Alternative splicing

Biosciences 741: Genomics
Fall, 2013
Week 6

Function(s) of RNA splicing

- Splicing of introns must be completed before nuclear RNAs can be exported to the cytoplasm. This led to early speculation that intron splicing may be required for gene function.
- However, genetically-engineered transgenes without introns worked perfectly well in a variety of species. Moreover, some normal genes lack introns and again work perfectly well. So introns are dispensable.
- The results of *in vitro* mutagenesis and gene fusion experiments led to the "first come, first served" model of RNA splicing, according to which any 5' splice donor will work with any 3' splice acceptor.
- The "first come, first served" model implies that RNA splicing is likely to be somewhat error-prone, and indeed the first splice variants to be specifically tested turned out to be dispensable (this is one type of a "non-functional" splice variant).
- Conserved splice variants are much more likely to be functional, but only about 20% of splice variants are conserved even between humans vs. mice! (how, specifically is conservation identified?)

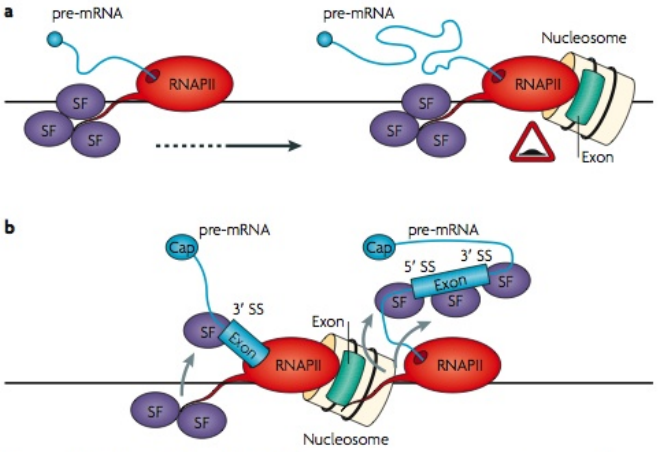


Figure 3 | **The 'speed bump' model.** Nucleosome occupancy marks exons and is coupled to transcription. **a** | RNA polymerase II (RNAPII), associated with different splicing factors (SFs), travels along the gene and transcribes it. When RNAPII reaches an area with high nucleosome occupancy and encounters specific histone modifications that mark an exon, it is slowed down. **b** | This panel shows RNAPII and the nucleosome at the point at which their coupling marks the exon boundaries for the splicing machinery. RNAPII transcribes the exon and SFs detach from the carboxy-terminal domain of RNAPII and bind to the 3′ splice site (3′ SS) region of the precursor mRNA (pre-mRNA). During transcription elongation, additional SFs bind intronic and exonic splicing regulatory elements and the 5′ SS.

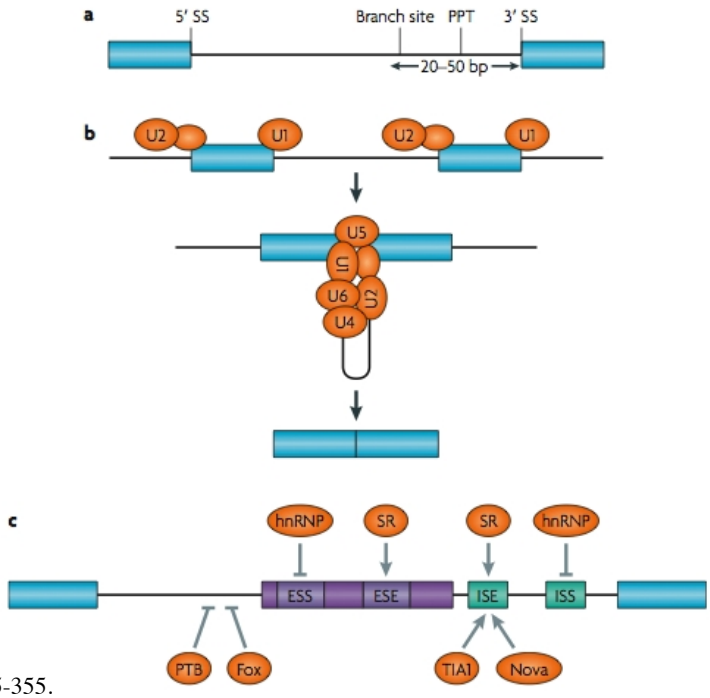
Keren et al. (2010)
Nat. Rev. Genet.
11, 345-355.

