Example 2: effect of a UG repeat near the 3’ss on splicing efficiency

Variation in the number of GT T repeats near the 3’ss in CFTR intron 8 is associated to Cystic Fibrosis

GT 9-13 **T 3-9**

TGTGTGTGTGTGTGTGTGTGTGTGTGTTTTTTTTTAACAG
TGTGTGTGTGTGTGTGTGTGTGTGTTTTTTTTTAACAG
TGTGTGTGTGTGTGTGTGTGTGTTTTTTTTTAACAG
TGTGTGTGTGTGTGTGTGTGTGTGTGTTTTTTTAACAG
TGTGTGTGTGTGTGTGTGTGTGTGTTTTTTTAACAG
TGTGTGTGTGTGTGTGTGTGTGTTTTTTTAACAG
TGTGTGTGTGTGTGTGTGTGTGTGTGTTTTTAACAG
TGTGTGTGTGTGTGTGTGTGTGTGTTTTTAACAG
TGTGTGTGTGTGTGTGTGTGTGTTTTTAACAG

TGTGTGTGTGTGTGTGTGTGTGTGTGTTTAACAG

intron 8 *exon 9*

The non-evolutionary conserved TG tract in CFTR intron 8 is primate specific and deviates from canonical 3′ splice sites

```
                                                           intron 9          exon 9
human      gaaaatatctgacaaactcatctttt-atttttgatgtgtgtgtgtgtgtgtgtgtgtttttttaacagGGATTTGGGGAATTATTTGAGAAA..
macaca     gaaaatatctgacaaactcatctttt-atttttgatgtgtgtgtgtgtgtgtgt---tttatttaacagGGATTTGGGGAATTATTTGAGAAA..
chimp      gaaaatatctgacaaactcatctttt-atttttgatgtgtgtgtgtgtgttt------ttttttaacagGGATTTGGGGAATTATTTGAGAAA..
baboon     gaaaatatctgacaaactcatctttt-atttttgatatgtgtgtgtgtgtgt-----tttatttaacagGGATTTGGGGAATTATTTGAGAAA..
lemur      gaaaatatctga----tctatctttt-atttttgat----------------------tttttaaacagGGATTTGGGGAATTATTTCAGAAA..
rabbit     gaaaatatctgaaaa---ca-ctttt-cttttcaat------------------------tttttatagGGATTTGGGGAATTATTTGAGAAA..
armadillo  aaaagtatctgaaaagcttgccactt-atttttatct----------------------ttctttatagCGATTTGGAGAATTATTTGAGAAA..
cat        gaaaatacttgaaaaatctacctttttgttttctctctcta-----------------ttttttttcagGGATTTGGGGAATTACTTGAGAAA..
horse      gaaaatatctgaaaaacttacctttt-attttctatgttttt---------------ttttttttacagGGATTTGGGGAATTATTTGAGAAA..
cow        gaaaatatctgagaaacataccttcttatttactgtgtttttt--------------tcttttttacagGGATTTAGTAAATTATTTGAGAAA..
hedgehog   gataaaacataagctattaaaagaatctgaaaacctttttttt--------------tcttttttacagGGTTTGGGGGATTATTTGAGAAA..
pig        gaaatgatctgaaaaacataccttcttatttttatgtttttt---------------t-ttttcaacagGGATTTGGGAAATTATTTGAGAAA..
dog        gaaaatacctgaaaagcctacctttttattttctctgtatcta-----------------ttttttcagGGATTTGGGGAATTACTTGAGAAA..
mouse      ggaaaa-tctgaaa-----gcattattatcttcaatgt---------------------gtatgaatagGGATTTGGGGAATTACTGGAGAAA..
rat        ggaaat-tctgaaa-----acattat---ctttaatgt---------------------ttataaatagGGATTTCAGGAATTACTGGAGAAA..
Consensus 3′ss                                   yyyyyyyyyyyyyyyyyyyyyyyyncagG                  Y=T/C
```

Exon 9 CFTR in humans had been subject to a retrotransposition event followed by amplification of the integration site (Rozmahel et al. 1997)

This insertion and rearrangement events placed fortuitous new *cis*-acting elements, including the TG tract in the proximity of its splice junction, that induce aberrant exon skipping

Insertion of an UG repeat sequence near the 3' ss inhibits splicing in an heterologous *in vitro* system

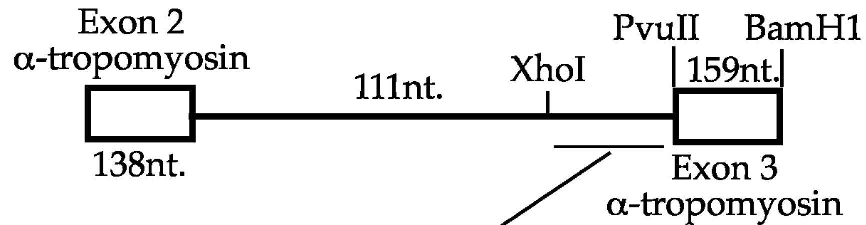
Exon 2
α-tropomyosin
111nt.
XhoI
PvuII
BamH1
159nt.
Exon3
tropomyosis
or
Exon 9
CFTR

PY7(wt)
3'ss exon
CTCGAGGCTTCATTTTTTCTTTTTCTTTTTCAG**CTGGAA**

PY7(UG12U5)
3'ss exon
CTCGAG<u>TGTGTGTGTGTGTGTGTGTGTG</u>TTTTTAACAG**CTGGAA**

PY7 (wt)
PY7 (UG12U5)
PY7 CFTR
PY7 CFTR (UG12U5)
(hours) 0 2 0 2 0 2 0 2
Tropo myosin Ex3
CFTR Ex9

In vitro splicing assay
Increasing the length of the UG repeat inhibits splicing



PY7(wt)

3'ss *exon*

CTCGAGGCTTCATTTTTTCTTTTTCTTTTTCAG**CTGGAA**

PY7(UG6U5)

3'ss *exon*

CTCGAGGCTTCATGTGTGTGTGTGTTTTTACAG**CTGGAA**

PY7(UG8U5)

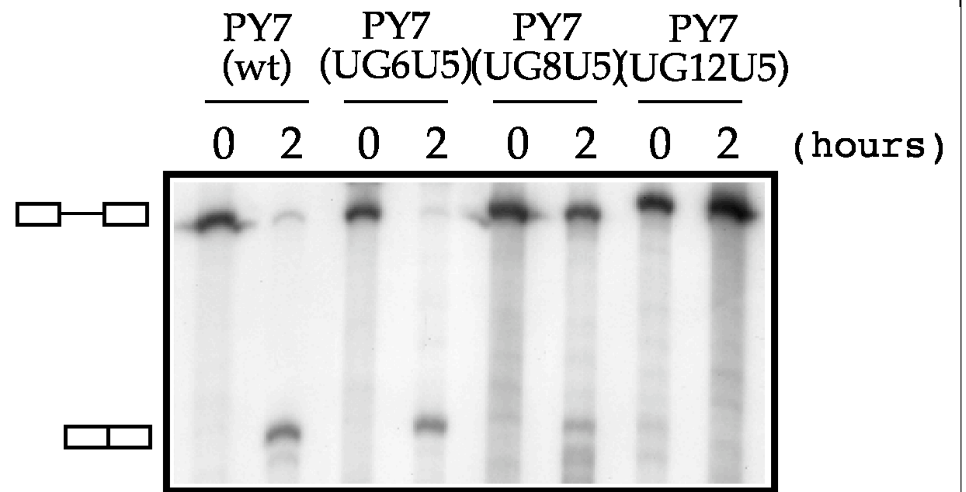
3'ss *exon*

CTCGAGTGTGTGTGTGTGTGTGTTTTTAACAG**CTGGAA**

PY7(UG12U5)

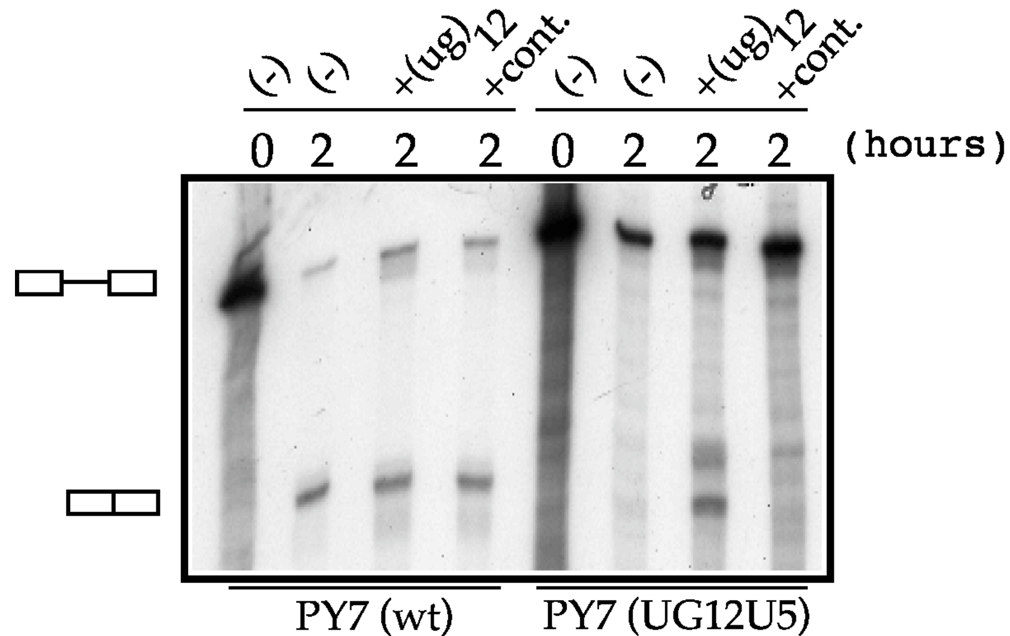
3'ss *exon*

CTCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTTTTTAACAG**CTGGAA**



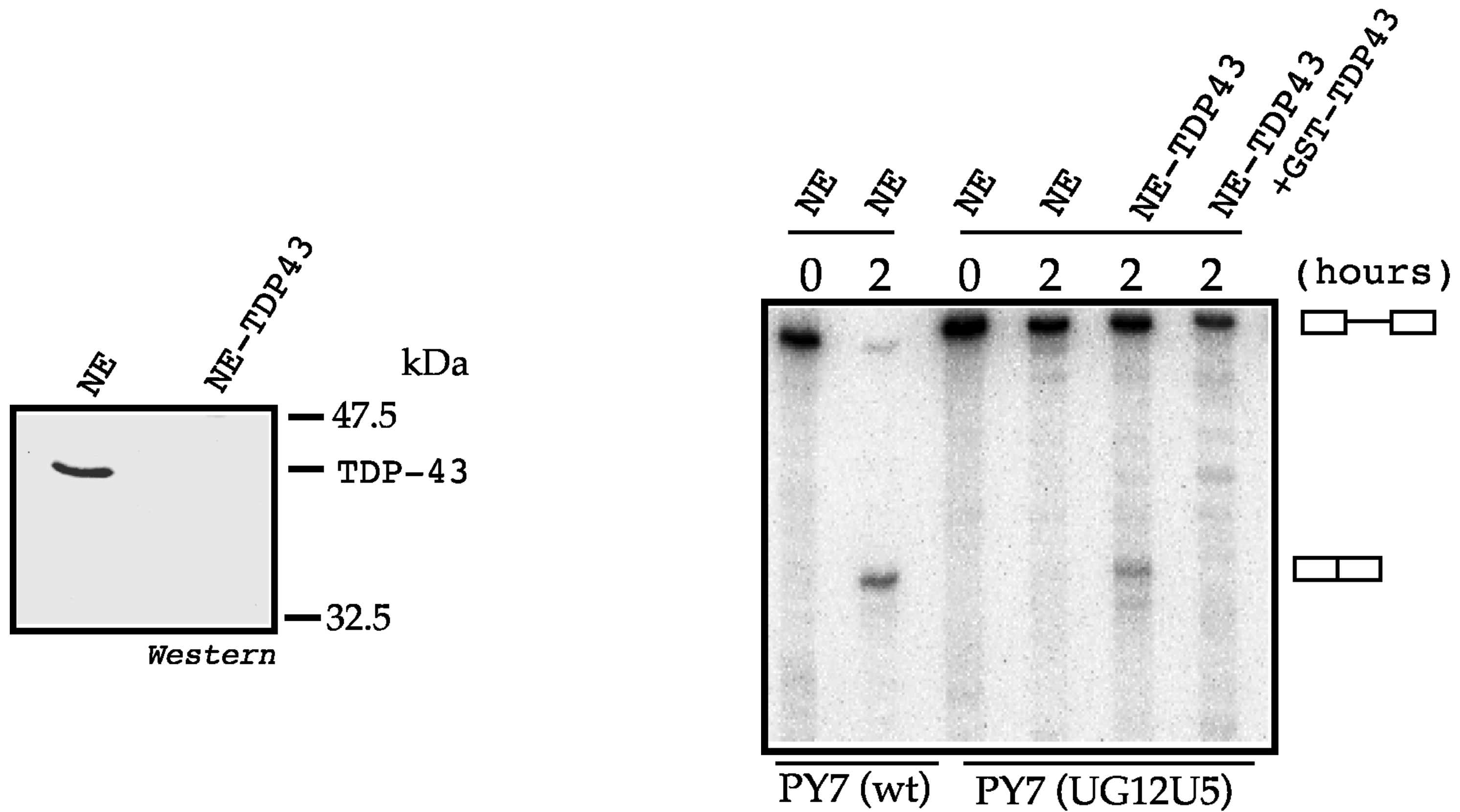
E.Buratti, A. Brindisi, F. Pagani and FE Baralle *Am J Human Genet 2004*

