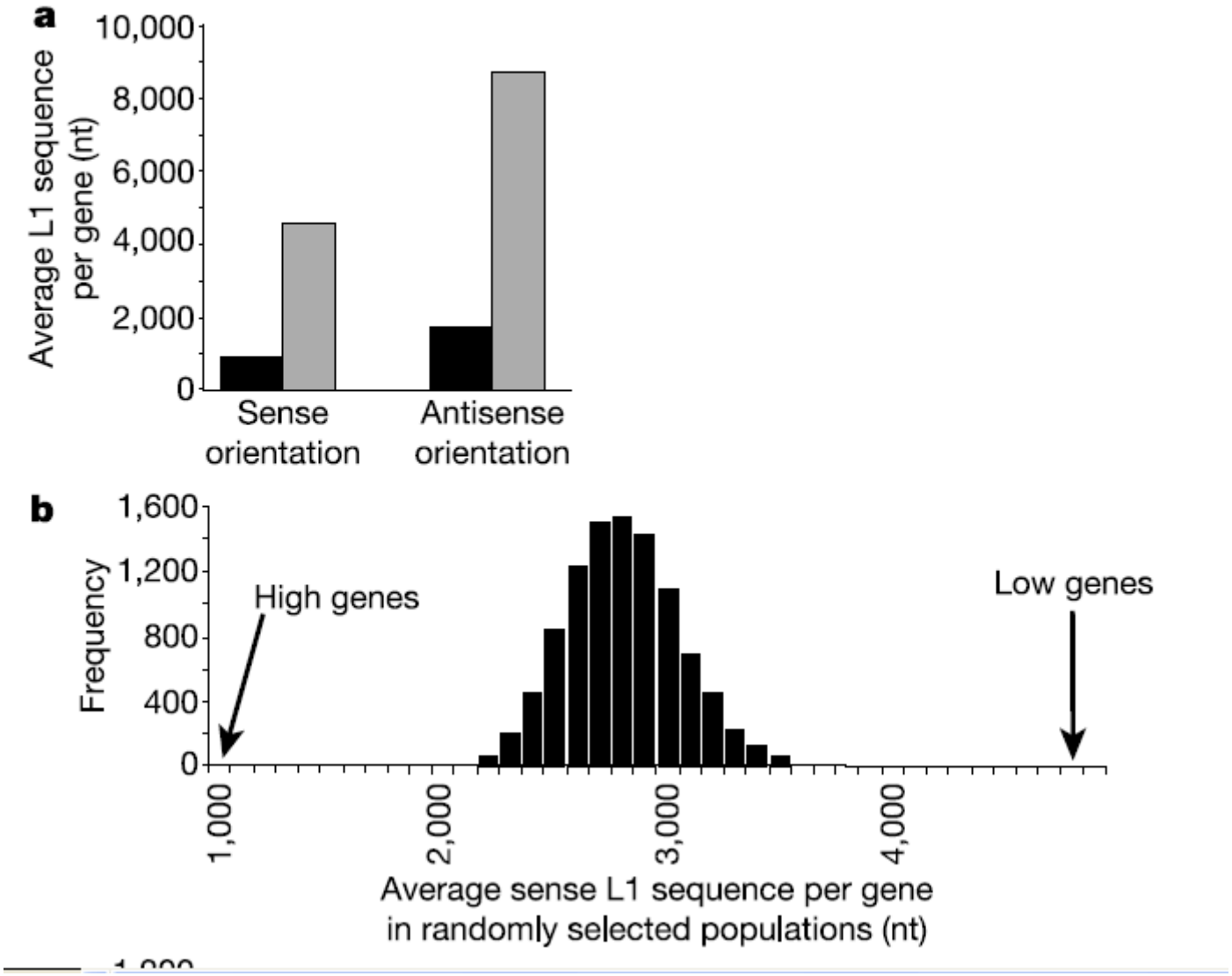
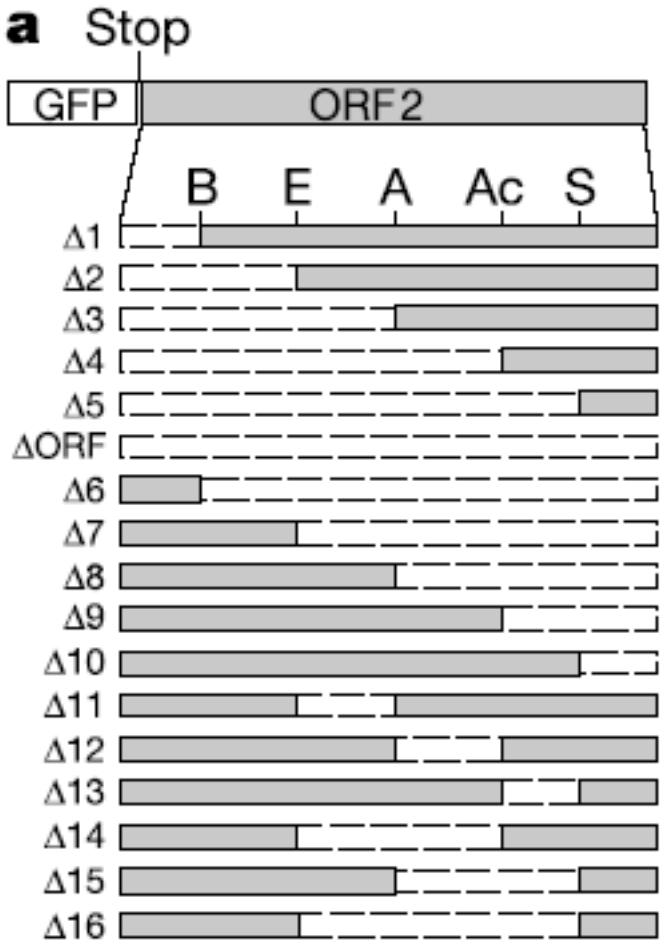


