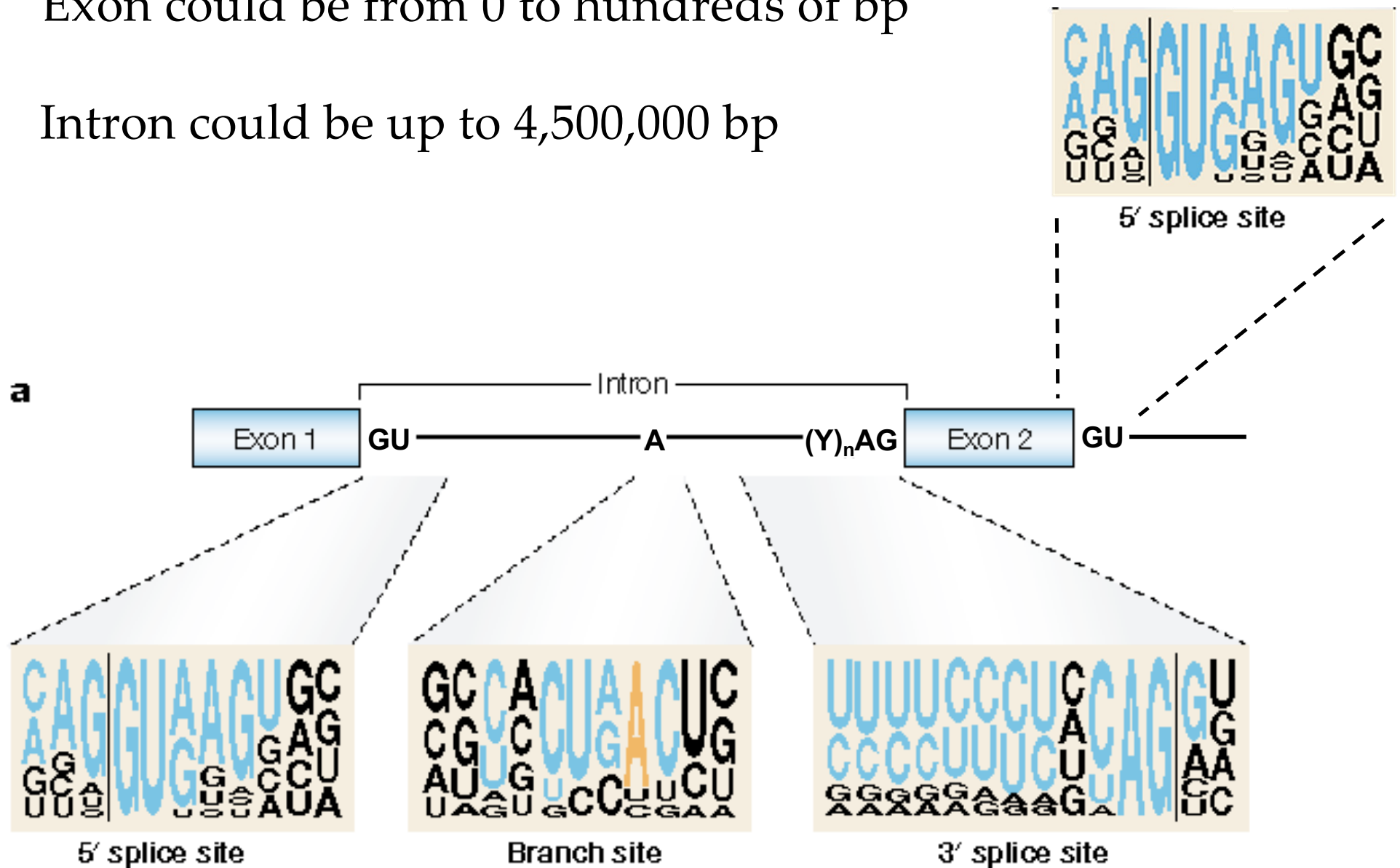


# Alternative Splicing

# How are splice sites identified?

Exon could be from 0 to hundreds of bp

Intron could be up to 4,500,000 bp

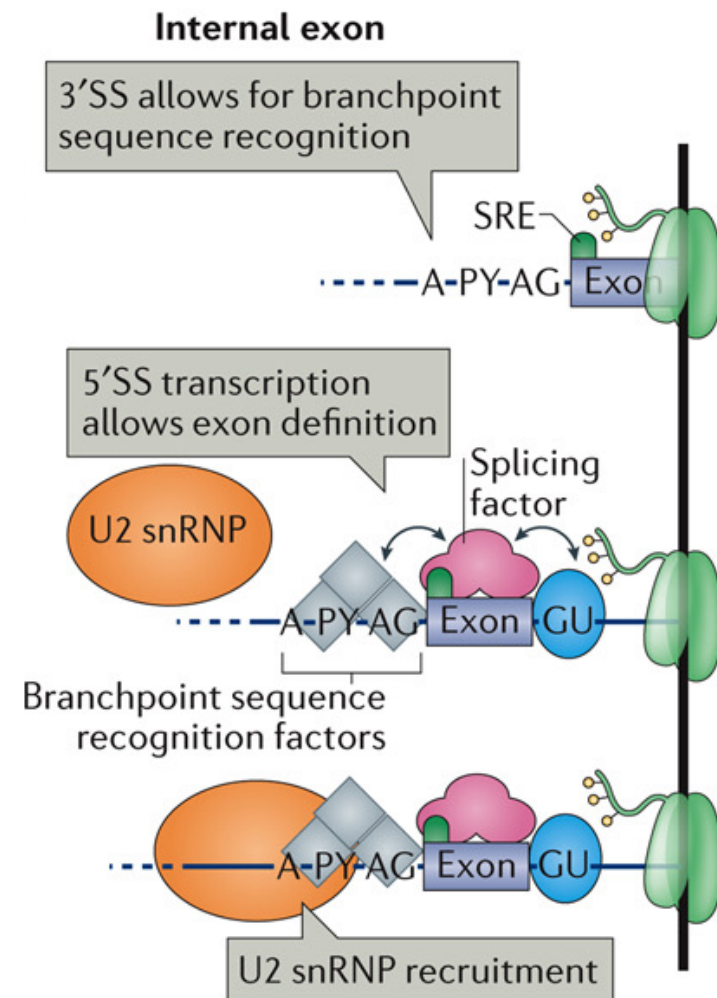


# Strategies for splice site identification

The complexity of gene architecture varies between phyla requiring different mechanisms to identify splice sites. Metazoan splice sites are short and poorly conserved, in contrast to budding yeast splice sites. This implies that the spliceosome might encounter frequent 'incorrect' splice sites along the transcripts.

In metazoa, accurate splice site selection depends on short conserved sequences, known as **splicing regulatory elements (SREs)**, that reside in introns or exons.

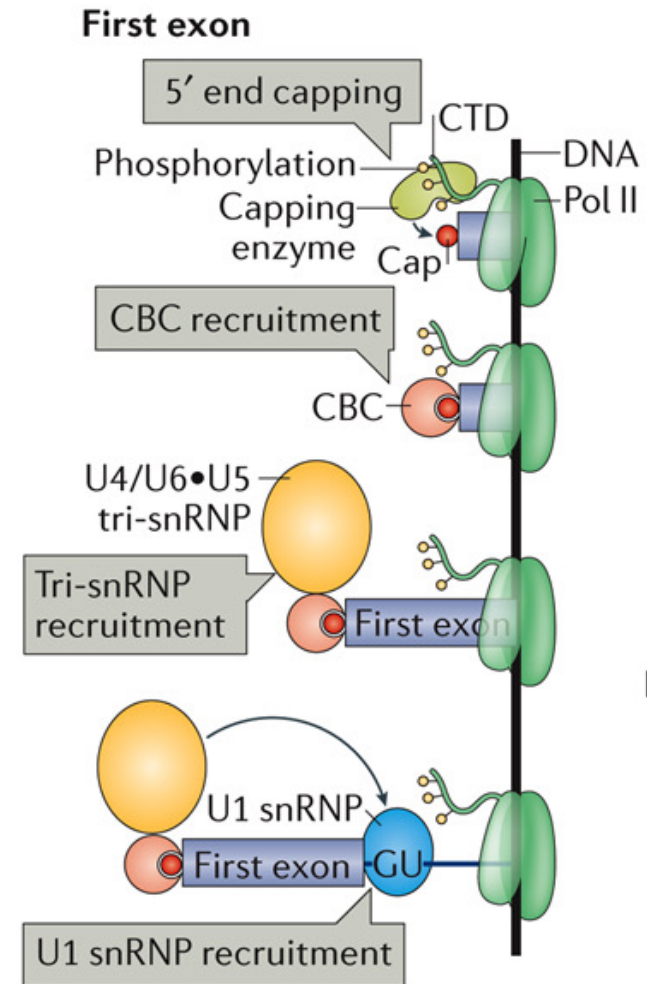
Regulatory proteins, such as **SR proteins** and **hnRNPs**, specifically bind to SREs and influence splice site recognition and/or spliceosome assembly.



# First exon definition

**First exon:** Splicing of the first intron depends on first exon definition . First exon boundaries consist of m7G cap structure at the 5' end of the transcript and the 5'SS of the first intron. The nuclear **cap-binding complex (CBC)** serves as a platform for interacting with factors that are involved in RNA processing.

CBC directly interacts with tri-snRNP protein components and its depletion impairs both U1 snRNP and tri-snRNP recruitment to the pre-mRNA .

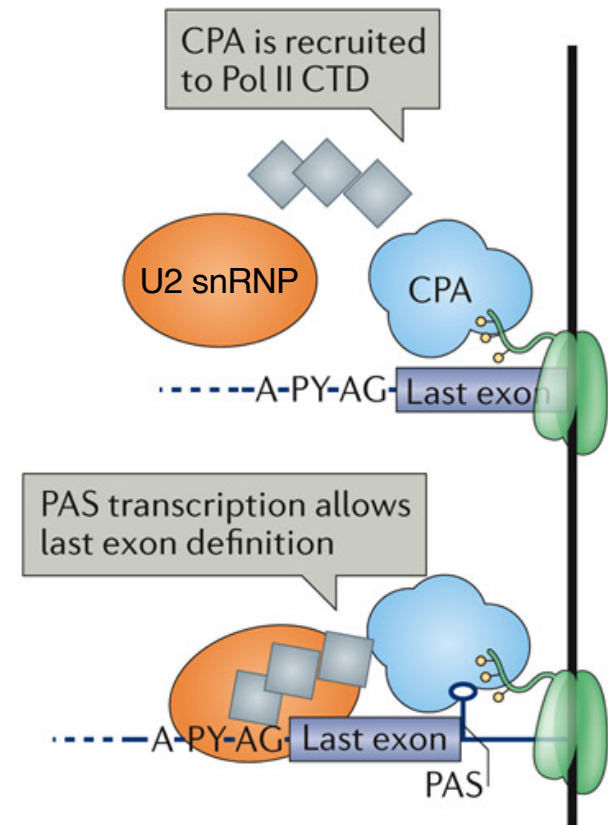


# Terminal exon definition

**Terminal exon:** The 3'SS of the last intron and the poly(A) site (PAS) set the terminal exon boundaries. PAS elimination results in the specific inhibition of last intron splicing, indicating that 3' end processing contributes to terminal exon definition. The U2 snRNP, U2AF65 and cleavage and polyadenylation specificity factor (CPSF) functionally and physically interact. Splicing and the regulation of 3' end processing is reciprocal, as the inactivation of the terminal 3'SS inhibits 3' end processing and transcription termination.

The mammalian U1 snRNP component U1 70k directly interacts with poly(A) polymerase and inhibits polyadenylation. This may reflect its role in protecting nascent RNA from premature 3' end processing. In such a model, the U1 snRNP binds to nascent RNA and suppresses the activity of adjacent cryptic PASs.

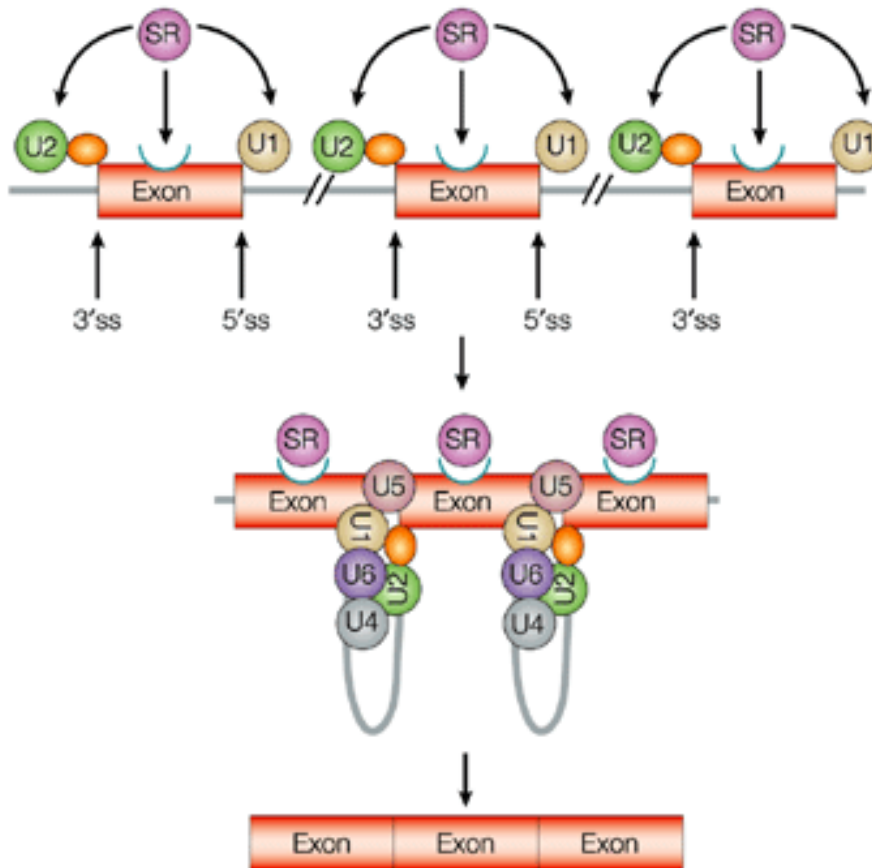
Terminal exon



# How Are Short Exons Flanked by Long Introns Defined and Committed to Splicing?

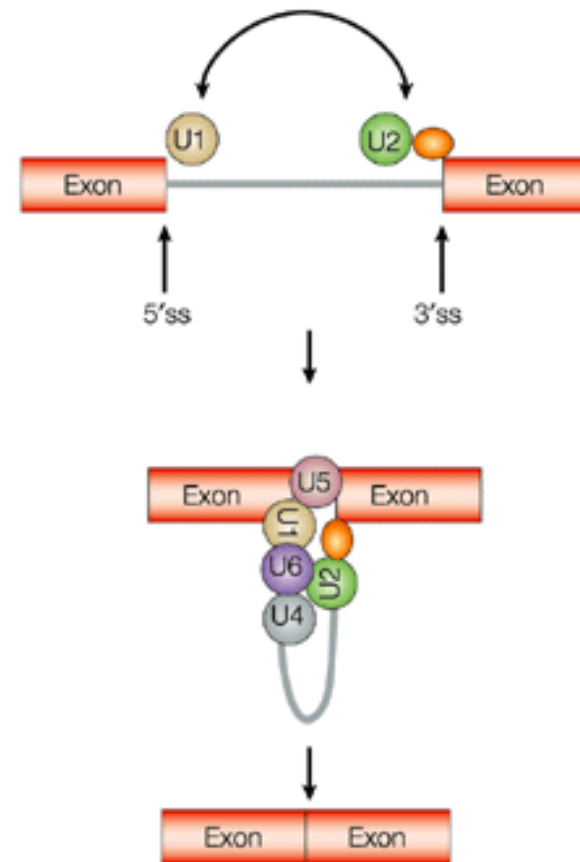
Intron > 250 nt

**a Exon definition**



Intron < 250 nt

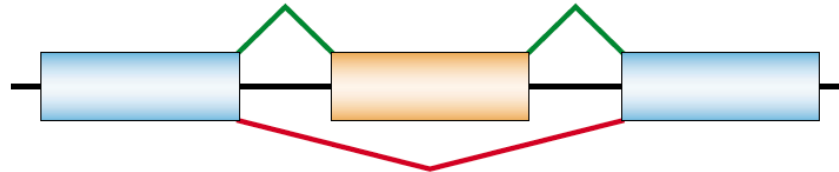
**b Intron definition**



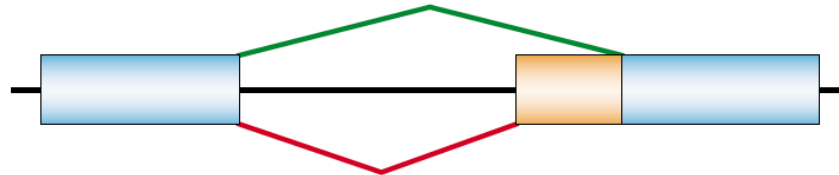
# Modes of alternative splicing

**b**

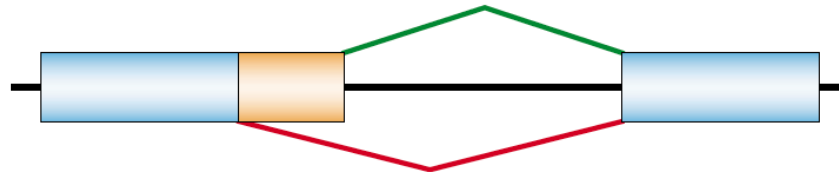
Exon skipping/inclusion



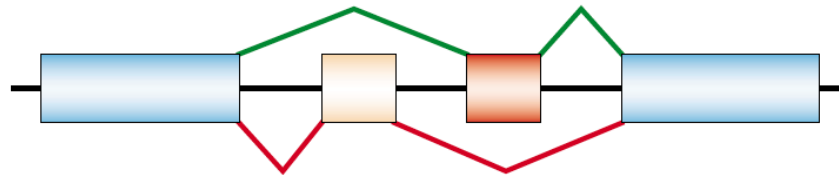
Alternative 3' splice sites



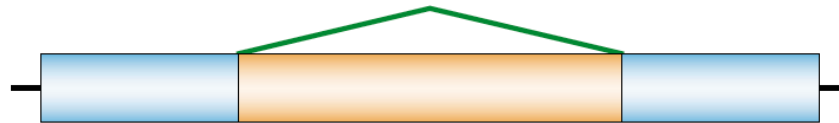
Alternative 5' splice sites



Mutually exclusive exons

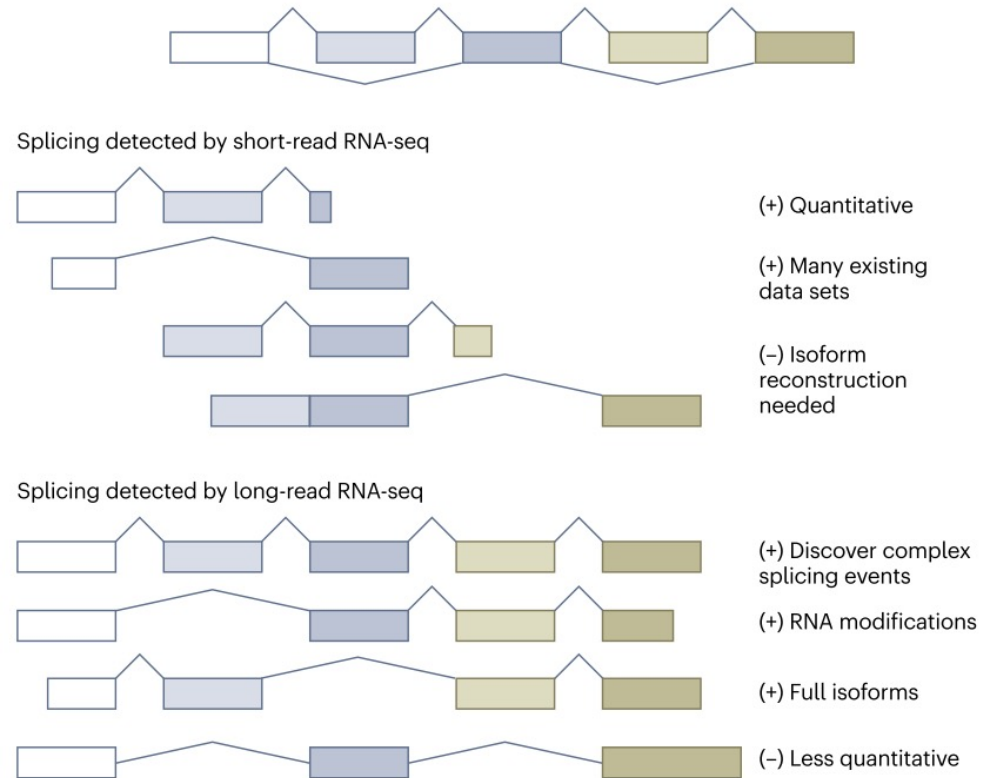


Intron retention

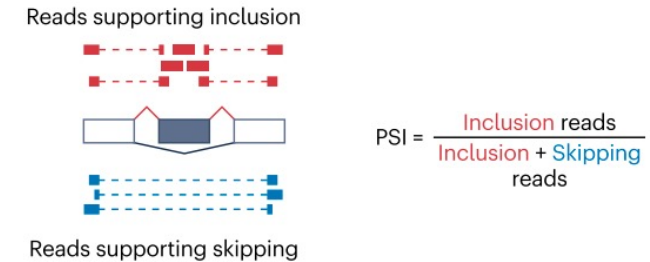


# How to detect and quantify differential splicing

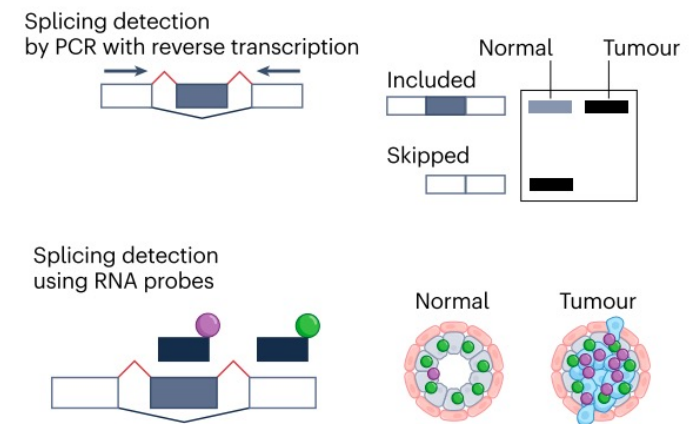
## a Transcriptome-wide detection of splicing