Addition of an UG competitor RNA restore normal splicing

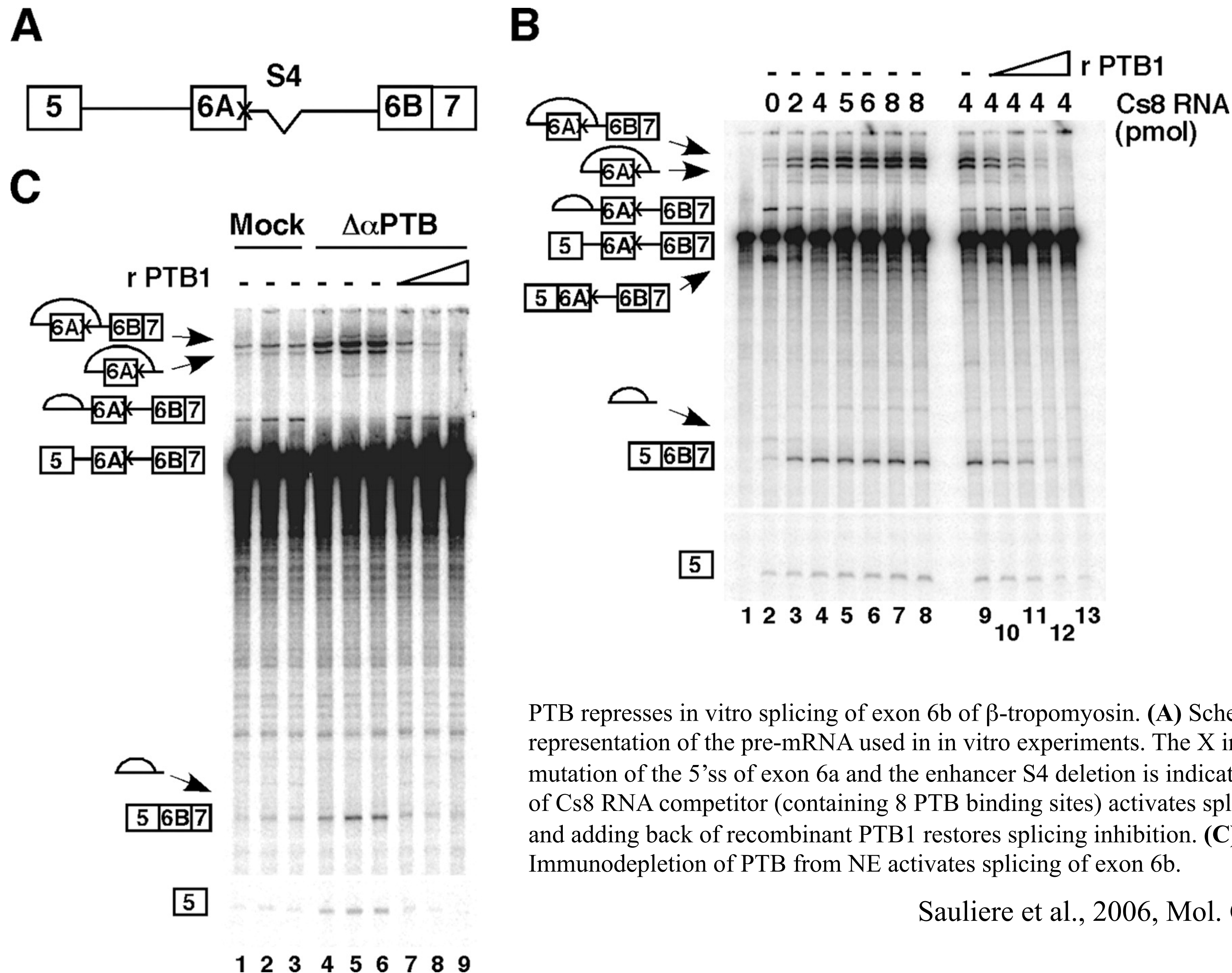


E.Buratti, A. Brindisi, F. Pagani and FE Baralle *Am J Human Genet 2004*

Depletion of TDP43 activates splicing



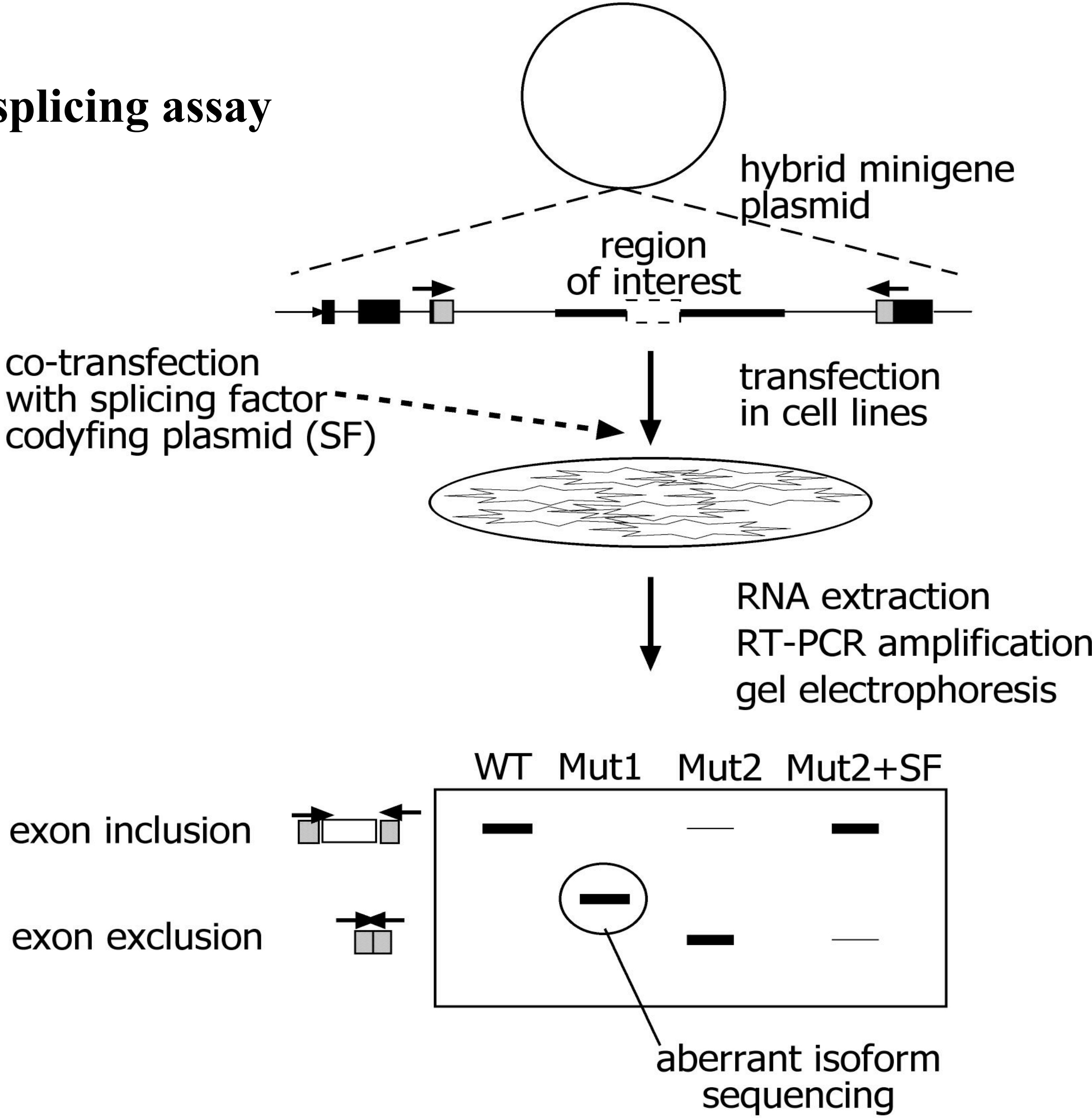
Example 3: sometimes you can see multiple splicing patterns



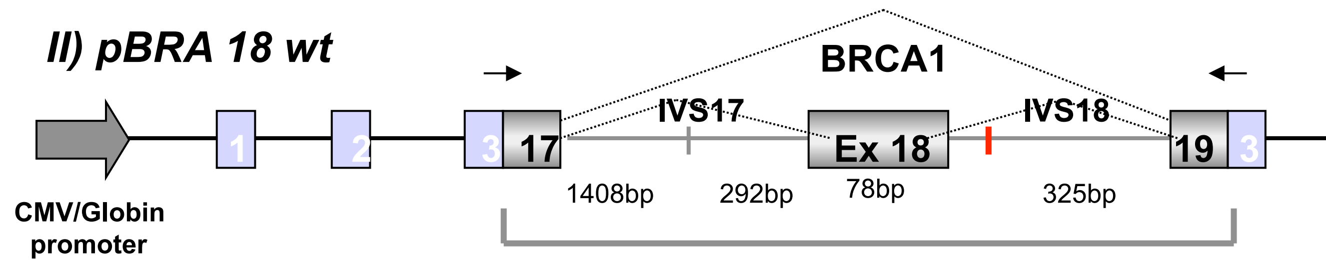
PTB represses in vitro splicing of exon 6b of β-tropomyosin. **(A)** Schematic representation of the pre-mRNA used in in vitro experiments. The X indicates the mutation of the 5'ss of exon 6a and the enhancer S4 deletion is indicated. **(B)** Addition of Cs8 RNA competitor (containing 8 PTB binding sites) activates splicing of exon 6b, and adding back of recombinant PTB1 restores splicing inhibition. **(C)** Immunodepletion of PTB from NE activates splicing of exon 6b.

Sauliere et al., 2006, Mol. Cell Biol.

The minigene splicing assay



Hybrid minigene for BRCA1 exon 18



Minigene splicing pattern

- Exon 18 Wt is normally included in BRCA1 mRNA
- The G6->T mutation causes inappropriate skipping of the entire constitutive exon *in vivo*.
- This skipping results in retention of the same reading frame and removal of 26 amino acids.

SELEX
ESE
SR proteins
ESE finder
RESCUE ESE

SELEX

Systematic evolution of ligands by exponential enrichment

Since its first description in 1990, the SELEX technology is widely applied as an in vitro selection method to evolve nucleic acid ligands, called aptamers, with new functionalities.

The term aptamer is derived from the Latin word "aptus"—which means fitting (Ellington and Szostak, 1990) and the Greek word "meros" meaning particle.

Aptamers are short single-stranded nucleic acid oligomers (ssDNA or RNA) with a specific and complex three-dimensional shape characterized by stems, loops, bulges, hairpins, pseudoknots, triplexes, or quadruplexes. Based on their three-dimensional structures, aptamers can well-fittingly bind to a wide variety of targets from single molecules to complex target mixtures or whole organisms (Fig. 1).

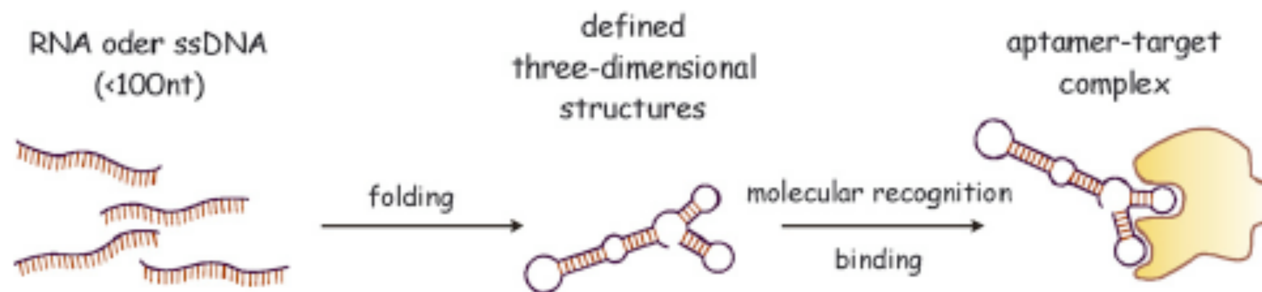


Fig. 1. Schematic representation of the functionality of aptamers.

Biomolecular Engineering 24 (2007) 381–403

Why to perform SELEX?

SELEX is a method to characterize the interaction of nucleic acids with proteins.

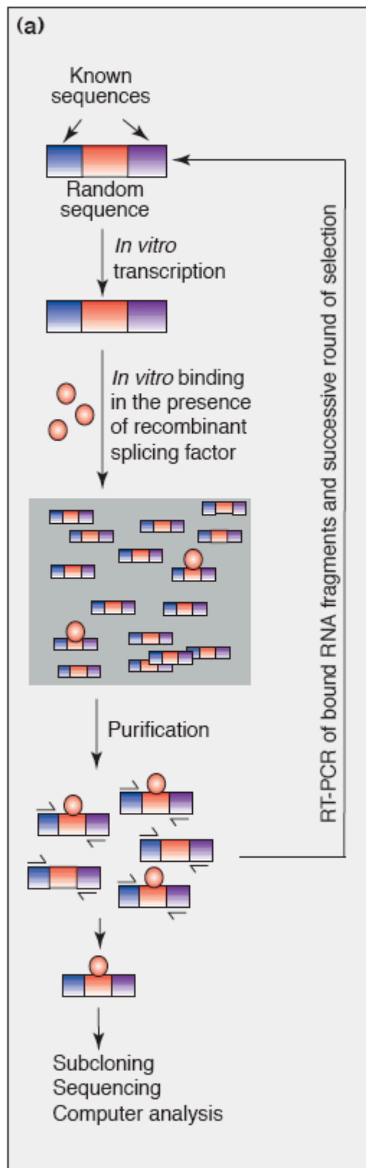
It can be used to generate new nucleic acids-binding species, called aptamers, that can be used in therapy.

SELEX is the process by which vast libraries of randomized oligonucleotide sequences are screened for their interactions with target molecules (proteins or small chemical entities). The SELEX process has been used to identify high-affinity ligands that bind to proteins, other nucleic acids, peptides, and small molecules (Gold et al., Curr Op Gen Dev, 1997).

Method principles

The method is based on the purification of large libraries of degenerate RNAs on affinity columns made with the protein of interest. The binding RNAs are amplified and several steps of selections are made, in order to enrich the population with true protein binding sites.

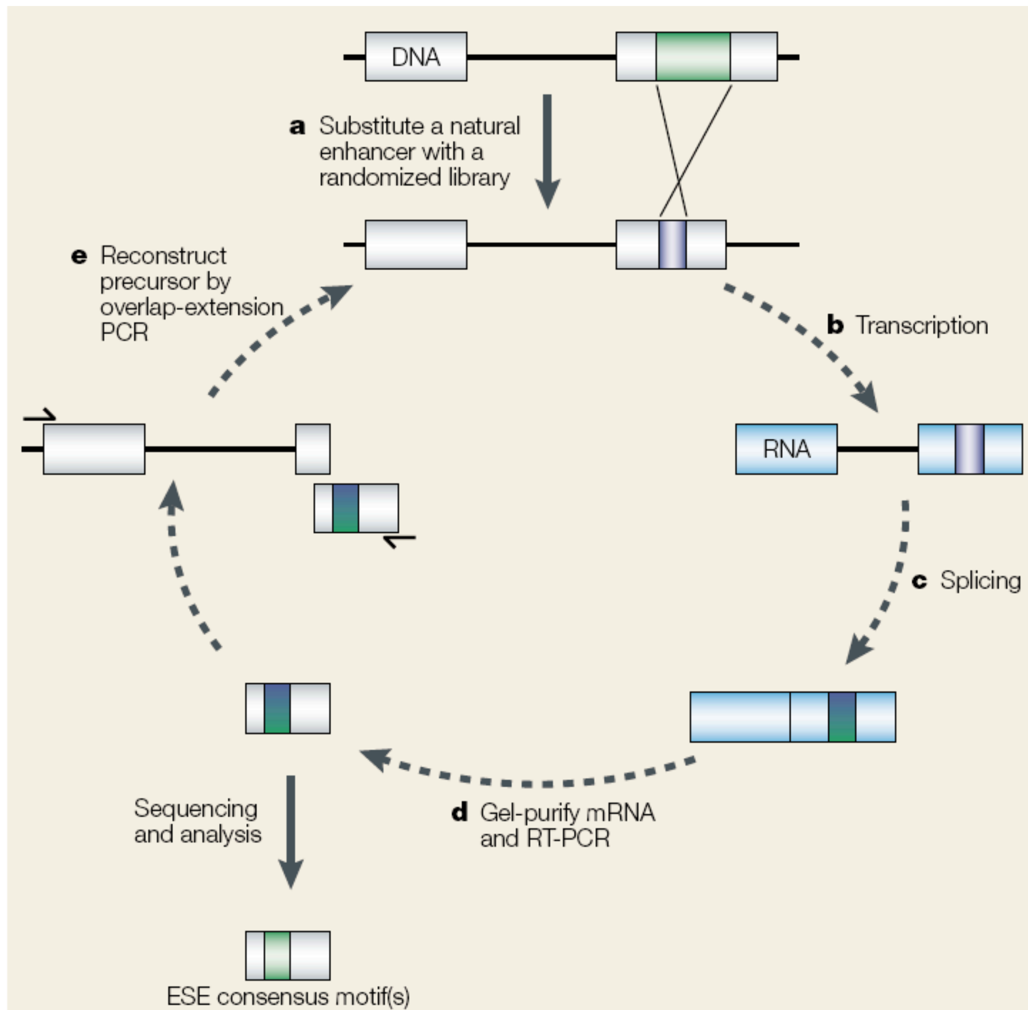
The classical SELEX approach: selection by protein binding



Building the RNA library: in the middle of two DNA fixed sequences, a stretch of 20-25 random nts is cloned and in vitro transcribed. It gives the possibility of having 10^{14}-10^{15} different RNA molecules (potential target sites). The library is put in contact with the recombinant protein of interest (in this case splicing factors), for example on an affinity column. The RNA sequences bound to the protein are collected, purified, amplified and they are put in contact again with the proteins. This cycle is repeated several times, until almost all the RNAs in the pool are bound to the proteins. The sequences are cloned, sequences and the consensus nucleotides are determined.