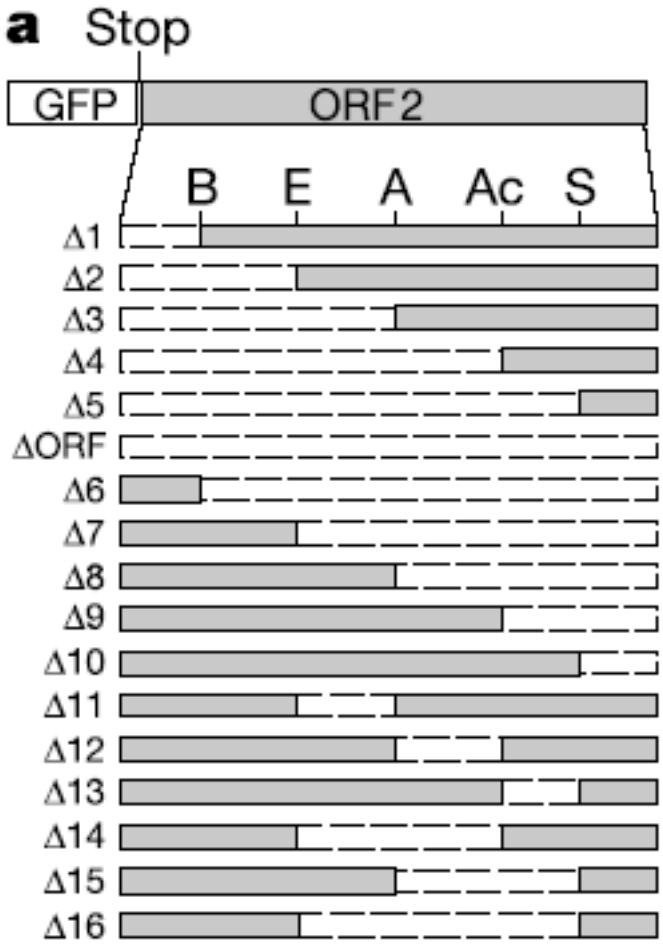
Figure 5 Bioinformatic analysis of L1 content in genes. **a**, Average L1 content of genomic loci of sets of highly (black bars) and poorly (grey bars) expressed genes (see Methods). **b**, Average L1 content in sets of randomly selected populations of genes (see Methods). Positions where the highly and poorly expressed genes would be (data superimposed from **a**) are indicated and are outside the random distribution ($P < 0.01$). **c**, Data from **a**, normalized to total intron content. **d**, Highly and poorly expressed genes were sorted into high GC, low L1 isochore or low GC, high L1 isochore[50] classes. The percentage of each expression class falling into each isochore is indicated. Subpopulations were analysed as described in **a** and **c**.

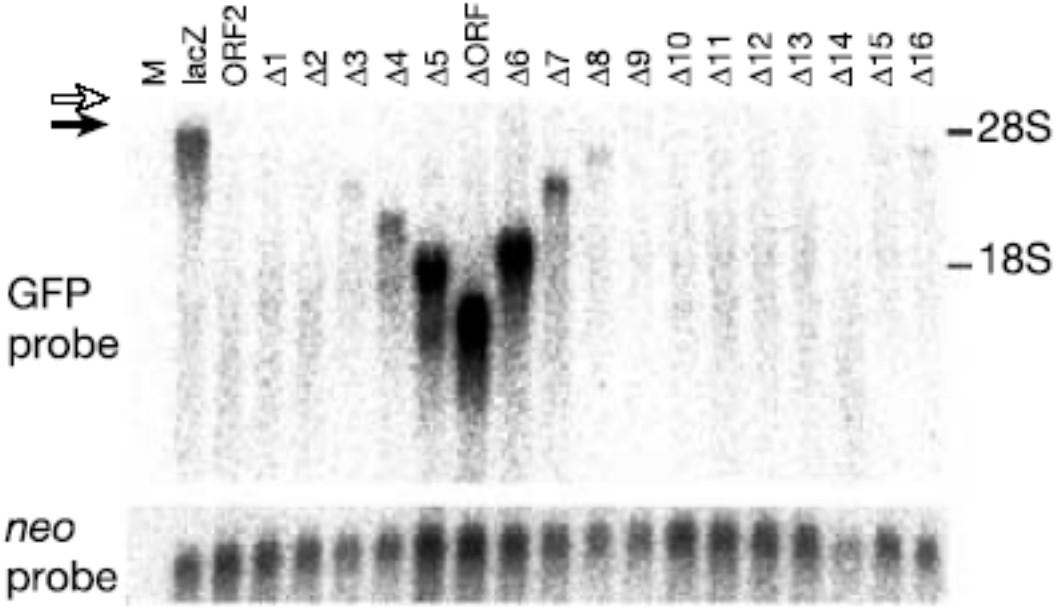
L1 mutants


a
Stop
GFP
ORF2
B
E
A
Ac
S
Δ1
Δ2
Δ3
Δ4
Δ5
ΔORF
Δ6
Δ7
Δ8
Δ9
Δ10
Δ11
Δ12
Δ13
Δ14
Δ15
Δ16

L1 mutants

RNA analysis





Decrease in expression depends on L1 length

Figure 3 Decrease in L1 expression is dependent on length. **a**, The left panel depicts the structures of deletion constructs. Hollow regions represent deleted sequences. B, *Bbv*CI; E, *Eco*RI; A, *Afl*II; Ac, *Acl*I; S, *Spe*I. The right panel shows a total RNA analysis of HeLa transfections. Lanes: M, mock; lacZ, pGFPstoplacZ; ORF2, pGFPstopORF2. Open and black arrows show the expected positions of GFPstopORF2 and GFPstoplacZ, respectively. **b**, The adenosine base composition of the sense strand, in 50-nucleotide windows, was plotted for each position in L1.2 with MacVector 6.5.3 (Oxford Molecular). **c**, The top panel shows the structures of GFPstopORF1, GFPstop4ORF1 and GFPstop5UTR. The 4ORF1 repeat is about 4,500 nucleotides long and the 5′ UTR repeat is about 4,000 nucleotides long. The bottom panel shows a total RNA analysis of HeLa transfections. Open, black and grey arrows show the expected positions of GFPstop4ORF1, GFPstop5UTR and GFPstopORF1, respectively.