## b Percent spliced in (PSI)



## c Targeted detection of splicing



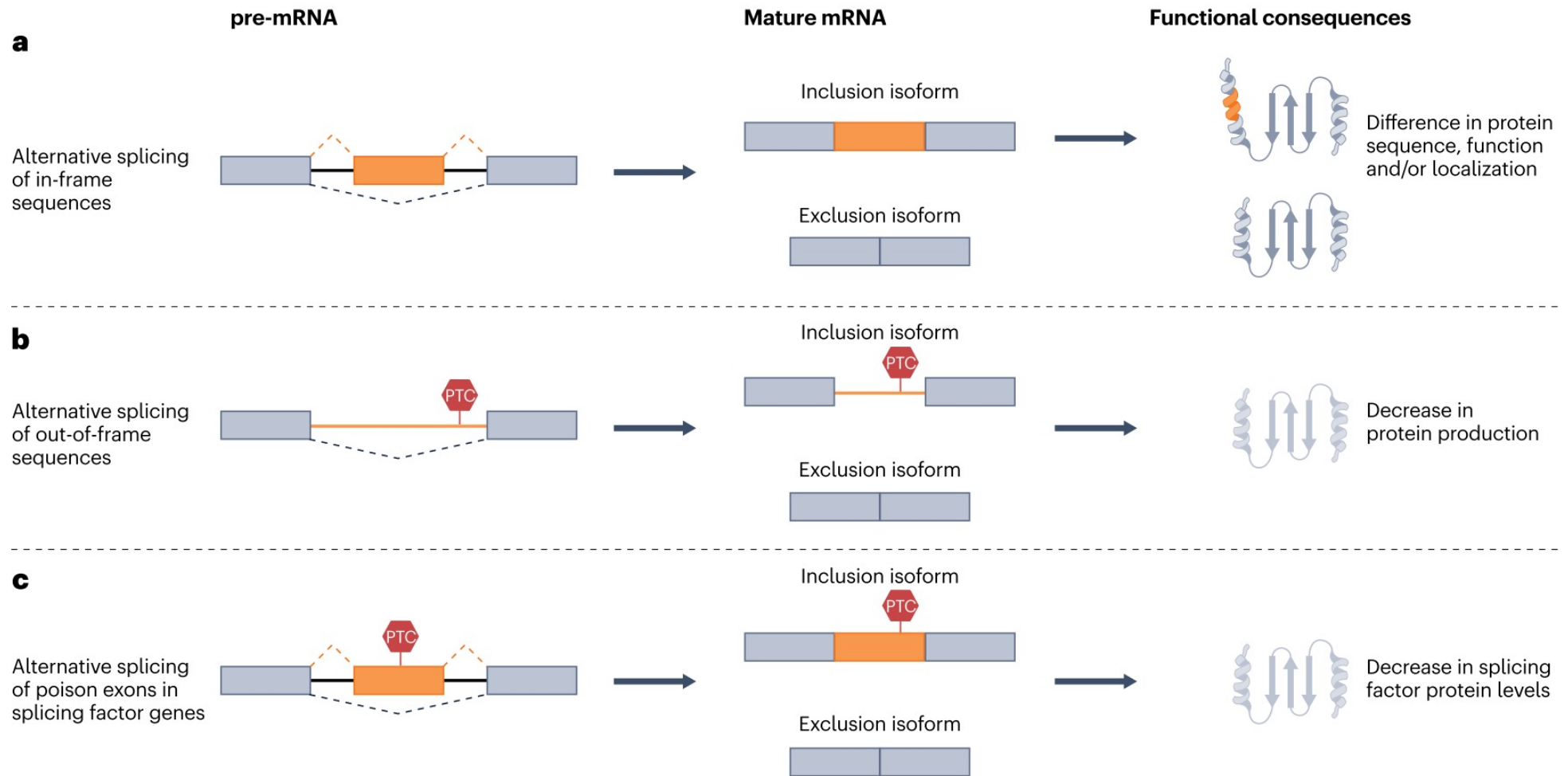
Nature Reviews Cancer

When individual splicing events are studied, alternative splicing is typically quantified using a 'percent spliced in' (PSI) or 'isoform fraction' value ranging from 0 to 100%, defined as expression of the isoforms that follow a splicing pattern of interest relative to the total expression of all transcripts of the gene

# The human transcriptome is composed of a vast RNA population that undergoes further diversification by splicing

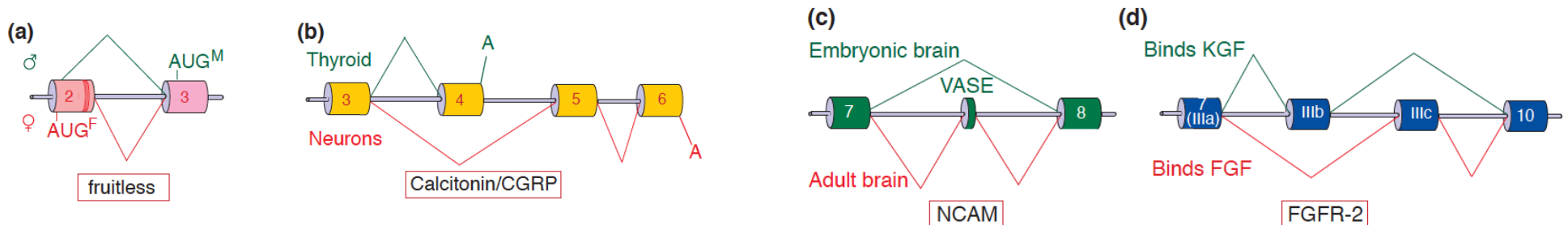
- **Alternative slicing** is an important mechanism for the production of different forms of proteins, called isoforms, by different types of cells.
- Alternative splicing can generate mRNAs that differ in their 3'-UTR or coding sequence.
- These differences might affect mRNA stability, localization or translation. Furthermore, some splicing mRNA isoforms could change the reading frame, resulting in the generation of different protein isoforms with diverse functions and/or localizations.
- 90%–95% of human genes undergo some level of alternative splicing. Out of the ~20,000 human protein coding genes, high resolution mass spectrometry analyses revealed that ~37% of them generate multiple protein isoforms, indicating that alternative splicing contributes to proteome complexity.

# Functional consequences



# Different modes of alternative splicing

- (a) Alternative splicing in the *Drosophila* gene *fruitless* governs sexual orientation and behaviour.
- (b) Alternative 3'-splice-site usage, associated with differential use of polyadenylation sites in the vertebrate gene for calcitonin and calcitonin-gene-related peptide (CGRP) generates a calcium homeostatic hormone in the thyroid gland or a vasodilator neuropeptide in the nervous system.
- (c) Differential inclusion or skipping of the variable alternatively spliced exon (VASE) in the gene for neural cell adhesion molecule (NCAM) in embryonic (green) versus adult (red) rat brain, represses or promotes axon outgrowth during development.
- (d) Mutually exclusive use of exons IIIb and IIIc in mammalian fibroblast growth factor receptor 2 (FGFR-2) changes its binding specificity for growth factors during prostate cancer progression. The pattern of splicing represented in green generates an mRNA encoding a receptor with high affinity for keratinocyte growth factor (KGF), whereas that in red generates a receptor with high affinity for fibroblast growth factor (FGF).



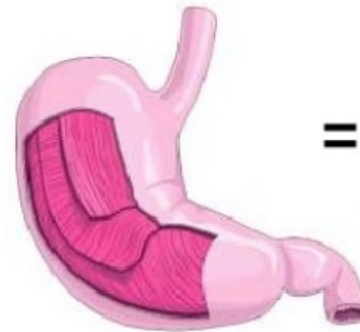
# Alternative splicing as a regulator of development and tissue identity

- It is now established that alternative splicing contributes to cell differentiation and lineage determination, tissue identity acquisition and maintenance, and organ development.

% of  
alternative  
spliced  
genes



= 48%



= 16%

- The biological importance of alternative splicing is highlighted by the large number of human diseases caused by mutations in *cis acting* sequence elements in premRNA (including 5' and 3' splice sites, and exonic and intronic enhancer or silencer sequences), *trans acting* splicing factors or other components of the spliceosome.

# Alternative splicing prediction

To predict splicing outcomes in specific cell types and conditions is very difficult, if not impossible, for several reasons:

- The same regulatory sequences on different or the same transcripts can be recognized by different RBPs depending on cellular and specific sequence contexts;
- RBPs can be regulated by other RBPs or mRNA modifications;
- the same RBP can exert positive or negative regulatory effects on different splicing events depending on the location of the binding motif for the RBP relative to the alternatively spliced regions in the transcript;
- we still lack a complete list of RBPs and their binding sites.

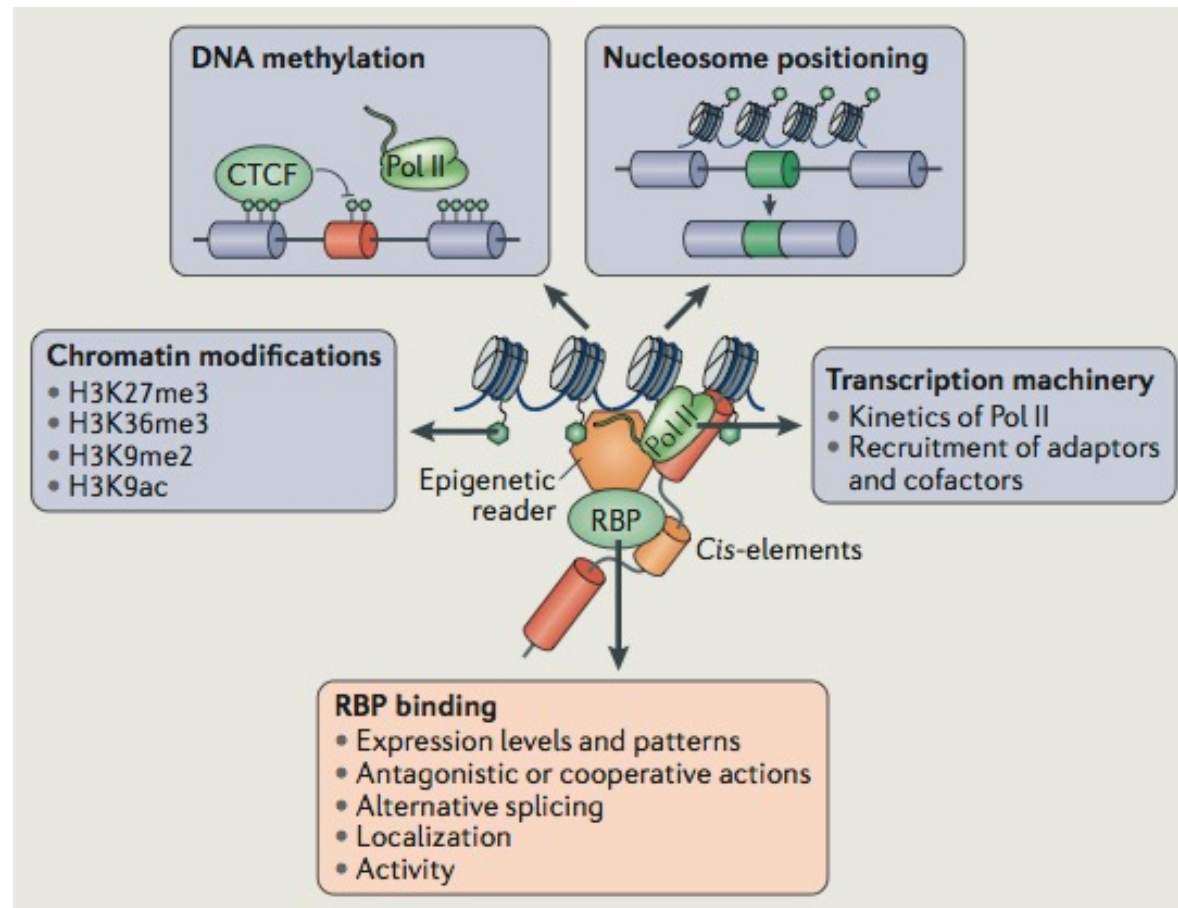
# Alternative splicing regulation

Alternative splicing outcomes are influenced by several factors:

1. **Transcription;**
2. **splice site strength;**
3. ***cis* regulatory sequences in pre-mRNAs that favour or impair exon recognition;**
4. **the expression levels of *trans acting* factors (RNA-binding proteins and splicing factors);**
5. **RNA modifications (editing, m<sup>6</sup>A)**

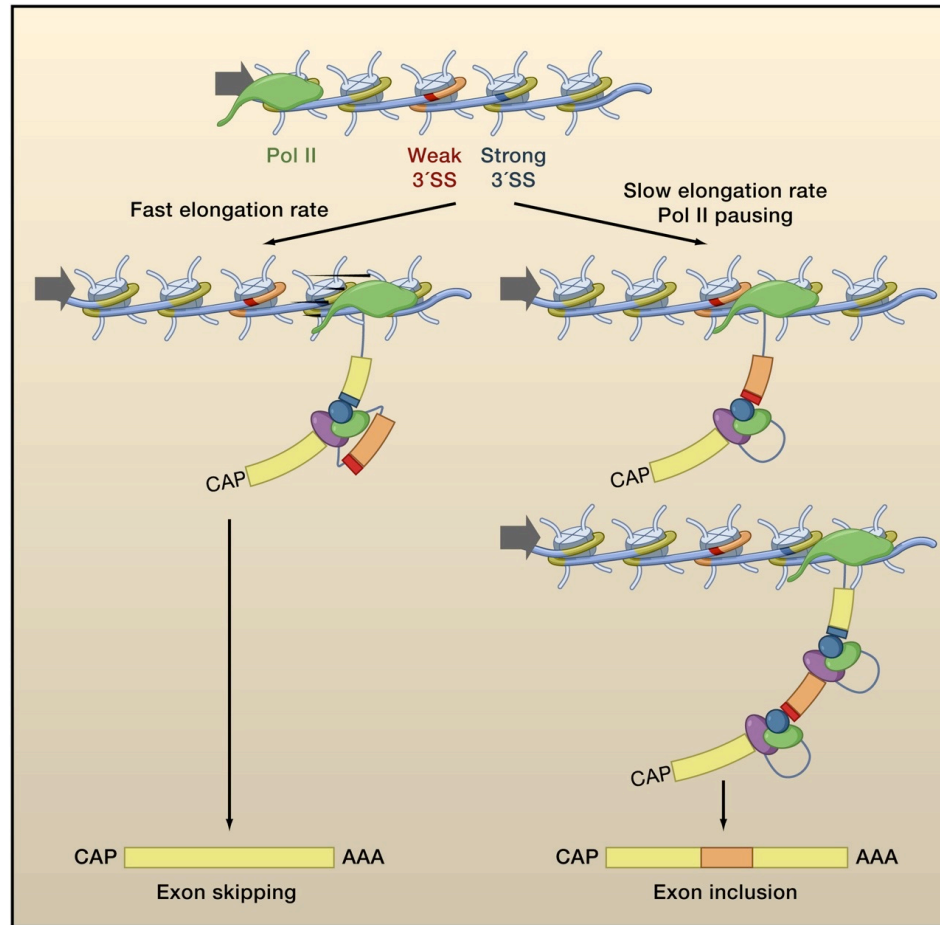
# Alternative splicing and transcription

Alternative splicing outcomes can be affected by RNA polymerase II (Pol II) kinetics, which can be controlled by chromatin modifications, chromatin structure and nucleosome occupancy.



# The rate of RNAPII elongation influences splicing

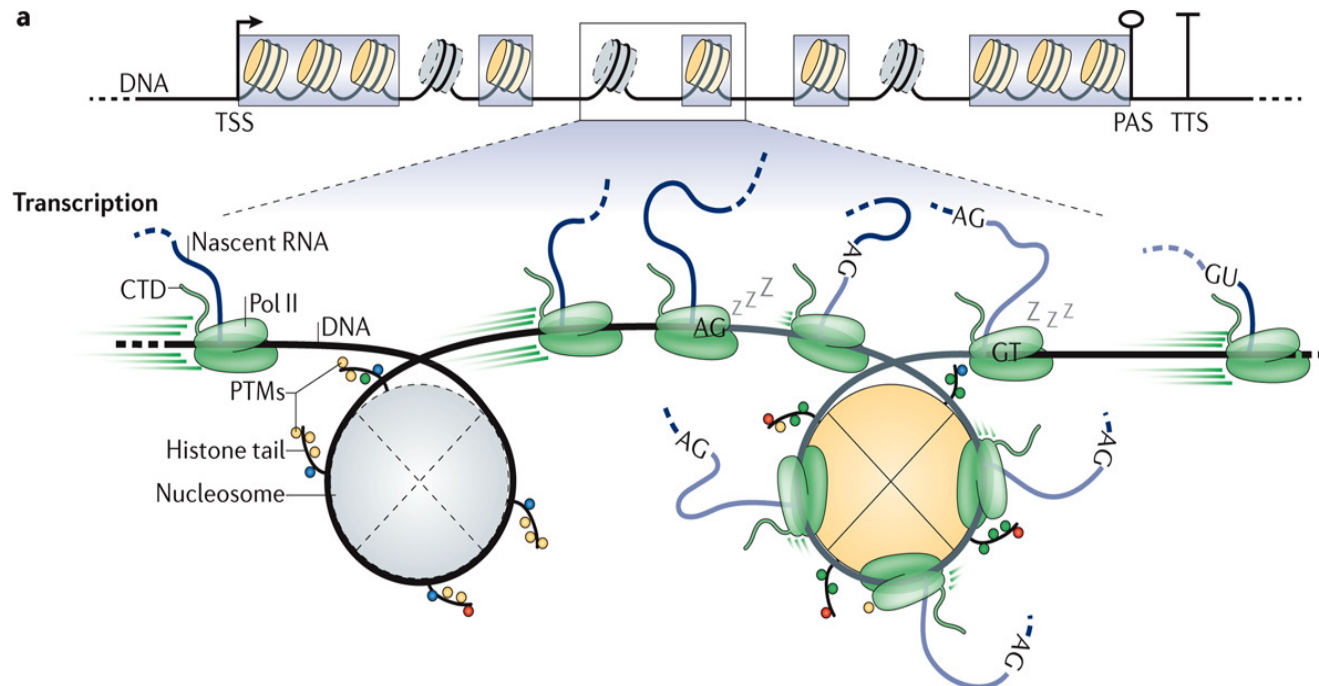
The rate of RNAPII elongation influences alternative splicing by affecting the step at which splice sites and regulatory sequences emerge in the nascent pre-mRNA during transcription. High elongation rates favor skipping, whereas low transcriptional elongation rates favor exon inclusion. Slow elongation causes preferential excision of the upstream intron (first served = first excised).