The "speed bump" model

- Highly used exons are more occupied by nucleosomes than rarely used exons, which are more occupied by nucleosomes than introns.
- This difference in nucleosome placement, density, and stability is believed to affect the rate of RNA polymerase transit through exons vs. introns, which in turn would affect the time available for splicing complexes to form. Slowing RNA polymerase transit through exons may facilitate the accuracy of splicing, as well as the differential splicing of specific exons.
- Exons are marked by characteristic histone modifications, including H3K36me3, H3K79me1, H4K20me1, and H2BK5me1. These marks are conserved, for example in humans and fish.
- The modulation of these histone modifications has been shown, in some cases, to result in splice site switching.
- Thus it is likely that RNA splicing is regulated, at least partially, at the epigenetic level.

Regulated alternative splicing

- More recently, genomic studies of the transcriptome expressed in different tissues and life stages have shown that approximately 30% to 60% of splice variants are developmentally regulated in *Drosophila* and *Caenorhabditis*.
- Regulated splice variants are far more likely to be conserved. In most cases, regulated splice variants are regulated by specific RNA binding proteins that either enhance or inhibit the splicing of weak (non-canonical) splicing signals.
- Algorithms are available both to recognize splicing signals, and to recognize binding sites for splicing regulators. However, these algorithms are not particularly accurate, and the majority of the predicted binding sites for splicing regulators are not functional. Why?
- Because some functional splice variants (as shown by knockout studies) are not conserved, another implication of these results is that new splice variants evolve moderately rapidly, particularly following the insertion or deletion of new exons or introns.
- How (by what process) are introns deleted in evolutionary time?



Keren et al. (2010)
Nat. Rev. Genet. 11, 345-355.

Figure 1 | **The splicing machinery.** Splicing is a conserved mechanism controlled by the spliceosome — a complex composed of many proteins and five small nuclear RNAs (U1, U2, U4, U5 and U6) that assemble with proteins to form small nuclear ribonucleoproteins (snRNPs). **a** | The four conserved signals that enable recognition of RNA by the spliceosome are: the exon–intron junctions at the 5′ and 3′ ends of introns (the 5′ splice site (5′ SS) and 3′ SS), the branch site sequence located upstream of the 3′ SS and the polypyrimidine tract (PPT) located between the 3′ SS and the branch site. **b** | The key steps in splicing are shown. Regulation of splicing can occur at the basic level of splice-site recognition by the spliceosome through the facilitation or interference of the binding of U1 and U2 snRNPs to the splice sites[7]. The unlabelled orange ovals represent other, unspecified components of the spliceosome. **c** | Exons and introns contain short, degenerate binding sites for splicing auxiliary proteins. These sites are called exonic splicing enhancers (ESEs), intronic splicing enhancers (ISEs), exonic splicing silencers (ESSs) and intronic silencing silencers (ISSs). Splice-site recognition is mediated by proteins that bind specific regulatory sequences, such as the serine/arginine (SR) proteins, heterogenous nuclear ribonucleoproteins (hnRNPs), polypyrimidine tract-binding (PTB) proteins, the TIA1 RNA-binding protein, Fox proteins, Nova proteins, and more[7,9,10]. Constitutive exons are shown in blue, alternatively spliced regions in purple, and introns are represented by solid lines.

Keren et al. (2010) Nat. Rev. Genet. 11, 345-355.

