RNA Processing: Splicing I (Basic)

Reviews:  Licatalosi & Darnell, Nat Rev Genet 11:75-87 (2010)
          Luco et al., Cell 144: 16-26 (2011)

Eukaryotic RNA Production and Processing

(A) EUCARYOTES

Figure 6–21 part 1 of 2. Molecular Biology of the Cell, 4th Edition.
Taking Care of the Ends: 5’-End Capping

5’ end of nascent RNA transcript
5’ pppNpNp 3’

- GTP

7-methylguanosine

5’-to-5’ triphosphate bridge

add methyl group to base

add methyl group to ribose (only on some caps)

Figure 6–22 part 2 of 2. Molecular Biology of the Cell, 4th Edition.
And Between the Ends... The Two-Step Splice of Life

Figure 6-26 part 1 of 2. Molecular Biology of the Cell, 4th Edition.

Figure 6-26 part 2 of 2. Molecular Biology of the Cell, 4th Edition.
Annealing RNA from virus-infected cells with viral DNA revealed the existence of multiple introns (transcribed regions of the DNA that are not included in mature mRNA).
The giant elastic protein titin is encoded by a single gene located in human and mouse on chromosome 2. The human titin gene is composed of 363 exons that code for a total of 38,138 amino acid residues (mol. mass 4200 kDa).
Some Clues About Splice Sites

Conserved sequence motifs at the 5’ (A) and 3’ (C) splice sites and the branch point region (B)

• Produce a labeled RNA transcript by in vitro transcription

• Incubate in HeLa nuclear extract with ATP
What Cellular Factors Are Required for Splicing

Discovery of snRNPs:
Lerner and Steitz (1979) Antibodies to small nuclear RNAs complexed with proteins are produced by patients with systemic lupus erythematosus, *PNAS* 76: 5495 (1979)

Figure 22.8 U1 snRNA has a base paired structure that creates several domains. The 5' end remains single-stranded and can base pair with the 5' splicing site.

![Diagram of RNA splicing](Figure 6-33 Molecular Biology of the Cell, 4th Edition.)
snRNP-mRNA and snRNP-snRNP Interactions

• How do you show that splicing requires U1 snRNA basepairing with pre-mRNA?
RNA-RNA Interactions are Essential to Splicing

Mutations were introduced into the 12S splice site, which disrupted basepairing with U1 and abolished splicing.

Compensatory mutations were introduced into U1, which restored basepairing and splicing.

Conclusion: basepairing between U1 and the 5' splice site is required for splicing.

snRNPs Are Reshuffled During Splicing
Non-snRNP Factors Are Also Required for Splicing

**Figure 22.14** Spliceosomes are ellipsoidal particles with several discrete regions. The bar is 50 nm. Photograph kindly provided by Tom Maniatis.

- **U2AF** – U2 Auxiliary Factor
- **RRM** – RNA Recognition Motif
- **SF1/BBP** – Splicing Factor 1/Branch point Binding Protein
- **RS** – aRginine/Serine rich motif
SR (or RS) Proteins

structural motifs in SR proteins

<table>
<thead>
<tr>
<th>RRM</th>
<th>X</th>
<th>RS</th>
</tr>
</thead>
<tbody>
<tr>
<td>RNA recognition motif</td>
<td>arginine/serine-rich domain</td>
<td></td>
</tr>
</tbody>
</table>

exon-dependent functions

SR proteins bind to exons and enhance splicing of adjacent introns

regulated 3’ splice site selection

(U2AF65 binds the polypyrimidine tract)

regulated 5’ splice site selection

splicing enhancer

Graveley RNA 6, 1197(2000)

exon-independent functions

(U2AF65 binds the polypyrimidine tract)
How Complex is the Spliceosome?

In vitro derived spliceosome assembly cycle

Answer: 5 U snRNPs and hundreds of proteins

One technique used to purify spliceosomes


Only One Type of Spliceosome?

