MAJOR MESSAGES

Splice site consensus sequences do have sufficient information to uniquely direct spliceosomal trans-esterification events

Splicing enhancers and silencers appear to provide additional recognition information

“Exon definition” can enable the recognition of introns that are large

Splicing enhancers and silencers enable alternative splicing

Recursive splicing involving zero-length exons breaks up seemingly giant introns into shorter introns

There is a second spliceosome in many species, including humans, the ‘minor spliceosome.’

Introns often encode important noncoding RNAs including snoRNAs and miRNAs

Stalled splicing is used to generate 3’ ends of telomerase RNAs in fungi and to target transposable element RNAs for RNAi-dependent silencing in C. neoformans
The question: Why are there no described free-living eukaryotes that lack introns?

Model 1: Intron gain mechanisms outpace exceeds intron loss mechanisms (nonadaptive evolution).

Model 2: Evolution has found uses for splicing (adaptive evolution).

ALTERNATIVE SPLICING

Additional information in introns and exons

*S. cerevisiae* 5’ splice site and branchpoint sequences conform closely to the optimal consensus sequence. Beyond its close relatives (the hemiascomycetes), other eukaryotes (from protists to man) do not share this property.

Additional sequences that influences 5’ splice site, branchpoint and 3’ splice site usage exist in pre-mRNAs. These splicing enhancers (ESEs, ISEs) and silencers (ESSs, ISSs) are bound by sequence-specific RNA binding proteins (RBPs). There about 1500 RBPs in humans. Perhaps ~100 are involved in splicing.

ESE, exonic splicing enhancer; ESS, exonic splicing silencer; ISE, intronic splicing enhancer; ISS, intronic splicing silencer.
The exon definition model
Berget and colleagues observed that the mutation of a 5' splice site could prevent the splicing of the upstream intron in vitro. Because exons are short (170 nts on average in humans) while introns are longer (median length of human introns is 1334 bp), she proposed that exons were the initial unit of recognition (“exon definition”) and that enabled subsequence recognition of the upstream intron (“intron definition”).

The SR protein family -- major activators of splicing
One major class of proteins that recognize ESEs in exons are so called SR proteins, named because they harbor serine-arginine dipeptides repeats in clusters (SRSF1-12). These are phosphorylated and seem to function as protein-protein interaction domains with other SR domains. They bind SR-domains in U2AF and U1 snRNP to enhance their binding to weak consensus sequences. There is no structural information on phosphorylated SR domains, but would appear to have the potential for ionic interactions.
Cross-exon complexes can be assembled in vitro and converted into cross-intron, splicing-competent complexes. This can be accomplished by adding a 5' splice site-containing oligonucleotide in trans, which can undergo the first chemical step of splicing. The reconfiguration of exon-definition complexes to catalytically active splicing complexes is associated with the loss of U1 snRNP and SR proteins and the gain of several proteins including some triple snRNP-specific proteins.
Splicing enhancers and silencers mediate alternative splicing regulation

Alternative splicing occurs in several categories with the cassette exon and retained intron being the most common events. Cassette exons often contain stop codons leading to nonsense-mediated decay (NMD), providing a mechanisms by which a splicing switch can lead to RNA turnover.
Polypyrimidine tract binding proteins as a paradigm for splicing silencers

Splicing enhancers and silencers enable tissue-specific alternative splicing. One of many examples involves silencer proteins PTPB1 and PTBP2. These form a circuit that controls alternative in developing central neurons in mammals. PTBPs appear to act through multiple mechanisms. One characterized mechanism by which PTBP1 silences a neuron-specific exon of the Src kinase. In vitro studies show that PTBP1 binds U1 snRNP and prevent exon definition, leading to exon skipping.
“Recursive” splicing of long introns
Analysis of RNA-seq data in flies and mammals revealed a sawtooth pattern across many long introns. This was found to reflect exons of size zero, which are undetectable in mRNA. These findings extended a handful of earlier examples identified in Drosophila. It has been suggested that such ‘recursive’ splicing enables the splicing of unusually ‘long’ introns.
The parallel universe of the ‘minor’ spliceosome
There is a second spliceosome in many species, including humans, the ‘minor spliceosome.’ In this spliceosome, all of the snRNPs except for U5 are replaced by different ones (U11, U12, U4ATAC, and U6ATAC). These introns have distinct splice sites but can be intermingled with ‘major spliceosome’ introns. About 0.5% of human introns are of this minor type. The minor spliceosome is widely distributed among eukaryotes arguing it was present in the eukaryotic ancestor but lost in many lineages. In humans mutations in U4ATAC cause microcephalic osteodysplastic primordial dwarfism (MOPD).
Genes in which the INTRONS encode functional molecules

Introns can encode noncoding RNAs. In eukaryotes from yeast to man, snoRNAs, which guide rRNA modifications, are typically processed from introns. In one case in yeast, multiple snoRNAs are encoded by the introns and the exons code for NOTHING.

Hundreds of microRNAs in humans are encoded in introns. These are termed 'mirtrons.' Disease-associated mutations in the spliceosome (e.g. acute leukemia) impact the expression of both mRNA-coding genes and mirtron-coding genes.
Biological roles for stalled spliceosomes

Telomerase RNA in *S. pombe* acquires its 3’ end through step 1 of splicing without step 2. This mechanism is widely conserved in fungi. It seems possible that step 1-only splicing is a widespread mechanism for the generation of RNA 3’ ends.

![Diagram of spliceosome steps](image)

**C**

*S. pombe*  
GUAUGU UACUAC A UAGAU GAGAG  
Wind distance  
Perfect branch site

*S. cryophilus*  
GUCAGU GACUAAC AG  
Hyperstable

*S. octosporus*  
GUCAGU GACUAAC AG  
Hyperstable

*A. niger (other Aspergilli sp.)*  
AUAAAGU GCCUAAC AG  
Disrupts interaction
In the intron-rich yeast *Cryptococcus neoformans*, stalled spliceosomes are used to distinguish transposable element pre-mRNAs from 'normal' cellular RNAs. Stalling enables a spliceosome-associated RNA-dependent RNA polymerase complex called SCANR to convert splicing precursors and/or intermediates into double-stranded RNA which is then processed to siRNAs used for transposon silencing.