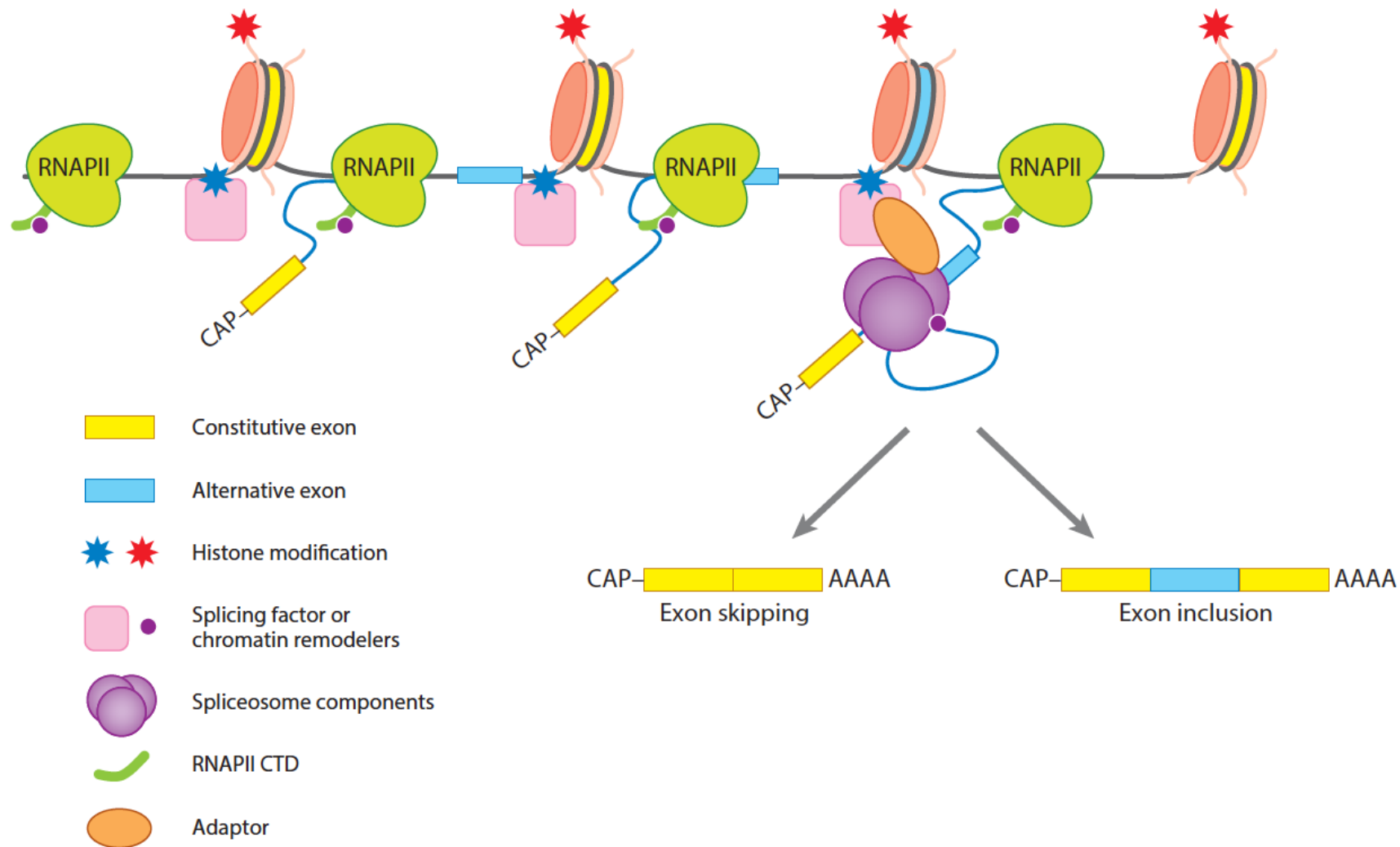
**Kinetic Coupling**

# Gene architecture, chromatin features and nascent RNA properties influence co-transcriptional splicing

Pol II elongation rate is influenced by a multitude of factors, such as the underlying DNA sequence, nucleosome position and histone modifications, which affect local chromatin structure, the activity of elongation factors, and the folding and processing of the nascent RNA.



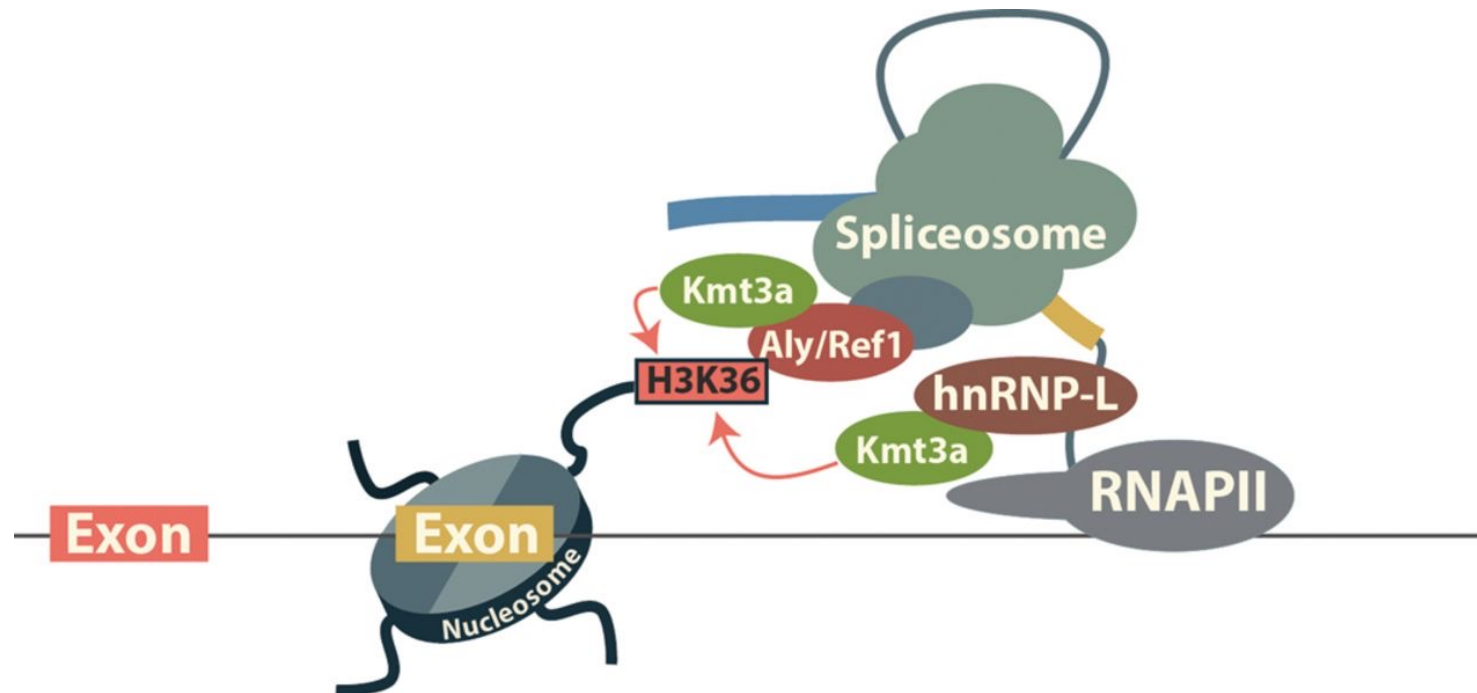
# When chromatin meets splicing: 'recruitment' model



H3K36me3 is strongly associated with exonic sequences and correlates with the level of gene expression and the rate of exon inclusion during alternative splicing. Splicing and H3K36me3 modification **can stimulate each other.**

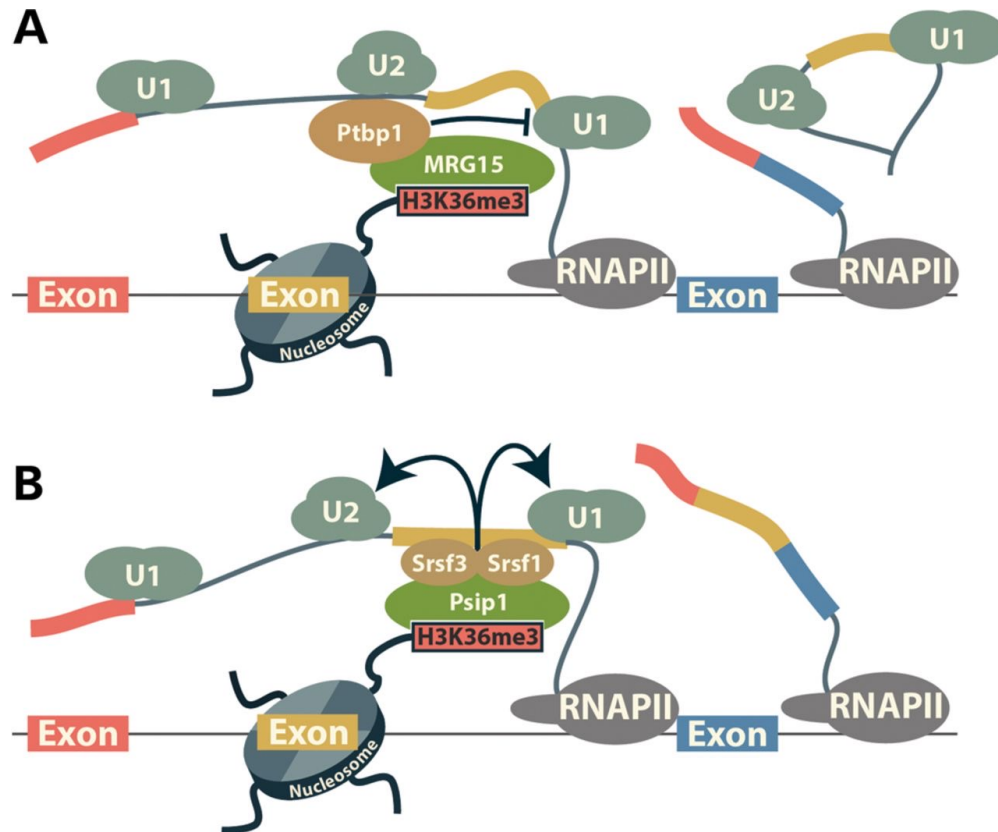
# Splicing stimulates H3K36me3 modification

There is a physical association between components of the splicing machinery and the **histone methyltransferases Kmt3**, which installs H3K36me3 modification .



In agreement with a mechanism that involves the recruitment of Kmt3a by the spliceosome, inhibition of splicing by splice site mutagenesis, pharmacological agents or depletion of SF3b3 (SAP130), a component of the core spliceosome, decrease Kmt3a recruitment to the chromatin and H3K36me3 levels.

# Recruitment model for the control of splicing by H3K36me3

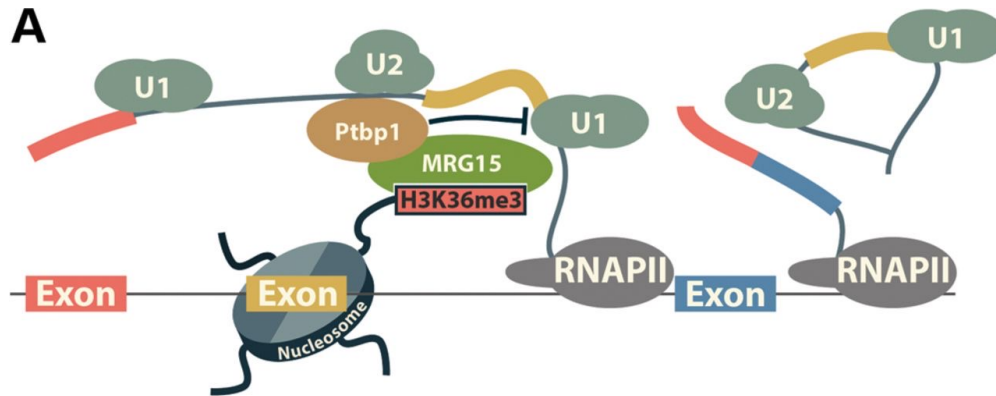


The MRG15 and Psip1 proteins 'read' the H3K36me3 mark by binding and can regulate alternative pre-mRNA splicing by recruiting splicing factors to the chromatin.

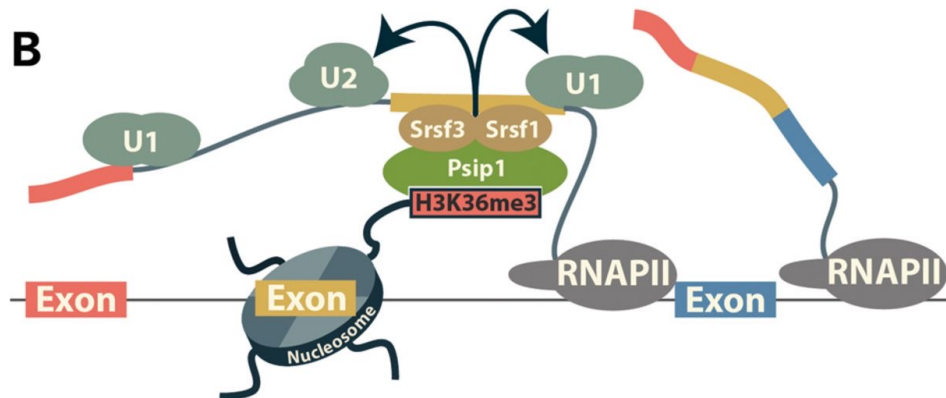
**MRG15** recognizes H3K36me3 and brings the **Ptbp1** splicing factor to the nascent transcripts .

**Psip1** interacts directly with H3K36me3 and associates with a large number of factors involved in pre-mRNA splicing that include SR and hnRNP proteins, DEAD-box helicases and core spliceosome components.

# Recruitment model for the control of splicing by H3K36me3

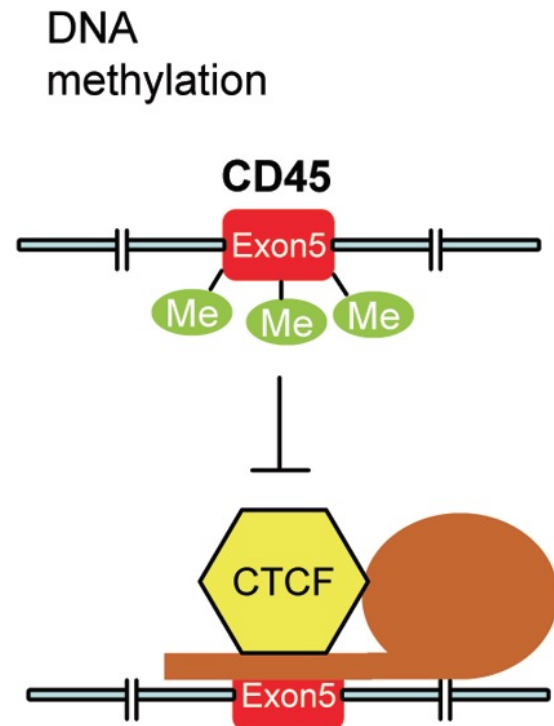


H3K36me3 can **repress** the inclusion of a subset of exons by recruiting Ptpb1 via its interaction with MRG15.



At the same time, the splicing of other exons is **facilitated** by the recruitment of Srsf1 and Srsf3 to H3K36me3 by a different reader, Psip1.

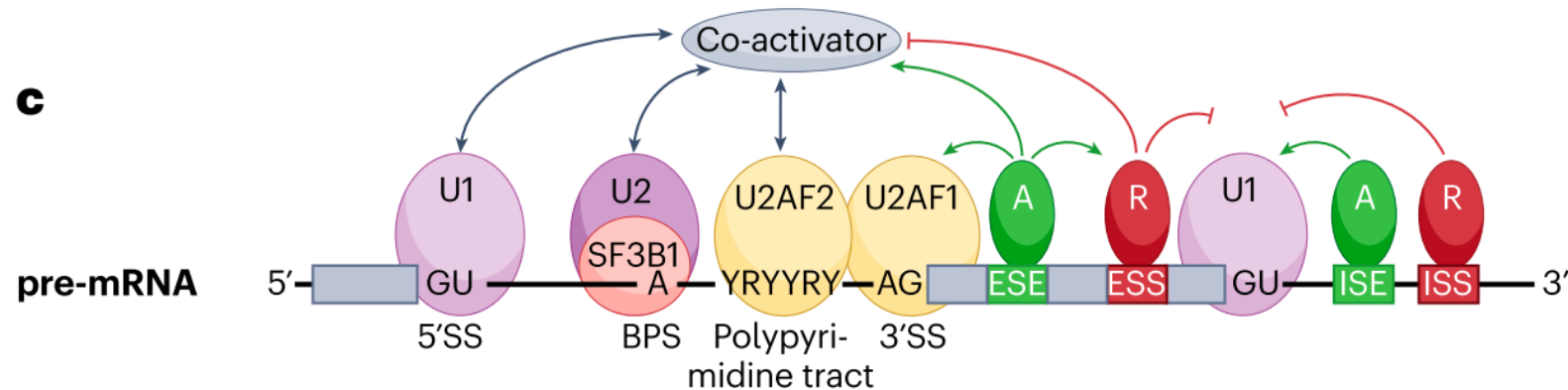
# DNA methylation and splicing



In the absence of CpG methylation CTCF binds to its binding site and facilitates the inclusion of CD45 exon 5 by inducing RNA pol II pausing at this exon.

This result provides evidence that DNA methylation influences splicing. Therefore, it is conceivable that changes in DNA methylation patterns during development or in disease, particularly in cancer, could determine alternative splicing outcome and subsequently affect transcriptome makeup.

# RNA splicing regulation



## ***Cis*-acting regulatory elements:**

ESE Exonic Splicing Enhancer

ESS Exonic Splicing Silencer

ISS Intronic Splicing Enhancer

ISE Intronic Splicing Silencer

## ***Trans*-acting factors** trans-Acting regulatory splicing factors act as splicing activator (A) or repressor (R):

SR proteins Serine arginine rich proteins

hnRNPs heterogeneous nuclear RiboNucleoprotein Particles (hnRNPs)

# RNA-binding proteins regulating alternative splicing

| RBP                            | Context where RBP regulates alternative splicing                                                                                                                                                                                                                         | Binding motif from CLIP-sequencing experiments / alternative splicing effect                                       |
|--------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------|
| <b>CELF1</b>                   | - Heart development <sup>1-3</sup><br>- Skeletal muscle development <sup>3</sup><br>- Myoblast differentiation (C2C12 cells) <sup>4</sup>                                                                                                                                | UGUU motifs / skipping or inclusion (position dependent) <sup>3,5,6</sup>                                          |
| <b>CELF2</b>                   | - T-cell activation <sup>7-9</sup><br>- Heart development <sup>3</sup>                                                                                                                                                                                                   | UGU-rich motifs <sup>8</sup>                                                                                       |
| <b>ELAVL</b>                   | - Brain development <sup>10</sup>                                                                                                                                                                                                                                        | U- and AU-rich motifs <sup>10-12</sup>                                                                             |
| <b>ESRP1</b><br><b>ESRP2</b>   | - Liver development (ESRP2) <sup>13</sup><br>- Epithelial-mesenchymal transition (ESRP1/2) <sup>14</sup><br>- Stomach smooth muscle development (ESRP1) <sup>15</sup><br>- Epidermis development (ESRP1/2) <sup>16</sup><br>- Kidney development (ESRP1/2) <sup>17</sup> | GU-rich motifs / skipping or inclusion (position dependent) <sup>18</sup>                                          |
| <b>HNRNPL</b>                  | - T-cell development <sup>19-21</sup>                                                                                                                                                                                                                                    | CA-rich motifs / skipping or inclusion <sup>20,22</sup>                                                            |
| <b>HNRNPLL</b>                 | - T-cell activation <sup>23</sup><br>- B-cell into plasma cell differentiation <sup>24</sup>                                                                                                                                                                             | CA-rich motifs / skipping or inclusion <sup>24</sup>                                                               |
| <b>MBNL1</b><br><b>MBNL2</b>   | - Heart development (MBNL1) <sup>1-3</sup><br>- Brain development (MBNL2) <sup>25</sup><br>- Erythropoiesis (MBNL1) <sup>26</sup>                                                                                                                                        | YGCY (preferred: UGCU) / skipping or inclusion (position dependent) <sup>25,27</sup>                               |
| <b>NOVA1</b><br><b>NOVA2</b>   | - Brain development (NOVA1/2) <sup>28,29</sup><br>- Vascular development (endothelial cells) (NOVA2) <sup>30</sup>                                                                                                                                                       | YCAY motifs / skipping or inclusion (position dependent) <sup>31</sup>                                             |
| <b>PTBP1</b><br><b>PTBP2</b>   | - Brain development <sup>32-34</sup><br>- Male germ cell development (PTBP2) <sup>35</sup><br>- Myoblast differentiation (C2C12 cells) <sup>36</sup><br>- Primary smooth muscle cell (aorta, bladder) differentiation (PTBP1) <sup>37</sup>                              | CU-rich motif / skipping <sup>33,38</sup>                                                                          |
| <b>QK</b>                      | - Myoblast differentiation (C2C12 cells) <sup>36</sup><br>- Vascular smooth muscle development and cell de-differentiation <sup>37,39</sup>                                                                                                                              | UAA-rich motifs / skipping or inclusion (position dependent) <sup>40</sup>                                         |
| <b>RBFOX1</b><br><b>RBFOX2</b> | - Brain development <sup>34,41-44</sup><br>- Heart development <sup>45,46</sup><br>- Skeletal muscle development <sup>46,47</sup><br>- Myoblast differentiation (C2C12 cells) <sup>48,49</sup>                                                                           | UGCAUG / inclusion (downstream binding) or skipping (upstream binding or within alternative exon) <sup>50,51</sup> |
| <b>RBM4</b>                    | - Pancreas development and pancreatic cell differentiation (AR42J cells) <sup>52</sup><br>- Neuronal differentiation (P19 cells) <sup>53</sup>                                                                                                                           | CGG or GTAACG / skipping or inclusion (position dependent) <sup>54,55</sup>                                        |
| <b>RBM20</b>                   | - Heart development <sup>56</sup>                                                                                                                                                                                                                                        | UCUU / skipping <sup>56</sup>                                                                                      |
| <b>RBM24</b>                   | - Heart and skeletal muscle development <sup>57</sup>                                                                                                                                                                                                                    | G(A/U)GUG <sup>58</sup>                                                                                            |
| <b>SAM68</b>                   | - Spermatogenesis <sup>59</sup><br>- T-cell activation <sup>60</sup><br>- Brain development <sup>61,62</sup><br>- Adipogenesis <sup>63</sup>                                                                                                                             | AU-rich motifs                                                                                                     |
| <b>SRRM4</b>                   | - Brain development and synaptogenesis <sup>64,65</sup>                                                                                                                                                                                                                  | UGC-rich motifs between polypyrimidine tract and 3'ss / inclusion <sup>66</sup>                                    |
| <b>SRSF1</b>                   | - Heart development <sup>67</sup>                                                                                                                                                                                                                                        | GGAGGA / inclusion <sup>68</sup>                                                                                   |
| <b>SRSF10</b>                  | - Heart development <sup>69,70</sup><br>- Skeletal muscle development, myoblast differentiation (C2C12 cells) <sup>70</sup><br>- Adipogenic differentiation <sup>71</sup>                                                                                                | GA-rich motifs / skipping or inclusion (position dependent) <sup>69,72</sup>                                       |

