Alternative Splicing

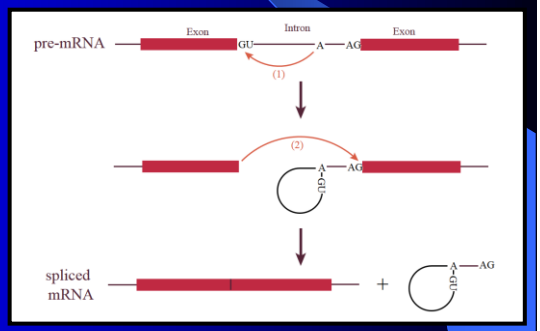
Fejlődés- és Molekuláris Genetika
2018

Splicing

Philip Sharp

Richard J. Roberts

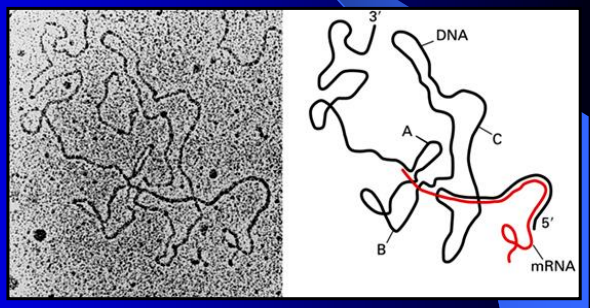
1993



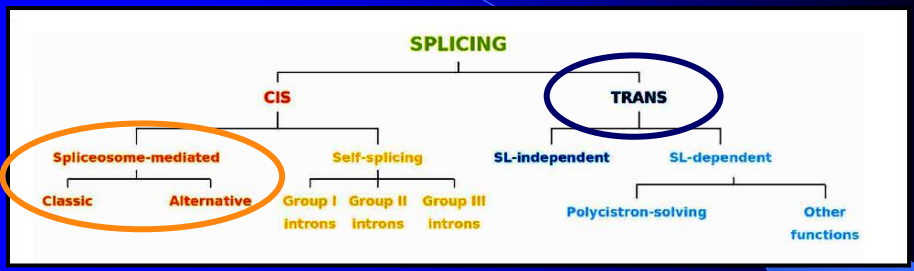
Intro to pre-mRNA Splicing

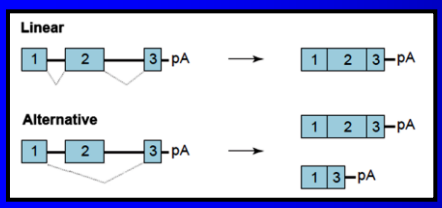
One of the earliest (1977) experiments showing that introns are present in genes is shown. In this experiment, a double-stranded DNA fragment containing most of the adenovirus hexon gene was denatured, hybridized with the hexon mRNA, and then viewed under the electron microscope.

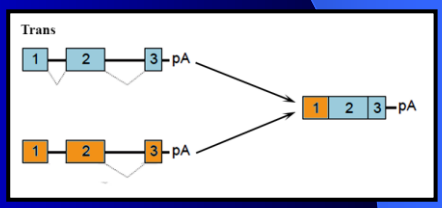
As shown in the micrograph and the schematic diagram on the right, DNA loop sequences corresponding to introns removed from the mRNA can be seen looping out from the DNA/RNA hybrid.



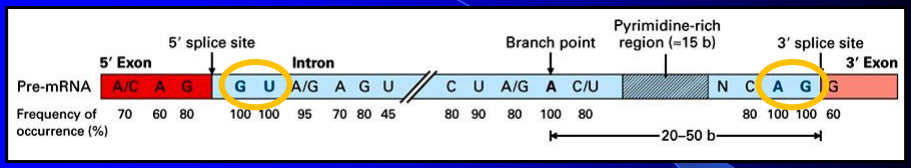
Splicing forms



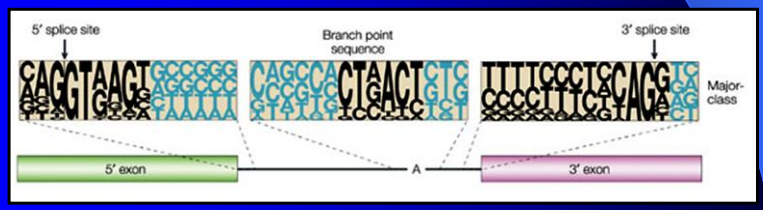




Splice Site Consensus Sequences



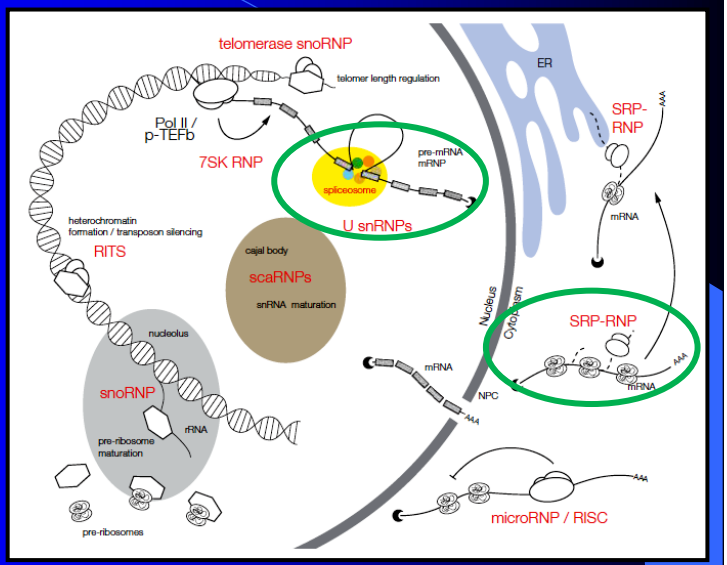
The GU dinucleotide at the 5' splice site of the intron and the AG dinucleotide at the 3' splice site are highly conserved. Also highly conserved within the intron is a branch point sequence containing the branch-point A residue located ~20-50 nucleotides upstream of the 3' splice site. The remaining central region of the intron (not shown) generally is unimportant for splicing.



Sequencing of the genome verified that the majority of splice sites did not match the consensus sequence; indeed, less than 5% of 50 splice sites matched the consensus, with greater than 25% having three or more mismatches from the 9 nt consensus. On the other hand, there are pseudoexons.

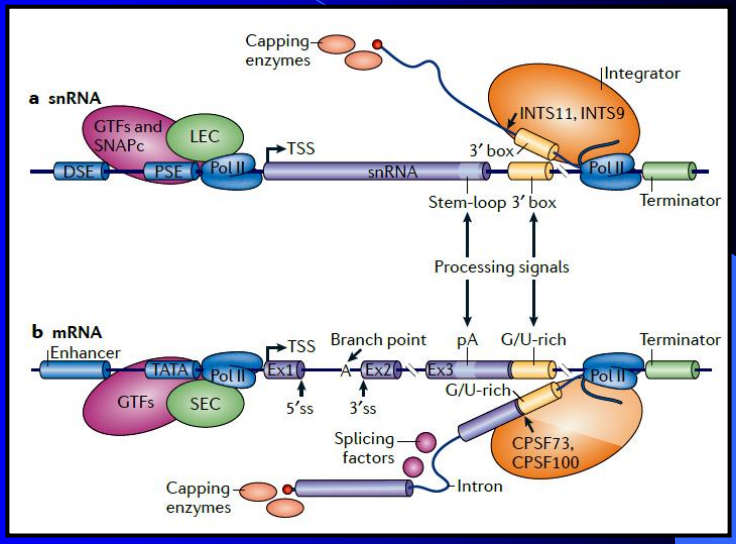
A simplified representation of RNPs

RITS: RNA-induced transcriptional silencing
snoRNP: small nucleolar RNP
UsnRNP: small nuclearRNP
p-TEFb: transcription elongation factor
SRP-RNP: signal recognition particle RNP
scaRNP: small Cajal body RNP
RISC: RNA-induced silencing complex



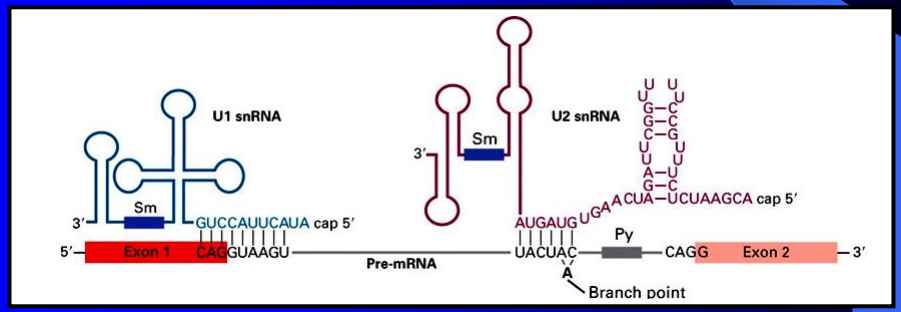
Transcription and processing of snRNAs and mRNAs.