Discovery of the AT-AC spliceosome


**A**

0 0.04 0.12 0.4 1.2 μM 2′-O-methyl oligo U2b

- pre-mRNA
- cryp (○)
- E1-E2 (●)

**B**
P120L

GAGCAGG auauccuugcagggcagaga (57 nt) → caguucuuaacgcccac AUGCCA

E1-E2 (●)

**Figure 1**

Spliceosome assembly and action. (a) The major spliceosome and (b) the AT–AC spliceosome are pictured at an early stage of spliceosome assembly and after the first reaction step (formation of the lariat intermediate) has occurred. U11 and U12 are pictured as entering the spliceosome as a two-snRNP complex. After joining of the U4–U6–U5 or U4atac–U6atac–U5 tri-snRNP, a conformational change occurs that loosens the association of U1 or U11 and U4 or U4atac with the spliceosome.

Is the Spliceosome an RNA Enzyme?

Yes: Splicing-related catalysis by protein-free snRNAs

RNA Processing: Splicing II (Alternative)

Why alternative splicing?

Figure 6–27. Molecular Biology of the Cell, 4th Edition.
How complicated can alternative splicing be?

38,000 potential Dscam isoforms
Types of Alternative Splicing

Faustino & Cooper, Genes & Dev. 17: 419 (2003)
Control of Splice Site Selection: Role of RNA Regulatory Regions
SR Proteins and Exonic Splicing Enhancers in Splice Site Selection

ESE: Exonic Splicing Enhancer
ESS: Exonic Splicing Silencer
ISE: Intronic Splicing Enhancer
ISS: Intronic Splicing Silencer

Role of Enhancers and Silencers in Alternative Exon Splicing

![Diagram of alternative exon splicing with enhancers and silencers]
How Does the Splicing Machinery Recognize Introns Versus Exons?

Pre-mRNA → Regulated Exon → Intron → Exon → mRNA

- Binding of U1 and U2 at Splice Sites and Initial Spliceosome Assembly Influenced by Regulatory Proteins
- Further Spliceosome Assembly, Structural Rearrangements, and Catalysis of Intron Excision

U1 snRNP → U2 snRNP and Auxiliary Factors → Positive Regulatory Protein (eg. SR Proteins)

Exon Definition Complex

Commitment Complex

Complexes across intron

Pre-spliceosome complex

Spliceosome

ATP

US, U4/U6 snRNPs

mGppG

AAAAA
Identification of Splicing Factors II

In vivo assays

Transfection analysis

Alternative Splicing in *Drosophila* Sex Determination

**tra pre-mRNA**

In males, the splicing factor U2AF binds to the proximal 3' splice site, leading to an mRNA containing a premature translational stop codon (UAG). In females, SXL binds to the proximal 3' splice site, thus preventing the binding of U2AF. Instead, U2AF binds to the distal 3' splice site, leading to an mRNA that encodes functional TRA protein.

**sxl pre-mRNA**

Alternative inclusion of exon 3 of *sxl* pre-mRNA is regulated by SXL protein. In both males and females, the first step of the splicing reaction results in lariat formation at the branchpoint sequence upstream from the 3' splice site preceding exon 3. Subsequently, the second-step splicing factor SPF45 binds to the AG dinucleotide of this splice site. In males, SPF45 promotes the second step of the splicing reaction, leading to the inclusion of exon 3. In females, SXL binds to a sequence upstream of the AG dinucleotide, interacts with SPF45 and inhibits its activity. This prevents the second step of the splicing reaction, leading to the exclusion of exon 3 and splicing of exon 2 to exon 4.

**dsx pre-mRNA**

Alternative splicing of *dsx* pre-mRNA is regulated by the assembly of heterotrimeric protein complexes on female-specific ESEs. The first three exons are constitutively spliced in both sexes. In males, the 3' splice site preceding exon 4 is not recognized by the splicing machinery, resulting in the exclusion of this exon, and splicing of exon 3 to exon 5. In females, the female-specific TRA protein promotes the binding of the SR protein RBP1, and the SR-like protein TRA2 to six copies of an ESE (indicated by green rectangles). These splicing enhancer complexes then recruit the splicing machinery to the 3' splice site preceding exon 4, leading to its inclusion in the mRNA. In females, polyadenylation (pA) occurs downstream of exon 4, whereas in males it occurs downstream of exon 6. ‘S’ designates the splicing machinery.
RNA Structures Are Also Important in Splicing Regulation