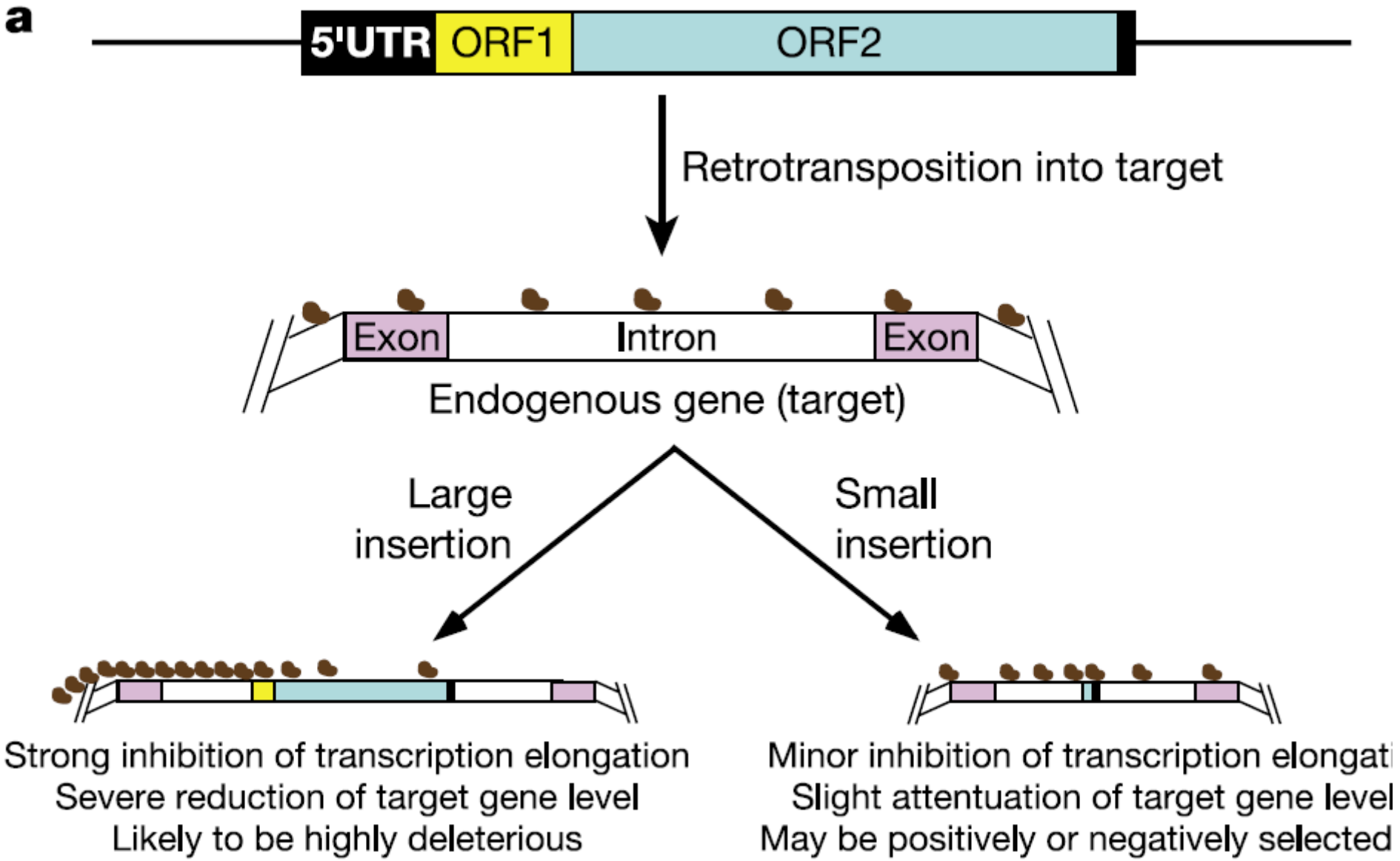

a
5'UTR
ORF1
ORF2
Retrotransposition into target
Exon
Intron
Exon
Endogenous gene (target)
Large insertion
Small insertion
Strong inhibition of transcription elongation
Severe reduction of target gene level
Likely to be highly deleterious
Minor inhibition of transcription elongati
Slight attentuation of target gene level
May be positively or negatively selected

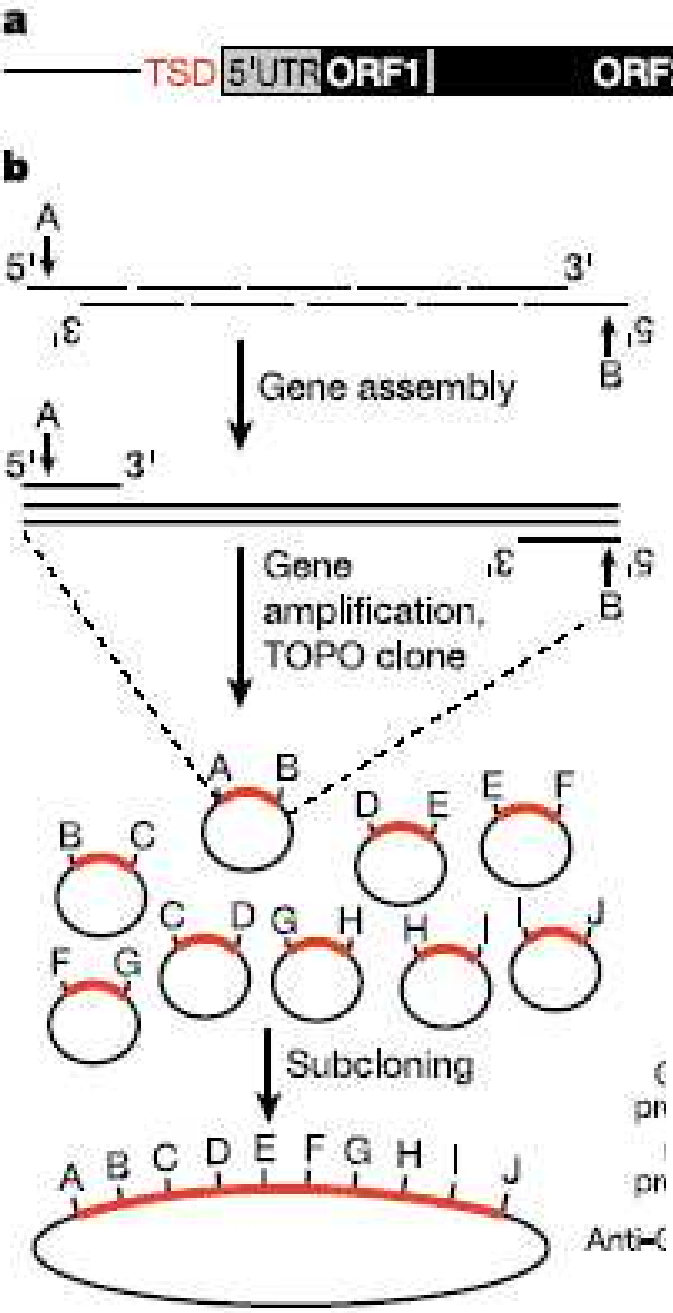
Figure 6 Models for L1-mediated modulation of gene expression/structure. **a**, Effects on transcription. Brown dots represent transcriptional complexes, which could be slowed, paused or dissociated from the templates on encountering significant amounts of L1 sequence. **b**, Effects on mRNA and protein structure. Left, hypothetical gene with three exons. Middle, intronic sense L1 insertions can produce a minor amount of prematurely polyadenylated mRNA, potentially giving rise to a truncated protein with additional, previously untranslated amino acids at the C terminus (white segment). Right, intronic antisense L1 insertions can produce a major amount of prematurely polyadenylated mRNA.