DSE/PSE: distal/proximal sequence element
LEC/SEC: little/super elongation complex
GTFs: general transcription factors
SNAPc: snRNA-activating protein complex
CPSF73/100: polyadenylation specificity factor
pA: polyA signal
INTS: integrator subunits

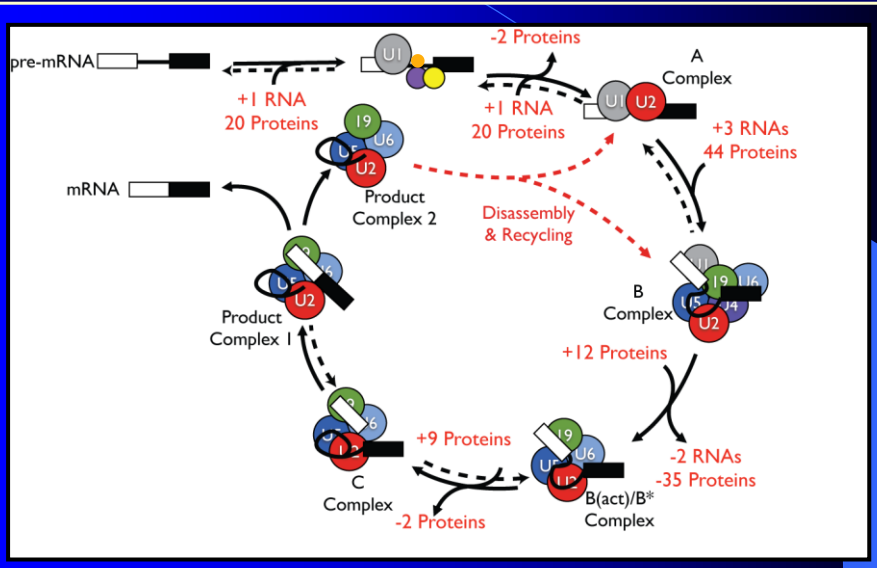


Small Nuclear RNAs (snRNAs) and Splicing

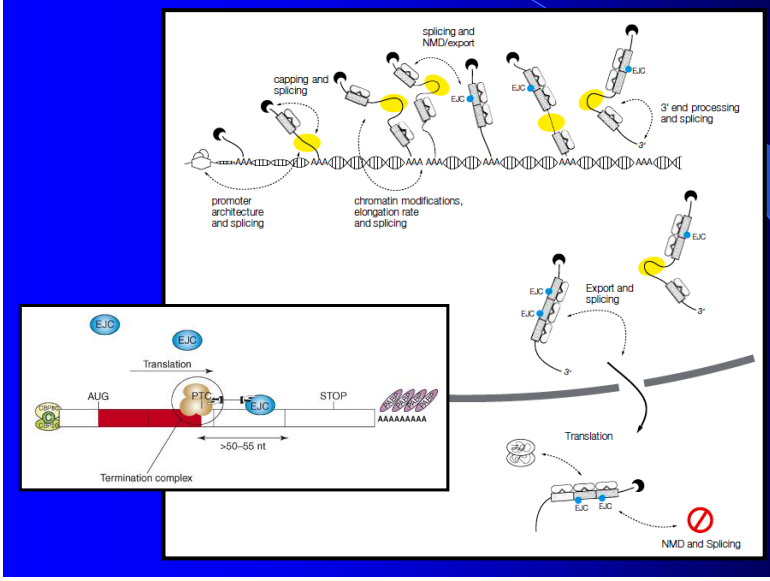
The splicing reaction requires 5 snRNAs (U1, U2, U4, U5, & U6) that range from about 100-200 nucleotides in length. Each snRNA forms a complex with 6-10 proteins. These snRNAs bind to pre-mRNA and each other within a larger splicing complex known as the spliceosome. Interactions between the U1 snRNA and the 5' splice site, and the U2 snRNA and the branch point sequence are crucial in selecting where splicing occurs.



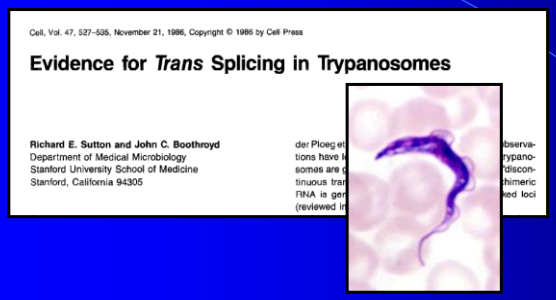
Splicing is dynamic, with sequential regulated alterations in RNA:RNA and RNA:protein interactions



Integration of splicing with other RNA-processing steps

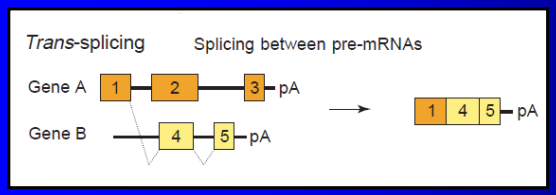


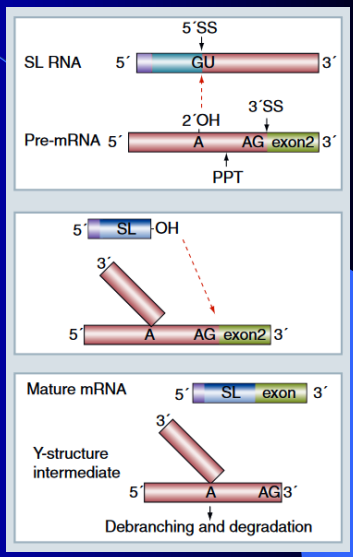
Trans-splicing


Cell, Vol. 47, 527–535, November 21, 1986, Copyright © 1986 by Cell Press

Evidence for *Trans* Splicing in Trypanosomes

Richard E. Sutton and John C. Boothroyd
Department of Medical Microbiology
Stanford University School of Medicine
Stanford, California 94305

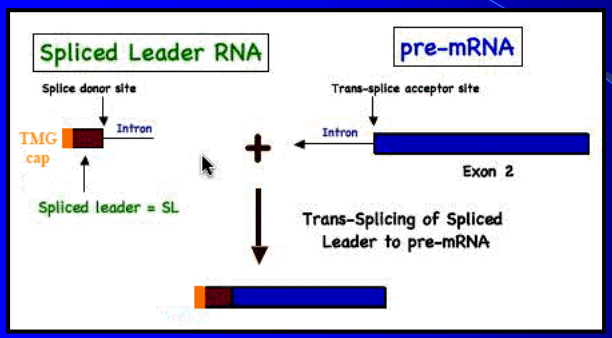




PPT: Polypyrimidine tract; SL: Spliced leader; SS: Splice site

Michaell, Fut. Microbiol., 2011

trans-Splicing Reactions Use snRNPs

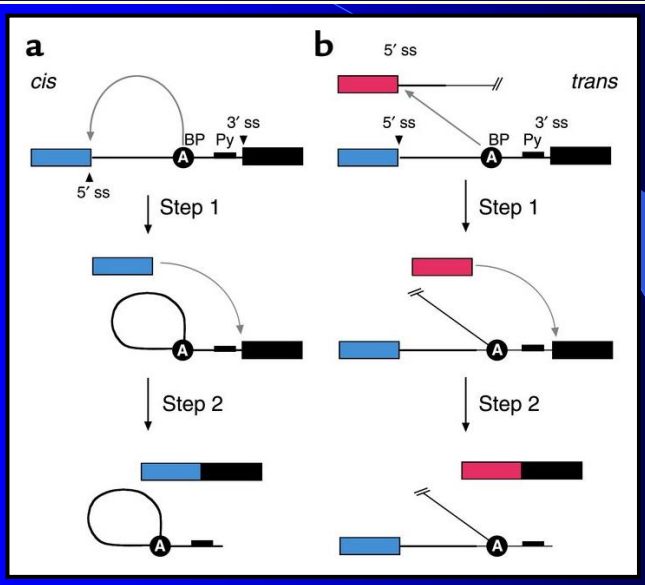


trans-splicing occurs in trypanosomes and worms where a short sequence (Splice leader (**SL) RNA**) is spliced to the 5′ ends of many precursor mRNAs.
In trypanosomes, all splicing is *trans*-splicing; all mRNAs begin with the SL, and genes do not contain introns. Transcription is polycistronic, and *trans*-splicing is responsible for separating the long polycistronic transcripts into monocistronic units.

