Alternative Splicing

Objective: To give a presentation of about 60 minutes at the end of the week covering the key aspects of the alternative splicing.

Alternative splicing (AS) was discovered in 1978, and in the subsequent 10-15 years was viewed mostly as a curiosity: an interesting way to generate several proteins from one gene (Ast, 2004). With the advent of large scale genome sequencing and EST determination, it has become clear that a very large percentage of genes are alternatively spliced. A key goal of bioinformatics is to predict as much as possible of the behaviour of a biological system by computational means using available knowledge. Presently, we can only predict coarse features of a gene from the sequence alone, although progress in this field is steady. Obviously, AS increases the variety of proteins encoded by the genome. For the researcher it creates the challenge of determining which of this is functional, which is tolerated noise, and which is directly detrimental or advantageous novelties. This is analogous to previous debates on the selective value of observed sequence variation within a species and molecular differences between species.

The figure below illustrates the relationship between transcripts and the ASG (Graph). The solid straight line represents a genic region of DNA. The dashed lines represent “jumps” that would not be part of a messenger RNA transcript, if that jump was selected; that is, the left endpoints of the dashed lines correspond to donor sites, and the right endpoints of the dashed lines corresponds to acceptor sites. A given ASG can generate all possible transcripts by traversing from left to right selecting all possible routes through the graph. Assuming that all intrinsic regions can be retained, the graph below can generate 18 different transcripts.

The Big Questions Are:
How widespread is alternative splicing?
What is its distribution on the tree of life?
What is its distribution within the genome?
How did it arise?
How does AS evolve?
How do you measure functional constraint on AS?
How much is accepted noise and how much functional diversity?
How do you measure the extent of AS?
How does one predict AS computationally?

Maximal Contents of Presentation
The scientific history of AS
The molecular biology of AS
AS on the tree of life
AS within different genes
The role of AS in disease
The present challenges in analyzing AS

Possible AS literature