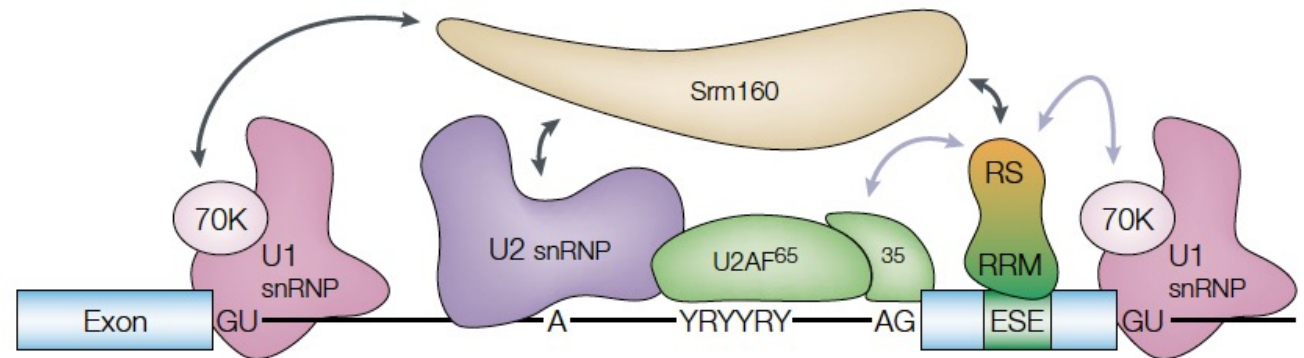
# SR protein action in ESE-dependent splicing

**Exonic enhancers** serve as binding sites for specific **serine/arginine-rich (SR) proteins**

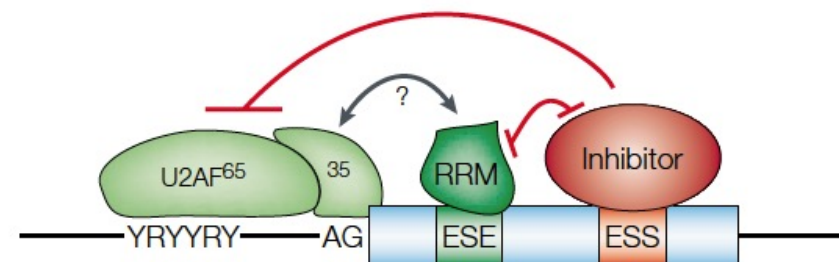
SR protein family is composed by a structurally related and highly conserved splicing factors that are characterized by the presence of RNA recognition motifs (RRM) and by a distinctive carboxy terminal domain that is highly enriched in Arg/Ser dipeptides (the RS domain).

SR proteins that are bound to ESEs can promote **EXON DEFINITION** by directly recruiting the splicing machinery through their RS domain and/or by antagonizing the action of nearby silencer elements

**a Recruiting function: RS-domain dependent**



**b Antagonist function: RS-domain independent**



# RNA motifs recognized by human SR proteins

| Protein | High-affinity binding site                                                                  | Ref.                                        | Functional ESE          |
|---------|---------------------------------------------------------------------------------------------|---------------------------------------------|-------------------------|
| SRp20   | WCWWC<br>CUCKUCY                                                                            | 112<br>14                                   | GCUCCUCUUCC<br>CCUCGUCC |
| SC35    | AGSAGAGUA<br>GUUCGAGUA<br>UGUUCSAGWU<br>GWUWCCUGCUA<br>GGGUAUGCUG<br>GAGCAGUAGKS<br>AGGAGAU | 32<br>32<br>112<br>112<br>112<br>112<br>112 | GRYYMCYR*<br>UGCYGY     |
| 9G8     | (GAC) <sub>n</sub><br>ACGAGAGAY<br>WGGACRA                                                  | 112<br>112<br>14                            |                         |
| SF2/ASF | RGAAGAAC<br>AGGACRRAGC                                                                      | 32<br>32                                    | CRSMSGW*                |
| SRp40   | UGGGAGCRGU YRGCUCGY                                                                         | 114                                         | YRCRKM*                 |
| SRp55   |                                                                                             |                                             | YYWCWSG*                |
| TRA2β   | (GAA) <sub>n</sub>                                                                          | 115                                         |                         |

# The SR proteins

## A. Human SR Proteins

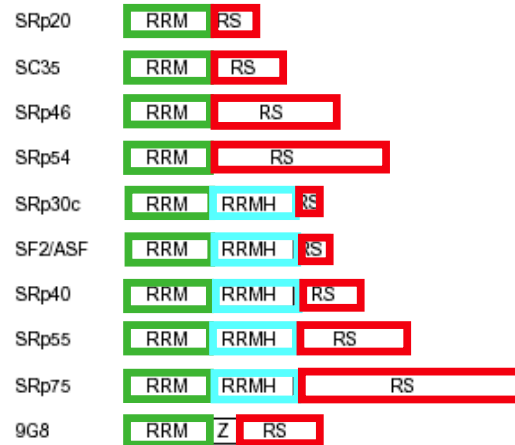
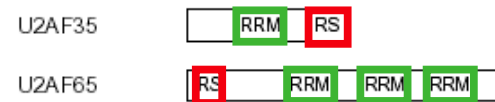


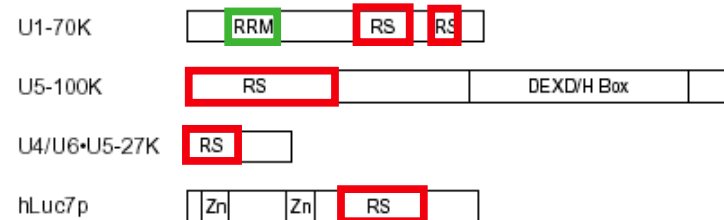
**FIGURE 2.** Schematic diagram of human SR proteins and SR related proteins. **A:** The domain structures of the known members of the human SR protein family are depicted. RRM: RNA recognition motif; RRMH: RRM homology; Z: zinc knuckle, RS: arginine/serine-rich domain. **B:** The domain structures for some of the human SRps that participate in pre-mRNA splicing are depicted. All proteins, with the exception of SRm300, are drawn to scale. RRM: RNA recognition motif; RS: arginine/serine-rich domain; Zn: zinc finger; DEXD/H Box: motif characteristic of RNA helicases.

## B. Human SR Related Proteins

### U2 Auxiliary Factor



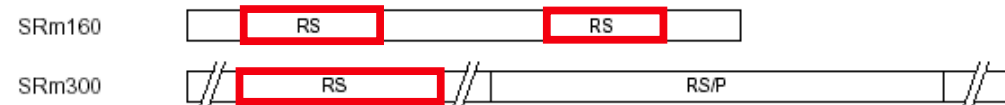
### snRNP Components



### Splicing Regulators



### Splicing Coactivators



### RNA Helicases

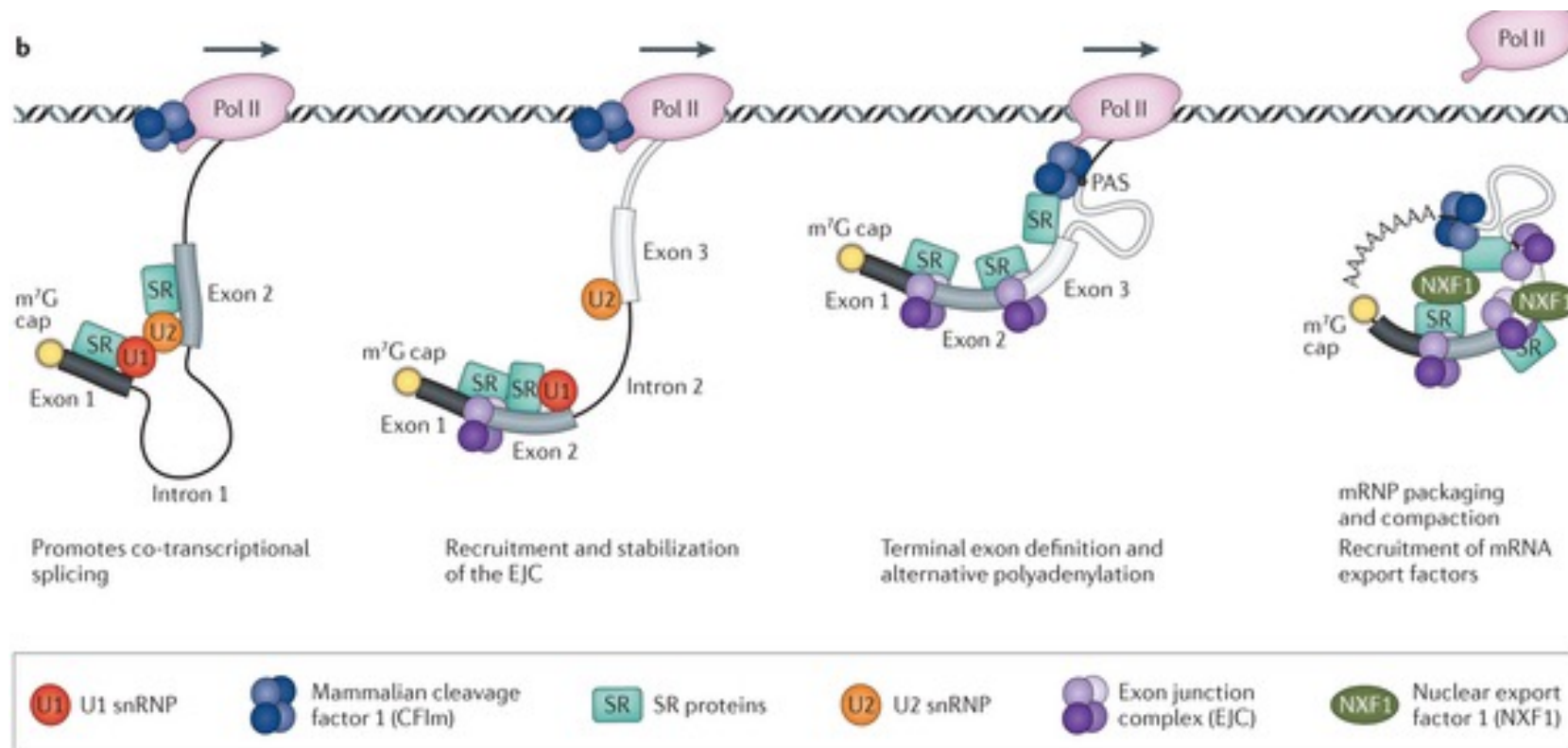


### Protein Kinases



# SR proteins couple co-transcriptional splicing to mRNA export

SR proteins bind to specific binding sites within nascent transcripts and aid in the recruitment of the U1 and U2 snRNPs, thereby promoting co-transcriptional splicing. SR proteins help in the packaging and compaction of mRNPs through tight interactions with EJC components and other SR proteins bound to the same transcript. Ultimately, they recruit nuclear RNA export factor 1 (NXF1) to the compact, mature mRNP to allow efficient nuclear export.



# Splicing silencers

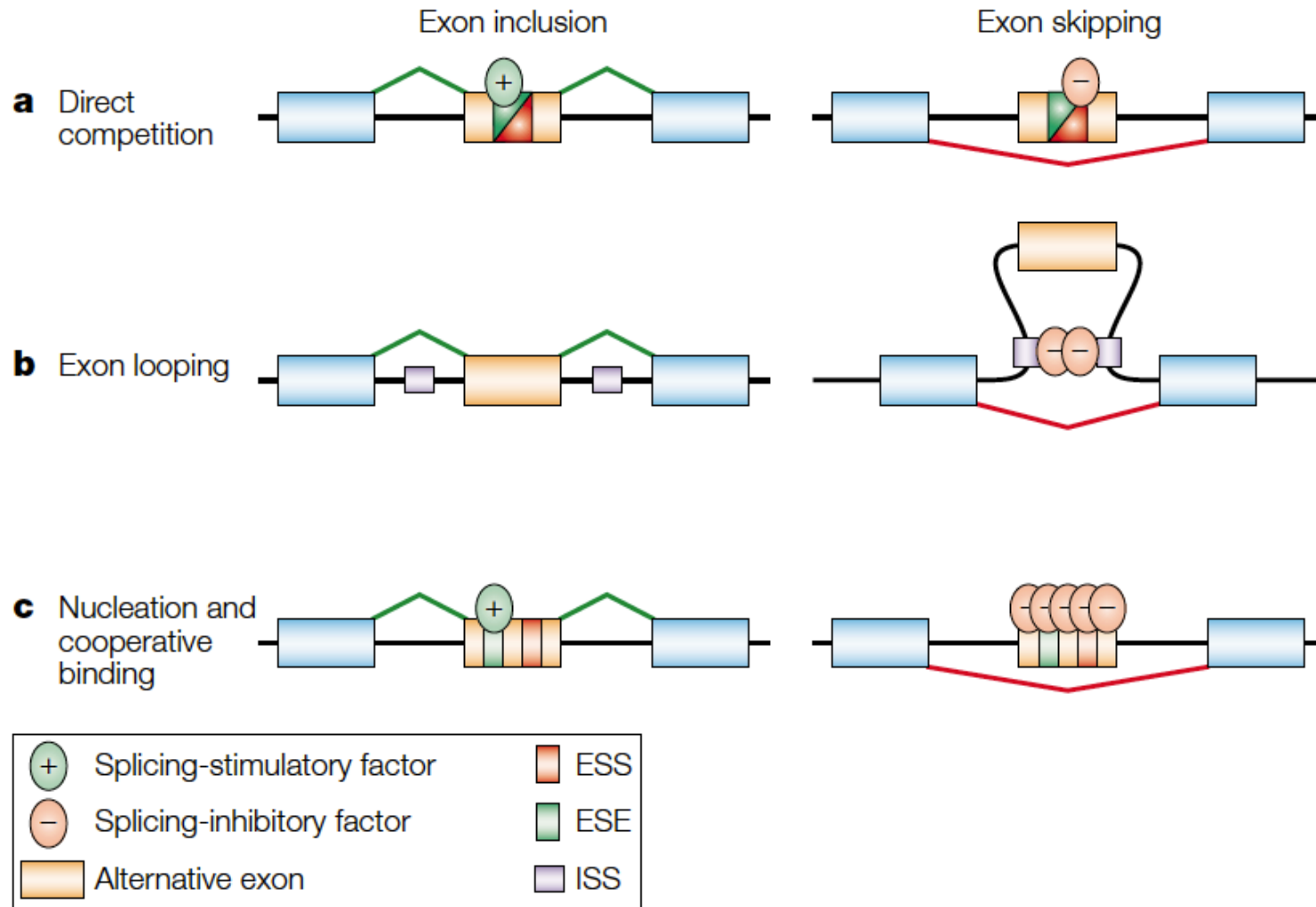
Splicing silencers are less well characterized than splicing enhancer.

Most described silencers are intronic elements, but several ESS elements have also been reported

Their mechanisms of action are still not fully understood. Silencers seem to work by interacting with negative regulators, which often belong to the **heterogeneous nuclear ribonucleoprotein (hnRNP)** family

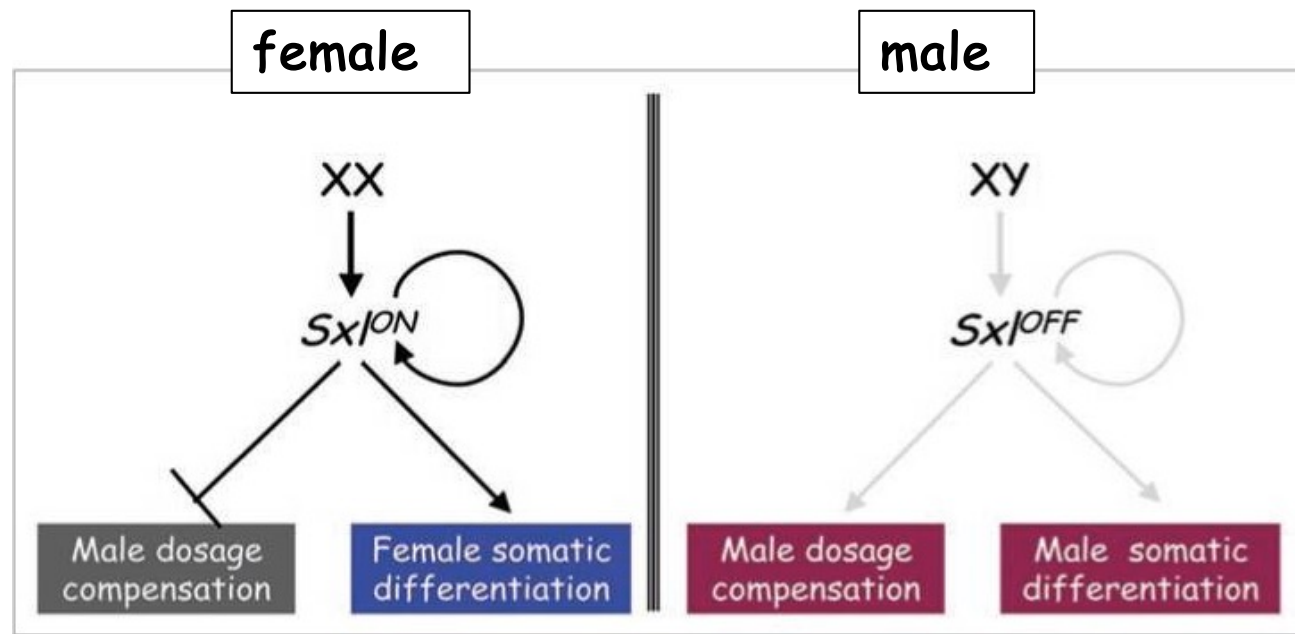
Similar to SR proteins, hnRNP proteins have a modular structure, which consists of one or more RNA-binding domains associated with an auxiliary domain that is often involved in protein–protein interactions.