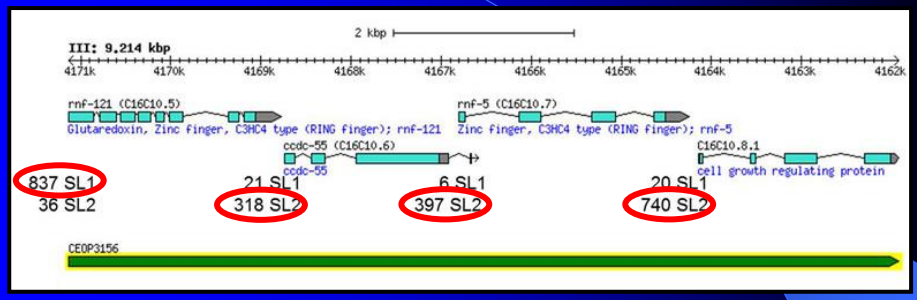
TGM cap: trimethyl-G cap

- In nematodes, the genes do contain introns, and the pre-mRNA products of many genes are not subject to *trans*-splicing.
- SL RNAs have a structure resembling the Sm-binding site of U-snRNAs.
- SL sequence wears an unique trimethylguanosine cap.
- SL sequence is required to binding to signal recognition particle receptor

Mechanism of trans-splicing



SL2-type operon



There are >1200 documented operons of this type in the *C. elegans* genome. The figure shows a four-gene operon with exons shown as colored boxes and introns as angled lines.

SL1 leaders are spliced mostly to pre-mRNAs from genes with outrons, intron-like sequences at the 5′-ends of the pre-mRNAs. In contrast, ***SL2*** leaders are nearly exclusively *trans-spliced* to genes that occur downstream in polycistronic pre-mRNAs produced from operons.

Reprogramming of tau alternative splicing by spliceosome-mediated RNA trans-splicing: Implications for tauopathies

Teresa Rodriguez-Martin*†‡, Mariano A. Garcia-Blanco§, S. Gary Mansfield§, Andrew C. Grover¶, Michael Hutton¶, Qingming Yu‖, Jianhua Zhou‖, Brian H. Anderton†‡, and Jean-Marc Gallo*†**

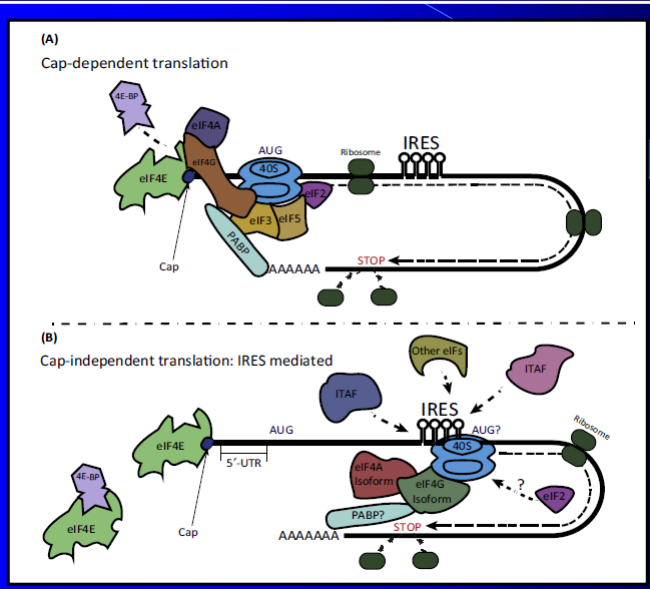
Trans-Splicing-Mediated Improvement in a Severe Mouse Model of Spinal Muscular Atrophy

Tristan H. Coady and Christian L. Lorson

Correction of tau mis-splicing caused by FTDP-17 *MAPT* mutations by spliceosome-mediated RNA *trans*-splicing

Teresa Rodriguez-Martin[1,2], Karen Anthony[1], Mariano A. Garcia-Blanco[3], S. Gary Mansfield[4], Brian H. Anderton[2] and Jean-Marc Gallo[1,*]

Translation of polycistronic gene



eIF4: eukaryote initiation factors
PABP: Poly-A Binding Protein
IRES: Internal Ribosome Entry Site
ITAF: IRES Trans-Acting Factor

Discovery of alternative splicing

First predicted by Walter Gilbert in 1978

First discovered for an immunoglobulin heavy chain gene in 1980
(Edmund Choi, Michael Kuehl & Randolph Wall, *Nature* **286**, 776 - 779)

Alternative splicing gives two forms of the protein with different C-termini:

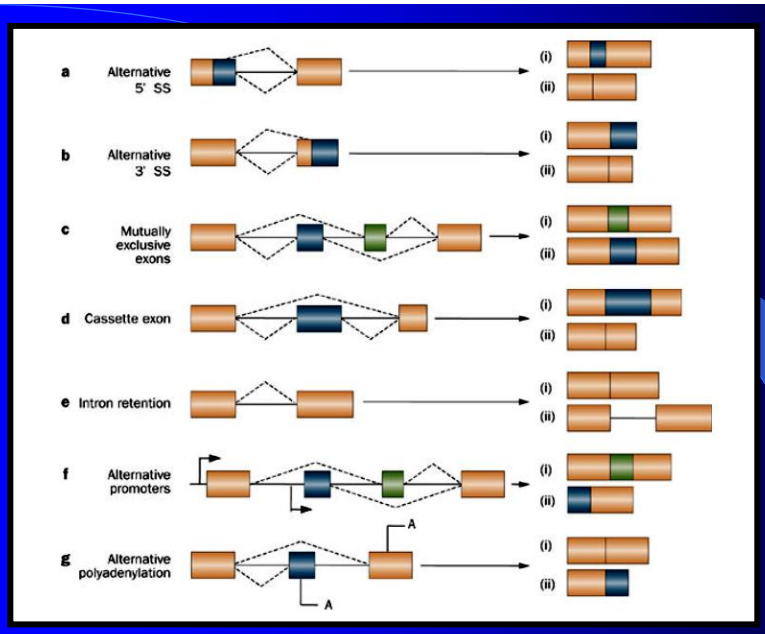
- First form is shorter and secreted
- Other stays anchored in the plasma membrane via its C-terminus

Advantages of alternative splicing

More different transcripts (and proteins) from a single gene

Diverse 3'UTR → Different
- stability
- localization
- translational control
- reading frame
- PT modifications

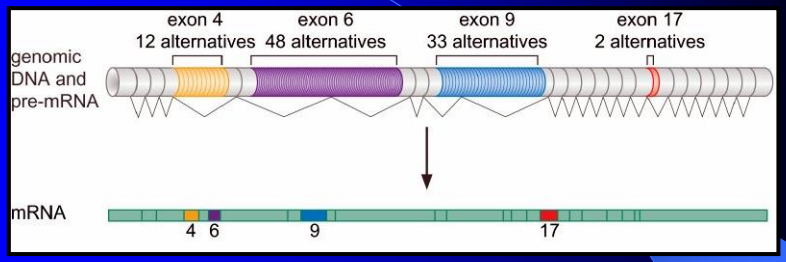
Genome-wide studies estimated that 90–95% of human genes undergo some level of alternative splicing and the transcript could be non-coding.



Types of alternative splicing

Down Syndrome Cell Adhesion Molecule

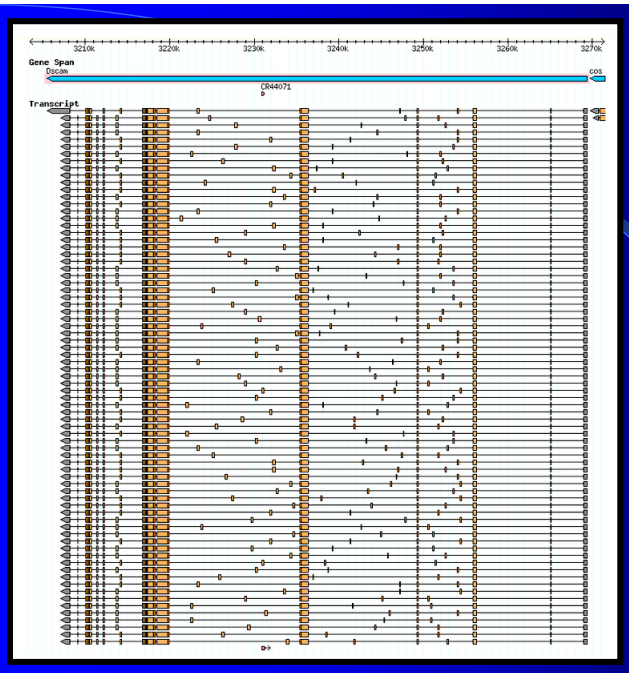
DSCAM maps to chromosome 21 in a region critical for the neurocognitive.