Buratti and Baralle, Trends Mol Med, 2005

Functional SELEX to identify Exonic Splicing Enhancer RNA motifs that bind to SR proteins

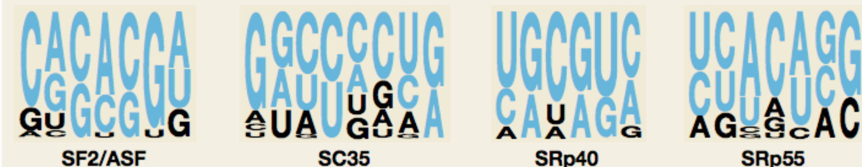


Cartegni et al., Nature reviews, 2002

To identify exonic splicing enhancer (ESE) motifs by functional *in vivo* or *in vitro* SELEX (systematic evolution of ligands by exponential enrichment[105]), a minigene is used that harbours ESE sequences that are required for the efficient splicing of its pre-mRNA. As shown in the accompanying figure, the natural enhancer (green box) is replaced by random sequences (blue) from an oligonucleotide library (**a**). The resulting pool of minigenes is then transfected into cultured cells, or is transcribed *in vitro*, to generate a pool of pre-mRNAs (**b**). Following *in vivo* or *in vitro* splicing (**c**), the pool of spliced mRNAs is gel purified and amplified by reverse-transcription (RT)-PCR (**d**). This pool of enhancer-enriched sequences is then used to reconstruct new minigene templates by OVERLAP-EXTENSION PCR[106] (**e**), to use in a new enrichment cycle. The iteration of this entire procedure yields a limited number of 'winners' — sequences that have good splicing enhancer activity[27,107].

To identify ESEs that are recognized by individual SR proteins, the splicing step was carried out in S100 extract complemented with one of four different SR proteins (SF2/ASF, SC35, SRp40 and SRp55)[13,30,31]. Transcripts were obtained from an immunoglobulin-μ (*IGHM*)-derived minigene, in which the natural enhancer was substituted with a pool of 20-nucleotide random sequences. After a few cycles of enrichment, spliced products were sequenced and aligned to derive a consensus motif. The frequencies of the individual nucleotides at each position were then used to calculate a score matrix, which can be used to predict the location of SR-protein-specific putative ESEs in exonic sequences[79,88] (TABLE 1).

The consensus motifs obtained with these four SR proteins are shown below; the height of each letter reflects the frequency of each nucleotide at a given position, after adjusting for background nucleotide composition[79]. At each position, the nucleotides are shown from top to bottom in order of decreasing frequency; blue letters indicate above-background frequencies.



SELEX in vivo
in cell lines

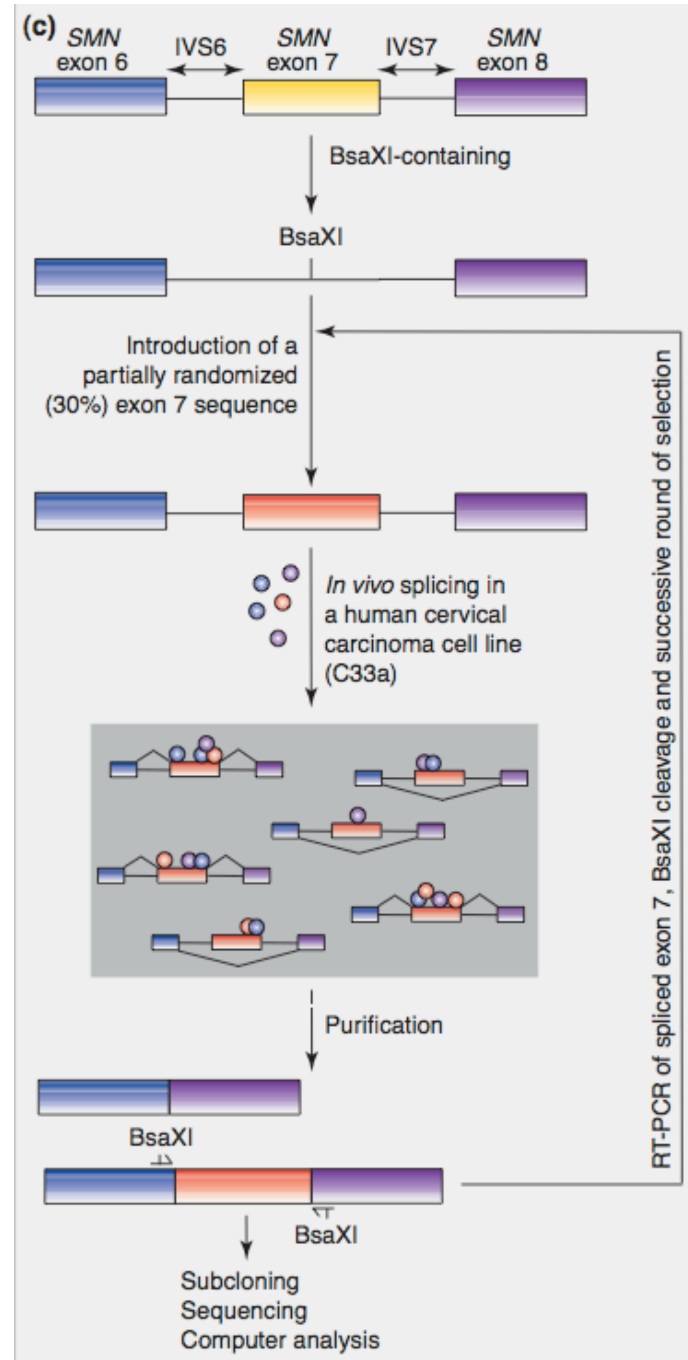


TABLE 1. RNA sequences identified as SR protein-binding sites by in vitro selection methods.

Protein	Binding site[a]	Method[b]	Reference
SF2/ASF	RGAAGAAC	SELEX	Tacke & Manley, 1995
	AGGACRRAGC		
	SRSASGA	Functional	Liu et al., 1998
SC35	AGSAGAGUA	SELEX	Tacke & Manley, 1995
	GUUCGAGUA		
	UGUUCSAGWU		
	GWUWCCUGCUA		
	GGGUAUGCUG	SELEX	Cavaloc et al., 1999
	GAGCAGUAGKS		
	AGGAGAU		
	GRYYCSYR	Functional	Liu et al., 2000
9G8	AGACKACGAY	SELEX	Cavaloc et al., 1999
	ACGAGAGAY		
SRp40	UGGGAGCRGUYRGCUCGY	SELEX	Tacke et al., 1997
	ACDGS	Functional	Liu et al., 1998
SRp55	USCGKM	SELEX	Liu et al., 1998
B52	GRUCAACCDNGGCGAACNG	SELEX	Shi et al., 1997
hTra2β	(GAA)n	SELEX	Tacke et al., 1998

[a]N: any nucleotide; R: purine; Y: pyrimidine; S: G or C; K: U or G; W: A or U; D: A, G, or U; M: A or C.

[b]"SELEX" indicates that the RNA sequence was determined to be a high affinity binding site for a purified SR protein; "functional" indicates that the RNA sequence was determined to function as an SR protein-specific splicing enhancer.

http://rulai.cshl.edu/cgi-bin/tools/ESE3/esefinder.cgi?process=home

Search | background | matrices & thresholds | input & output | caveats

Go to ESEfinder2.0

Matrix infomration

Please choose a matrix library: SRProteins

Choose the matrix and the threshold to be used:

	Matrices (select one or more)	Threshold
☑	SF2/ASF (SF2/ASF round 3 winner)	1.956
☑	SF2/ASF (IgM-BRCA1) (Smith06-HMG-matrix)	1.867
☑	SC35 (SC35 round 3 winner)	2.383
☑	SRp40 (SRp40 round 3 winner)	2.67
☑	SRp55 (SRp55 round 3 winner)	2.676
		Reset thresholds

Sequence infomration

Enter here your input data in FASTA or MULTI-FASTA format (<5000nt, accept both 'T' and 'U' as being equivalent)
(please read important information about search format description)

alternatively, upload a text file: Choose File no file selected

Protein	Matrix	Logo	Threshold

SF2/ASF — Threshold 1.956

	[1]	[2]	[3]	[4]	[5]	[6]	[7]
A	-1.14	0.62	-1.58	1.32	-1.58	-1.58	0.62
C	1.37	-1.1	0.73	0.33	0.94	-1.58	-1.58
G	-0.21	0.17	0.48	-1.58	0.33	0.99	-0.11
T	-1.58	-0.5	-1.58	-1.13	-1.58	-1.13	0.27

SF2/ASF (IgM-BRCA1) — Threshold 1.867

	[1]	[2]	[3]	[4]	[5]	[6]	[7]
A	-1.58	0.15	-0.97	0.74	-1.19	-0.75	0.43
C	1.55	-0.53	0.79	0.33	0.72	-0.62	-0.99
G	-1.35	0.44	0.41	-0.98	0.51	1.03	0.00
T	-1.55	-0.28	-1.28	-0.92	-1.09	-0.52	0.20

SC35 — Threshold 2.383

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
A	-0.88	0.09	-0.06	-1.58	0.09	-0.41	-0.06	0.23
C	-1.16	-1.58	0.95	1.11	0.56	0.86	0.32	-1.58
G	0.87	0.45	-1.36	-1.58	-0.33	-0.05	-1.36	0.68
T	-1.18	-0.2	0.38	0.88	-0.2	-0.86	0.96	-1.58

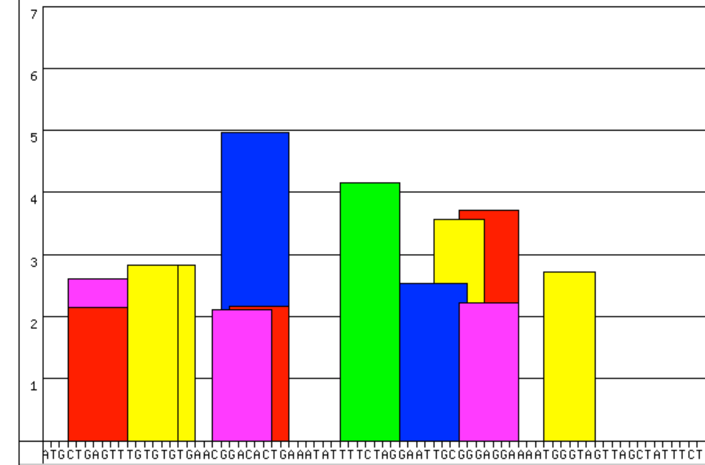
SRp40 — Threshold 2.670

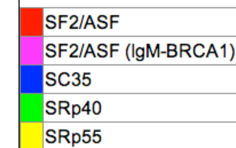
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
A	-0.13	-1.58	1.28	-0.33	0.97	-0.13	-1.58
C	0.56	0.68	-1.12	1.24	-0.77	0.13	-0.05
G	-1.58	-0.14	-1.33	-0.48	-1.58	0.44	0.8
T	0.92	0.37	0.23	-1.14	0.72	-1.58	-1.58

SRp55 — Threshold 2.676

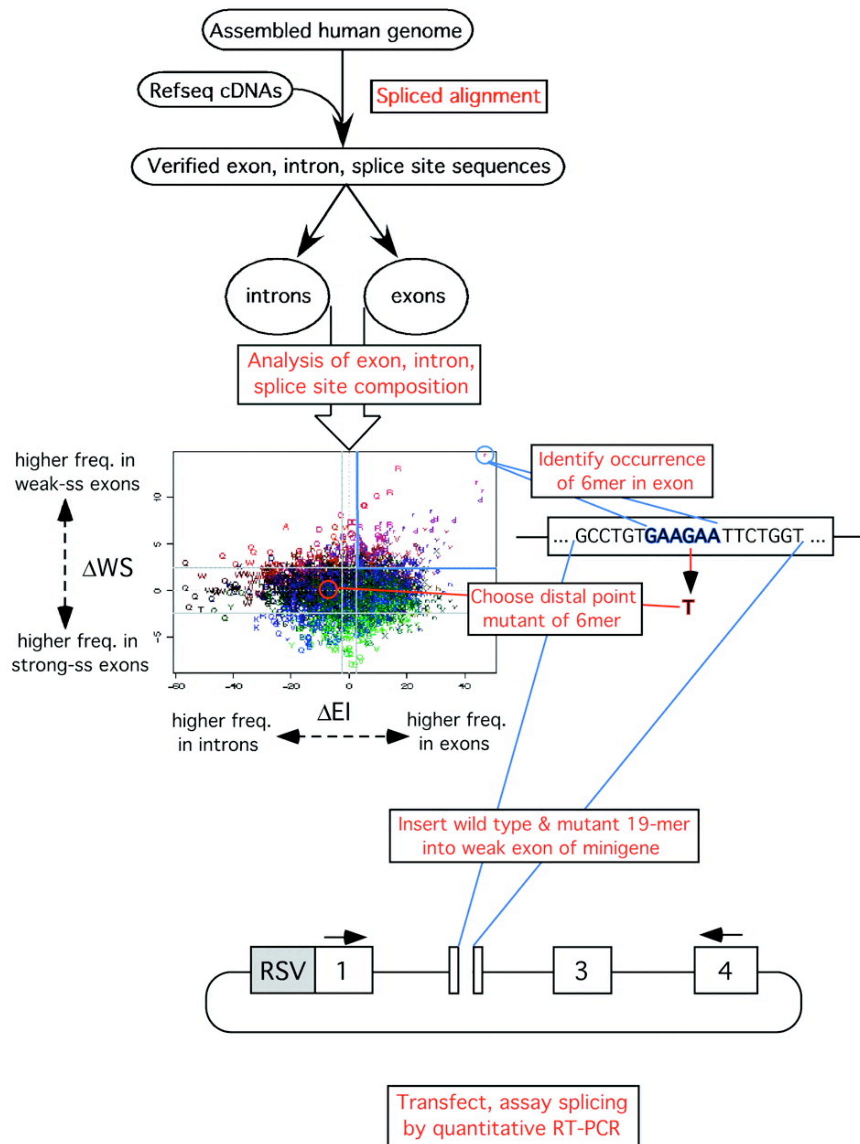
	[1]	[2]	[3]	[4]	[5]	[6]
A	-0.66	0.11	-0.66	0.11	-1.58	0.61
C	0.39	-1.58	1.48	-1.58	-1.58	0.98
G	-1.58	0.72	-1.58	0.72	0.21	-0.79
T	1.22	-1.58	-0.07	-1.58	1.02	-1.58