A highly active synthetic mammalian retrotransposon

Jeffrey S. Han & Jef D. Boeke

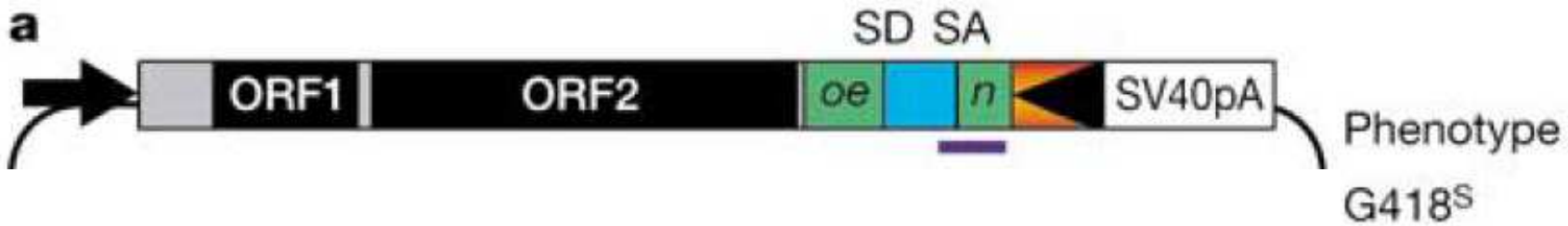
Department of Molecular Biology and Genetics and High Throughput Biology Center, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, USA

LINE-1 (L1) elements are retrotransposons that comprise large fractions of mammalian genomes[1]. Transcription through L1 open reading frames is inefficient owing to an elongation defect[2], inhibiting the robust expression of L1 RNA and proteins, the substrate and enzyme(s) for retrotransposition[3–5]. This elongation defect probably controls L1 transposition frequency in mammalian cells. Here we report bypassing this transcriptional defect by synthesizing the open reading frames of L1 from synthetic oligonucleotides, altering 24% of the nucleic acid sequence without changing the amino acid sequence. Such resynthesis led to greatly enhanced steady-state L1 RNA and protein levels. Remarkably, when the synthetic open reading

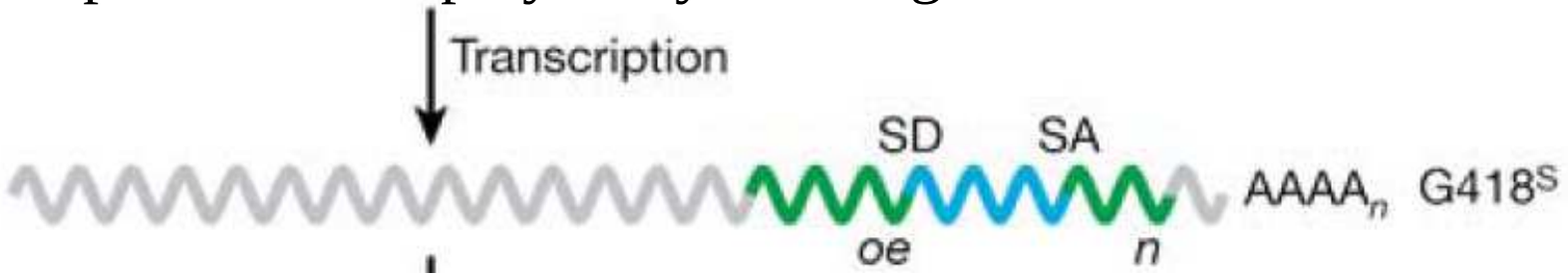


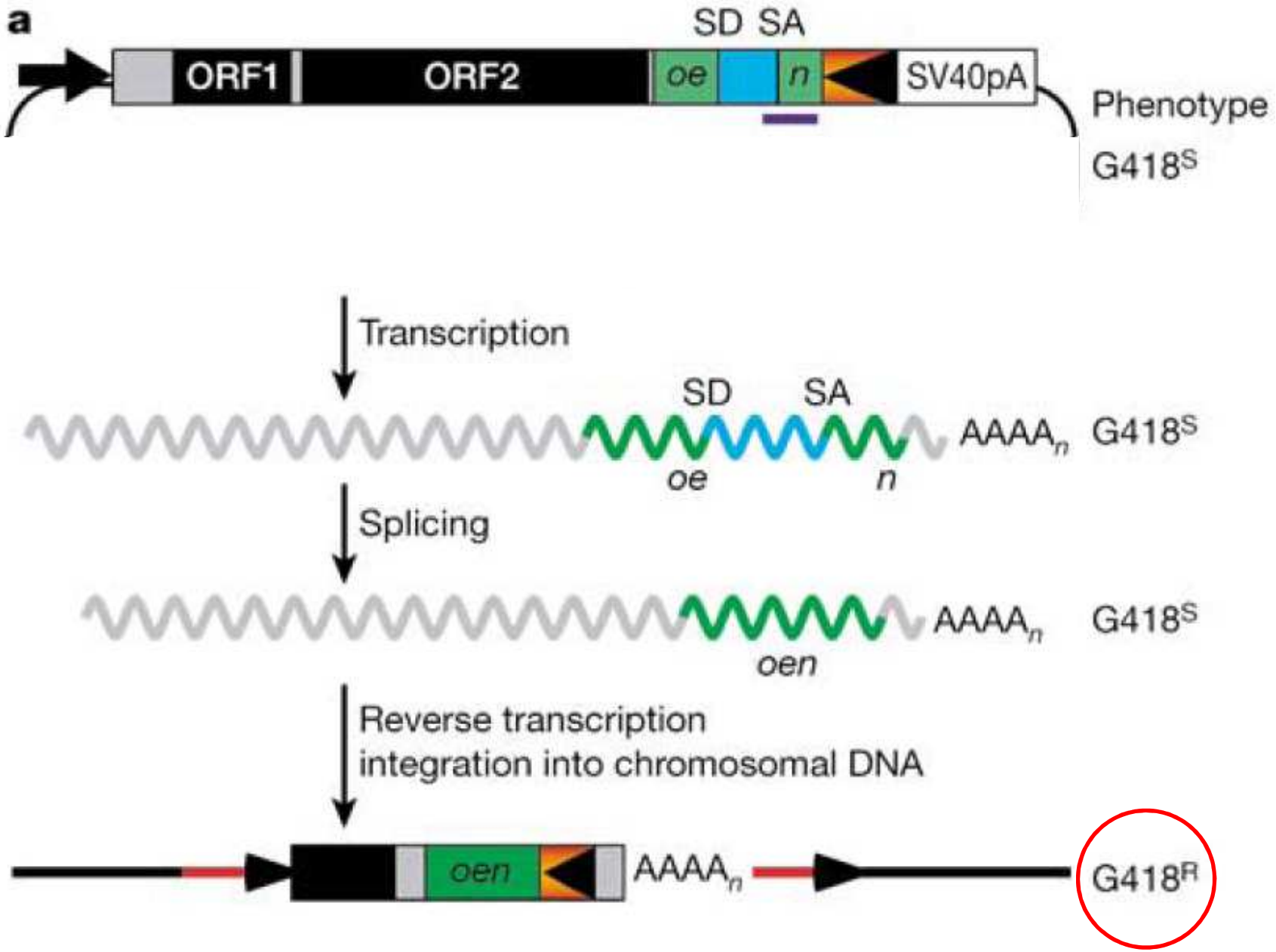
Overview of gene synthesis. Oligonucleotides encoding each fragment were mixed and subsequently used as template amplification.