# Models of splicing silencing

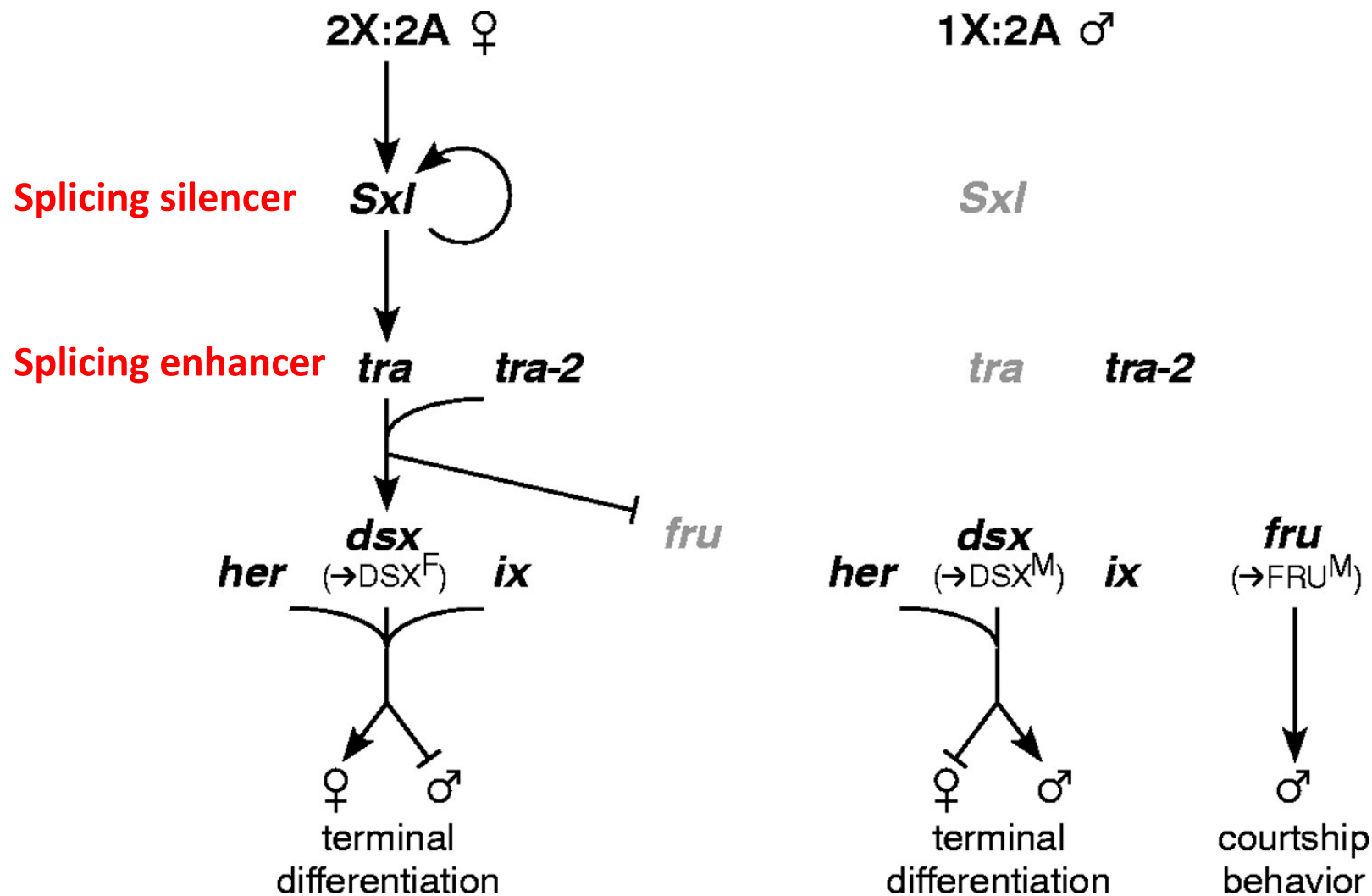


# Sex determination in *Drosophila*

In *D. melanogaster* males, dosage compensation is orchestrated by *sex lethal* (SXL) protein



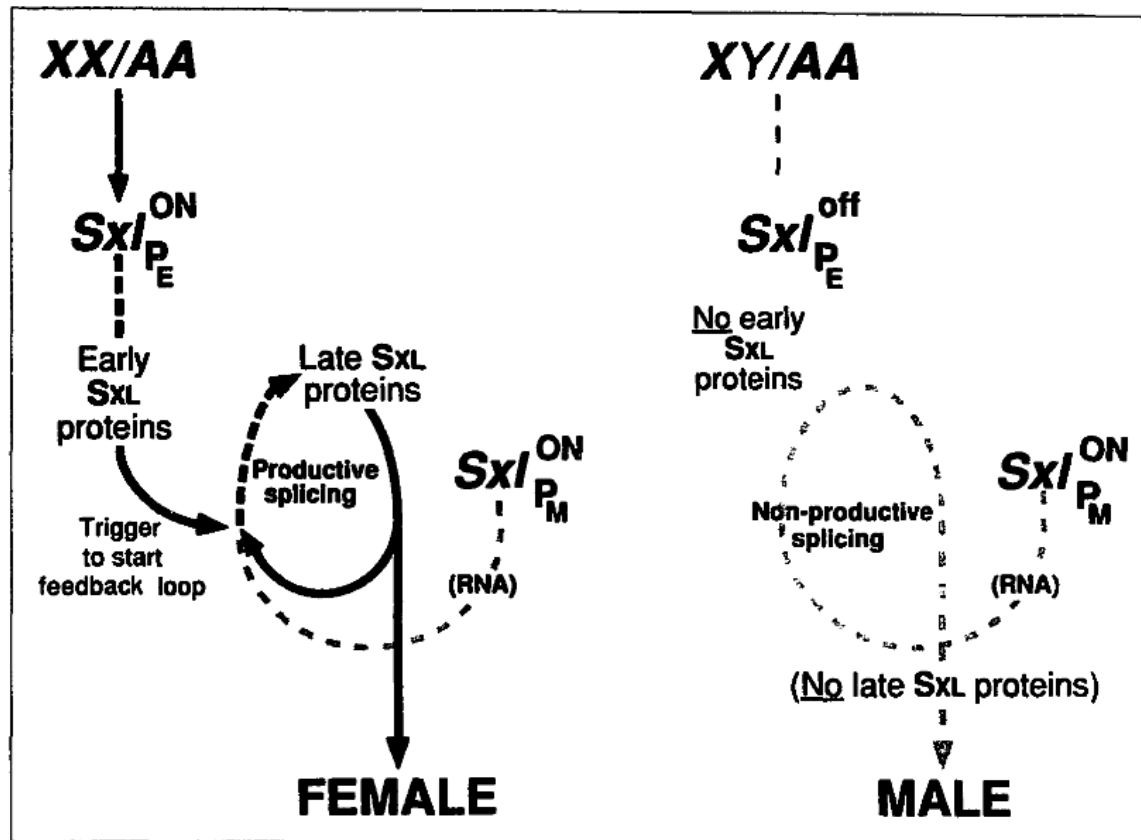
# Sex determination in *Drosophila*



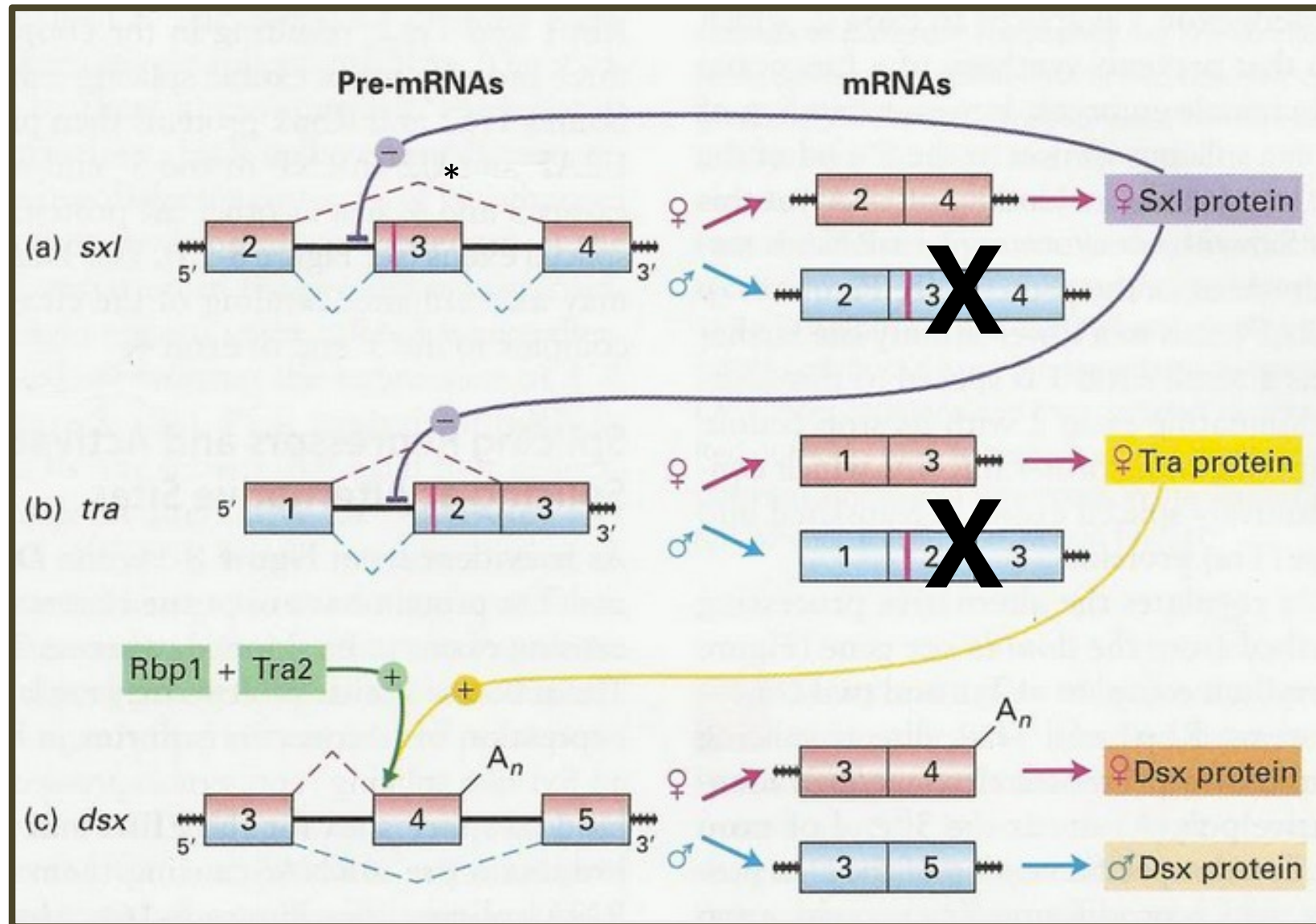
DSX-F and DSX-M are transcription factors that determine somatic sexual characterization

# Sex determination in *Drosophila*

Promoter Early ( $P_E$ ) is active only in female while Promoter maintaince ( $P_M$ ) in both sexes but Sxl mRNA maturation required Sxl protein for correct splicing



# Sex determination in *Drosophila*



\*m6A modification is required for female-specific alternative splicing of Sxl

# The pre mRNA splicing pathology

- A considerable number of disease-causing mutations in exons or introns may disrupt previously unrecognised splicing regulatory elements
- Variability in the basal splicing machinery among different cell types cause cell-specific sensitivities to individual splicing mutations
- Exon sequence variation at CERES elements (Composite Exonic Regulatory Elements of Splicing, which represents a physical overlap of enhancer/silencer activity) may represent a frequent disease-causing mechanism
- Even the most benign looking polymorphism in an exon (or in an intron) cannot be ignored as it may affect the splicing process

# Pathologies resulting from aberrant splicing can be grouped in three major categories

- Mutations affecting proteins that are involved in splicing

Examples:

Spinal Muscular Atrophy

Cancer

Retinitis Pigmentosa

Myotonic Dystrophy

- Mutations affecting a specific messenger RNA and disturbing its normal splicing pattern

Examples:

Duchenne Muscular Dystrophy

$\beta$ -Thalassemia

Cystic Fibrosis

Frasier Syndrome

Frontotemporal Dementia and Parkinsonism

- Single point mutations in coding exons

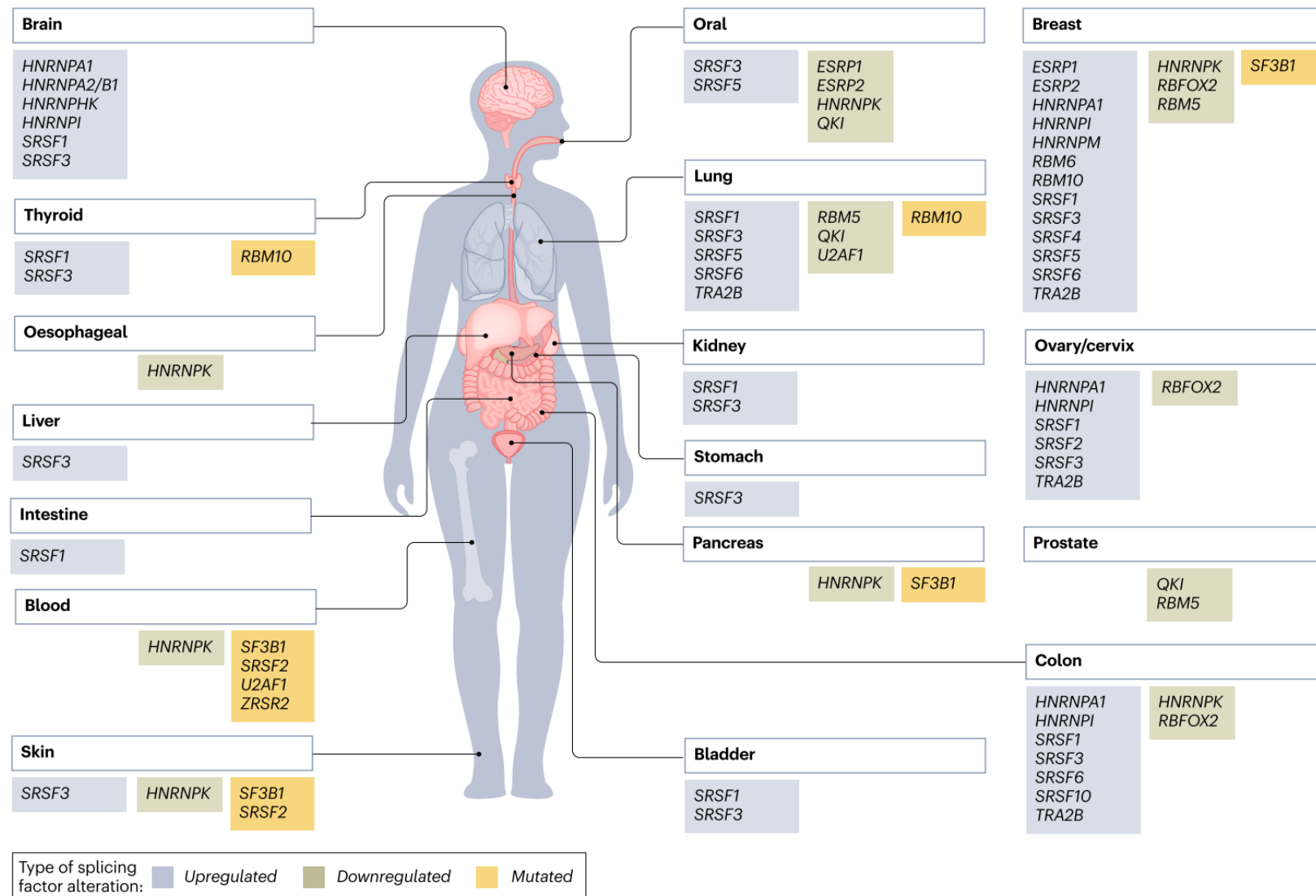
|                    | Probability of<br>effect on<br>encoded<br>protein | Probability of<br>nonsense-<br>mediated<br>decay | Probability of<br>effects on<br>splicing |
|--------------------|---------------------------------------------------|--------------------------------------------------|------------------------------------------|
| Nonsense mutations | +                                                 | +                                                | +/-                                      |
| Missense mutations | +/-                                               | -                                                | +/-                                      |
| Silent mutations   | -                                                 | -                                                | ++                                       |

Table 1 | Disease-associated splicing alterations

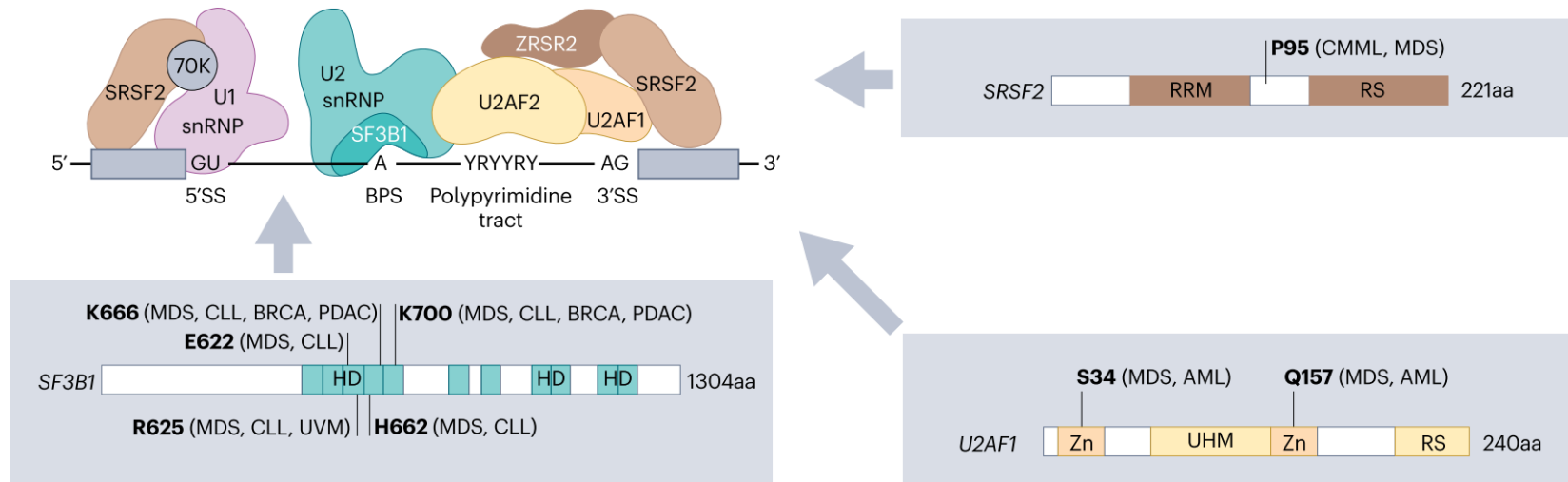
| Disease                                                           | Gene (mutation)                                                                                               | Mechanism                                                      | Splicing effect                                                          | Inheritance           |
|-------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------|--------------------------------------------------------------------------|-----------------------|
| <b>Cis</b>                                                        |                                                                                                               |                                                                |                                                                          |                       |
| Limb girdle muscular dystrophy type 1B (LGMD1B)                   | <i>LMNA</i> <sup>24</sup> (c.1608 + 5G>C)                                                                     | 5'ss mutation                                                  | Intron 9 retention resulting in NMD                                      | Dominant              |
| Familial partial lipodystrophy type 2 (FPLD2)                     | <i>LMNA</i> <sup>25</sup> (c.1488 + 5G>C)                                                                     | 5'ss mutation                                                  | Intron 8 retention resulting in NMD                                      | Dominant              |
| Hutchinson–Gilford progeria syndrome (HGPS)                       | <i>LMNA</i> <sup>26</sup> (c.1824C>T)                                                                         | Alternative 5'ss                                               | 150 nt deletion in exon 11, resulting in progerin generation             | Dominant              |
| Dilated cardiomyopathy (DCM)                                      | <i>LMNA</i> <sup>28</sup> (c.640-10A>G)                                                                       | Alternative 3'ss                                               | Extension of exon 4 adding 3 amino acids to lamin A/C                    | Dominant              |
| Familial dysautonomia (FD)                                        | <i>IKBKAP</i> <sup>128</sup> (c.2204 + 6T>C)                                                                  | Decreased U1 recruitment                                       | Exon 20 skipping                                                         | Recessive             |
| Duchenne muscular dystrophy (DMD)                                 | <i>DMD</i> <sup>129</sup> Exon 45–55 deletions are common                                                     | Exon deletions and skipping                                    | Frameshift resulting in NMD                                              | X-linked              |
| Becker muscular dystrophy (BMD)                                   | <i>DMD</i> <sup>130</sup> (c.4250T>A)                                                                         | ESS creation                                                   | Exon 31 partial in-frame skipping                                        | X-linked              |
| Early-onset Parkinson disease (PD)                                | <i>PINK1</i> (REF. 131) (c.1488 + 1G>A)                                                                       | U1 5'ss mutation                                               | Cryptic splice site usage, resulting in exon 7 skipping                  | Recessive             |
| Frontotemporal dementia with parkinsonism chromosome 17 (FTDP-17) | <i>MAPT</i> <sup>132</sup> (c.892A>G)                                                                         | ESS mutation                                                   | Increased exon 10 inclusion                                              | Dominant              |
| X-linked parkinsonism with spasticity (XPDS)                      | <i>ATP6AP2</i> (REF. 133) (c.345C>T)                                                                          | Novel ESS creation                                             | Increased exon 4 exclusion                                               | X-linked              |
| <b>Spliceosome</b>                                                |                                                                                                               |                                                                |                                                                          |                       |
| Retinitis pigmentosa (adRP)                                       | <i>PRPF6</i> (REF. 134) (c.2185C>T)                                                                           | Abnormal nuclear localization                                  | Decreased U4/U6 interaction affecting spliceosome assembly and recycling | Dominant              |
|                                                                   | <i>SNRNP200</i> (REF. 135) (c.3260C>T), (c.3269G>T)                                                           | • Decreased helicase activity<br>• Decreased proof-reading     | Compromised splice site recognition, leading to mis-spliced mRNAs        | Dominant              |
| Myelodysplastic syndromes (MDS)                                   | <i>U2AF1</i> (REF. 46) (c.101G>A)                                                                             | Altered 3'ss preference                                        | Increased alternative 3'ss usage                                         | Somatic               |
| Microcephalic osteodysplastic primordial dwarfism type 1 (MOPD I) | <i>RNU4ATAC</i> <sup>54–56</sup> (g.30G>A), (g.50G>A), (g.50G>C), (g.51G>A), (g.53C>G), (g.55G>A), (g.111G>A) | 5' and 3' stem loop mutations & secondary structure disruption | Compromised minor spliceosome activity                                   | Recessive             |
| <b>Trans</b>                                                      |                                                                                                               |                                                                |                                                                          |                       |
| Spinal muscular atrophy (SMA)                                     | <i>SMN1</i> (REFS 136,137) (c.922 + 6 T/G), deletion                                                          | Loss of SMN full-length protein                                | Altered RNP biogenesis <sup>98</sup>                                     | Recessive             |
| Amyotrophic lateral sclerosis (ALS)                               | <i>TARDP</i> <sup>77</sup> (c.991C>A), (c.1009A>G)                                                            | C-terminal mutations alter protein-protein interactions        | TDP-43 target mis-splicing                                               | Sporadic and Dominant |
|                                                                   | <i>FUS</i> <sup>138</sup> (c. 1566C>T), (c. 1561T>G)                                                          | • Decreased U1 interaction<br>• Increased SMN binding          | FUS target mis-splicing                                                  | Dominant              |
| Dilated cardiomyopathy (DCM)                                      | <i>RBM20</i> (REF. 139) (c.1962T>G)                                                                           | Altered R/S RNA binding domain                                 | <i>TTN</i> mis-splicing                                                  | Dominant              |
| Limb-girdle muscular dystrophy 1G (LGMD1G)                        | <i>HNRPDL</i> <sup>140</sup> (c. 1667G>A), (c. 1667G>C)                                                       | Altered import of HNRPDL into nucleus                          | HNRPDL target mis-splicing                                               | Dominant              |
| Autosomal dominant leukodystrophy (ADLD)                          | <i>LMNB1</i> (REF. 141) duplication                                                                           | Increased <i>RAVER2</i> expression                             | PTBP1 target mis-splicing mediated by <i>RAVER2</i>                      | Dominant              |

*ATP6AP2*, ATPase, H<sup>+</sup> transporting, lysosomal accessory protein 2; *DMD*, dystrophin; ESS, exonic splicing silencer; *HNRPDL*, heterogeneous nuclear ribonucleoprotein D-like; *IKBKAP*, inhibitor of  $\kappa$ -light polypeptide gene enhancer in B cells, kinase complex-associated protein; *LMNA*, lamin A; *MAPT*, microtubule-associated protein tau; NMD, nonsense-mediated decay; *PRPF6*, pre-mRNA processing factor 6; PTBP1, polypyrimidine tract binding protein 1; RNP, ribonucleoprotein; *SMN1*, survival of motor neuron 1; ss, splice site.