- Tau is a microtubule (MT)-associated protein (MAPT is the gene) required for MT polymerization/stability and neuronal axonal transport
- Six tau isoforms are produced by alternative splicing of exons 2, 3 and 10
- In the normal human brain, the 4R/3R ratio is ~1:1
- Fronto-temporal dementia with Parkinsonism linked to chr 17 (FTDP-17) is sometimes caused by mutations that alter exon 10 splicing and the normal 4R/3R ratio

**Six tau isoforms:**

**ESE, ESS and U1 binding site mutations**
Epigenetics and Splicing

chromatin-adaptor model

chromatin-adaptor complexes

Cell 144: 16-26 (2011)
Epigenetics and Splicing

RNA

RNA secondary structure

RNA motifs

Splicing factor abundance and modifications

RNA Pol II elongation rate

Splicing factor recruitment by RNA Pol II CTD

RNA Pol II

Chromatin-adaptor complex

Transcription factors and chromatin remodelers

Nucleosome positioning

CHROMATIN
Global Detection of Alternative Splicing: Microarrays

**a Microarray Screen**

- Microarray screen
- Exon probes
- Exon-junction probe for predicted alternative splice

**Sample 1 = Sample 2**
**Sample 1 > Sample 2**
**Sample 1 < Sample 2**

**b Validation by RT-PCR**

- Validation by RT-PCR
- Sample 1
- Sample 2
- + exon
- – exon

---

Nature Reviews | Genetics
Outline of RNA-Seq procedure.

(a) After two rounds of poly(A) selection, RNA is fragmented to an average length of 200 nt by magnesium-catalyzed hydrolysis and then converted into cDNA by random priming. The cDNA is then converted into a molecular library for Illumina/Solexa 1G sequencing, and the resulting 25-bp reads are mapped onto the genome. Normalized transcript prevalence is calculated with an algorithm from the ERANGE package. (b) Primary data from mouse muscle RNAs that map uniquely in the genome to a 1-kb region of the Myf6 locus, including reads that span introns. The RNA-Seq graph above the gene model summarizes the quantity of reads, so that each point represents the number of reads covering each nucleotide, per million mapped reads (normalized scale of 0–5.5 reads). (c) Detection and quantification of differential expression. Mouse poly(A)-selected RNAs from brain, liver and skeletal muscle for a 20-kb region of chromosome 10 containing Myf6 and its paralog Myf5, which are muscle specific. In muscle, Myf6 is highly expressed in mature muscle, whereas Myf5 is expressed at very low levels from a small number of cells. The specificity of RNA-Seq is high: Myf6 expression is known to be highly muscle specific, and only 4 reads out of 71 million total liver and brain mapped reads were assigned to the Myf6 gene model.
Nova-dependent Splicing Regulation: Towards a Splicing Code?

RNA map predicts Nova-dependent splicing regulation in 30/30 exons

Splicing enhancers

Splicing silencers

Nova regulates assembly of the early spliceosomal complex

RNA map predicts asymmetric Nova action in 19/19 pre-mRNAs
Splicing and Disease

A. SMN1 and SMN2

<table>
<thead>
<tr>
<th>Gene / pre-mRNA</th>
<th>mRNA / Protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMN1</td>
<td>SMN Normal</td>
</tr>
<tr>
<td>SMN2</td>
<td>SMNΔ7 Spinal muscular atrophy (SMA)</td>
</tr>
</tbody>
</table>

B. Dystrophin

<table>
<thead>
<tr>
<th>Dystrophin</th>
<th>mRNA / Protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>Duchenne muscular dystrophy (DMD)</td>
</tr>
</tbody>
</table>

C. MAPT

<table>
<thead>
<tr>
<th>MAPT</th>
<th>mRNA / Protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>4R tau = 3R tau Normal</td>
<td>4R tau &gt; 3R tau Frontotemporal dementia (FTDP-17)</td>
</tr>
</tbody>
</table>

D. CFTR

<table>
<thead>
<tr>
<th>CFTR</th>
<th>mRNA / Protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cystic fibrosis (mild)</td>
<td>Cystic fibrosis (severe)</td>
</tr>
</tbody>
</table>
Therapies for Splicing-mediated Disease

A. Antisense oligonucleotides (AOs)

B. snRNAs as vehicles for antisense RNA

C. RNA interference (RNAi)