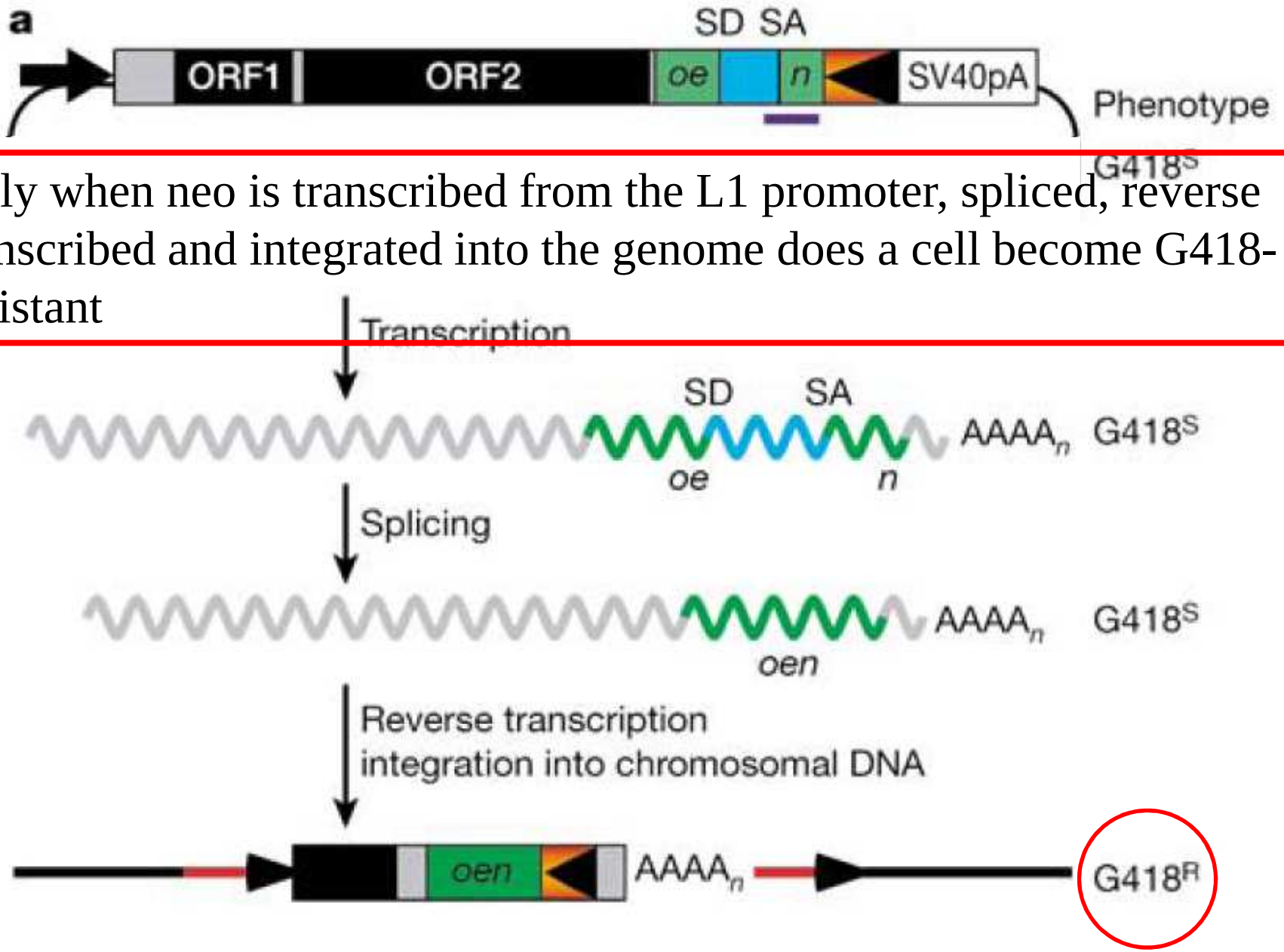
products were cloned and ligated together with unique restriction sites (labelled A to J)



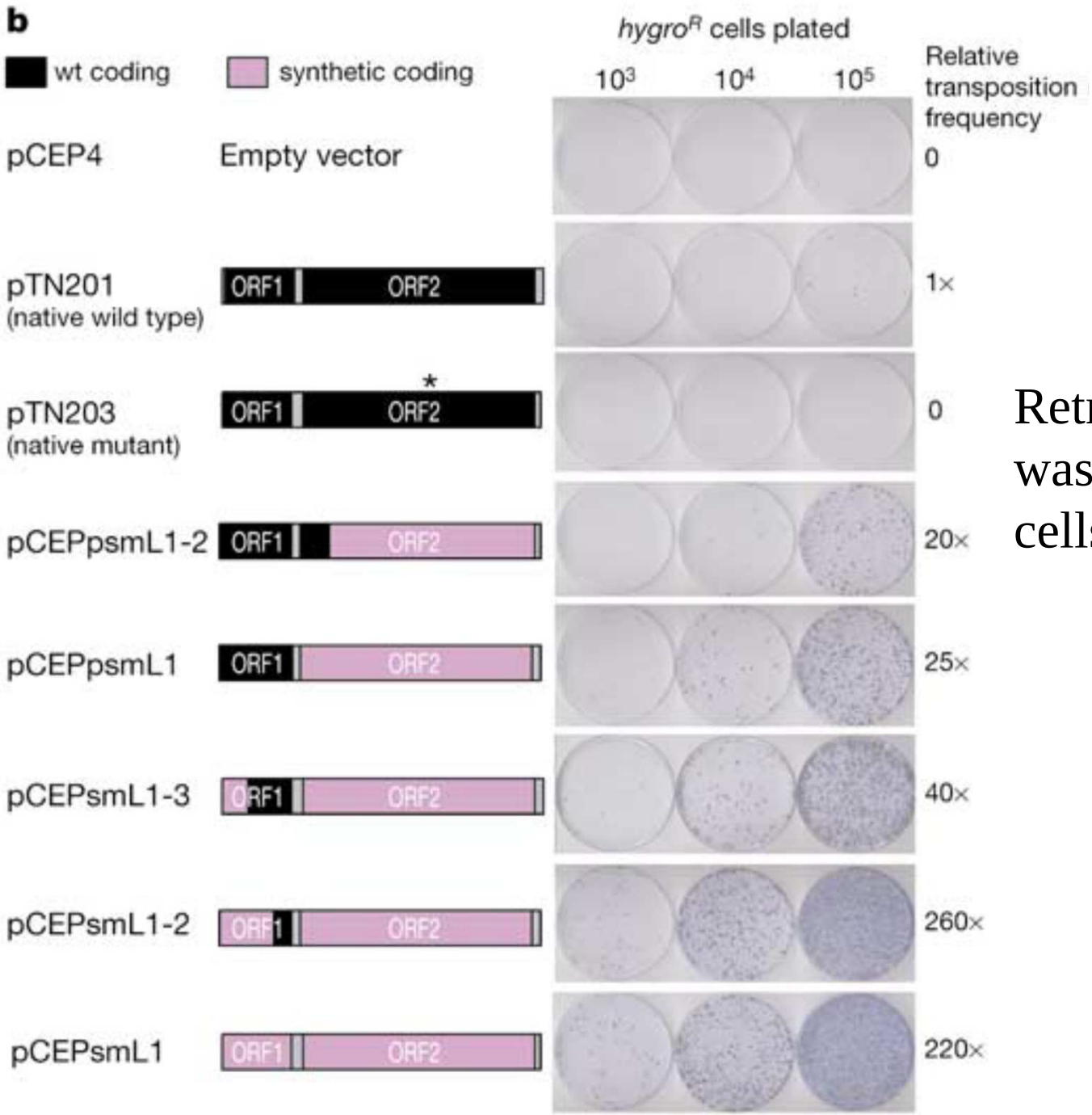
an intron-interrupted neo reporter in the 3’ untranslated region with its own promoter and polyadenylation signal