DSCAM gene ortholog (CG17800) in *D. melanogaster* can be spliced in 38,016 alternative ways.

How is this level of discrimination possible?

The level of DSCAM expression is increased by more than 20% in the DS brain.

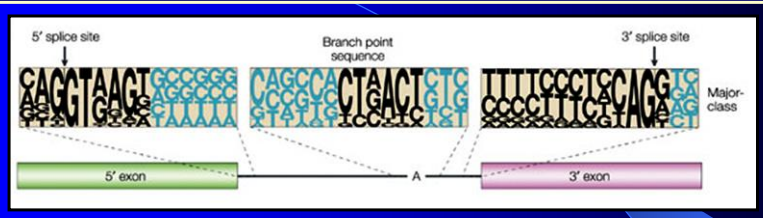


CG17800 on Flybase, 2013

What makes splicing alternative?

- *Cis* elements:
  Enhancers and silencers
- *Trans* factors:
  Regulator proteins

Specifications of exons and introns



Sequencing of the genome verified that the majority of splice sites did not match the consensus sequence; indeed, less than 5% of 50 splice sites matched the consensus, with greater than 25% having three or more mismatches from the 9 nt consensus. On the other hand, there are pseudoexons.

Splicing Regulatory Elements

Exon Splicing Enhancers ESE
Exon Splicing Silencers ESS

Intron Splicing Enhancers ISE
Intron Splicing Silencers ISS

Exon Splicing Enhancers (ESEs)

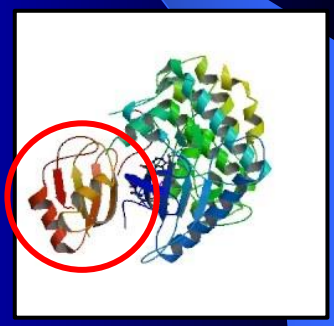
They bond SR (ser/arg rich) proteins.

SR proteins have an RS (arg/ser) domain and one or two RNA recognition motifs (RRMs).

An example: **ASF/SF2**
alternative splicing factor 1 (ASF1), pre-mRNA-splicing factor SF2

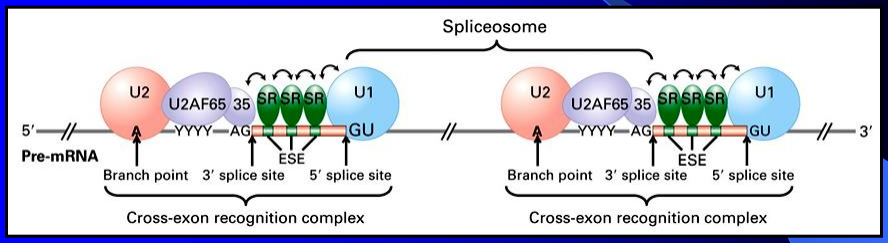
Alternative splicing is affected by ASF/SF2 in a concentration-dependent manner; differing concentrations of ASF/SF2 is a mechanism for alternative splicing regulation.

ASF/SF2 can act as an oncoprotein; it can alter the splicing patterns of crucial cell cycle regulatory genes and suppressor genes. ASF/SF2 controls the splicing of various tumor suppressor genes, kinases, and kinase receptors.



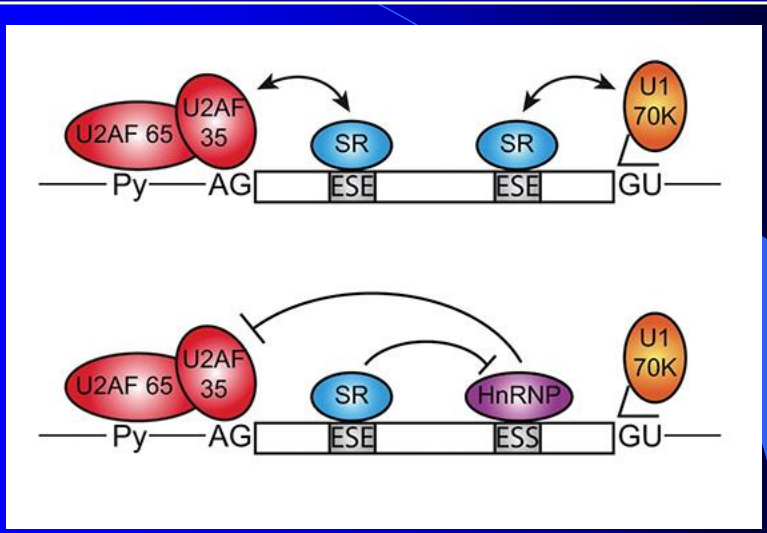
Exon Recognition in Long Pre-mRNAs

The average human intron is ~3,500 nucleotides in length, while the average exon is only ~150 nucleotides long. The longest introns are 500 kb in length.



Exons contain exonic splicing enhancers (**ESE**s) that bind SR proteins which recruit the U2 snRNP & U2AF factor to 3' splice sites, and the U1 snRNP to 5' splice sites flanking exons. These assemblies are known as *cross-exon recognition complexes*.

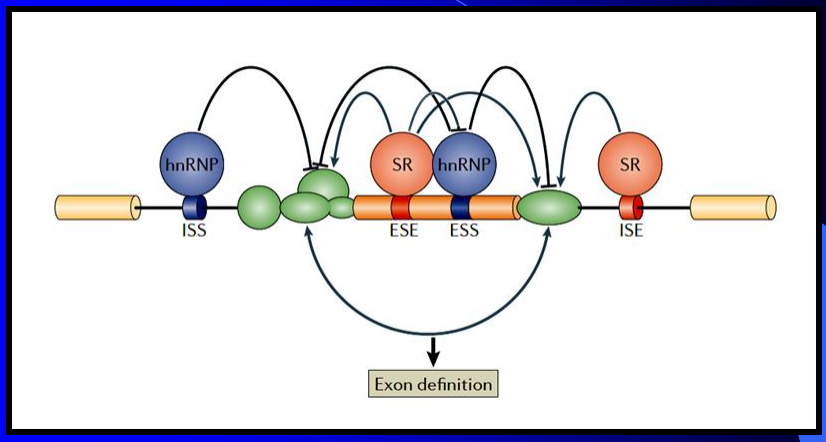
SR concentration and alternative splicing



Long and Caceres, Biochem J., 2008

hnRNP: Heterogeneous nuclear ribonucleoproteins

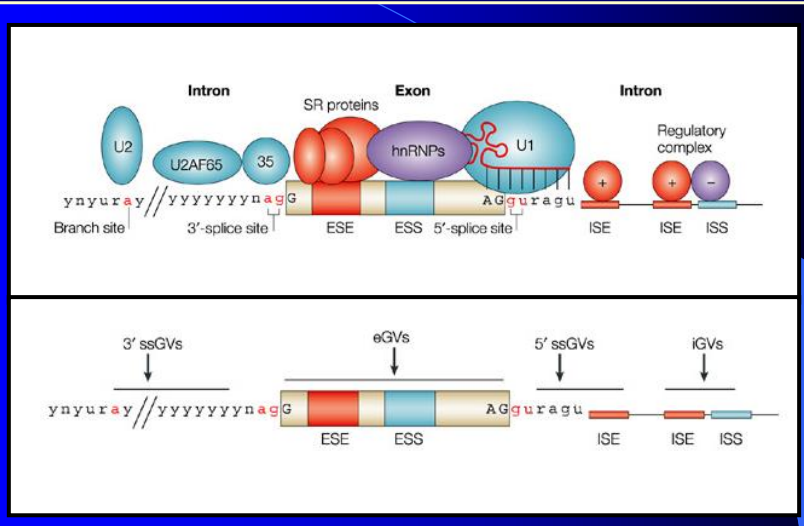
Splicing regulatory elements on introns



Splicing regulatory elements (SREs):
- ESE/ISE: Exonic/intronic splicing enhancer
- ESS/ISS: Exonic/intronic splicing silencer

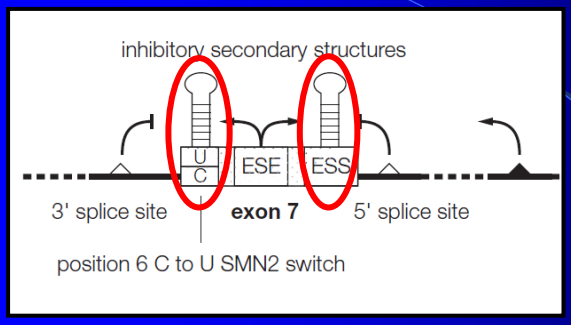
Splicing factors:
- SR: ser/arg-rich proteins
- hnRNP: heterogeneous nuclear ribonucleoproteins
- hnRNP1 = polypyrimidine track binding protein