Protein		[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	Logo	Threshold
SF2/ASF	A	-1.14	0.62	-1.58	1.32	-1.58	-1.58	0.62			1.956
	C	1.37	-1.1	0.73	0.33	0.94	-1.58	-1.58			
	G	-0.21	0.17	0.48	-1.58	0.33	0.99	-0.11			
	T	-1.58	-0.5	-1.58	-1.13	-1.58	-1.13	0.27			
SF2/ASF (IgM-BRCA1)	A	-1.58	0.15	-0.97	0.74	-1.19	-0.75	0.43			1.867
	C	1.55	-0.53	0.79	0.33	0.72	-0.62	-0.99			
	G	-1.35	0.44	0.41	-0.98	0.51	1.03	0.00			
	T	-1.55	-0.28	-1.28	-0.92	-1.09	-0.52	0.20			
SC35	A	-0.88	0.09	-0.06	-1.58	0.09	-0.41	-0.06	0.23		2.383
	C	-1.16	-1.58	0.95	1.11	0.56	0.86	0.32	-1.58		
	G	0.87	0.45	-1.36	-1.58	-0.33	-0.05	-1.36	0.68		
	T	-1.18	-0.2	0.38	0.88	-0.2	-0.86	0.96	-1.58		
SRp40	A	-0.13	-1.58	1.28	-0.33	0.97	-0.13	-1.58			2.670
	C	0.56	0.68	-1.12	1.24	-0.77	0.13	-0.05			
	G	-1.58	-0.14	-1.33	-0.48	-1.58	0.44	0.8			
	T	0.92	0.37	0.23	-1.14	0.72	-1.58	-1.58			
SRp55	A	-0.66	0.11	-0.66	0.11	-1.58	0.61				2.676
	C	0.39	-1.58	1.48	-1.58	-1.58	0.98				
	G	-1.58	0.72	-1.58	0.72	0.21	-0.79				
	T	1.22	-1.58	-0.07	-1.58	1.02	-1.58				





A computational method for identifying sequences with ESE activity



Hexanucleotide sequences are identified as candidate ESEs on the basis that they have both significantly higher frequency of occurrence in exons than in introns and also significantly higher frequency in exons with weak (non-consensus) splice sites than in exons with strong (consensus) splice sites.

Some were experimentally validated in heterologoous splicing systems

The RESCUE-ESE approach identified 238 hexamers as candidate ESEs

Fairbrother WG, et al Predictive identification of exonic splicing enhancers in human genes. Science. 2002 ;297:1007-13

A computational method for identifying sequences with ESE activity

RESCUE-ESE Web Server

RESCUE-ESE output for 32097:6128:11:16:45

RESCUE-ESE v 1.0 run date: 11/13/2010 time: 11:16:45
SPECIES: human
sequence length 83
total matches 17
unique matches 16

```
                                                              TGTGGA
                                                          GAACTG
                                                         AGAACT
                                                        AAGAAC
                                 GAAAAA                GAAGAA
                                AGAAAA                AGAAGA
                               CAGAAA                GAGAAG               GAAAAA
                              TCAGAA                TGAGAA              CTGAAA
                             TTCAGA      CAGATC                        CCTGAA
ATTTTGTTCAAATGTTTGTTTTGAGGGGATTTCAGAAAAACTCAGATCTGAGAAGAACTGTGGAGCCTGAAAAAATGCCGCAG
.........|.........|.........|.........|.........|.........|.........|.........|...
        10        20        30        40        50        60        70        80
```

Fairbrother WG, et al Predictive identification of exonic splicing enhancers in human genes. Science. 2002 ;297:1007-13

PUM2, a novel murine puf protein, and its consensus RNA-binding site

White EK, Moore-Jarrett T, Ruley HE. RNA. 2001 Dec;7(12):1855-66.

Binding site for a "puf" protein, implicated in mRNA degradation

SELEX II

UUCCCGACU NNNNNNNNNNNNNNNNNNNNN GGAAGCUUC

C8-1 UGCUGUACAUAGUGCAUCC
C8-3 GUGUACAUAACGCGCGUGCC
C8-5 UGUACAUGUCUAACCCCGCCC
C8-6 UGUAAGUAGUCCCCCCGGCCC
C8-7 UGUAAAUACAAAGUGCGCCC
C8-9 AGAUUGUGUAGUUUAGUGCGU
C8-10 UGUAAAUAGCCAGGGUGCGCC
C8-11 UGUAGAUAGCACCUGACCCCG
C8-13 UGUAGAUAAUCGUUUGUGCCG
C8-14 UGUAAAUAGAAACCGGCCCCG
C8-18 UGUAUACUAGUGACCCCCUCG
C8-22 UGUAAGAUAUGUAUCGUCC
C8-23 UGUAGAUAAGUCCCGUCGCCC
C8-26 UGUAAAUUGGUGCCUCCCCG
C8-27 UGUAAAUAACUGUUCUCGCCC
C8-28 UGUAAAUAACUGGGCCCGUCU
C8-29 UGUAGAUAGCUCAGCCCUCG
C8-30 UGUAGAUAACUCAUGCGCCC
C8-34 UGUACAAAGAUAACCGUGCCG
C8-35 UGUAGUUUAGCGCGCUCCGU
C8-38 UGUACAUACAGAGGGCUCGCC
C8-39 UGUAAUAUGGGUGAUGUGCUG
C8-40 AGUGUAAGAUCAAGGCCUGU

C8-16 GUGGAUAUGUUCUCUACUGU

Consensus: UGUANAUARNNNNBBBBSCCS lass 17

Nucleic acid degenerate base abbreviations

Code	Integer	Base Name	Meaning	Complement
A	1	**A**denine	A	T
C	2	**C**ytosine	C	G
G	3	**G**uanine	G	C
T	4	**T**hymine	T	A
U	4	**U**racil	U	A
R	5	(Pu**R**ine)	G|A	Y
Y	6	(P**Y**rimidine)	T|C	R
K	7	(**K**eto)	G|T	M
M	8	(A**M**ino)	A|C	K
S	9	**S**trong interaction (3 H bond)	G|C	S
W	10	**W**eak interaction (2 H bonds)	A|T	W
B	11	Not-A (B follows A)	G|T|C	V
D	12	Not-C (D follows C)	G|A|T	H
H	13	Not-G (H follows G)	A|T|C	D
V	14	Not-T (or U) (V follows U)	G|A|C	B
N, X	15	A**N**y nucleotide	G|A|T|C	N
-	16	Gap of indeterminate length	Gap	-

A novel SELEX approach to label in vivo RNAs

RNA Mimics of Green Fluorescent Protein

Jeremy S. Paige,[1] Karen Y. Wu,[1] Samie R. Jaffrey[1,2]*

Green fluorescent protein (GFP) and its derivatives have transformed the use and analysis of proteins for diverse applications. Like proteins, RNA has complex roles in cellular function and is increasingly used for various applications, but a comparable approach for fluorescently tagging RNA is lacking. Here, we describe the generation of RNA aptamers that bind fluorophores resembling the fluorophore in GFP. These RNA-fluorophore complexes create a palette that spans the visible spectrum. An RNA-fluorophore complex, termed Spinach, resembles enhanced GFP and emits a green fluorescence comparable in brightness with fluorescent proteins. Spinach is markedly resistant to photobleaching, and Spinach fusion RNAs can be imaged in living cells. These RNA mimics of GFP provide an approach for genetic encoding of fluorescent RNAs.

The fluorophore in green fluorescent protein (GFP) is formed from three residues in the nascent protein, Ser^{65}-Tyr^{66}-Gly^{67}, that undergo an autocatalytic intramolecular cyclization. The resulting fluorophore, 4-hydroxybenzlidene imidazolinone (HBI) (Fig. 1A), is encased within the protein, enabling its fluorescence (*1*). Chemically synthesized HBI is nonfluorescent (*2*), as is denatured GFP (*3*). However, upon refolding, the fluorescence of GFP is recovered (*3*). The folded GFP protein forms specific contacts with the fluorophore that prevent intramolecular motions, making fluorescence the major pathway available to dissipate the energy of the excited state fluorophore (*4*).

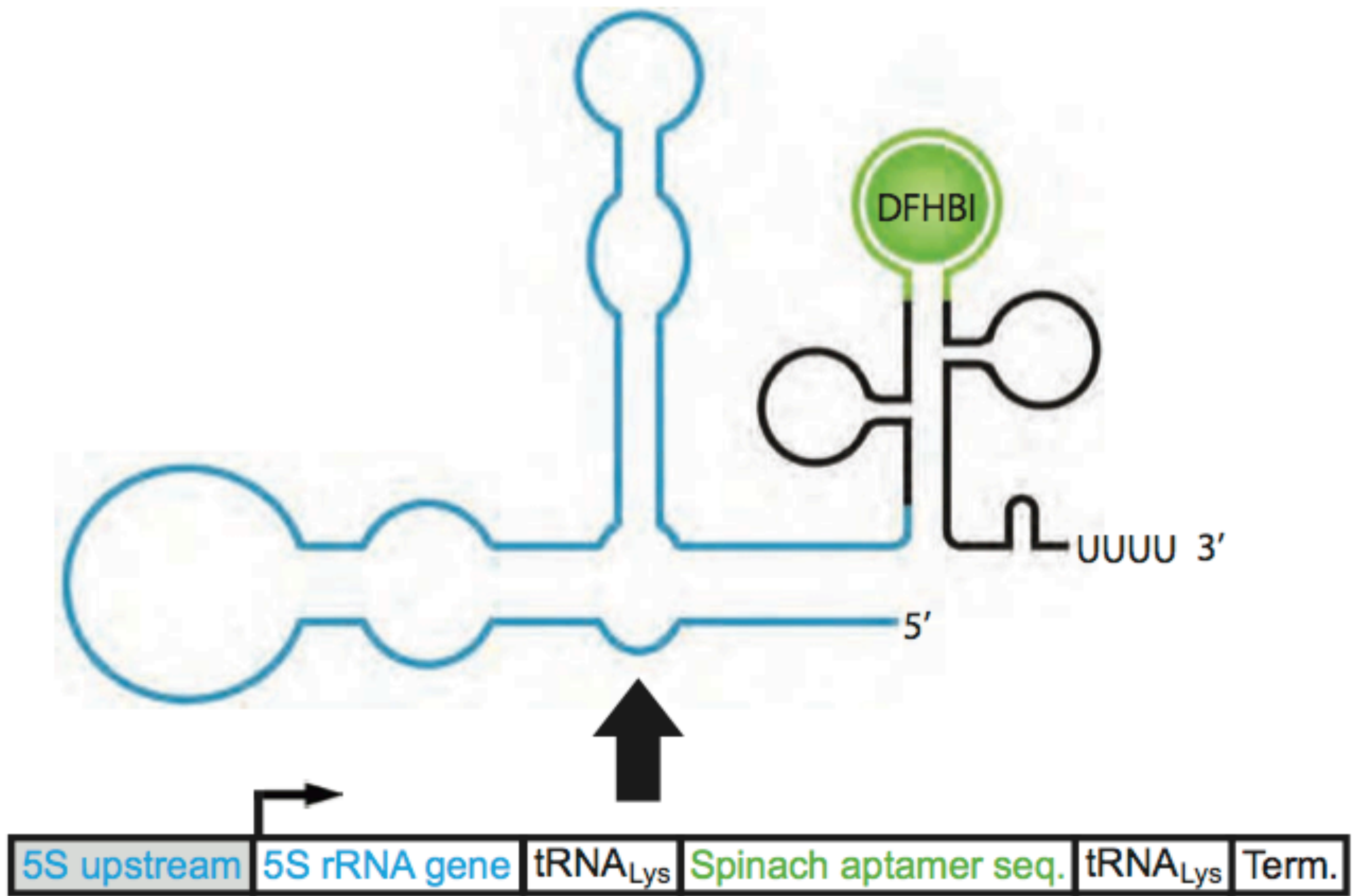
The ability to confer GFP-like functionality to RNA would facilitate studies of RNA biology and advance RNA-based applications. An RNA sequence with GFP-like properties should exhibit

[1]Department of Pharmacology, Weill Medical College, Cornell University, New York, NY 10065, USA. [2]Tri-Institutional Program in Chemical Biology, Weill Medical College, Cornell University, New York, NY 10065, USA.

*To whom correspondence should be addressed. E-mail: srj2003@med.cornell.edu

24-2

(Spinach)


DFHBI
UUUU 3′
5′
5S upstream
5S rRNA gene
tRNA Lys
Spinach aptamer seq.
tRNA Lys
Term.