a
SD SA
ORF1
ORF2
oe
n
SV40pA
Phenotype
G418S
Transcription
SD
SA
AAAAn
G418S
oe
n
Splicing
AAAAn
G418S
oen
Reverse transcription
integration into chromosomal DNA
oen
AAAAn
G418R



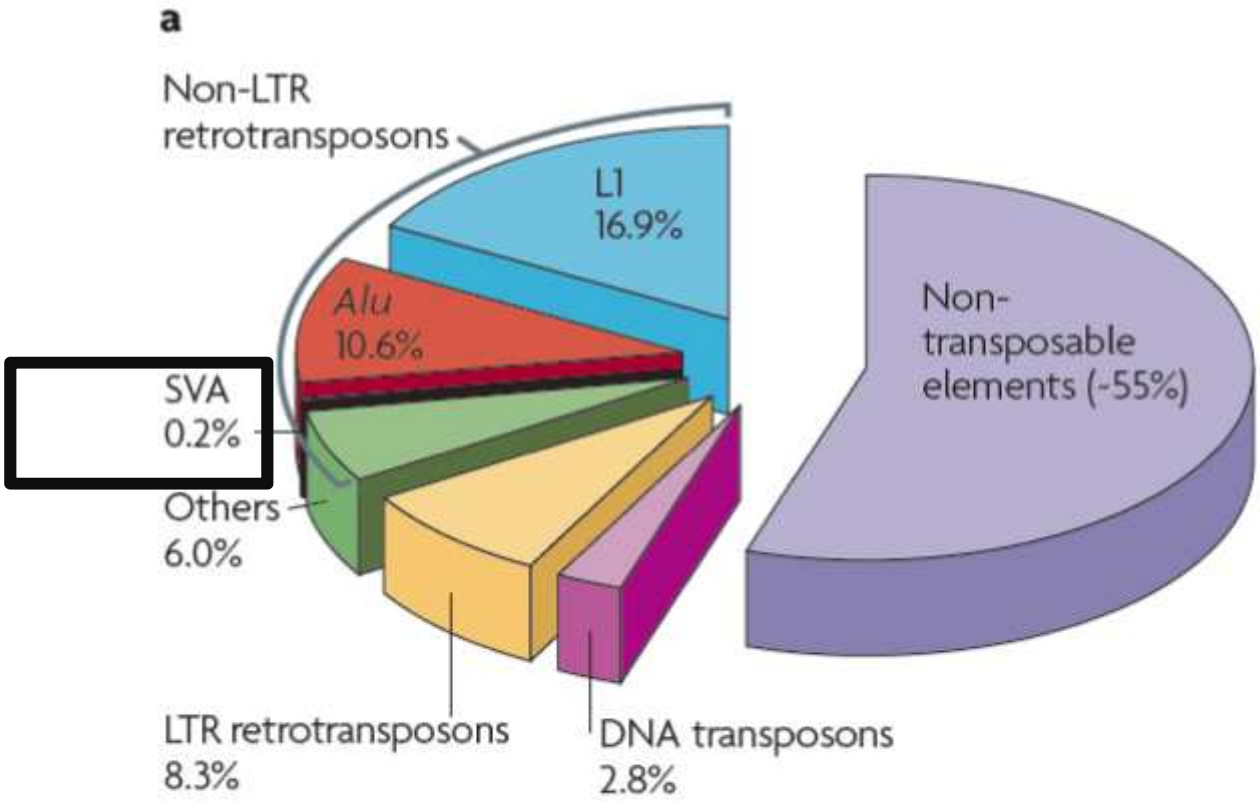
Only when neo is transcribed from the L1 promoter, spliced, reverse transcribed and integrated into the genome does a cell become G418-resistant



Retrotransposition was assayed in HeLa cells

Figure 2 Retrotransposition of synthetic mL1. **a**, The retrotransposition assay. The L1 element contains an intron-interrupted *neo* reporter in the 3′ untranslated region with its own promoter and polyadenylation signal. Only when *neo* is transcribed from the L1 promoter, spliced, reverse transcribed and integrated into the genome does a cell become G418-resistant[4]. Blue lines represent probes for RNA analysis (Fig. 4). SD, splice donor; SA, splice acceptor. **b**, Retrotransposition was assayed in HeLa cells ($N = 3$). pTN201 contains only wild-type native mouse L1 sequence, and pTN203 contains wild-type native mouse L1 sequence with a D709Y reverse transcriptase point mutation[22]. The average absolute number of colonies for pTN201 was 440 events per 10^6 transfected cells.