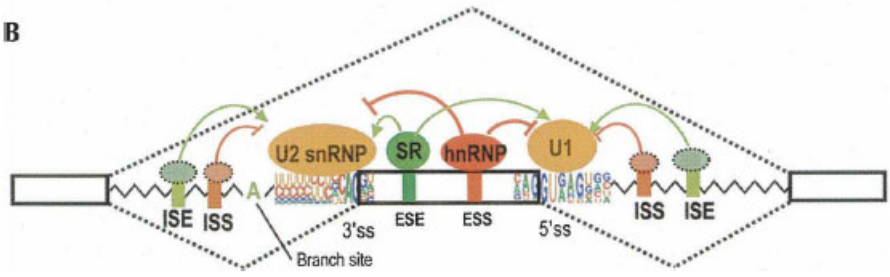
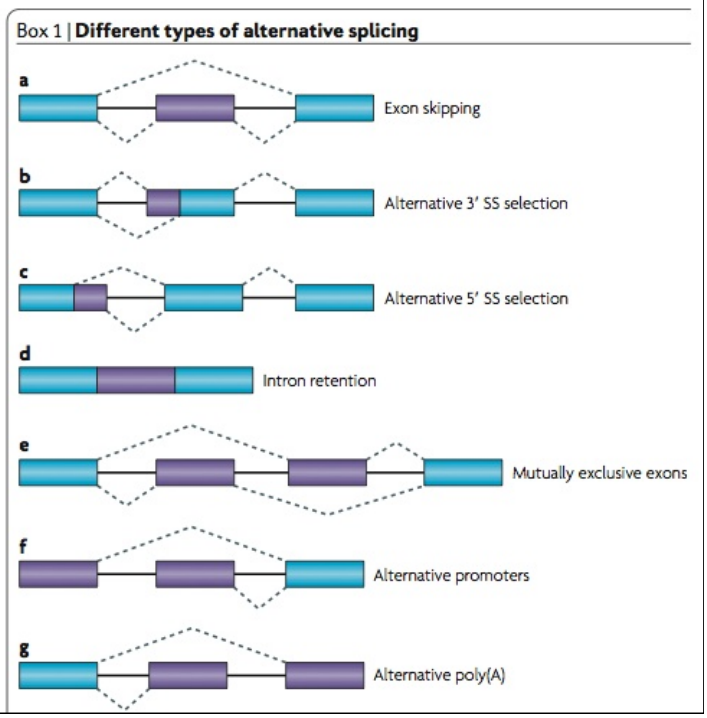


FIGURE 1. (*A*) Major forms of alternative splicing. In many cases, these common forms can be combined to generate more complicated alternative splicing events. (*B*) A schematic of regulated splicing. (Open boxes) Exons, (jagged lines) introns, (brackets) splice sites (ss). The consensus motifs of ss are shown in pictogram, and the branch point adenosine is indicated. (Dashed lines) Two alternative splicing pathways, with the middle exon either included or excluded. Splicing is regulated by *cis*-elements (ESE, ESS, ISS, and ISE) and *trans*-acting splicing factors (SR proteins, hnRNP, and unknown factors).



Keren et al. (2010) Nat. Rev. Genet. 11, 345-355.

There are several different types of alternative splicing (AS) events, which can be classified into four main subgroups. The first type is exon skipping, in which a type of exon known as a cassette exon is spliced out of the transcript together with its flanking introns (see the figure, part **a**). Exon skipping accounts for nearly 40% of AS events in higher eukaryotes[17,111] but is extremely rare in lower eukaryotes. The second and third types are alternative 3′ splice site (3′ SS) and 5′ SS selection (parts **b** and **c**). These types of AS events occur when two or more splice sites are recognized at one end of an exon. Alternative 3′ SS and 5′ SS selection account for 18.4% and 7.9% of all AS events in higher eukaryotes, respectively. The fourth type is intron retention (part **d**), in which an intron remains in the mature mRNA transcript. This is the rarest AS event in vertebrates and invertebrates, accounting for less than 5% of known events[17,19,98,111]. By contrast, intron retention is the most prevalent type of AS in plants, fungi and protozoa[19]. Less frequent, complex events that give rise to alternative transcript variants include mutually exclusive exons (part **e**), alternative promoter usage (part **f**) and alternative polyadenylation (part **g**)[12,19,112]. Another rare form of AS involves reactions between two primary transcripts in *trans*[113] (not shown).