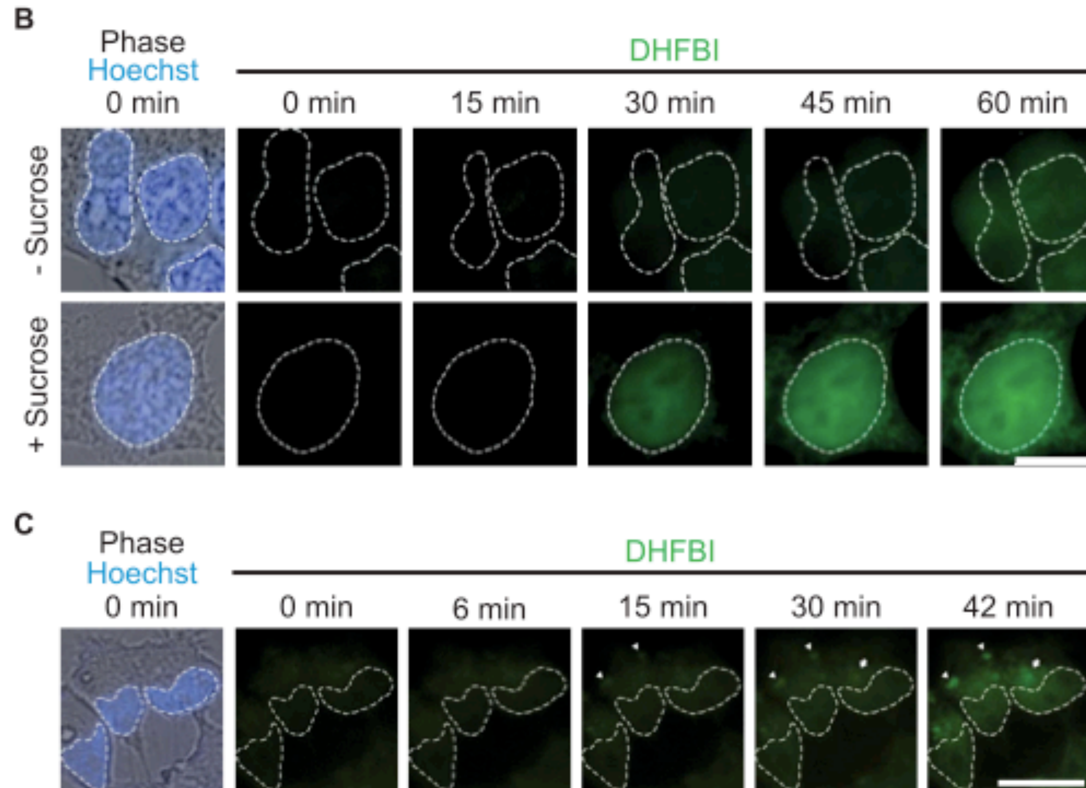
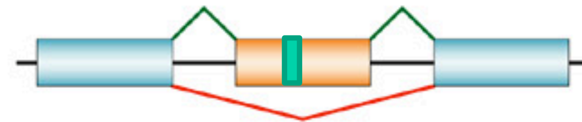


Fig. 4. Live-cell imaging of Spinach fusion RNAs. (**A**) Live-cell imaging of Spinach-tagged 5*S* RNA. Fluorescence and phase images of HEK293 T cells expressing 5*S* tagged with either Spinach or Lambda, a control RNA. Fluorescence is detected in 5*S*-Spinach–expressing cells in the presence of 20 μM DFHBI, with granule formation present in cells treated with 600 mM sucrose for 30 min (↑Suc). White dashed lines indicate nuclear borders assessed by means of Hoescht 33342 staining. (**B**) 5*S*-Spinach RNA induction in response to stress. 5*S*-Spinach–expressing HEK293 T cells were pretreated with 30 nM ML-60128 for 16 hours and then treated with vehicle or 600 mM sucrose for 60 min. Treatment of cells with sucrose resulted in a rapid induction of 5*S*-Spinach RNA and an increase in total 5*S*-Spinach levels compared with control cells. (**C**) 5*S*-Spinach RNA localization into granules. 5*S*-Spinach–expressing HEK293 T cells were stimulated with 600 mM sucrose in order to monitor the rate of formation of 5*S*-Spinach–containing granules. Arrowheads indicate granules that formed earliest, and arrows indicate granules that developed later during the time course of treatment. Scale bar, 10 μm.

EXONIC e INTRONIC SPLICING SILENCERS

Sono meno ben caratterizzati rispetto agli ESE (circa un terzo di sequenze casuali clonate all'interno dell'esone centrale di un minigene costituito da 3 esoni hanno un effetto inibitorio sullo *splicing*)



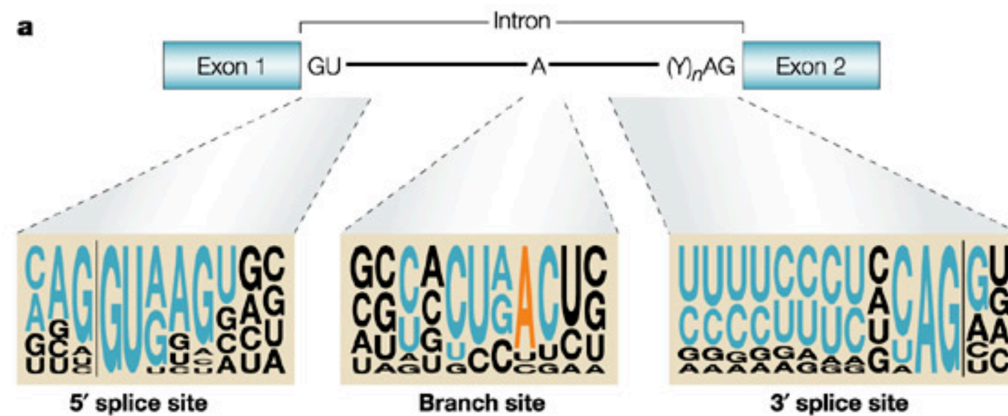
Sembra che interagiscano con elementi regolativi negativi dello *splicing*, che in molti casi appartengono alla famiglia delle **ribo-nucleo-proteine eterogenee nucleari (hnRNP)**, una classe di proteine leganti l'RNA che si associano ai trascritti nascenti.

Le hnRNP, come le proteine SR, hanno:
-un dominio di legame all'RNA
-un dominio di interazione proteina-proteina

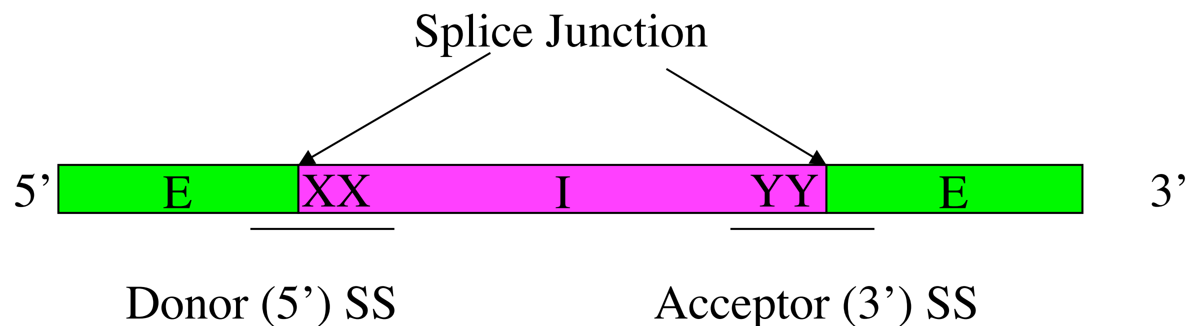
Heterogen**n**ous **r**ibo**n**ucleoprotein **p**articles (hnRNP) proteins

- In nucleus nascent RNA transcripts are associated with abundant set of proteins
- hnRNPs prevent formation of secondary structures within pre-mRNAs
- hnRNP proteins are multidomain with one or more RNA binding domains and at least one domain for interaction with other proteins
- some hnRNPs contribute to pre-mRNA recognition by RNA processing enzymes
- The two most common RNA binding domains are RNA recognition motifs (RRMs) and RGG box (five Arg-Gly-Gly repeats interspersed with aromatic residues)

Spliceosoma alternativo AT AC

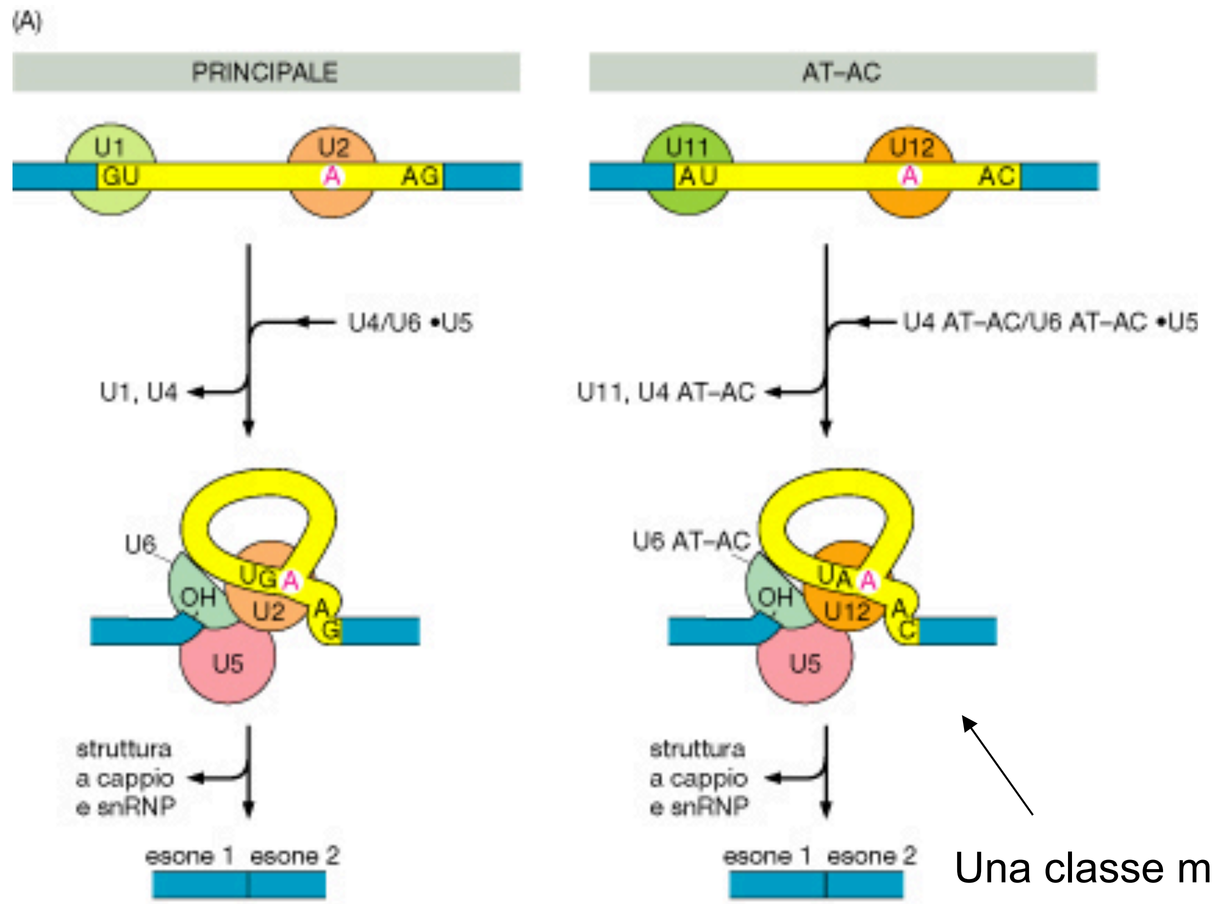


Splice Site Conservation



Class	XX	YY
U2_GT_AG (13289)	GT	AG
U2_GC_AG (1085)	GC	AG
U12_GT_AG (688)	GT	AG
U12_AT_AC (187)	AT	AC

Uno spliceosoma alternativo



Una classe minoritaria di introni con **sequenze consensus AT-AC** viene processata da uno spliceosoma “alternativo” con snRNP leggermente diverse nelle loro sequenze di interazione con il pre-mRNA.

Si tratta di una classe di introni rappresentata nell’uomo da un migliaio di introni

Lo splicing autocatalitico

Lo *splicing* autocatalitico

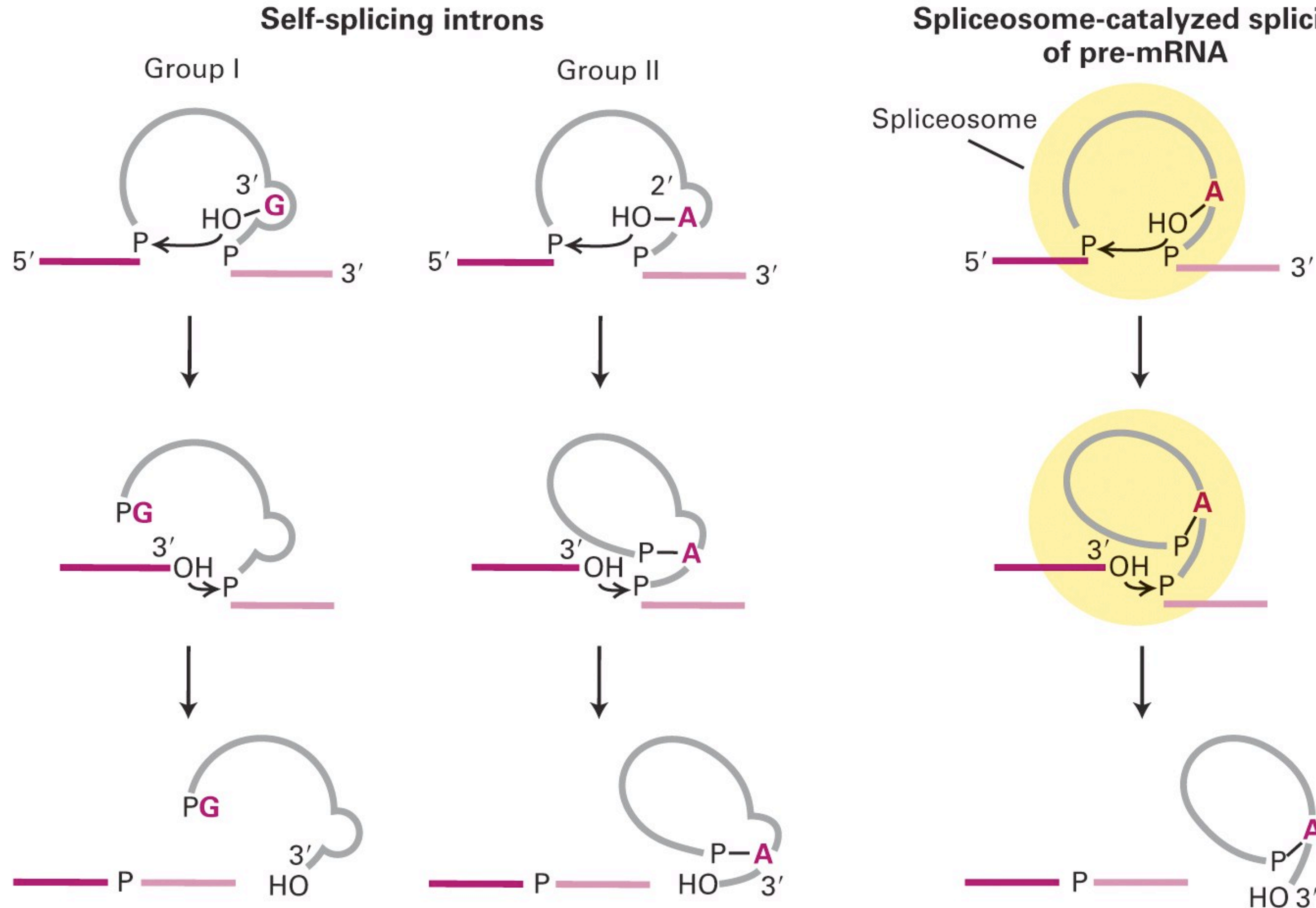
Oltre allo *splicing* dei pre-mRNA esistono almeno altri due tipi di *splicing*, definiti autocatalitici, essi coinvolgono:

Introni *self-splicing* di gruppo I scoperti per la prima volta studiando lo splicing dell'**rRNA 26S del protozoo cicliato Tetrahymena.**
Sono piuttosto rari:
- rRNA nucleari di alcuni eucarioti
- geni degli organelli
-qualche gene procariotico
Lo *splicing* avviene con un percorso diverso rispetto a quello dei pre-mRNA nucleari.