# Recurrent splicing factor alterations in cancer

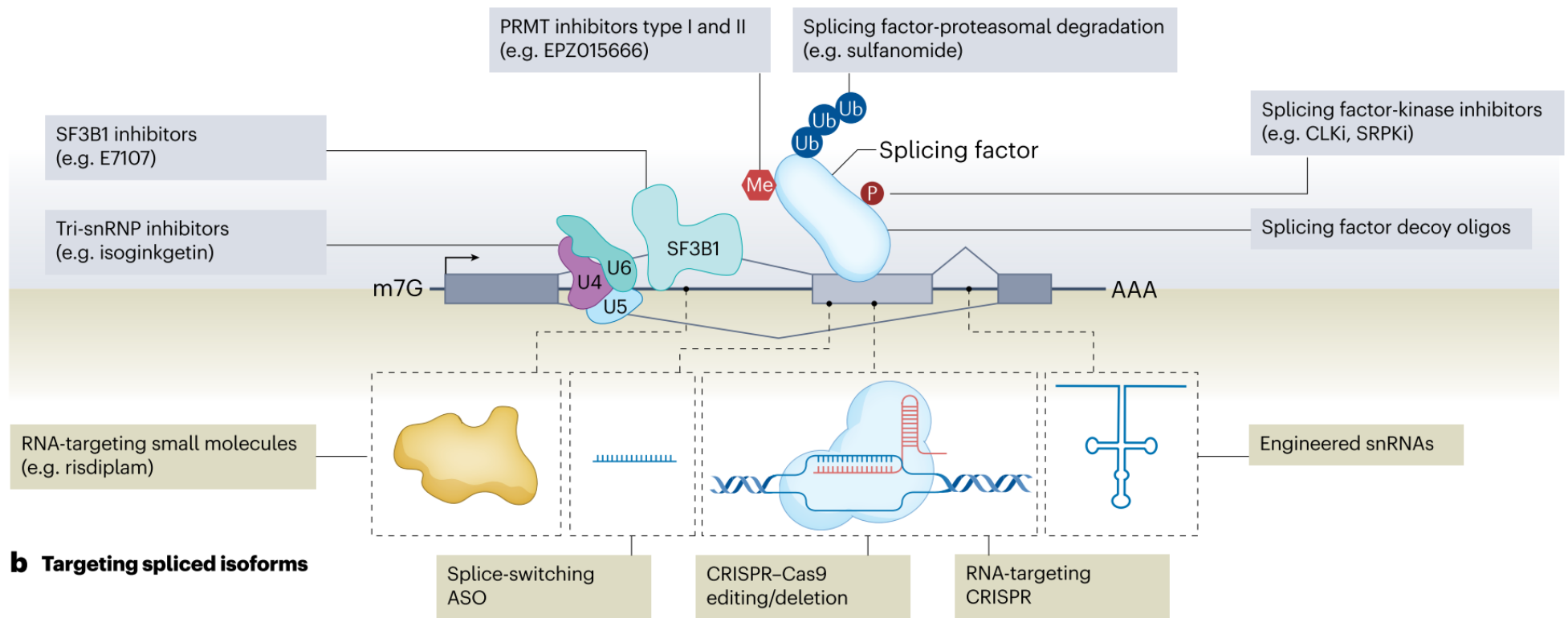


# Recurrent hot-spot mutations in human malignancies



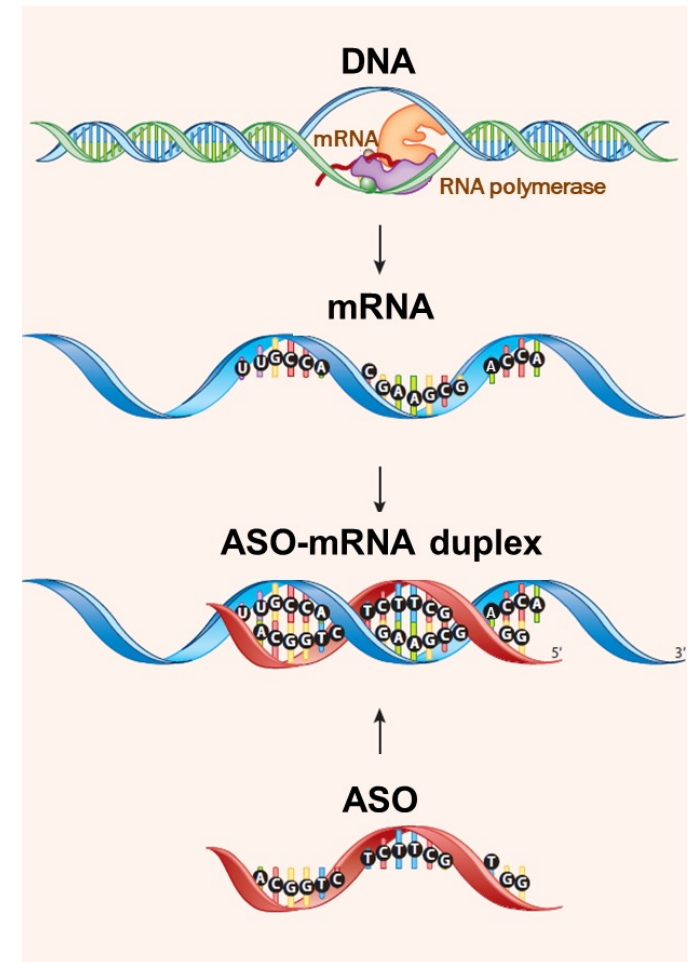
# Therapeutic approaches to target splicing

## a Targeting the spliceosome and splicing factors



# Antisense oligonucleotides (ASOs)

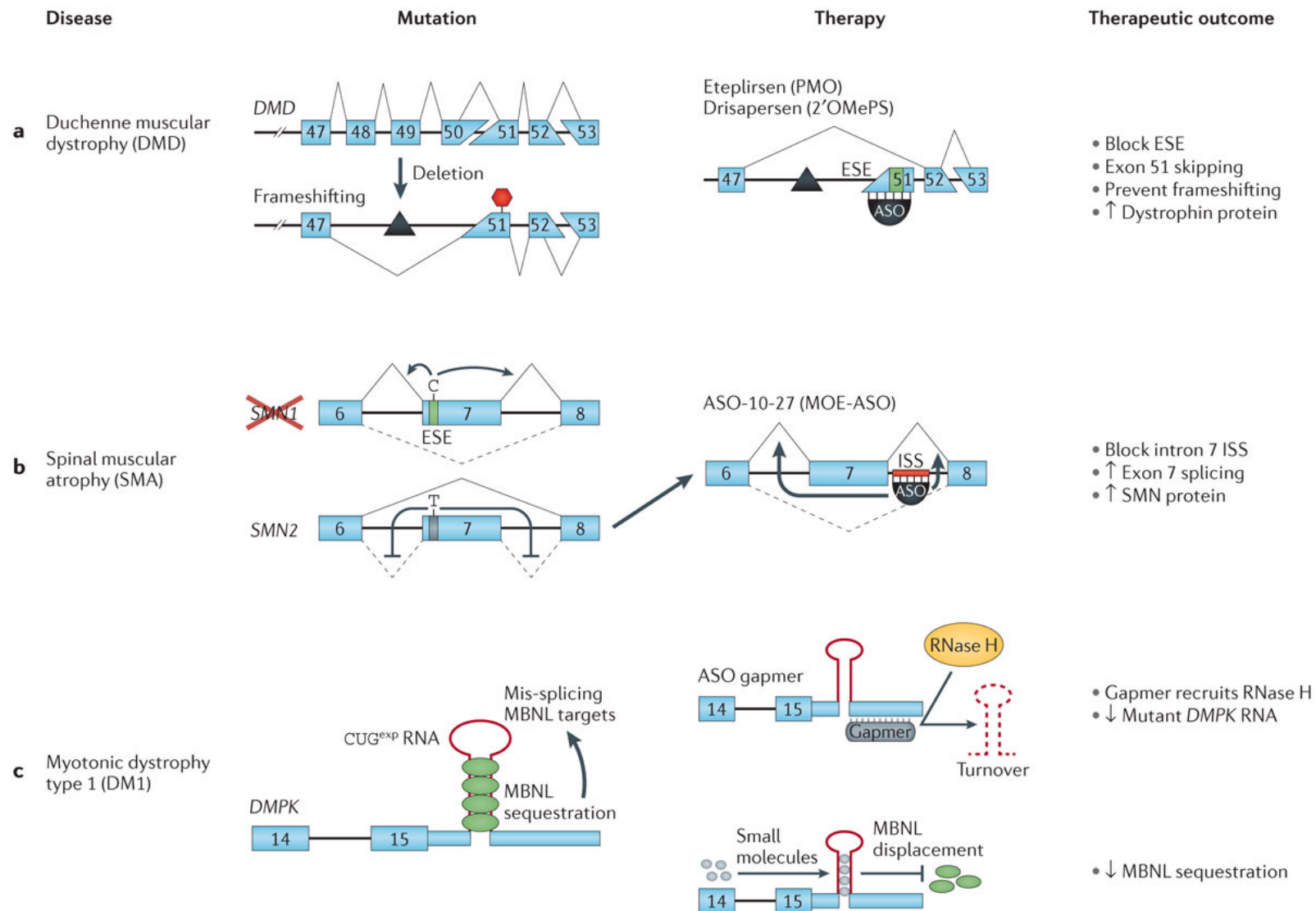
- ASOs are short, synthetic, single-stranded nucleic acids that bind to mRNA by base pairing and can modulate protein levels through several mechanisms
- Numerous chemical classes; most comprise a phosphorothioate backbone, plus one or more ribose-sugar modifications and a base modification
- The sequence of an ASO determines what RNA it will bind, and where along the RNA sequence
- By binding to a pre-mRNA in the nucleus, an ASO can affect the resulting splicing pattern
- The size and chemical characteristics of ASOs prevent blood-brain barrier penetration



Bennett (2019) *Annu Rev Med* 70: 307-32.

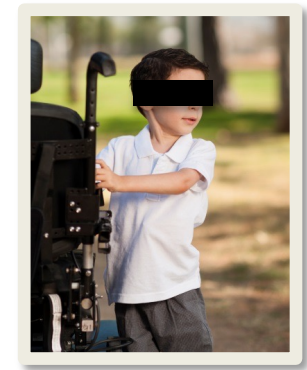
# Therapeutic strategies

Examples of therapies based on antisense oligonucleotide (ASO) and small molecule approaches:



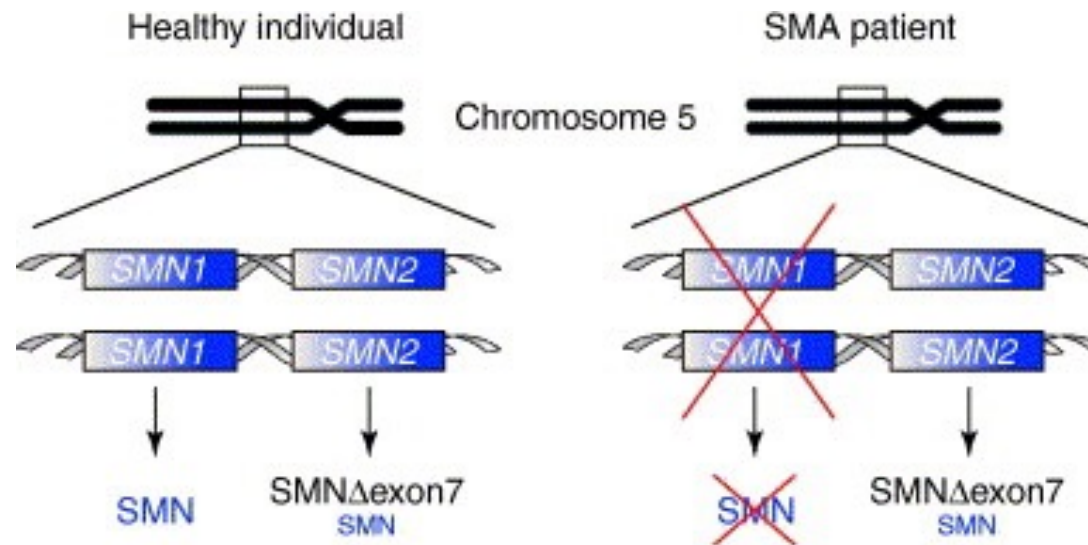
# Spinal Muscular Atrophy

- Pediatric neuromuscular disorder, autosomal recessive
- Degeneration of  $\alpha$ -motor neurons in the spinal cord and lower brainstem
- 1 in ~10,000 newborns
- Loss-of-function mutations in the *SMN1* gene, which codes for SMN protein
  - SMN functions in snRNP assembly and axonal mRNA transport
- Closely related *SMN2* gene (unique to humans) provides partial function
- Variable severity (SMA type 1-4) inversely proportional to *SMN2* copy number



# Spinal muscular atrophy: the RNP connection

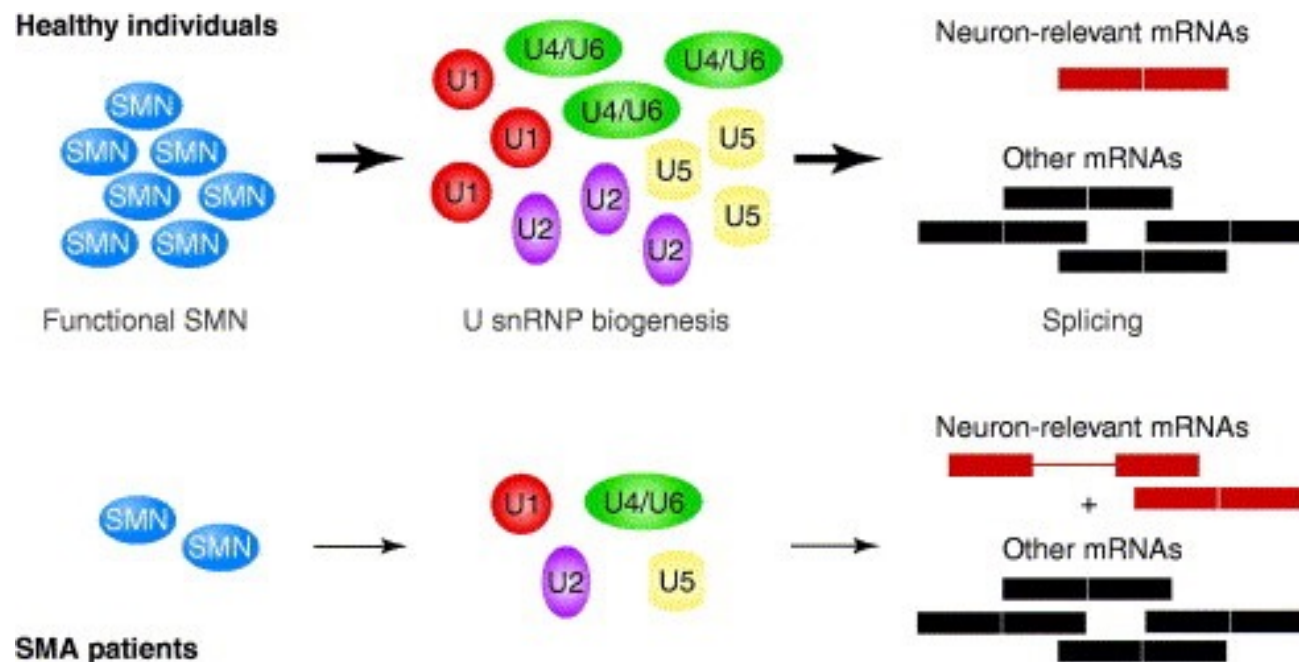
Degenerated motor neurons in the spinal cord are the pathological hallmark of **spinal muscular atrophy (SMA)**. SMA is caused by mutations in the ubiquitously expressed **survival motor neuron 1 (SMN1)** gene, which lead to reduced levels of functional SMN protein. Two nearly identical copies of the *SMN* gene are located on the long arm of chromosome 5 (5q13). Functional SMN protein is predominantly produced from *SMN1*, whereas the major product of *SMN2* is a truncated and non-functional protein (*SMN $\Delta$ exon7*). Mutations that cause disease inactivate *SMN1*, leaving *SMN2* as the only source of functional protein.



TRENDS in Molecular Medicine

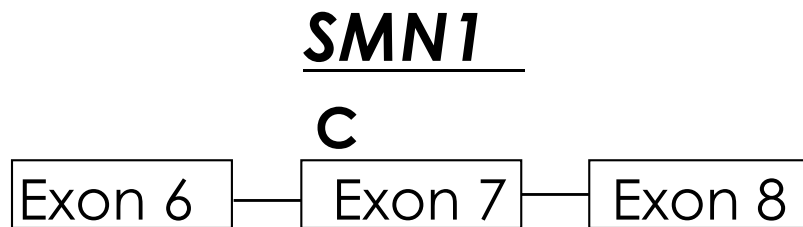
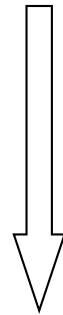
# Spinal muscular atrophy: the RNP connection

SMN deficiency results in reduced production of spliceosomal U snRNPs. As a consequence, processing of mRNAs with sub-optimal splice sites (e.g. tissue-specific transcripts) would be compromised. Therefore, inefficient splicing of neuron-relevant mRNAs is the basis for the tissue-specific phenotype in SMA.