In the figure, constitutive exons are shown in blue and alternatively spliced regions in purple. Introns are represented by solid lines, and dashed lines indicate splicing options.

Keren et al. (2010) Nat. Rev. Genet. 11, 345-355.

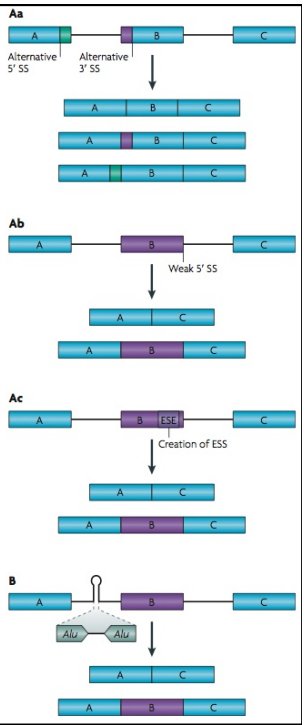


Figure 2 | **Transition from constitutive to alternative splicing.** There are two mechanisms by which a constitutive exon can become an alternative exon. **A** | Mutations that lead to suboptimal recognition of the exon and result in exon skipping. **Aa** | Mutations can lead to a new alternative 5′ splice site (5′ SS) or 3′ SS. **Ab** | Mutations can lead to suboptimal recognition of the 5′ SS. **Ac** | Mutations in exons (or introns) can disrupt an exonic splicing enhancer (ESE) (or intronic splicing enhancer (ISE)) or may create an exonic splicing silencer (ESS) (or intronic silencing silencer (ISS)). **B** | A secondary structure, usually formed between two *Alu* elements in opposite orientation, can interrupt exon recognition. The resulting isoforms are represented for each pathway. Constitutive exons are shown in blue, alternatively spliced regions are shown in purple, and introns are represented by solid lines.

Keren et al. (2010) Nat. Rev. Genet. 11, 345-355.

Alternative splicing vs. transcriptional regulation

Table 1 | **Functional enrichment in genes undergoing alternative splicing or transcriptional changes during physiological transitions***

Tissue type, cell type or process	Gene ontology terms enriched in genes regulated through alternative splicing	Gene ontology terms enriched in genes regulated through transcript levels	Refs
Brain or neural tissue	GTPase-based signalling, cell–cell signalling, cytoskeletal organization and biogenesis, vesicular-mediated transport, transmission of nerve impulse and neurophysiologic processes	Synaptic function, nerve impulse and transmission, nervous system development, cytoskeletal organization and biogenesis and secretory pathways	163
Heart development	Cell structure and motility, cytoskeletal remodelling, cell signalling, RNA splicing, muscle specification, excitation–contraction coupling and cell cycle control	Signal transduction and oxidative (lipid and steroid) metabolism, cell adhesion, cytoskeletal organization and biogenesis, nucleic acid metabolism and cell signalling	26
Skeletal muscle differentiation	Cytoskeletal organization, actin binding, cell junction and nucleotide kinase, integrin signalling pathway, nucleic acid metabolism and RNA splicing	Muscle contraction, muscle development, cytoskeletal organization, cell signalling, cell cycle, transcription, nucleic acid and protein metabolism, cell adhesion and ion transport	27, 164
Epithelial-to-mesenchymal transition	Cytoskeleton structure, cell adhesion, polarity, cell migration and RNA splicing	Cell cycle inhibition, apoptotic inhibition, cytoskeletal organization and biogenesis, cell structure and motility and cell adhesion	77, 165
T cell activation	Interphase of mitotic cell cycle (affected early); cell division (affected late)	Cell adhesion, immune defence response, cytoskeletal protein binding (affected early); cell cycle (affected late)	30
Ca^{2+}-induced cell excitation	Cell signalling (affected early); RNA splicing, transcription, cell cycle, apoptosis, lipid and carbohydrate metabolism (affected late); Ca^{2+} binding, cell adhesion, plasma membrane, and extracellular matrix (affected throughout the time course)	Lipid and carbohydrate metabolism (affected late); transcription, Ca^{2+} binding and retrograde vesicle-mediated transport from the Golgi to the ER (affected throughout the time course)	31

*Several studies probing alternative splicing and steady-state mRNA level changes during physiological transitions have found that genes undergoing alternative splicing and/or transcript level changes are enriched in an overlapping yet distinct set of gene ontology annotation terms.