EVOLUZIONE RECENTE DEI TRASPOSONI

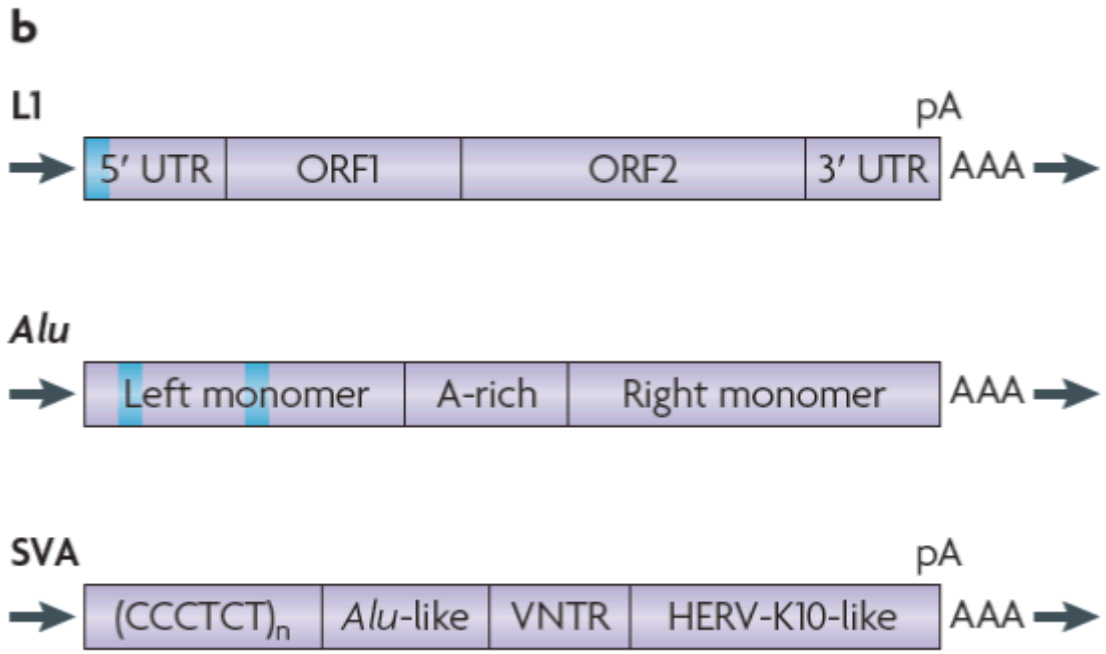

a
Non-LTR retrotransposons
L1 16.9%
Alu 10.6%
SVA 0.2%
Others 6.0%
Non-transposable elements (~55%)
LTR retrotransposons 8.3%
DNA transposons 2.8%

SVA

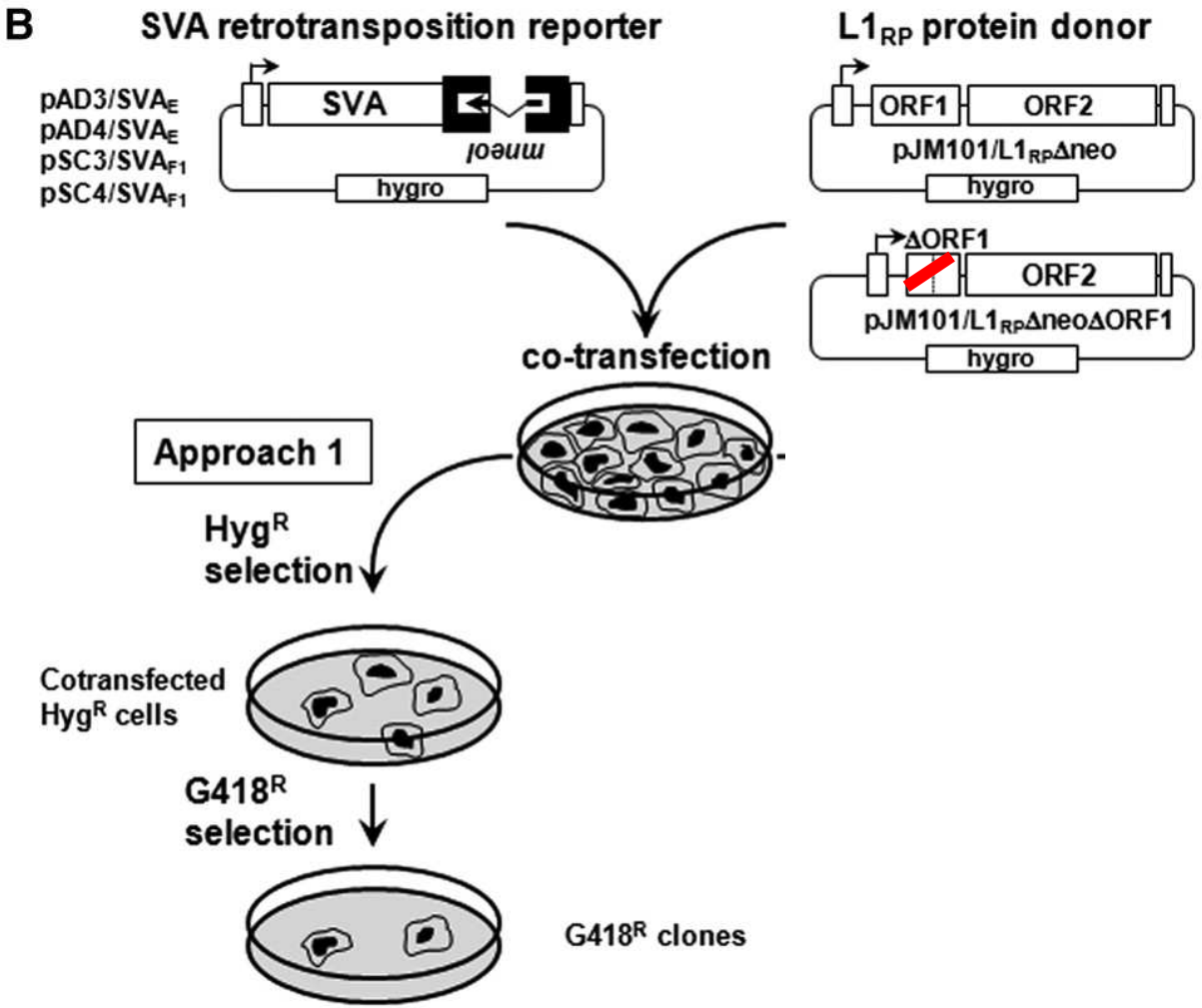
SINE-VNTR-Alu (SVA) elements are nonautonomous, hominid-specific non-LTR retrotransposons

composite mobile elements.

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


b
L1
pA
5′ UTR
ORF1
ORF2
3′ UTR
AAA
Alu
Left monomer
A-rich
Right monomer
AAA
SVA
pA
(CCCTCT)n
Alu-like
VNTR
HERV-K10-like
AAA

Rationale of the SVA trans-mobilization assay.