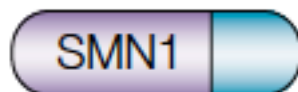
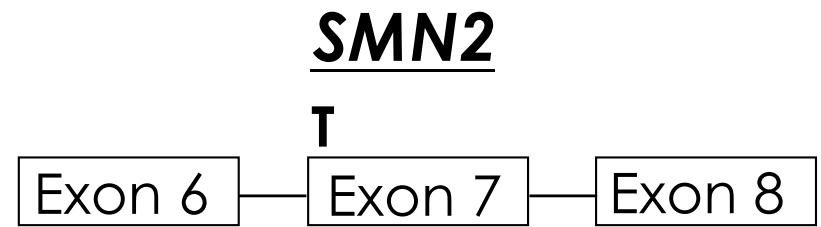
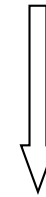
# Alternative Splicing of the *SMN* Genes

*SMN2*, a nearly identical copy of *SMN1*, fails to compensate for *SMN1* loss due to a critical C to T transition at the 6th position in exon 7 that leads to exon 7 skipping during pre-mRNA splicing of *SMN2* (9); as a consequence, a truncated, dysfunctional and rapidly degraded protein (*SMN*Δ7) is produced



Exon 6   Exon 7   Exon 8

**90% of transcripts**



**90% of transcripts**

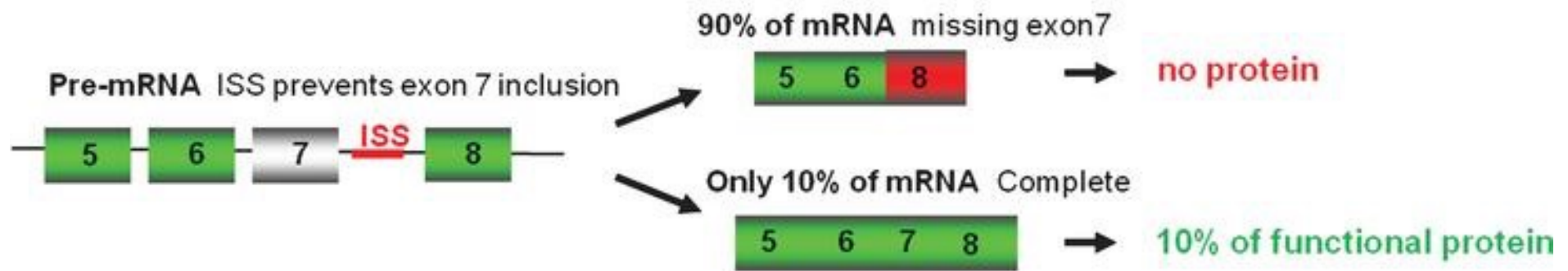
Exon 6   Exon 7   Exon 8

**10% of transcripts**

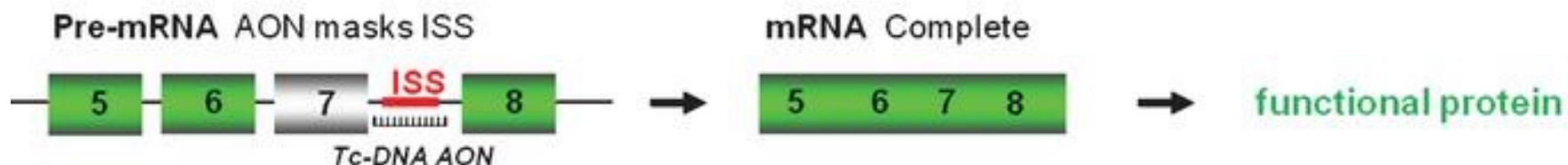
# Targeting RNA-Splicing for SMA Treatment

Modulation of alternative splicing (exon 7 inclusion) in SMN2 by antisense oligonucleotide is targeted for SMA treatment. Systemic administration of an antisense oligonucleotide (ASO-10-27) to neonates with SMA robustly rescues mice with severe SMA and extended median lifespan of animals with SMA 25-fold.

## Spinal muscular atrophy – SMN2



## Exon inclusion approach



# Clinical classification of SMA

| SMA Type | Age of Onset    | Highest Achieved Motor Function | Natural Age of Death | Typical Number of <i>SMN2</i> Copies |
|----------|-----------------|---------------------------------|----------------------|--------------------------------------|
| 0        | Prenatal/fetal  | None                            | <6 months            | 1                                    |
| I        | <6 months       | Sit with support only           | <2 years             | 1-3                                  |
| II       | 6–18 months     | Sit independently               | >2 years             | 2-3                                  |
| III      | >18 months      | Walk independently              | Adulthood            | 3-4                                  |
| IV       | Adult (20s-30s) | Walk through adulthood          | Adult                | ≥4                                   |

Adapted from Table 1 of Verhaart et al. 2017.<sup>6</sup>

Number of SMN2 copies based on Calucho et al. 2018.<sup>11</sup>

## Spinal muscular atrophy frequency

A common orphan disease

**1 in 10,000**

born with spinal muscular atrophy

**1 in 50**

is a carrier

Patients worldwide

North America

**14,000**

Europe  
**10,000**

Japan  
**1,000**

## Types of SMA

**Type I** Severe

Life expectancy <2 years

Never sits independently

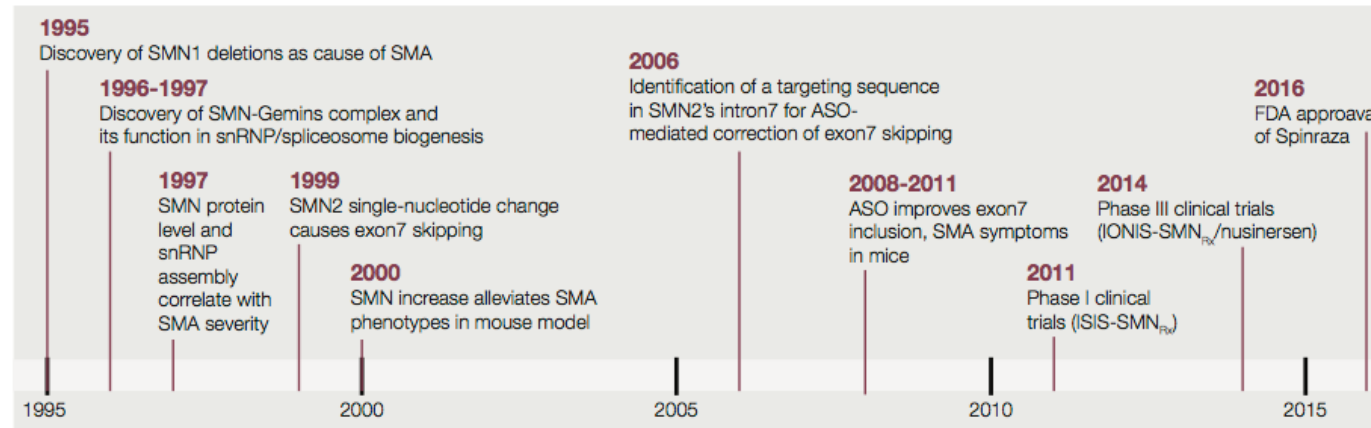
**Type II** Intermediate

Life expectancy >2 years

Sits but never walks independently

**Type III** Mild

Able to stand and walk limited steps



NAME: **Spinraza** (nusinersen, IONIS-SMN, ISIS-SMN)

APPROVED FOR: Spinal muscular atrophy (SMA) in pediatric and adult patients

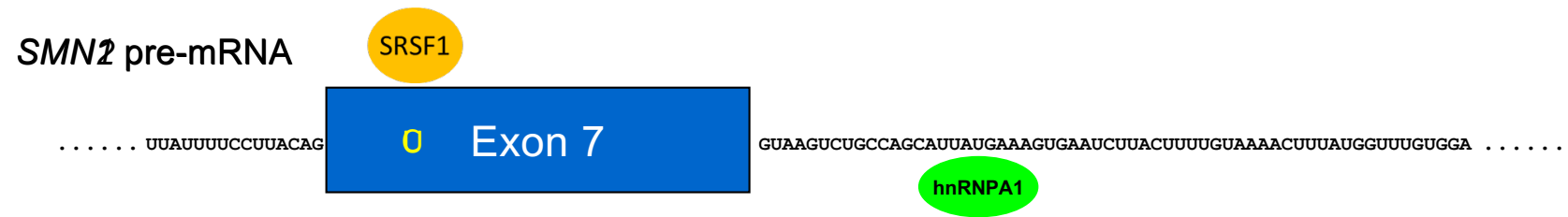
TYPE: 2'-O-methoxyethyl phosphorothioate-modified antisense oligonucleotide (ASO) delivered by intrathecal injection into the cerebrospinal fluid

MOLECULAR TARGETS: SMN2 pre-mRNA intron7 ISS

EFFECTS ON TARGETS: Masking SMN2 pre-mRNA intron7 ISS with ASO enhances exon7 splicing inclusion, boosting full-length SMN mRNA and protein level in the spinal cord, and improving achievement of motor milestones

PRICE: \$125,000 per injection, which adds up to \$750,000 for the first year of treatment and \$375,000 after that.

# Nusinersen (Spinraza)



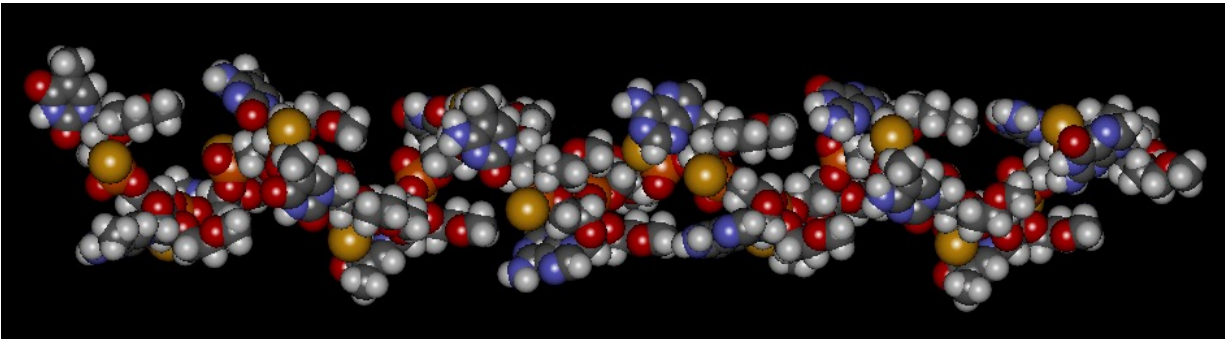
Hua, Vickers, Okunola, Bennett & Krainer (2008) *Am J Hum Genet* 82: 834

# Nusinersen (Spinraza)

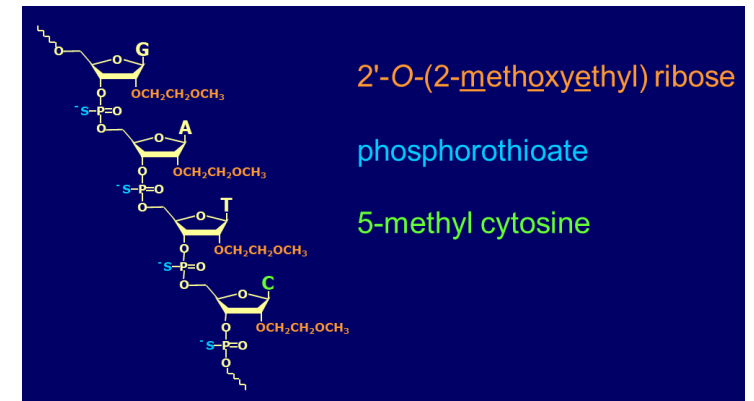
SMN2 pre-mRNA



SRSF1

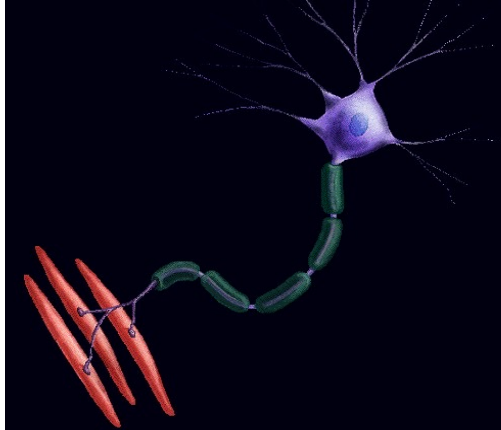


18mer ASO; MW 7127

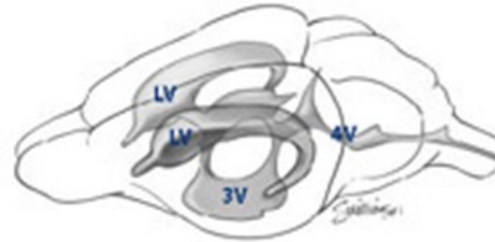
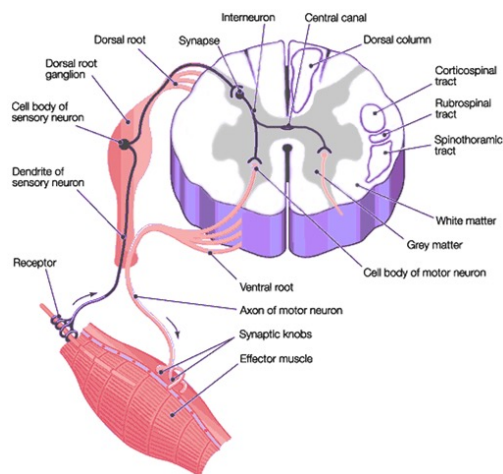


Hua, Vickers, Okunola, Bennett & Krainer (2008) *Am J Hum Genet* 82: 834

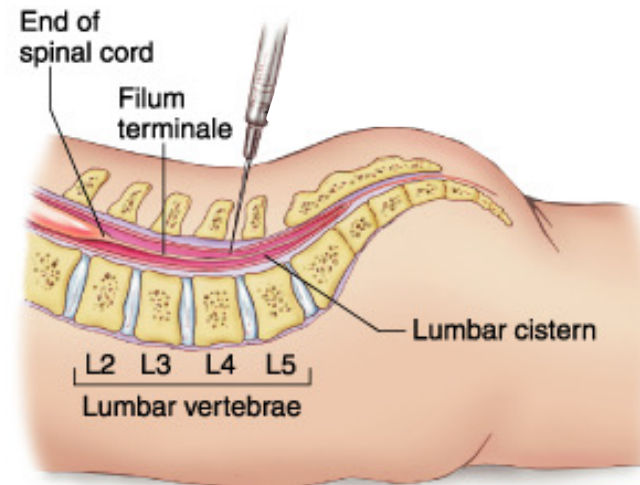
# Intrathecal ASO delivery



[www.uofaweb.ualberta.ca](http://www.uofaweb.ualberta.ca)



<http://www.glittra.com/yvonne/neuropics.html>



<http://www.mdguidelines.com/lumbar-puncture>

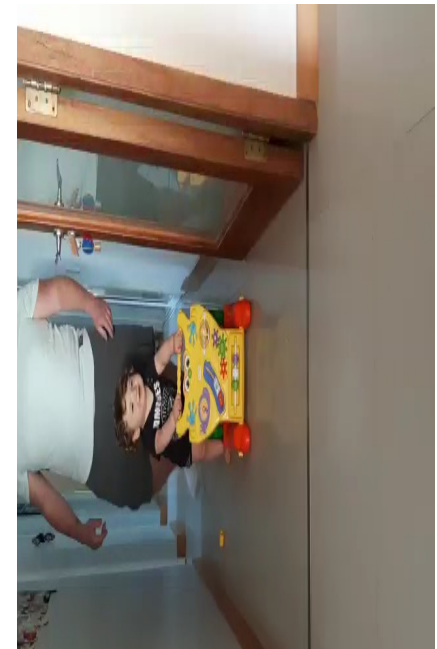
Long half-life of nusinersen in the CNS  
 Loading doses: 12 mg @ 2 weeks x 4  
 Maintenance doses: 12 mg @ 4 months

# Spinraza updates

- September 30, 2022: >11,000 patients currently on nusinersen worldwide, including commercial, early-access, and clinical-trial settings
- Nusinersen approved in >50 countries; reimbursed in >30
- NURTURE trial: genetically diagnosed SMA infants treated before onset of symptoms show unprecedented survival and motor-function gains for up to 5 years
- 46 states in the U.S. currently do newborn screening for SMA: 97% of newborn babies; implementation studies in other countries
- New DEVOTE randomized clinical trial currently evaluating higher nusinersen doses



Two other approved SMA therapies:  
Zolgensma (Novartis 2019) & Evrysdi (Roche 2020)



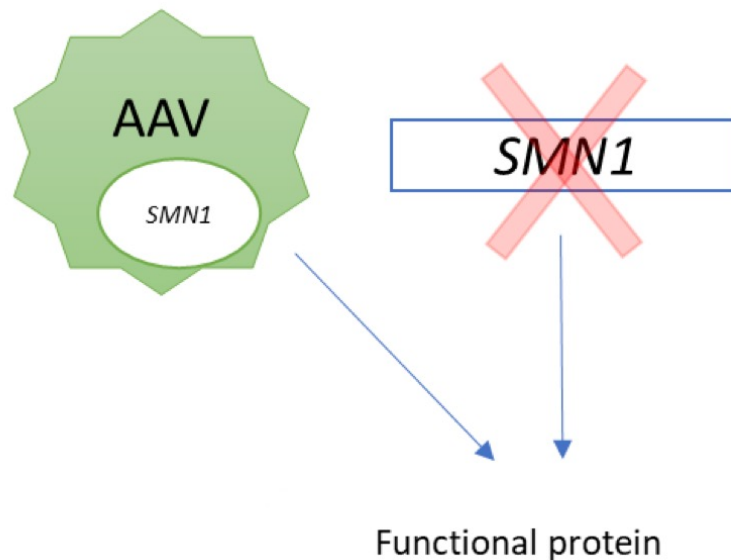
# Novartis' gene therapy for SMA - Zolgensma

PRICE: \$2.125 million

ZOLGENSMA is a prescription gene therapy used to treat children less than 2 years old with spinal muscular atrophy (SMA).

ZOLGENSMA is given as a one-time infusion into the vein.

ZOLGENSMA was not evaluated in patients with advanced SMA.

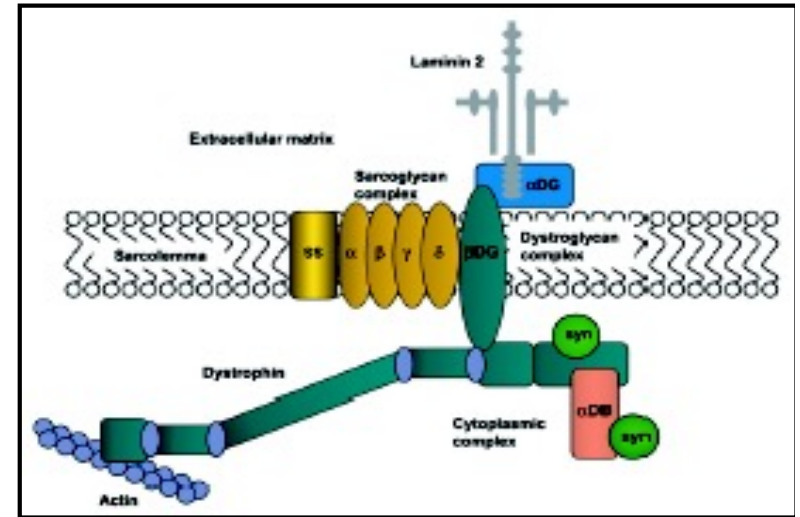


May-2020  
approved in Europe

# **The exon skipping strategy in the therapeutic treatment of Duchenne Muscular Dystrophy**

# Duchenne Muscular Dystrophy (DMD)

- X-linked recessive disorder
- affects 1 in 3500 live males
- DMD muscles degenerate with activity
- leads to *death* by the third decade of life



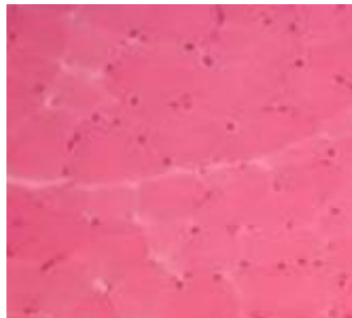
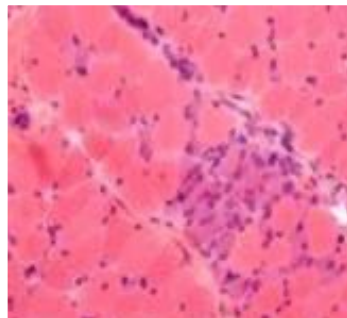
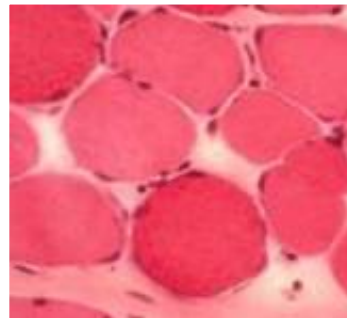
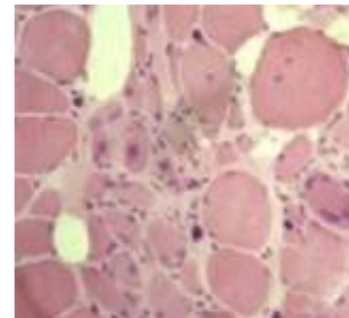
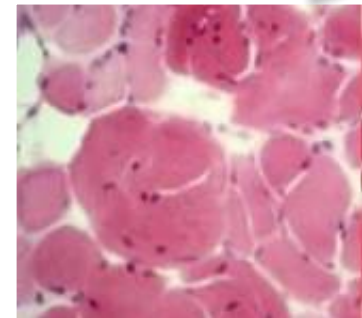
The gene is too big for a classical  
gene therapy intervention