Effects of genomic variants



eGVs: exonic genomic variants, iGVs: intronic genomic variants
ssGVs: splice site genomic variants

Survival of Motor Neuron (SMN) switch

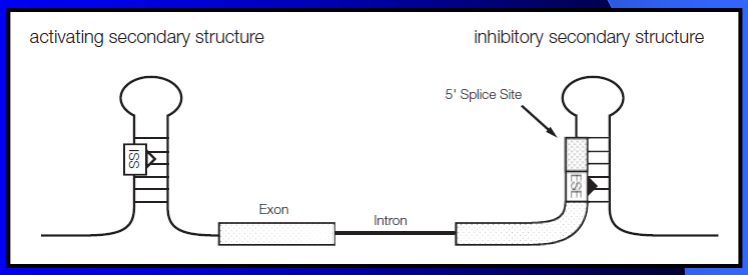


Two almost identical genes code for the proteins SMN1 (functional) and SMN2 (mostly nonfunctional). In the SMN1 transcript, the *C* at position 6 is in a proposed ASF/SF2 enhancer binding site promoting exon 7 inclusion. In SMN2, the *C to U* transition is also proposed to create an hnRNP A1 binding site that favors the exclusion of exon 7.

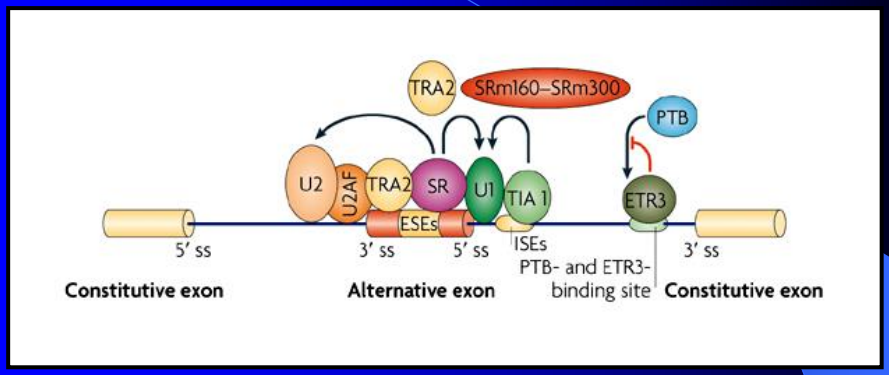
The role of RNA secondary structure in splicing.

↕

DEAD-box helicases

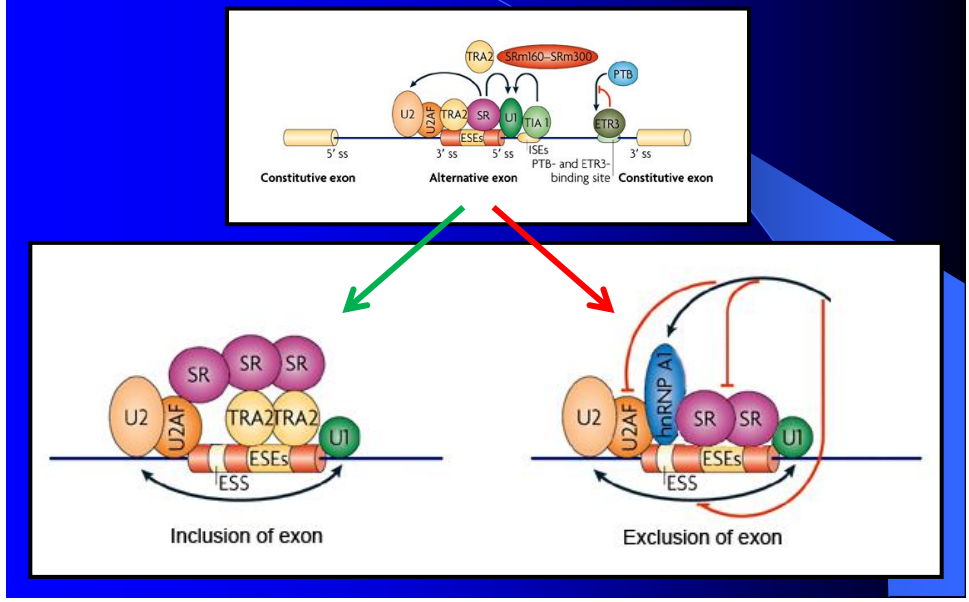


Alternative exon

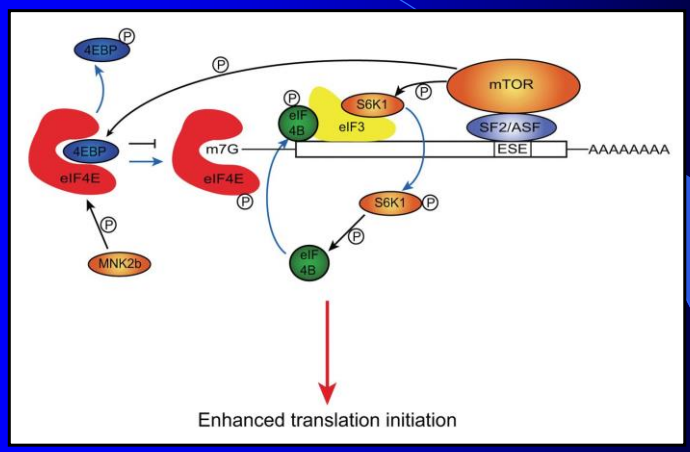


TRA2: transformer (splicing co-activators)
SRm: SR-related nuclear matrix protein
TIA1: T cell restricted intracellular antigen 1
ETR3: elav-type RNA-binding protein 3
PTB: polypyrimidine-tract binding protein

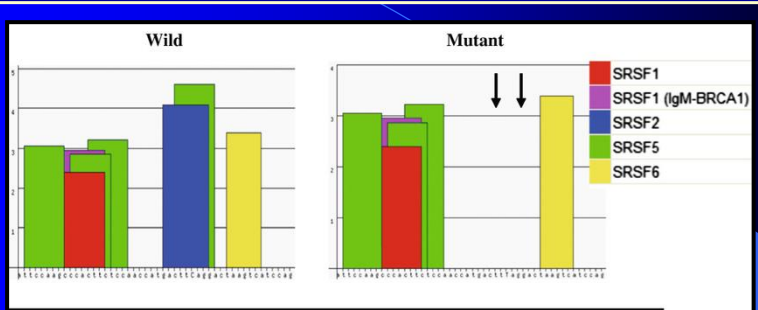
An example of the alternative splicing



Role of ESE and ASF/SF2 in translation



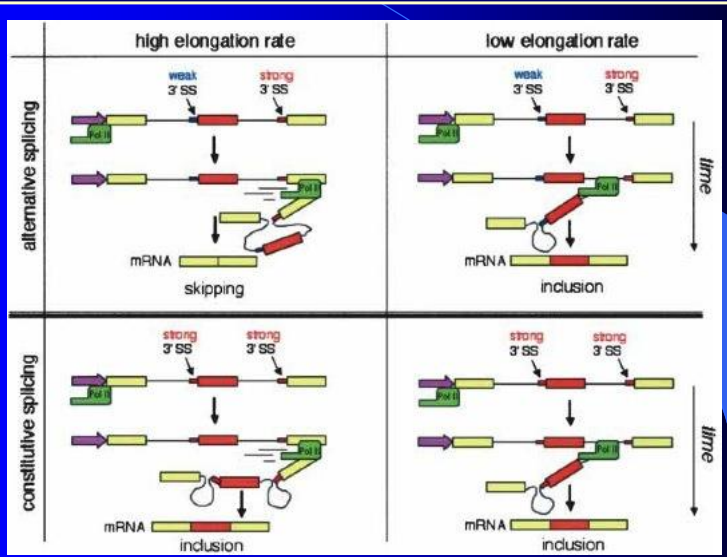
ESE motifs



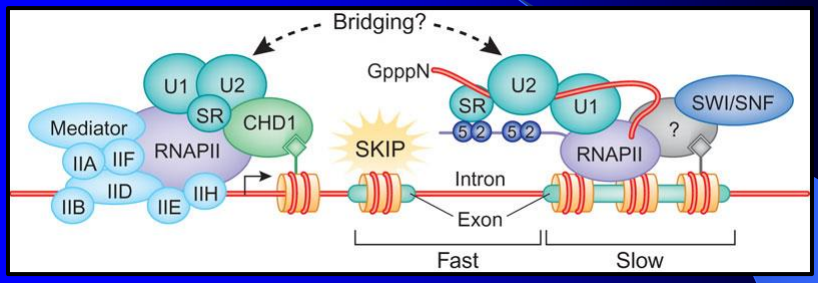
Splicing factor	Binding site	Score (wild / mutant)
SRSF1	CCCACTT	2.39 / 2.39
SRSF1 (IgM-BRCA1)	CCCACTT	2.95 / 2.95
SRSF2	GACTTCAG	4.08 / - -
SRSF5	TTCCAAG	3.05 / 3.36
SRSF5	CCACTTC	2.86 / 2.86
SRSF5	CTTCTCC	3.21 / 2.80
SRSF5	CTTCAGG	4.61 / - -
SRSF6	TAAGTC	3.39 / 3.39

Wang et al, BMC Genomics, 2014

Does splice site strength affect alternative splicing?