Introni *self-splicing* di gruppo II
Sono piuttosto rari:
- geni degli organelli
- qualche gene procariotico
Lo splicing si caratterizza per la presenza di un intermedio a lariat come per gli mRNA

Gli introni *self-splicing* non sono dei veri e proprio enzimi: mediano solo un ciclo di maturazione dell'RNA (l'RNA è anche il substrato della reazione stessa)

Splicing mechanisms in group I and group II self-splicing introns and spliceosome-catalyzed splicing of pre-mRNA; discovery of self-splicing of ribosomal RNA in *Tetrahymena*



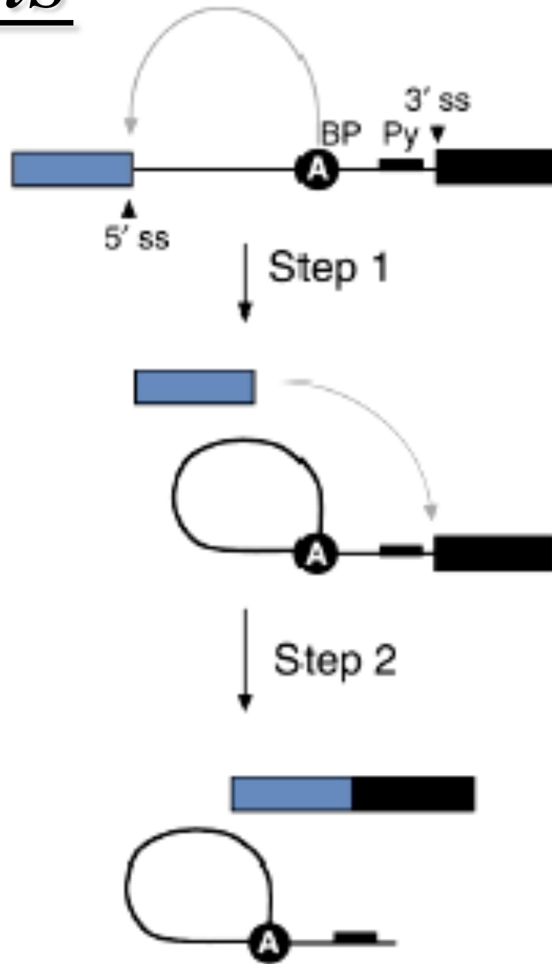
Occurs in pre-rRNAs from *Tetrahymena* and mitochondria & chloroplast origins. The 1st cleavage is carried out by an external cofactor guanosine (G, 3'-OH). The intron is released in a linear form.

Group II self-splicing introns are found in mitochondria and chloroplast pre-mRNAs. The 1st cleavage is carried out by the 2'-OH of A within the intron. The intron is released in the form of a lariat.
Common to all: two transesterification reactions are involved.

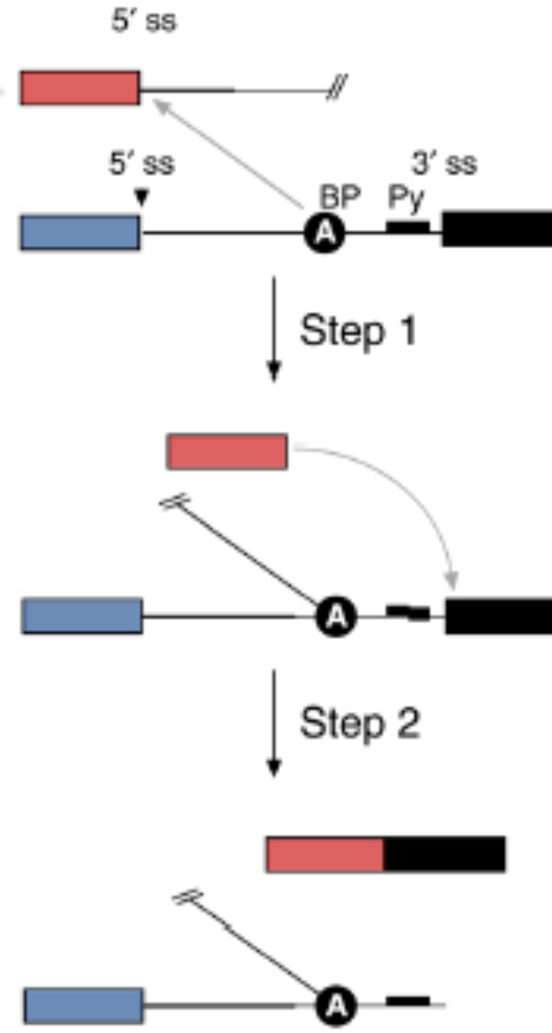
Trans splicing

Trans-Splicing

cis



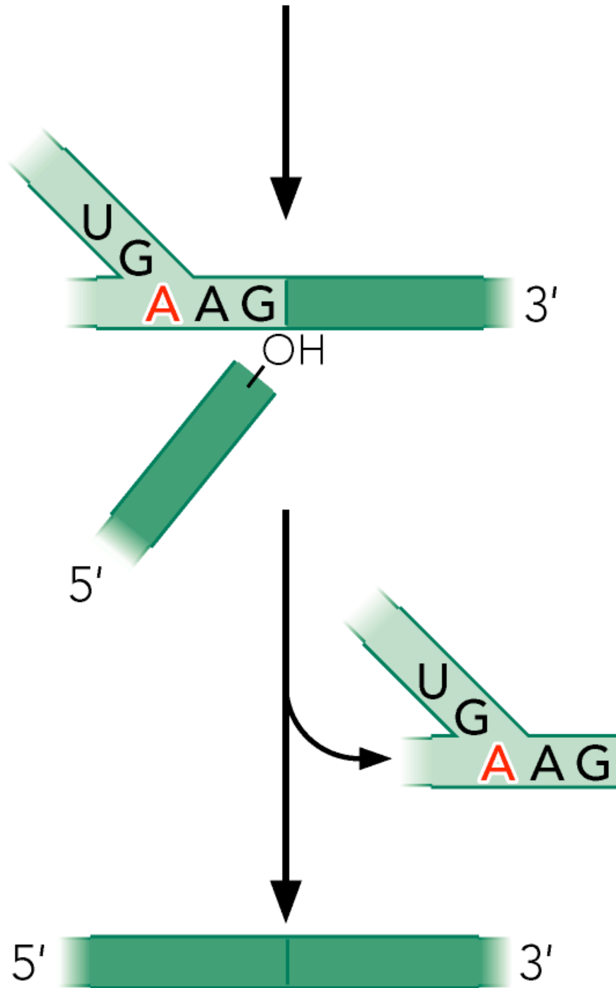
trans



Trans-Splicing

- Intermolecular splicing of pre-mRNAs
- First discovered in African **trypanosomes**, a disease(African Sleeping Sickness)-causing parasitic protozoan.
- The mRNAs had 35 nt not encoded in the main gene – called the spliced leader sequence.
- Spliced leader (SL) is encoded separately, and there about 200 copies in the genome .
- SL <u>primary</u> transcript contains ~100 nt that resemble the 5’ end of a mRNA intron.

Trans-splicing



La chimica della reazione è la stessa di quella dello *splicing* GT-AG. L'unica differenza è che l'introne rimosso in questo caso ha una struttura a "Y" anziché a laccio

(A)

PRINCIPALE

U1
GU
U2
A
AG
U4/U6 •U5
U1, U4
U6
OH
UG
A
U2
A
G
U5
struttura
a cappio
e snRNP
esone 1 esone 2

AT–AC

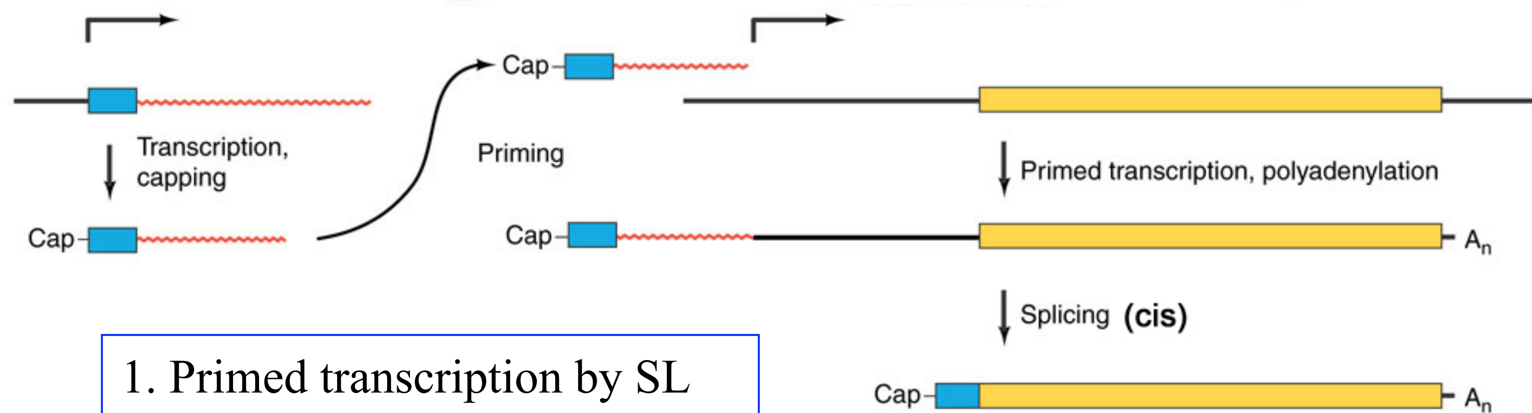
U11
AU
U12
A
AC
U4 AT–AC/U6 AT–AC •U5
U11, U4 AT–AC
U6 AT–AC
OH
UA
A
U12
A
C
U5
struttura
a cappio
e snRNP
esone 1 esone 2

TRANS

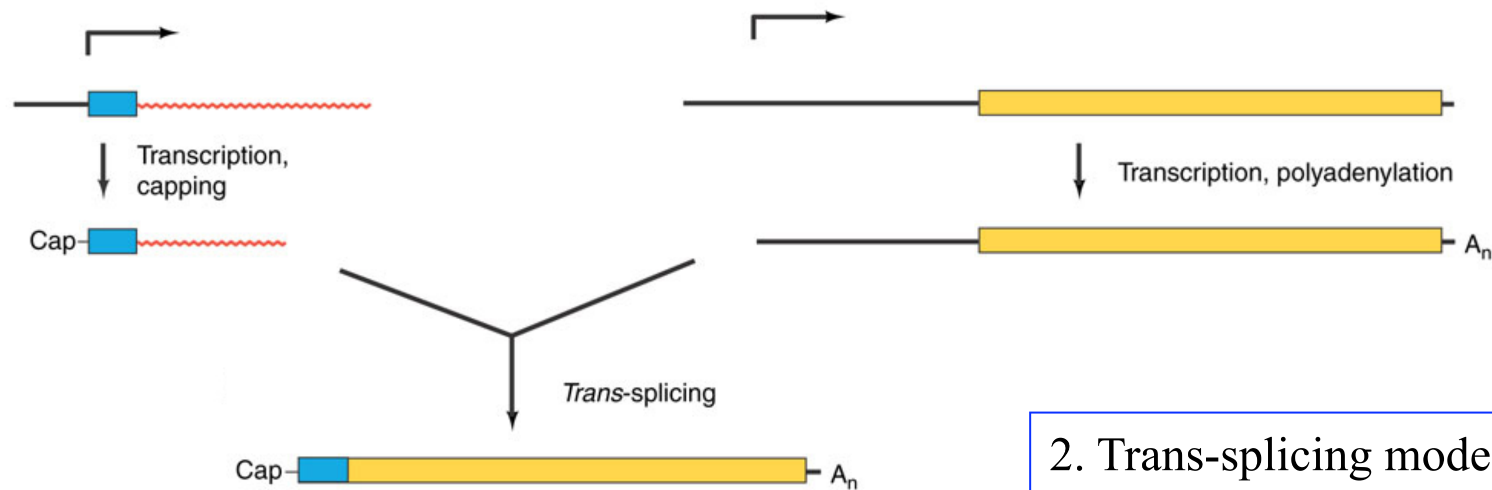
SL
U4
U5
GU
U6
SL RNP
U2
A
AG
U4
U6
SL
OH
UG
A
U2
A
G
U5
struttura
ramificata
e snRNP
esone 1 esone 2
SL

(B)

2 possible models to explain the joining of the SL to the coding region of a mRNA