## Dystrophin

- protein= 427 KDa
- DNA= 2,5 Mb
- cDNA= 14 Kb

# Duchenne Muscular Dystrophy (DMD)

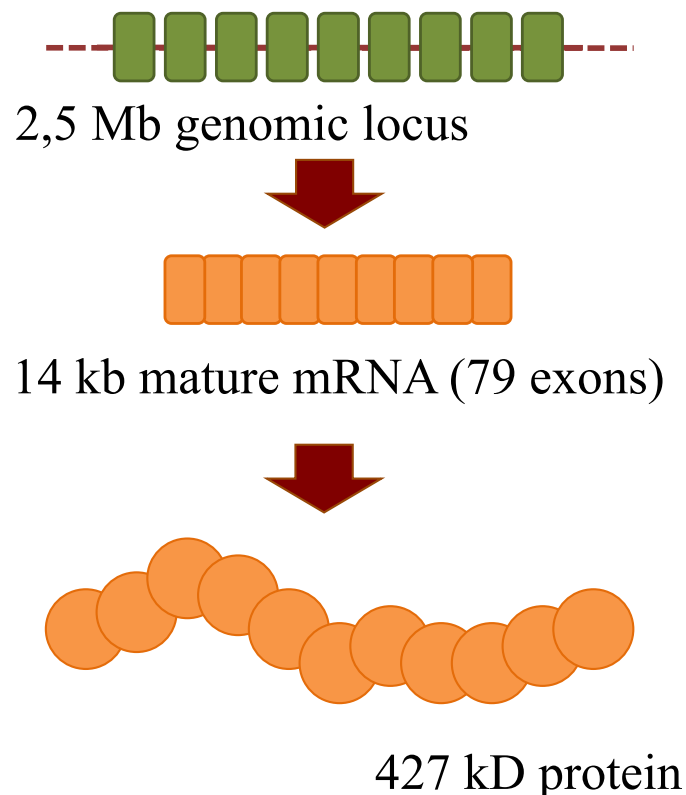
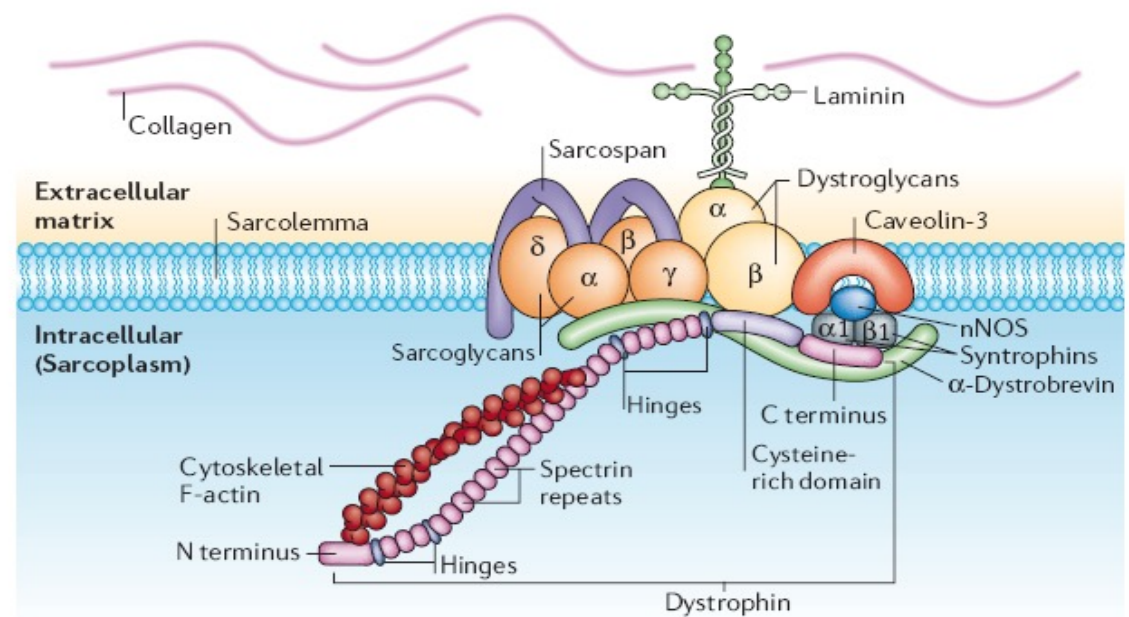
- is a severe disorder characterized by rapid progression of muscle degeneration, leading to loss of ambulation and death.
- X-linked recessive disorder that affects 1 in 3500 live males



Histopathology of a Duchenne muscle

# Dystrophin-The LONGEST GENE of our genome

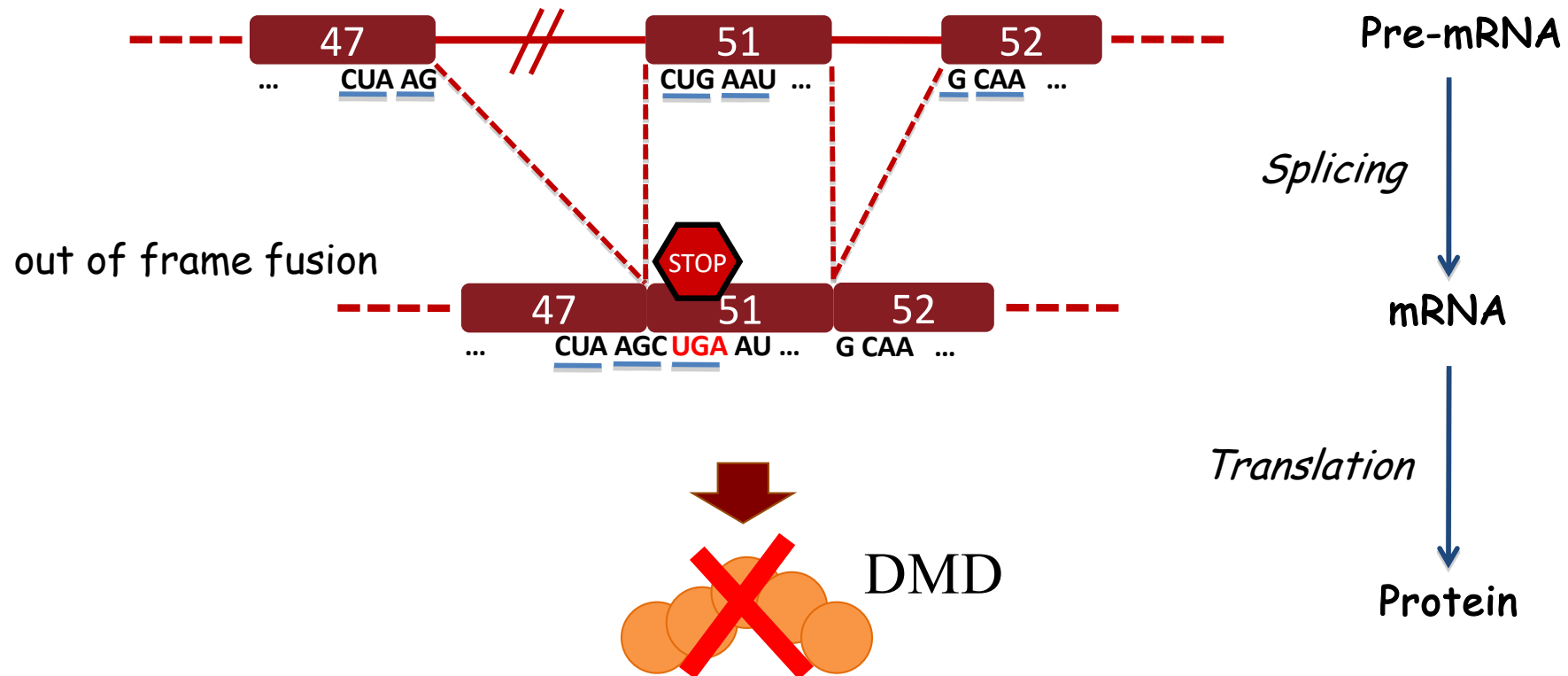
- Patients with DMD are deficient in dystrophin, a protein that connects the cytoskeleton of a muscle fiber to the surrounding extracellular matrix
- this deficiency causes sarcolemmal instability and muscle degeneration



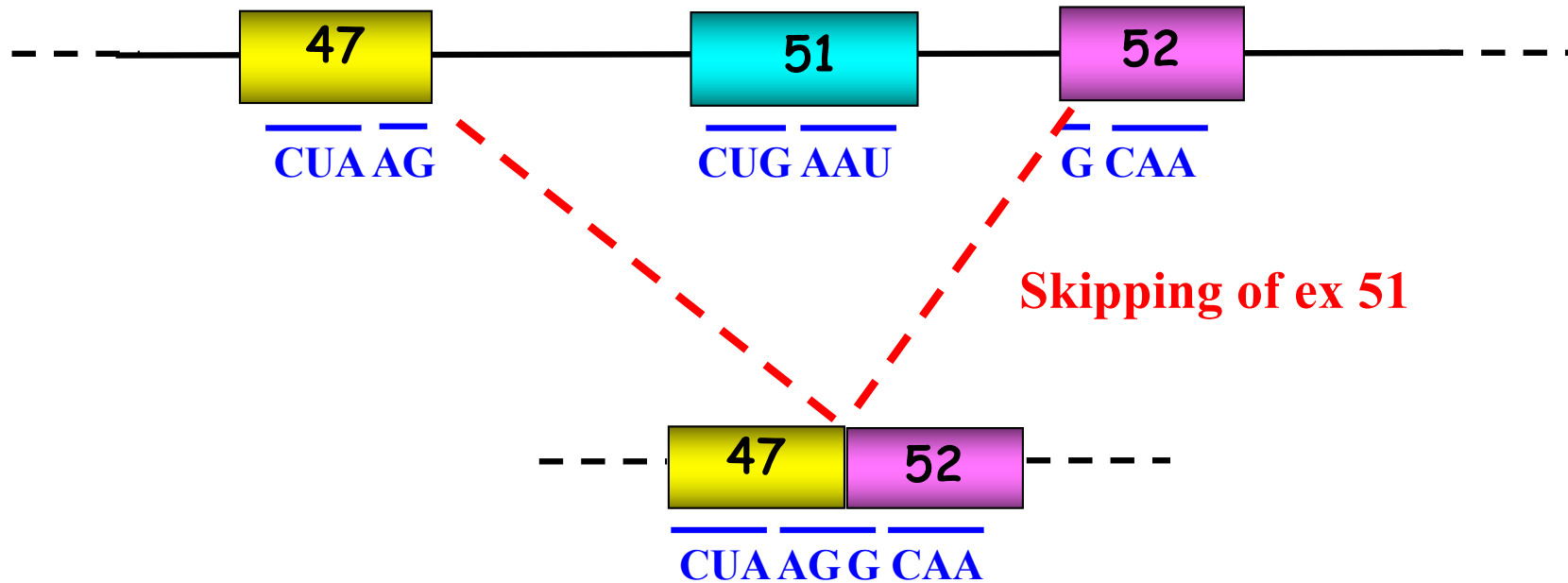
# Duchenne Muscular Dystrophy

## - the 48-50 deletion -

- DMD is caused by mutations in the Dystrophin gene that alter the pre-mRNA splicing and disrupt the open reading frame of the proteins, producing premature stop codons and mRNA degradation



# Exon skipping can revert the phenotype



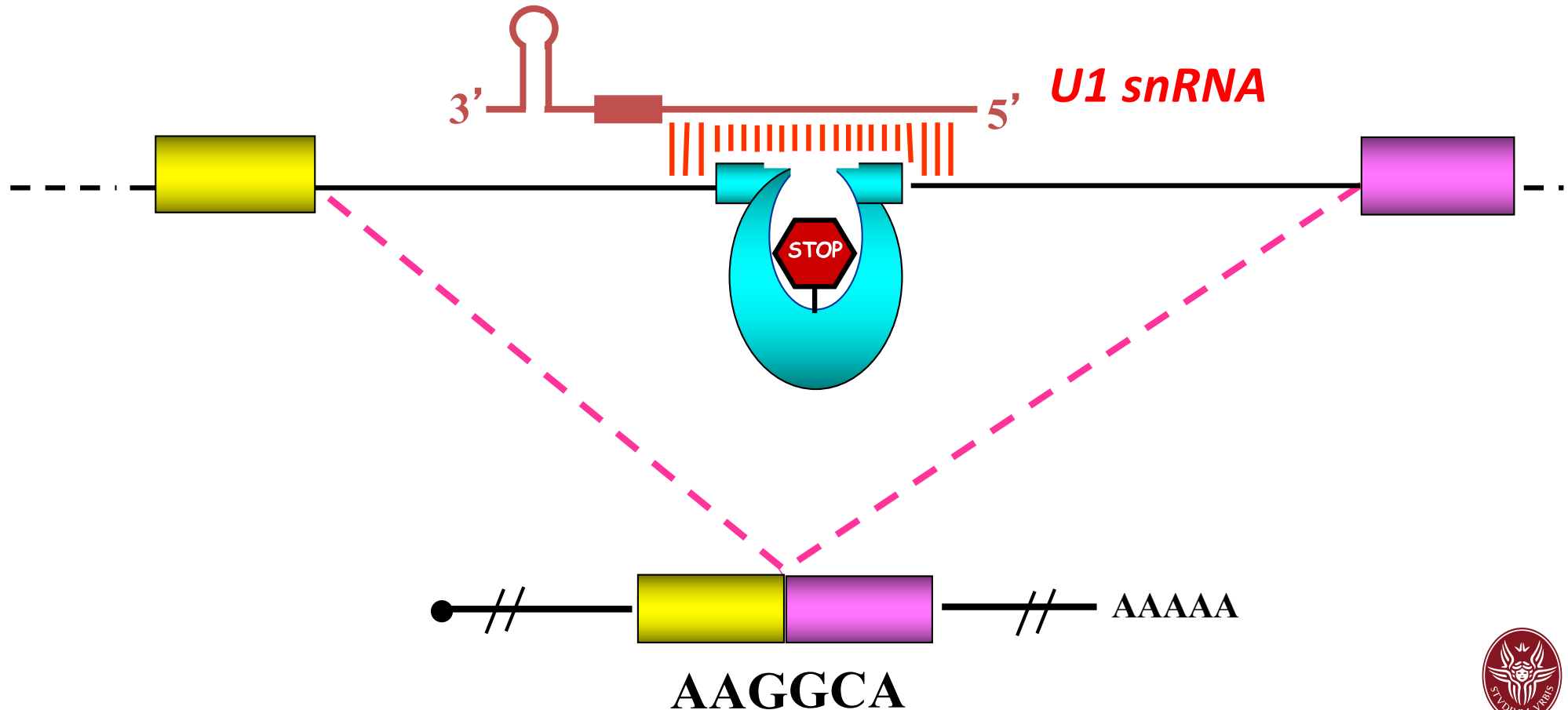
**In-frame mRNA** → translation of a shorter but still functional protein  
- **Becker-type** -

75% of all known dystrophin mutations can be cured by exon skipping  
skipping of ex 51 - 18%

# Antisense RNA technology applied to the correction of DMD mutations

## *U1 snRNA*

- nuclear RNA with specific recognition for splice junctions
- is matured in the cytoplasm and then reimported in the nucleus
- few transduced nuclei in the muscle fiber can provide chimeric antisense molecules to the entire fiber



# Exon skipping for the cure of DMD entered clinical trials (*Eteplirsen*)

van Deutekom, JC, *et al.* (2007).

Local dystrophin restoration with antisense oligonucleotide PRO051.

*N Engl J Med* **357**: 2677–2686.

Goemans, NM, *et al.* (2011).

Systemic administration of PRO051 in Duchenne's muscular dystrophy.

*N Engl J Med* **364**: 1513–1522.

Cirak et al.. (2011)

Exon skipping and dystrophin restoration in patients with Duchenne muscular dystrophy after systemic phosphorodiamidate morpholino oligomer treatment: an open-label, phase 2, dose-escalation study

*The Lancet* **378**: 595–605

## Exon skipping for the cure of DMD entered clinical trials (*Eteplirsen*)

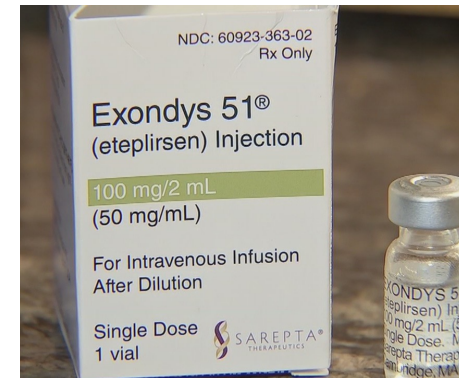
On September 19, 2016, *Eteplirsen* received accelerated conditional approval by the US FDA for boys with DMD deletion amenable to exon 51 skipping with limited data on only 12 cases and from just dystrophin protein as a surrogate marker in muscle biopsies, without proof of clinical improvement.

*Eteplirsen* (*ExonDys 51*) is reported to cost in the order of \$300,000-400,000/year per patient for life.

# Five FDA-approved splice-switching ASOs since 2016



SMA; Biogen  
Approved 2016



DMD Exon 51; Sarepta  
Approved 2016



DMD Exon 53; Sarepta  
Approved 2019



DMD Exon 53; NS Pharma  
Approved 2020

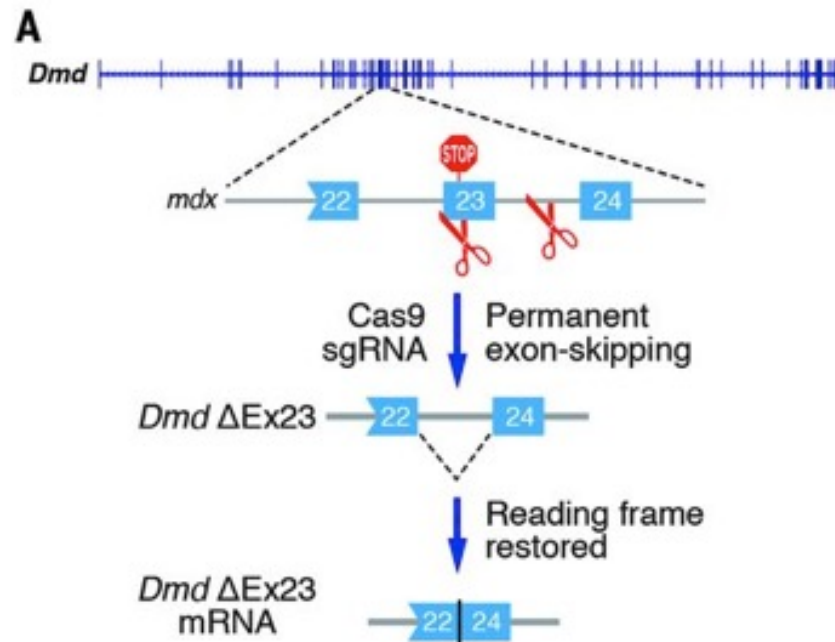


DMD Exon 45; Sarepta  
Approved 2021

# Exon skipping in the CRISPR/CAs9 Era

## In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy

Christopher E. Nelson,<sup>1,2</sup> Chady H. Hakim,<sup>3</sup> David G. Ousterout,<sup>1,2</sup>  
Pratiksha I. Thakore,<sup>1,2</sup> Eirik A. Moreb,<sup>1,2</sup> Ruth M. Castellanos Rivera,<sup>4</sup>  
Sarina Madhavan,<sup>1,2</sup> Xiufang Pan,<sup>3</sup> F. Ann Ran,<sup>5,6</sup> Winston X. Yan,<sup>5,7,8</sup>  
Aravind Asokan,<sup>4</sup> Feng Zhang,<sup>5,9,10,11</sup> Dongsheng Duan,<sup>3,12</sup> Charles A. Gersbach<sup>1,2,13\*</sup>



## Postnatal genome editing partially restores dystrophin expression in a mouse model of muscular dystrophy

Chengzu Long,<sup>1,2,3\*</sup> Leonela Amoasii,<sup>1,2,3\*</sup> Alex A. Mireault,<sup>1,2,3</sup> John R. McNally,<sup>1,2,3</sup>  
Hui Li,<sup>1,2,3</sup> Efrain Sanchez-Ortiz,<sup>1,2,3</sup> Samadrita Bhattacharyya,<sup>1,2,3</sup> John M. Shelton,<sup>4</sup>  
Rhonda Bassel-Duby,<sup>1,2,3</sup> Eric N. Olson<sup>1,2,3†</sup>

## In vivo gene editing in dystrophic mouse muscle and muscle stem cells

Mohammadsharif Tabebordbar,<sup>1,2\*</sup> Kexian Zhu,<sup>1,3\*</sup> Jason K. W. Cheng,<sup>1</sup>  
Wei Leong Chew,<sup>2,4</sup> Jeffrey J. Widrick,<sup>5</sup> Winston X. Yan,<sup>6,7</sup> Claire Maesner,<sup>1</sup>  
Elizabeth Y. Wu,<sup>1†</sup> Ru Xiao,<sup>8</sup> F. Ann Ran,<sup>6,7</sup> Le Cong,<sup>6,7</sup> Feng Zhang,<sup>6,7</sup>  
Luk H. Vandenberghe,<sup>8</sup> George M. Church,<sup>4</sup> Amy J. Wagers<sup>1†</sup>

Science

2016