Epigenetic modifications and alternative splicing



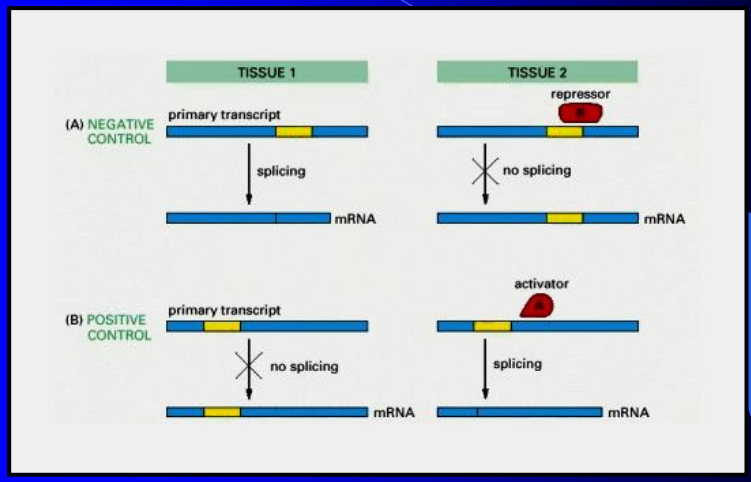
H3K4me3

H3K36me3

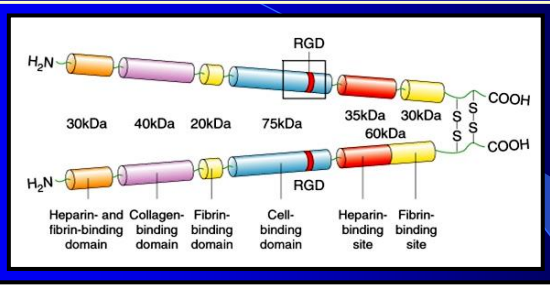
CHD1: Chromodomain-helicase-DNA-binding protein 1

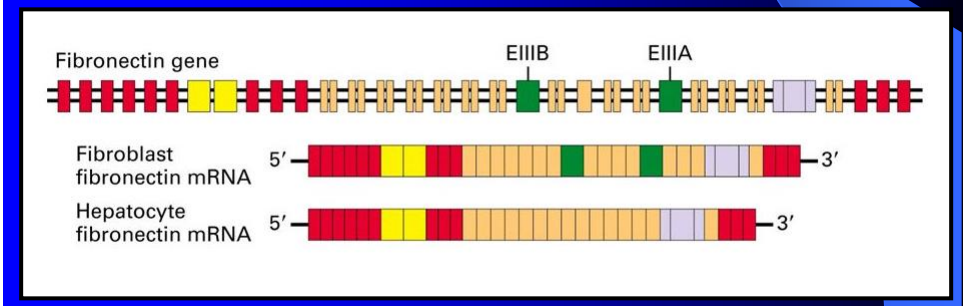
SWF/SNF: nucleosome remodeling complex

Tissue-specific splicing

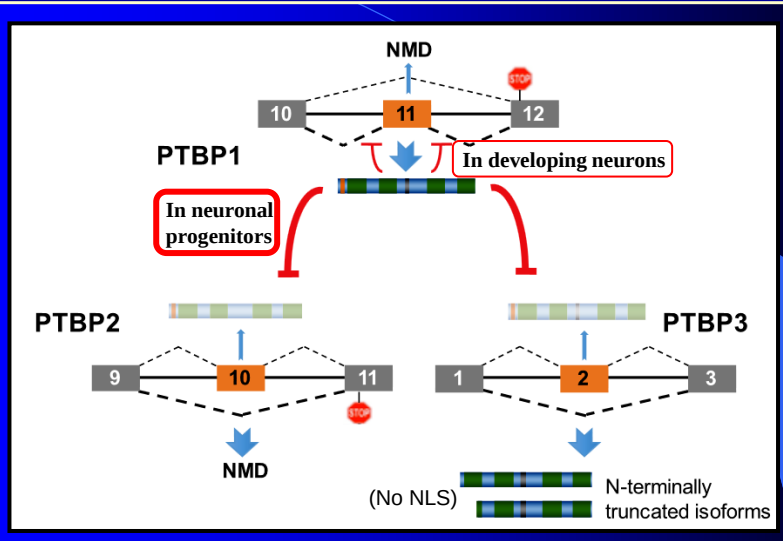


Fibronectin





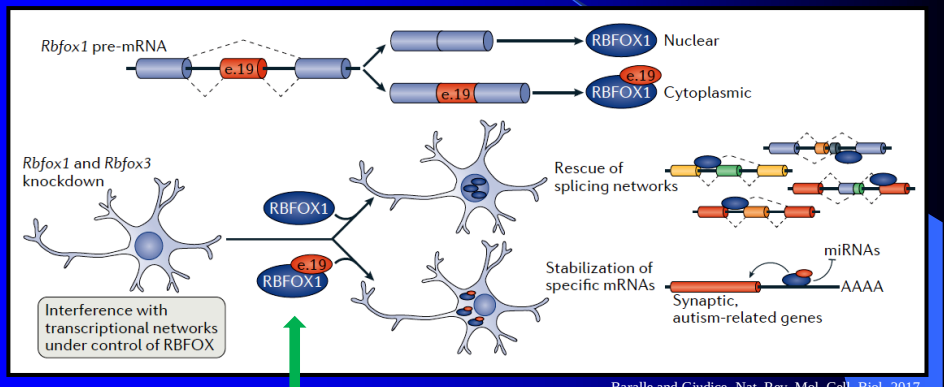
Splicing of splicers



Coelho et al., Bio. Soc. Trans. 2014

Skipping of PTBP1 exon 11 and PTBP2 exon 10 is frame-shifting, leading to insertion of a premature termination codon and NMD.

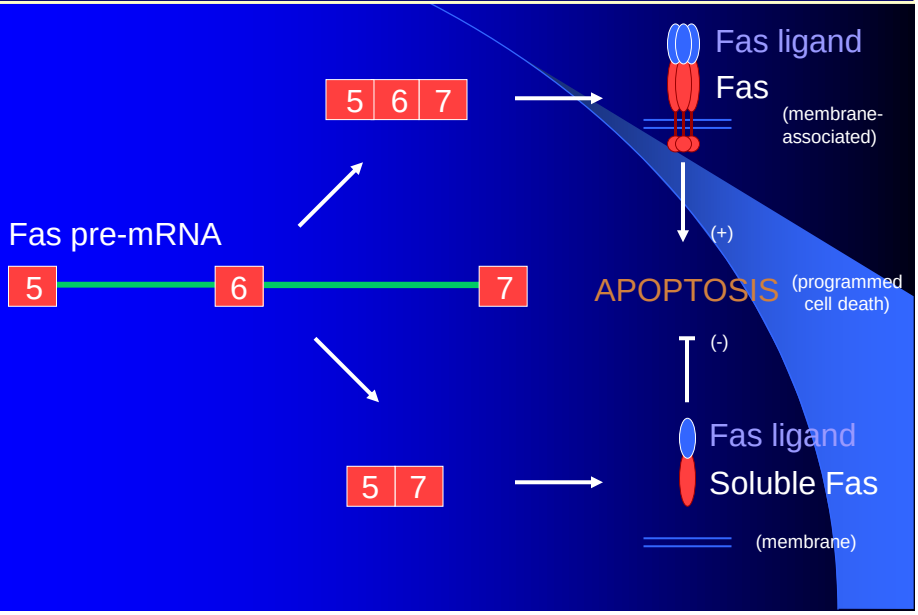
RNA-binding protein FOX1 homologue 1 (*Rbfox1*)