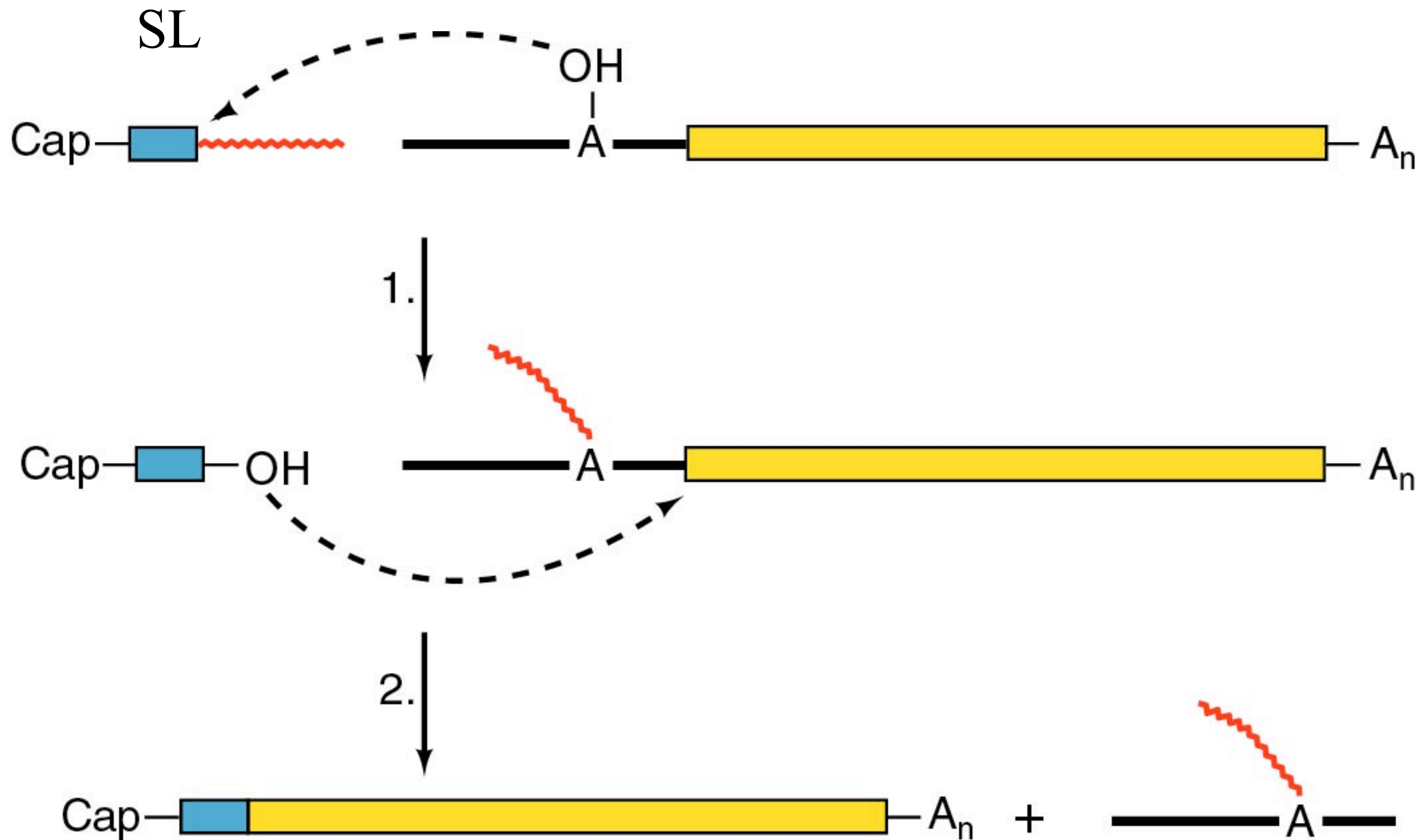


1. Primed transcription by SL



2. Trans-splicing model

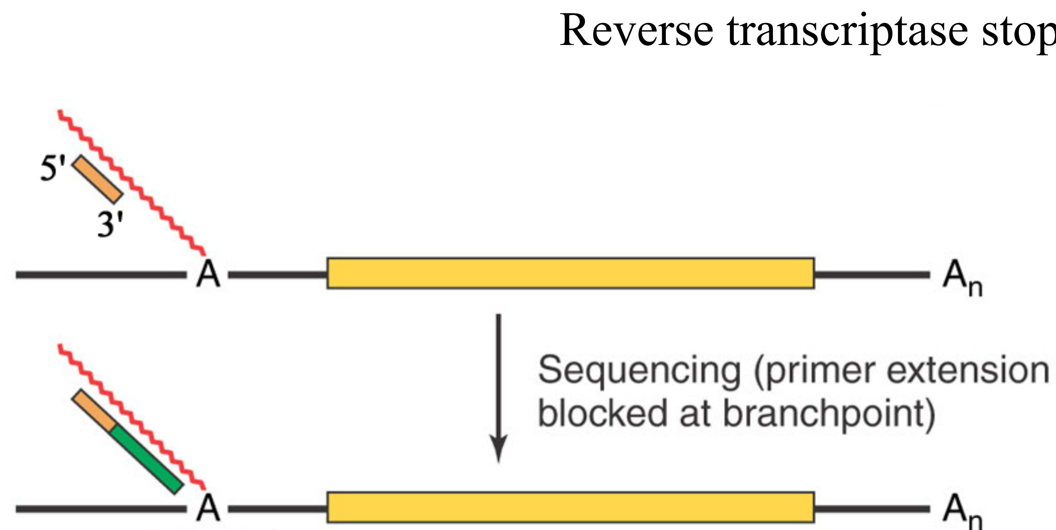
Trans-splicing in Trypanosomes



Trans-splicing should yield some unique "Y –shaped" intron-exon intermediates containing the SL half-intron.

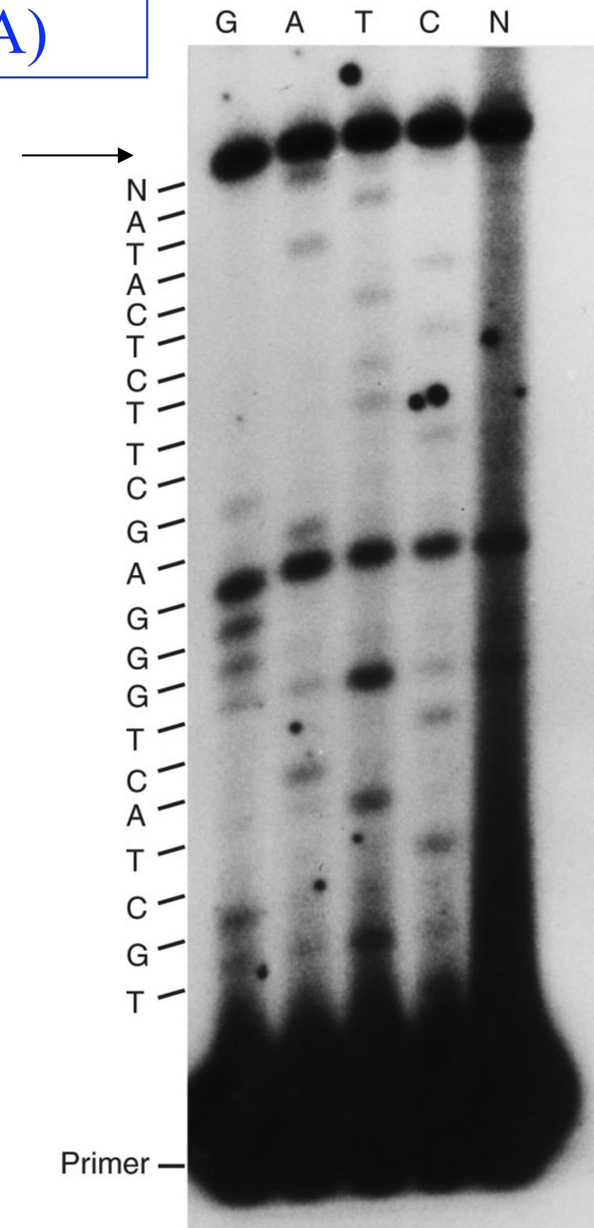
SL half-intron is attached to polyA+ RNA (mRNA)

Reverse transcriptase stopped



Primer anneals to SL half intron, which is extended with reverse transcriptase (RT) in the presence of a limiting ddNTP.

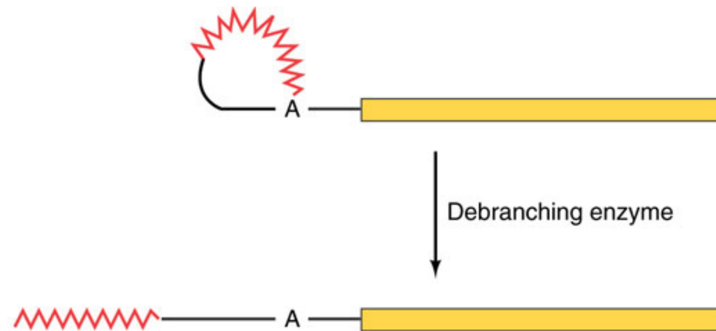
PolyA+ RNA (mRNA) was used as the RNA template. This would detect the intron-exon splicing intermediate, and the nt that becomes part of the branch (the RT falls off).

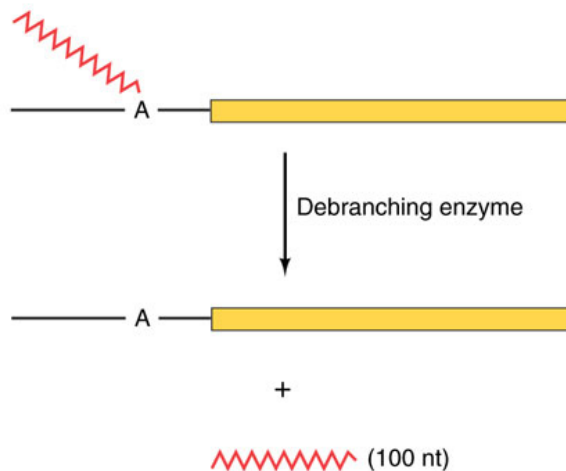


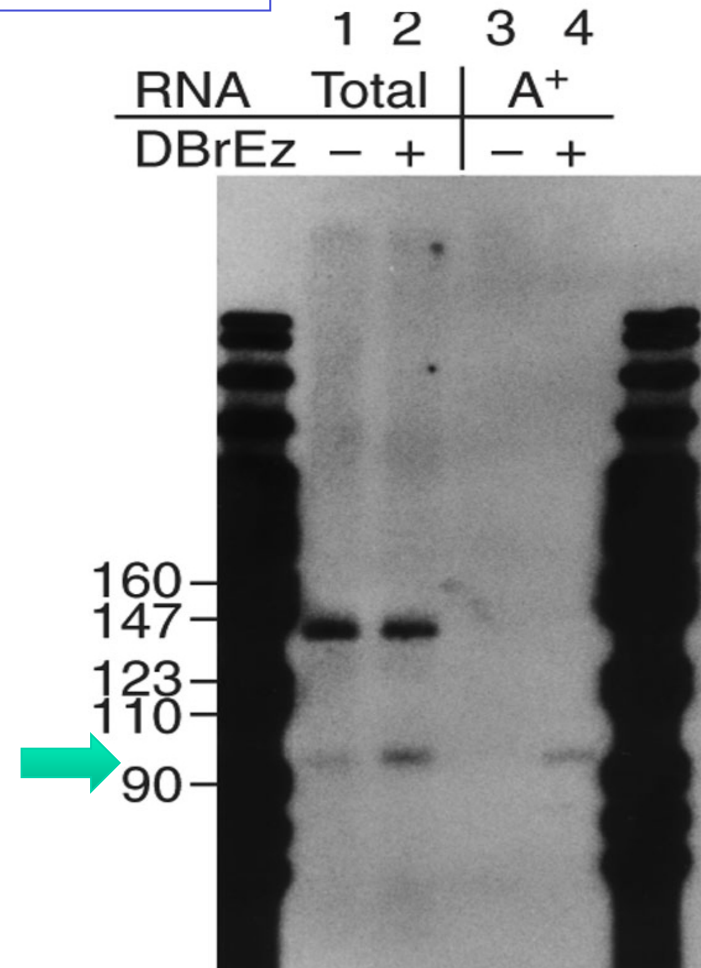
Murphy et al., Identification of a novel Y branch structure as an intermediate in tyrpanosome mRNA processing: Evidence for Trans splicing. *Cell* 47 (21 Nov 1986) p. 520, f. 4. Reprinted by permission of Elsevier Science.

Fig. 16.14, 2ed

Release of the SL half-intron from larger RNAs by a debranching enzyme.







Murphy et al., Identification of a novel Y branch structure as an intermediate in tyrpanosome mRNA processing: Evidence for Trans splicing. *Cell* 47 (21 Nov 1986) p. 521, f. 5.
Reprinted by permission of Elsevier Science

^{32}P-labeled total or polyA RNA was treated with debranching enzyme and then electrophoresed. The 147-nt RNA is probably 5S rRNA. The RNA of ~95-100 nt is the SL RNA.

This result is consistent with a *trans*-splicing model rather than a *cis*-splicing mechanism.

(A)

PRINCIPALE

U1 GU U2 A AG

U4/U6 •U5

U1, U4

U6 UG A OH U2 A G U5

struttura a cappio e snRNP

esone 1 esone 2

AT–AC

U11 AU U12 A AC

U4 AT–AC/U6 AT–AC •U5

U11, U4 AT–AC

U6 AT–AC UA A OH U12 A C U5

struttura a cappio e snRNP

esone 1 esone 2

TRANS

SL U4 GU U6 U5 U2 A AG

SL RNP

U4

U6 UG A OH SL U2 A G U5

struttura ramificata e snRNP

esone 1 esone 2

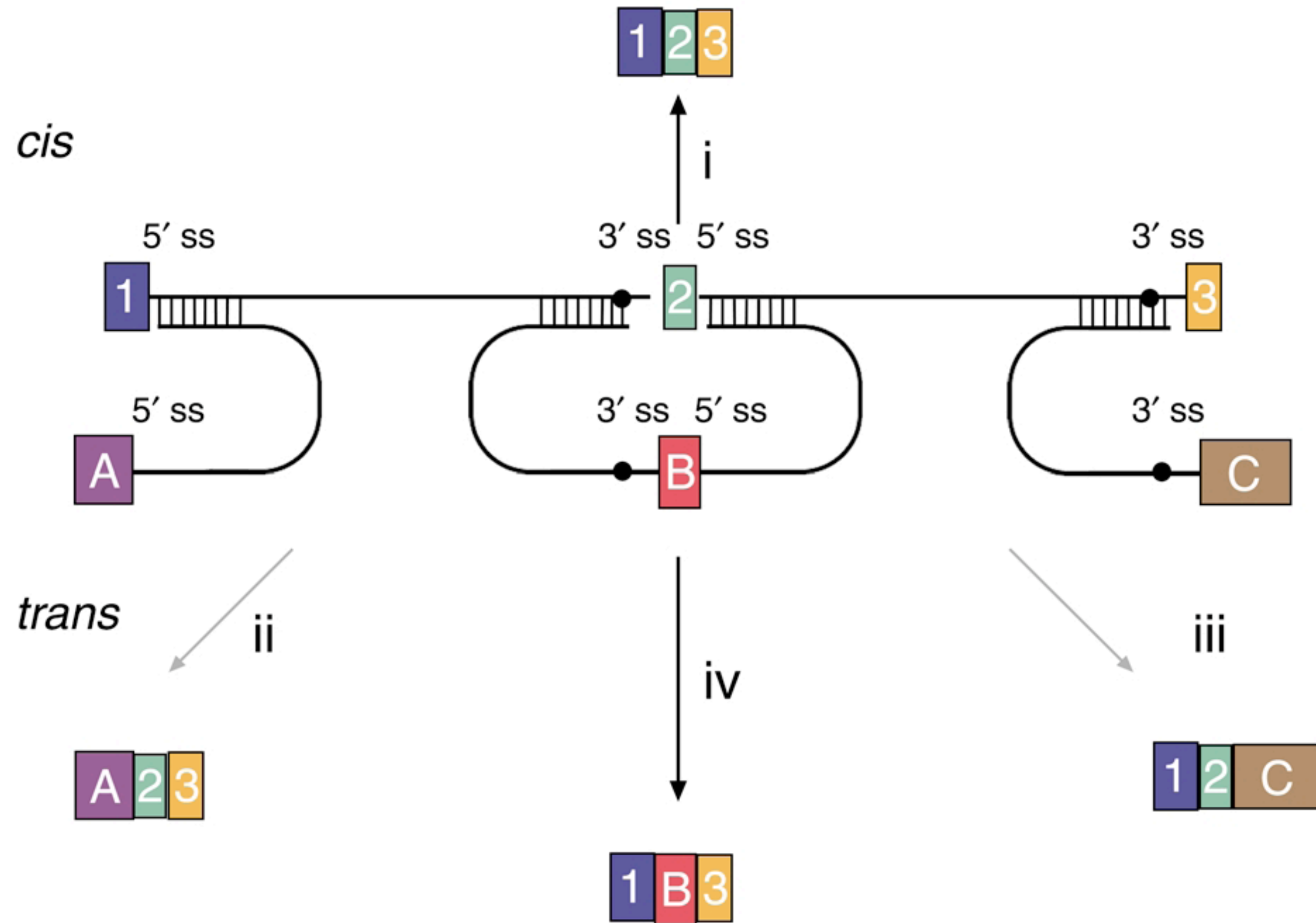
SL

(B)