Patterns of regulated RNA splicing

- Genes regulated by RNA splicing generally fall in similar (not always identical) functional categories, in comparison to genes regulated at the transcriptional level at the same developmental time and tissue. In other words, alternative splicing is functionally integrated with transcriptional regulation.
- In most cases, a gene regulated by RNA splicing is not regulated by transcriptional activation, or vice versa.
- RNA splicing can be used to turn a gene on or off, or to modulate its specificity (interactions with other proteins), or to modulate its targets (transcription factors), or to alter its protein-level regulation (protein regulatory cassettes).
- RNA splicing "cascades" are used to regulate all-or-nothing biological responses, such as sexual identity in Drosophila, or apoptosis.

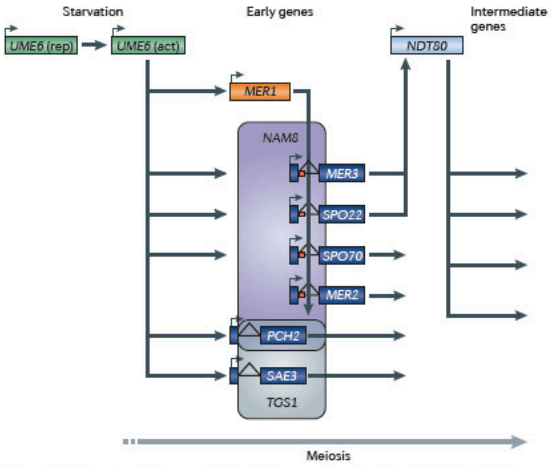
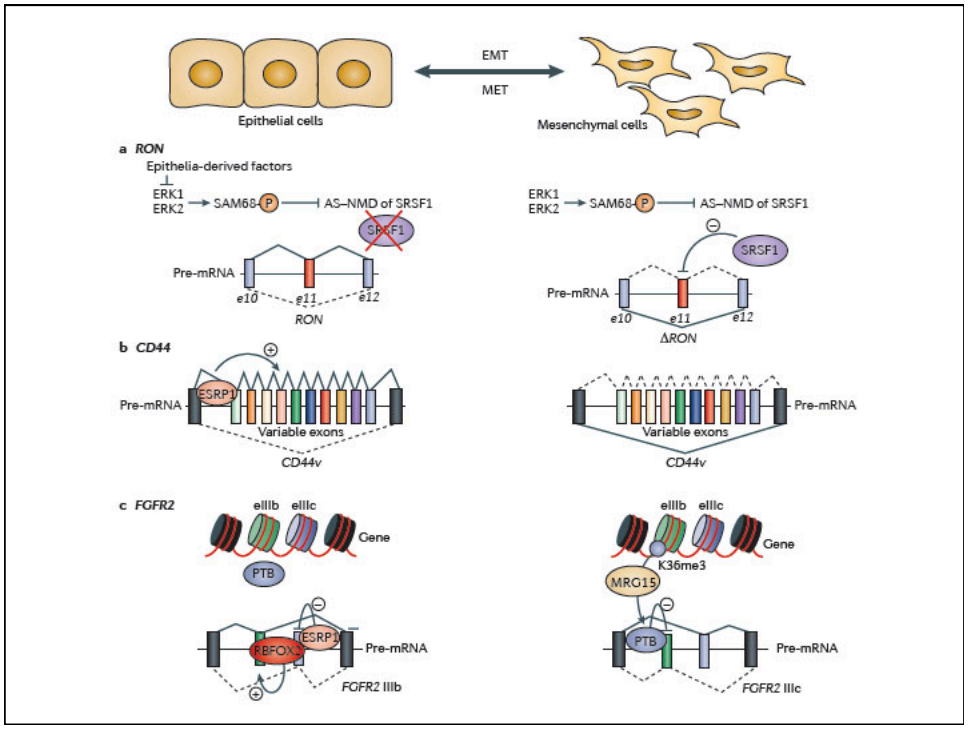
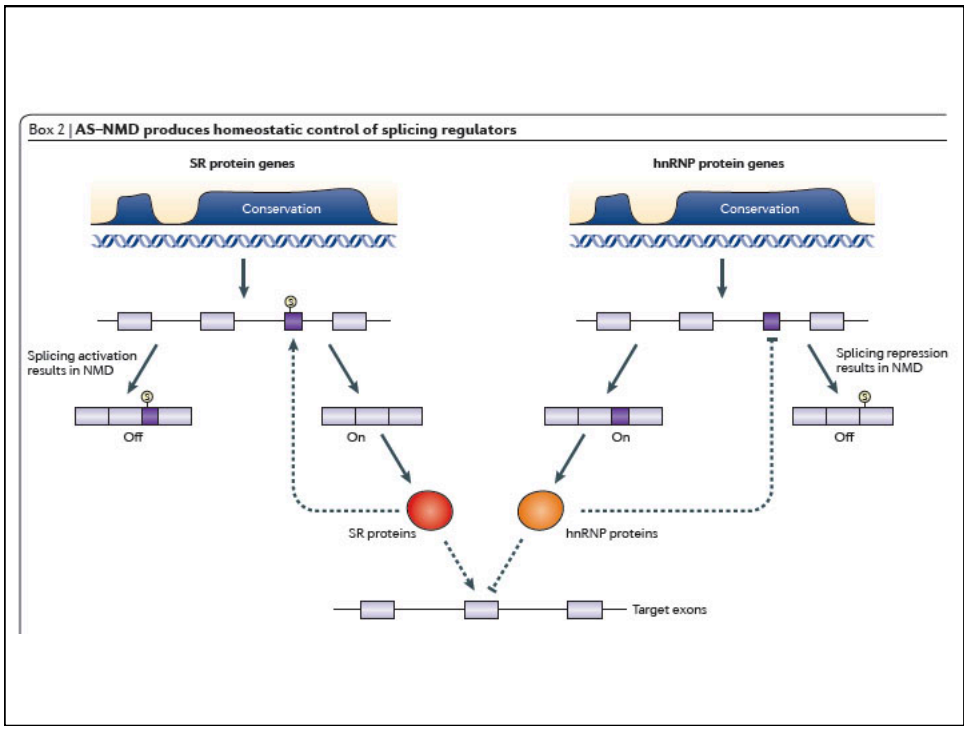
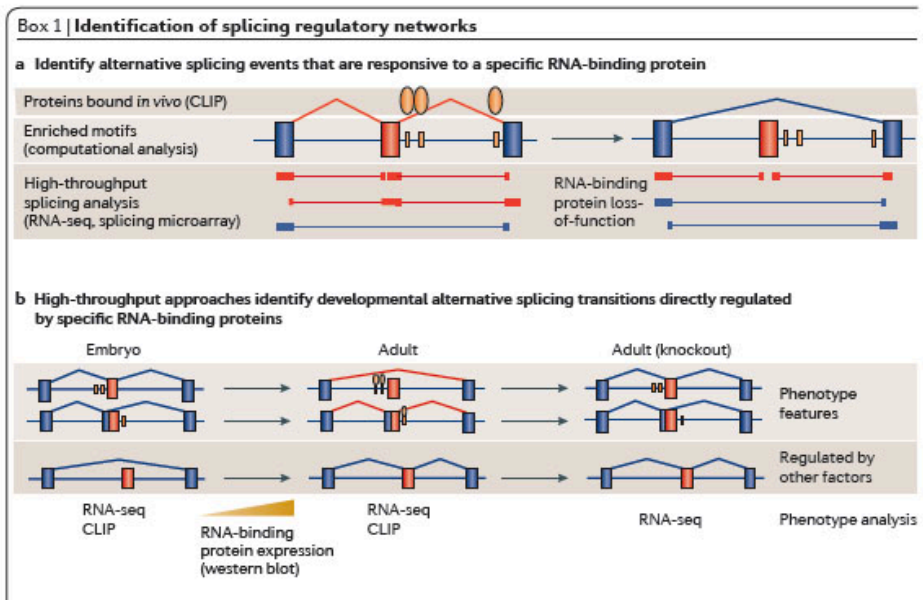