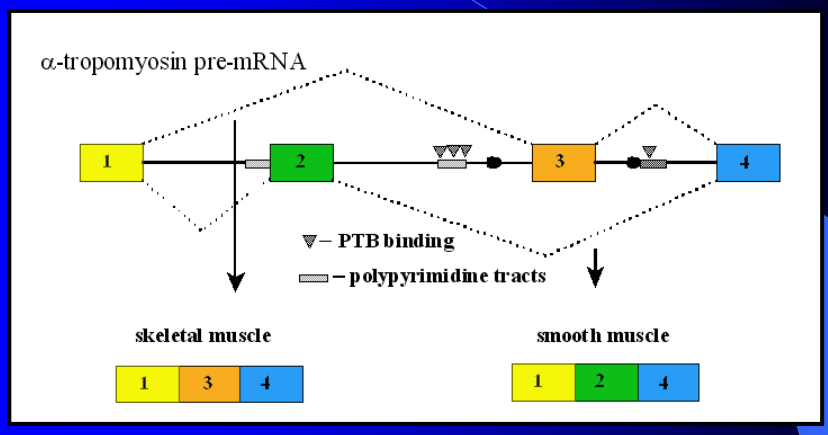
Baralle and Giudice, Nat. Rev. Mol. Cell. Biol, 2017

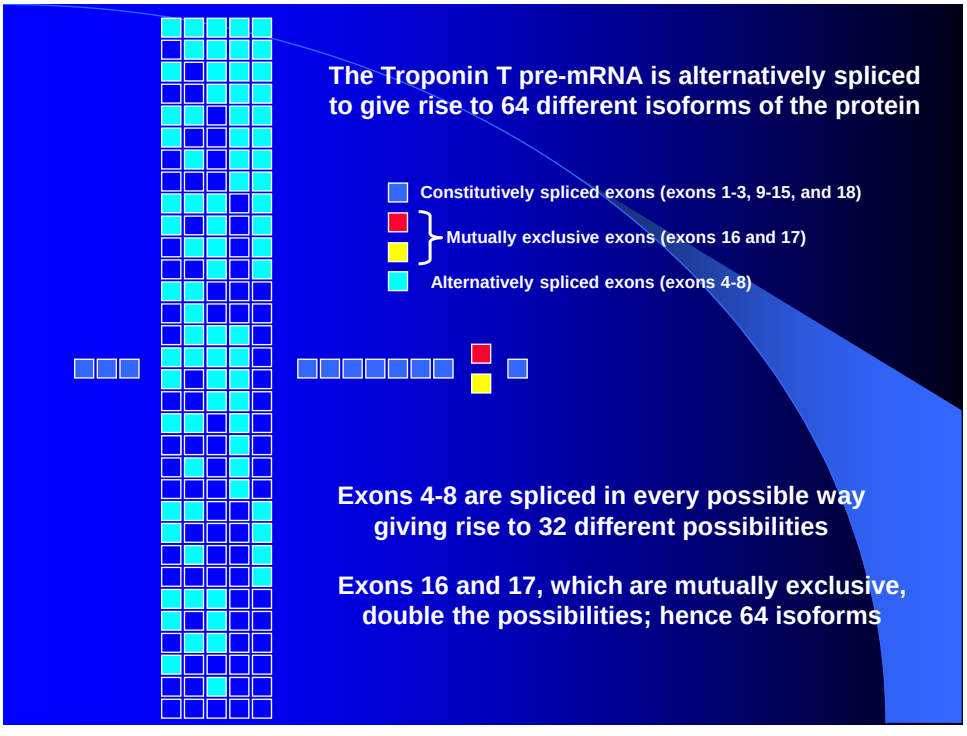
introduction of exogenous Rbfox1

Alternative splicing can generate mRNAs encoding proteins with opposite functions

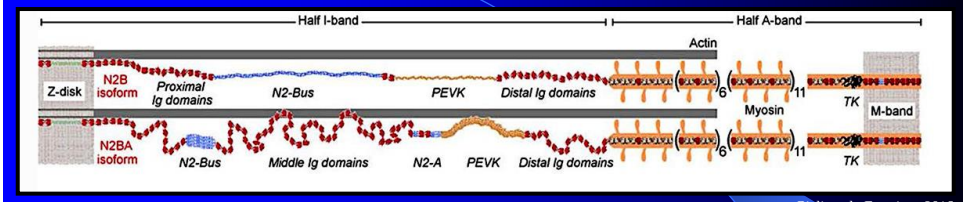


Splicing of tropomyosin mRNA



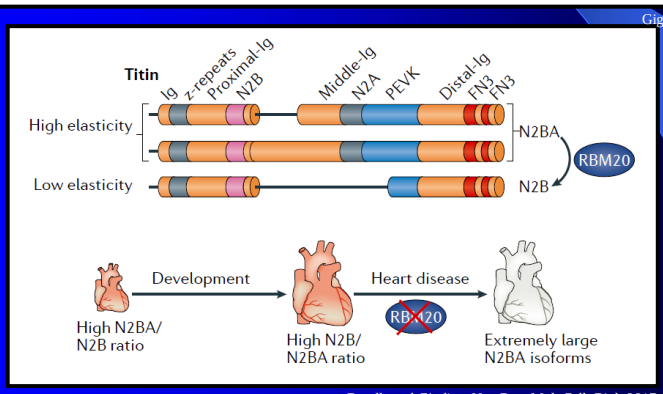


Heart development and titin splicing



Gigli et al., Frontiers, 2016

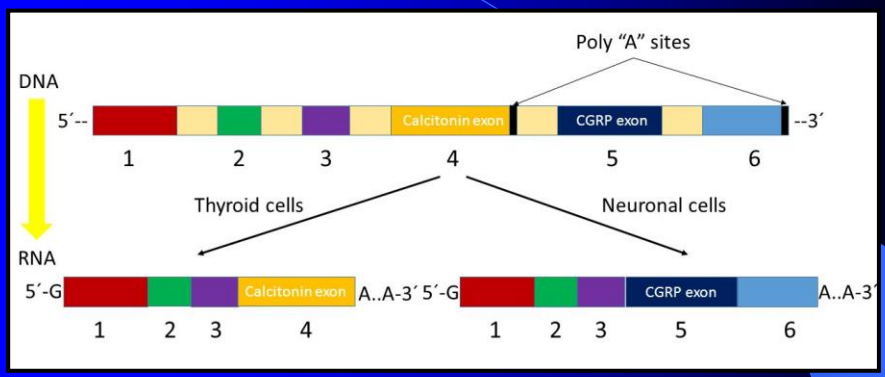
Titins:
N2BA: fetal
N2B: adult



RBM20:
RNA-binding protein 20

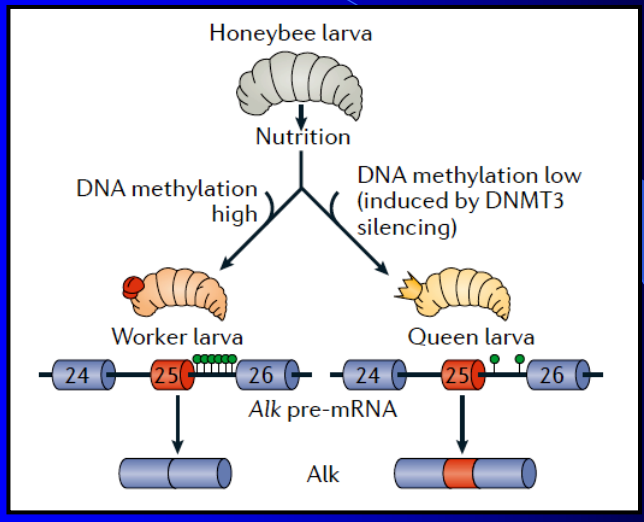
Baralle and Giudice, Nat. Rev. Mol. Cell. Biol, 2017

Alternative splicing of the CALCA gene



Calcitonin: 32 aa, lowers Ca^{2+} in blood
CGRP: 37 aa, acts as vasodilatator and functions in transmission of pain

alternative splicing and DNA methylation would be closely associated



Baralle and Giudice, Nat. Rev. Mol. Cell. Biol, 2017

Alk (anaplastic lymphoma kinase):
a metabolic regulator with the capacity to enable growth in a nutrient-independent manner