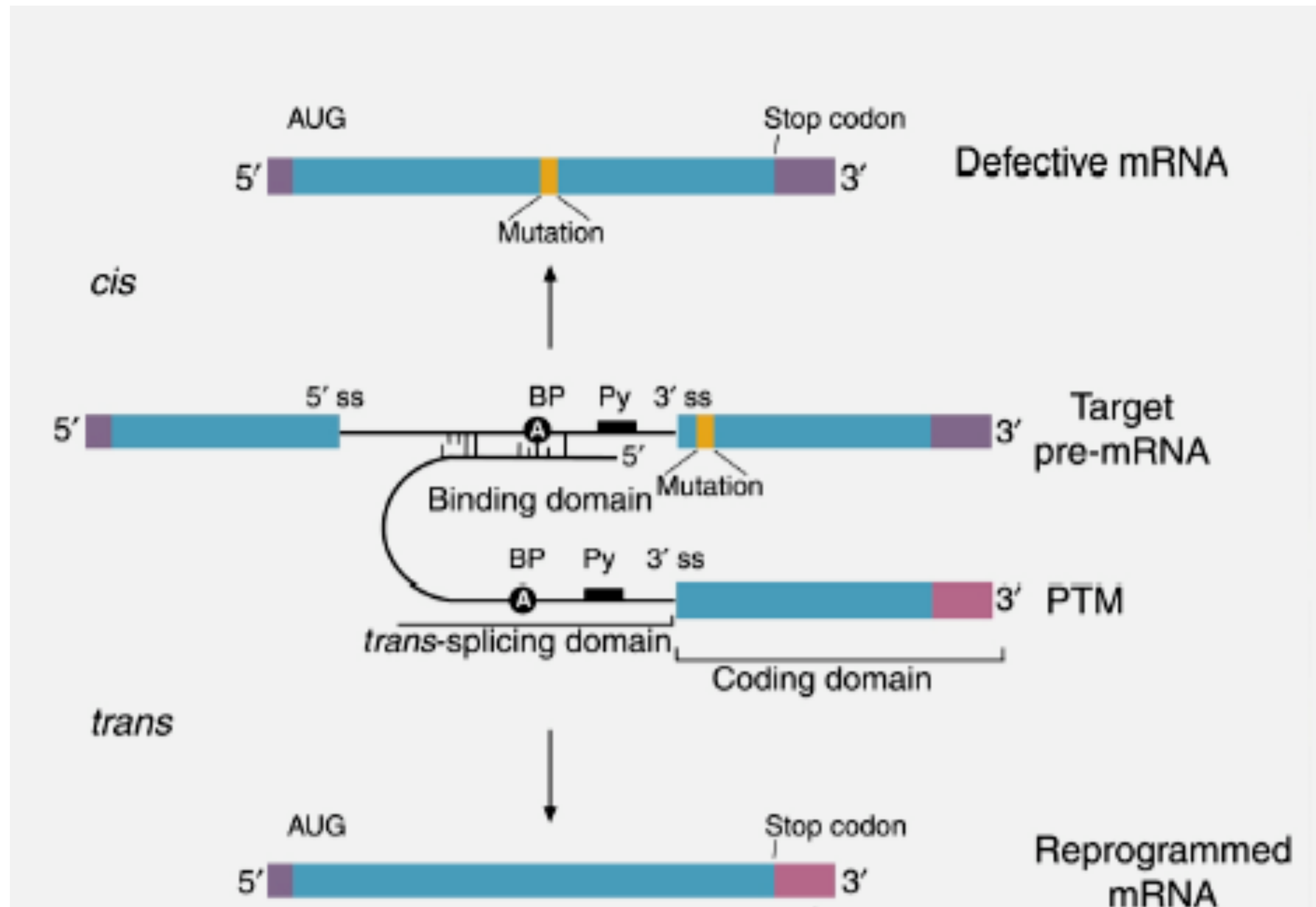
Trans-splicing can be explored as a therapeutic strategy to correct a splicing defect

Factor VIII, hemophilia A

SMaRT: spliceosome mediated RNA *trans*-splicing

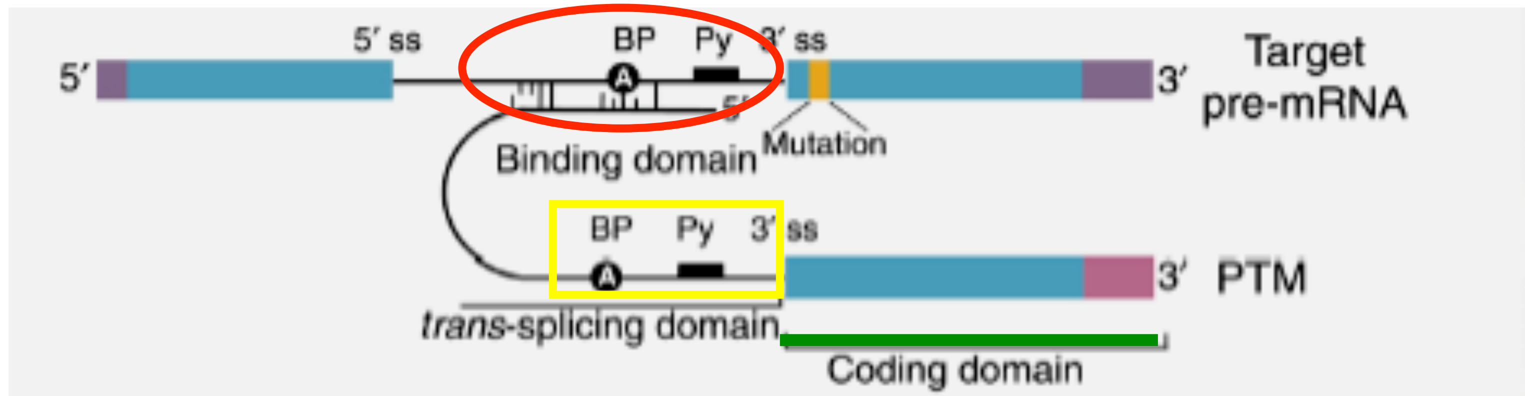


SMaRT: spliceosome mediated RNA *trans*-splicing



PMT: pre-*trans*-splicing molecule

PTMs structure



Prototype PTMs comprise three domains:

Binding domain

complementary to the target intron

Splicing domain

containing necessary splicing elements

Donor; Branch point; Polypirimidin tract; Acceptor

TSD

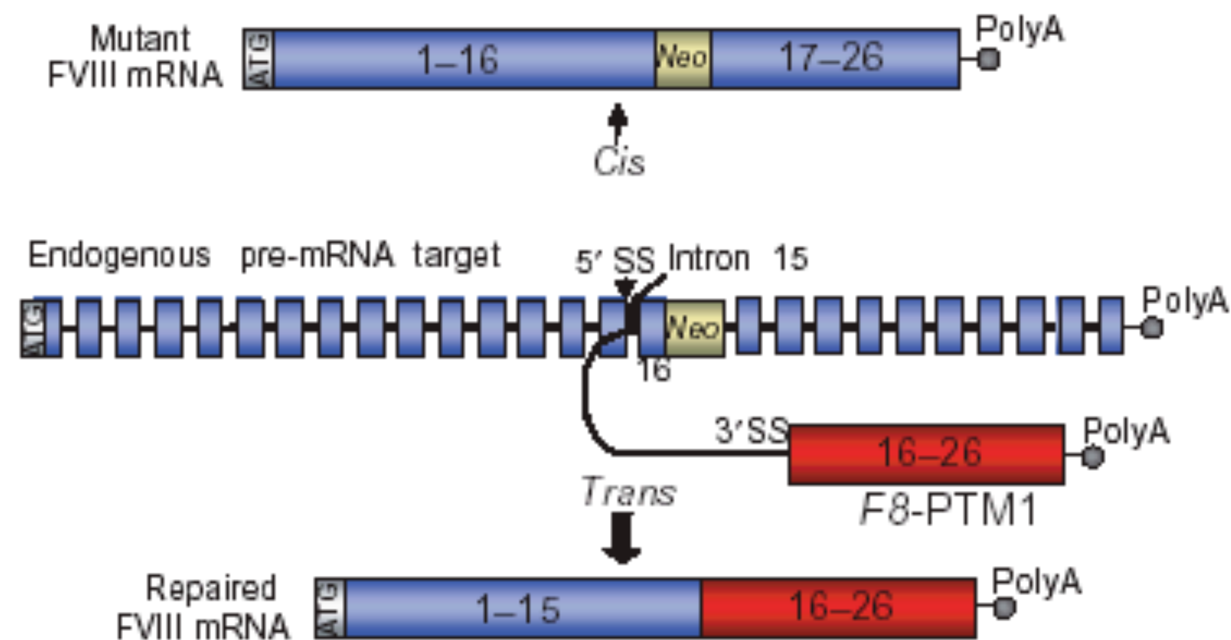
Trans-splicing domain

Coding domain

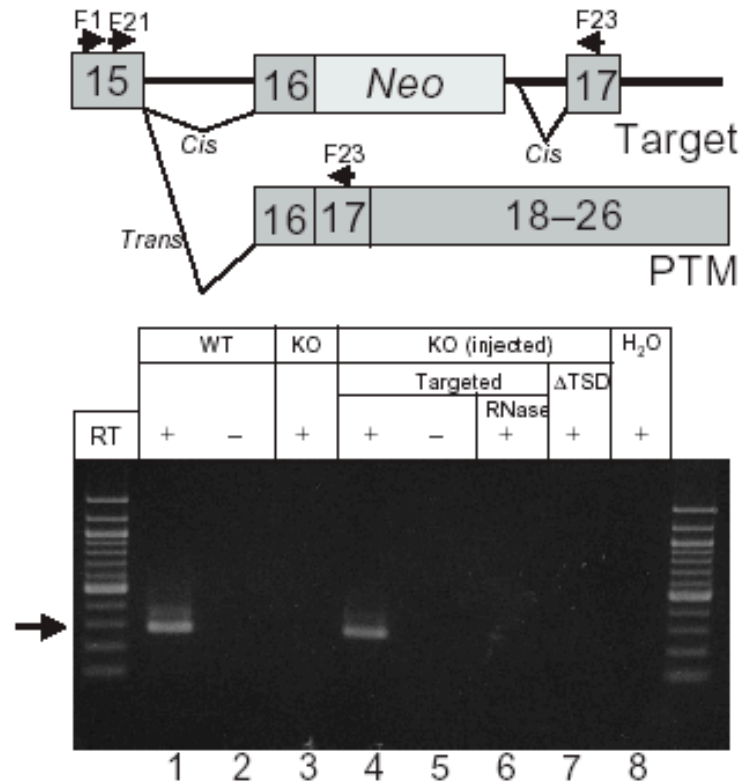
to be trans-spliced to the target

Phenotype correction of hemophilia A mice by spliceosome-mediated RNA *trans*-splicing

Hengjun Chao[1,3], S Gary Mansfield[2,3], Robert C Bartel[2], Suja Hiriyanna[2], Lloyd G Mitchell[2], Mariano A Garcia-Blanco[2] & Christopher E Walsh[1]

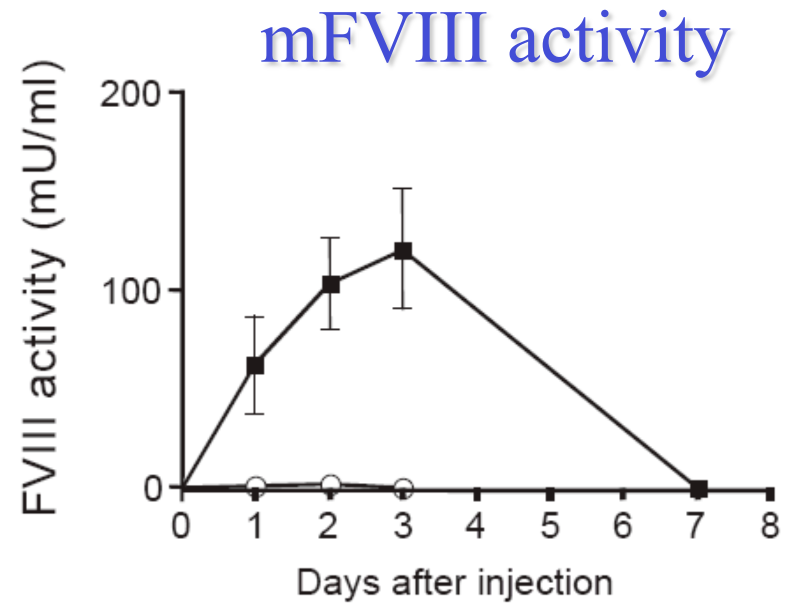


Correction of HA in E-16 mice (Naked DNA injection)



mFVIII mRNA

mFVIII activity



Trans-splicing advantages..........

- Reduction of transgene size

  replacement only of mutated portion

- Gene expression under the control of endogenous regulatory elements

- Tissue specific expression

  Repaired products are limited by the presence of the endogenous target

...........and disadvantages

- Low efficiency

Readings
pre mRNA splicing

- Mechanism of Alternative pre mRNA splicing Douglas Black Annu Rev Biochem 2003
- Another step forward for SELEXive splicing.Buratti E, Baralle FE.Trends Mol Med. 2005 Jan;11(1):5-9. Review.
- Listening to silence and understanding nonsense: exonic mutations that affect splicing.Cartegni L, Chew SL, Krainer AR. Nat Rev Genet. 2002 Apr;3(4):285-98
- Determinants of SR protein specificity.Tacke R, Manley JL.Curr Opin Cell Biol. 1999 Jun;11(3):358-62. Review.
- Nat Struct Mol Biol. 2009 Sep;16(9):902-3.When chromatin meets splicing.Kornblihtt AR, Schor IE, Allo M, Blencowe BJ.

In red those papers that are part of the exam (i.e. one question will be to comment a figure)

Model of spliceosome-mediated splicing of pre-mRNA

U4-snRNP si distacca dal complesso mentre

L'U1-snRNP si lega al

stretta vicinanza e U6-

snRNP si dissocia e

snRNP catalizza le due

viene sostituito da U6-snRNP

reazioni di transesterificazione