Figure 1 | **Role of splicing regulation during early meiosis in *Saccharomyces cerevisiae*.** Starvation induces a switch in the Ume6 transcription factor from a repressor (*UME6* (rep)) in vegetative cells to an activator (*UME6* (act)) to initiate meiosis during sporulation. Ume6 activates multiple early meiotic genes including *MER1* and 13 intron-containing genes. *MER1* encodes an RNA-binding protein that binds a motif near the 5′ splice site (orange box) and activates splicing of four intron-containing genes specifically during meiosis. Two Mer1 targets, *MER3* and *SPO22*, are required for activation of *NDT80*, the transcriptional regulator of intermediate meiotic genes. Transcriptional regulation of *MER1* and three of its four targets by *UME6* delays expression of Mer1-regulated genes compared to other *UME6* targets (for example, *PCH2* and *SAE3*), creating a lag period prior to *NDT80* activation[35]. Mutation analyses demonstrated separate but overlapping meiotic-dependent splicing requiring *MER1*, *NAM8* (purple shaded area) or *TGS1* (grey shaded area)[41].


EMT
MET
Epithelial cells
Mesenchymal cells
a RON
Epithelia-derived factors
ERK1
ERK2
SAM68
P
AS–NMD of SRSF1
SRSF1
Pre-mRNA
e10
e11
e12
RON
ΔRON
b CD44
ESRP1
Variable exons
CD44v
c FGFR2
eIIIb
eIIIc
Gene
PTB
RBFOX2
Pre-mRNA
FGFR2 IIIb
K36me3
MRG15
FGFR2 IIIc


Box 2 | AS–NMD produces homeostatic control of splicing regulators
SR protein genes
hnRNP protein genes
Conservation
Splicing activation results in NMD
Splicing repression results in NMD
Off
On
SR proteins
hnRNP proteins
Target exons

Finding targets of splicing regulators by CLIP and gene knockouts



Alternate promoter use

- Alternate promoter use is quite common. It is typically used to provide a mechanism for gene expression in different developmental times, or tissues, through entirely separate promoters that drive the same gene.
- Another use of alternate promoters is to provide multiple transcriptional mechanisms of physiological regulation of gene expression. These mechanisms may exist in parallel with each other (all expressed in the same cell type) or may be tissue-specific (i.e., calcium regulated in some tissues but not others).
- Another use of alternate promoters is to provide multiple N-terminal regulatory domains, all linked to the same C-terminal catalytic domain. This is useful for protein transport to different cytological or cytoskeletal portions of the cell, together with different regulation by protein kinases etc., and yet all having the same catalytic domain.

Discussion Questions

- How, specifically, are conserved introns and conserved splice variants identified? When you find a conserved splice variant, what is the significance of your discovery (how is it interpreted)?
- Why are "non-functional" splice variants usually constitutively spliced? What is meant by "non-functional" in this context? What is "constitutive splicing" in this context? Do these non-functional splice variants produce a protein? Why or why not?
- The majority of the predicted RNA binding sites for splicing regulators are not functional. What specific experiments lead us to this conclusion? Why (biochemically) are these sites non-functional? Why has evolution allowed this to happen?
- Describe the seven main types of alternative splicing, including an outline of the mechanism(s) by which each type is regulated. What is the biological significance of each type, in terms of specific change(s) in gene function caused by the splice variant?

Discussion Questions (continued)

- Discuss the "speed bump" model of RNA splicing. What is the evidence in favor of this model?
- Discuss the regulation of meiosis in yeast by RNA splicing.
- Discuss the regulation of the EMT transition at the level of alternative RNA splicing.
- Discuss the biological significance of alternative promoter usage. Is this regulated by RNA binding proteins, transcription factors, both, or neither? Why?