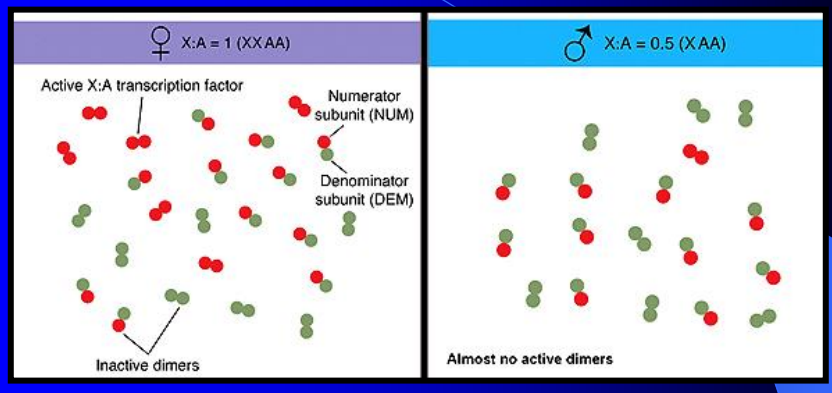
Evolutionary overview of alternative splicing

- Introns unlikely to have been derived from ancient genes
- Multi-intron genes probably predated alternative splicing
- Most eukaryotes have introns but alternative splicing prevalent only in *multicellular organism*
- S.cerevisiae has only 253 introns (3% of its genes) and only *6 genes* have *2 introns*
- S. pombe: 43% of its genes have introns (usually 40-75 nt)
- S.cerevisiae and S. pombe have NO alternative splicing

Somatic sex determination:
X : A ratio,
Chromosomal

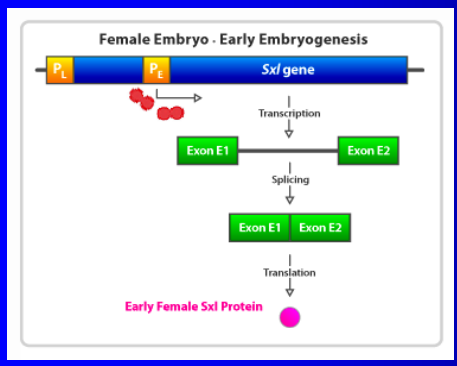
Kariotype	Caenorhabditis	Drosophila	Homo
	XX : X0	XX : XY	XX : XY
XX : 2A (1,0)	hermaphrodite	female	female
XY : 2A (0,5)		male	male
X : 2A (0,5)	male	*male (sterile)*	*female*
XXX : 2A (1,5)	hermaphrodite	female (sterile)	female
XXY : 2A (1,0)		*female*	*male*
XX : AAA (0,67)	male	intersex	
XXX : AAAA (0,75)	hermaphrodite	intersex	

How can the fly count to two?

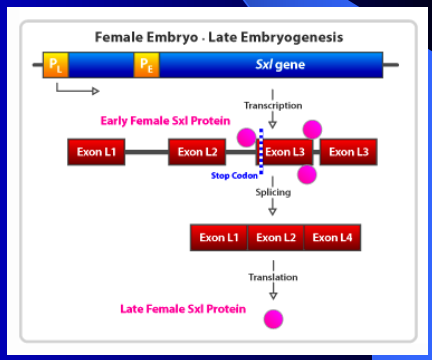


Numerators and denominators work together.

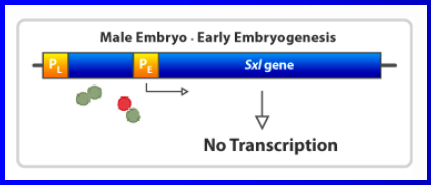
Numerators and *sex lethal* (*sxl*), the master switch gene



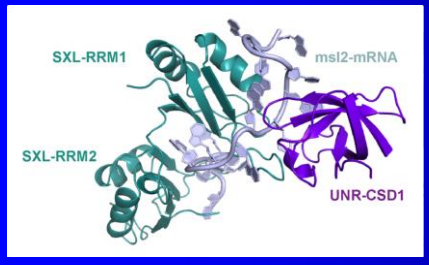
Females (XX)

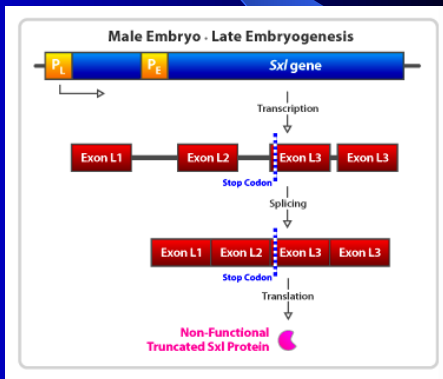


Numerators and *sex lethal* (*sxl*), the master switch gene

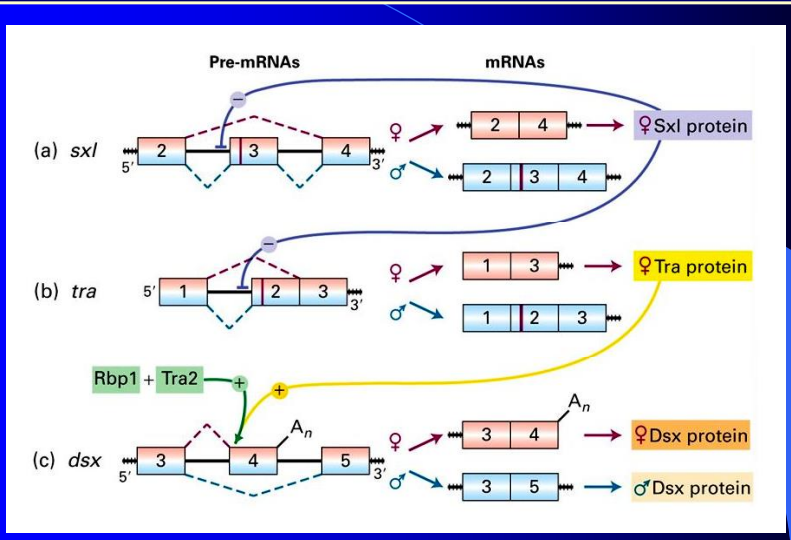


Males (XY)

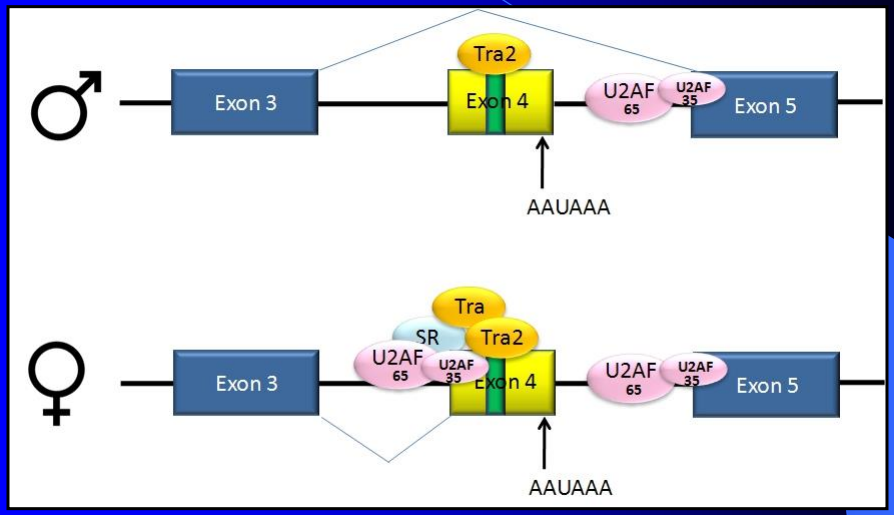


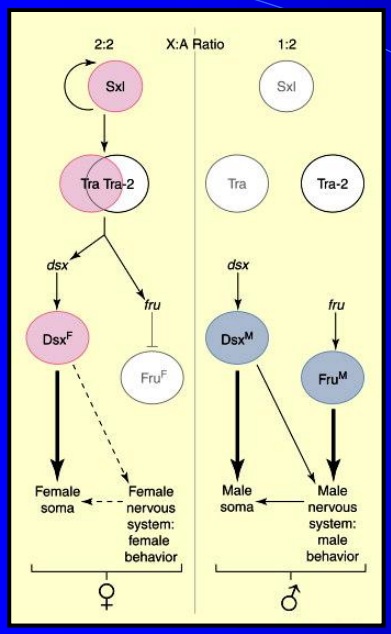


The pattern of sex-specific RNA splicing



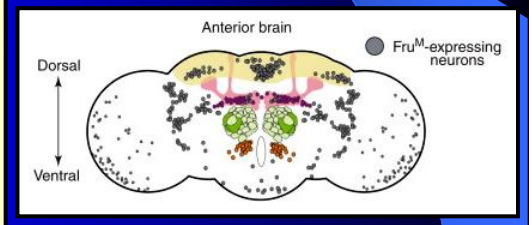
Alternative splicing for doublesex





Billeter et al., Curr. Biol. 2006

Tra controls the male sexual behavior



The End