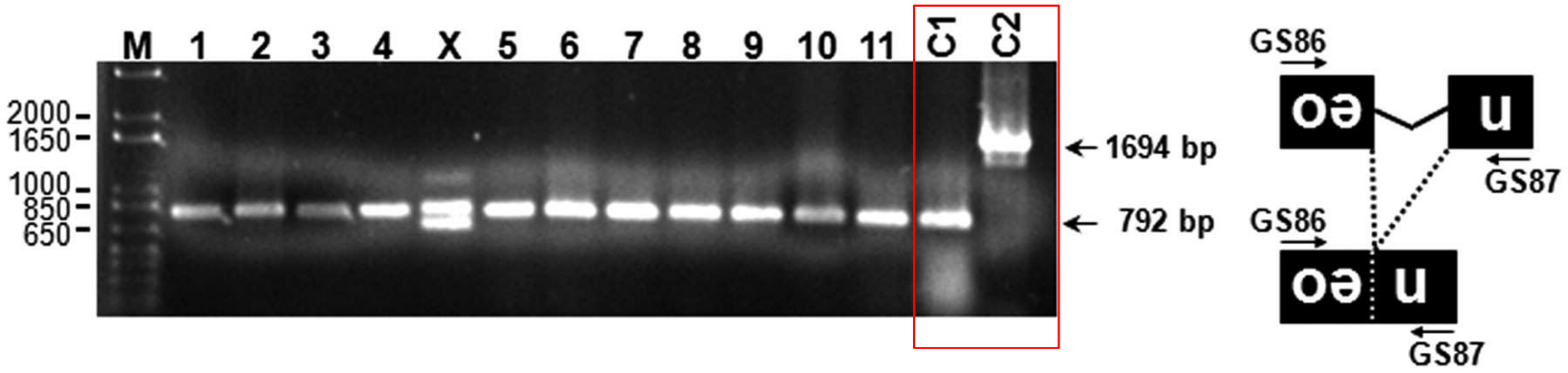


Raiz J et al. Nucl. Acids Res. 2011;nar.gkr863

Nucleic Acids Research

Diagnostic PCR to test for correct splicing of the intron from the mneoI indicator cassette.

PCR to test for correct splicing of the intron from the mneoI cassette



integration into the genome via authentic “retro”transposition

Raiz J et al. Nucl. Acids Res. 2011;nar.gkr863

Nucleic Acids Research

Intact L1 trans-mobilization.

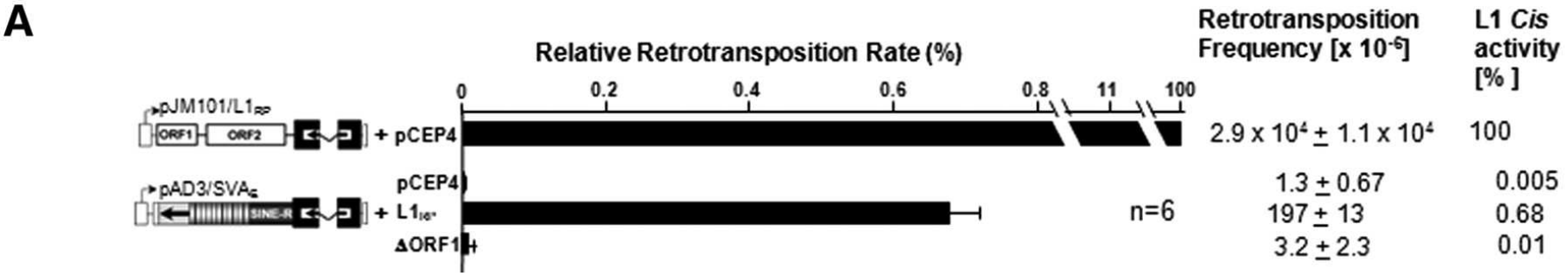
pJM101/L1RP =L1 cis activity Controllo positivo 100%

pCEP4 empty vector

Raiz J et al. Nucl. Acids Res. 2011;nar.gkr863

Nucleic Acids Research

L1 ORF1p is required for trans-mobilization of SVA reporter elements.



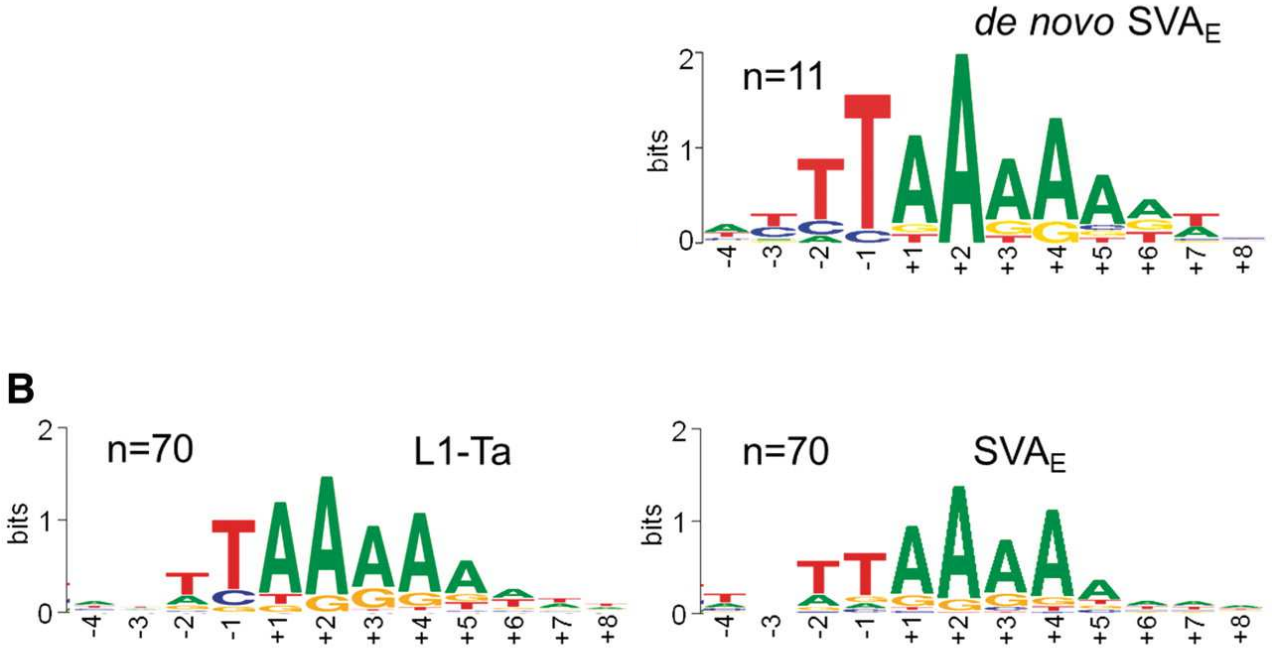
pJM101/L1RP =L1 cis activity Controllo positivo 100%

intact (L1RP) and mutant (ΔORF1) L1 protein donor plasmid
pCEP4 empty vector

Raiz J et al. Nucl. Acids Res. 2011;nar.gkr863

Nucleic Acids Research

The nucleotide profile of SVAE de novo insertion sites resembles the consensus target sequence of pre-existing human-non-LTR retrotransposons.



Raiz J et al. Nucl. Acids Res. 2011;nar.gkr863

Nucleic Acids Research