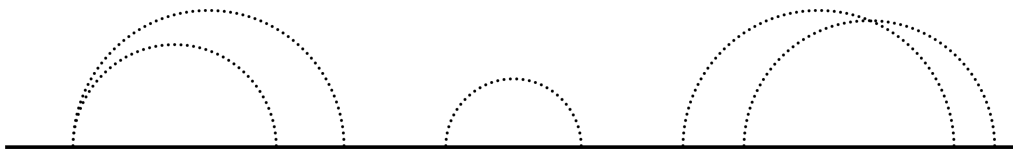


Alternative Splicing

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 intronic 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

- Ast, G. (2004) How did alternative splicing evolve? *Nat. Rev. Genet.* **5**(10): 773-82.
- Boue, S., Vingron, M., Kriventseva, E. and Koch, I. (2002) Theoretical analysis of alternative splice forms using computational methods. *Bioinformatics* **18** suppl. 2: S65-73.
- Bray, N. and Pachter, L. (2004) MAVID: constrained ancestral alignment of multiple sequences. *Genome Res.* **14**(4): 693-9.
- Chatterji, S. and Pachter, L. (2006) Reference based annotation with GeneMapper. *Genome Biology* **7**: R29.
- Mo Chen and James L. Manley (2009) Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches NATURE REVIEW S *Molecular cell Biology* **10**.11
- Jenkins, P., Lyngso, R.B. and Hein, J. (2006) How Many transcripts does it take to reconstruct the alternative splicing graph. In press.
- Kan, Z., Rouchka, E.C., Gish, W.R. and States, D.J. (2001) Gene Structure Prediction and Alternative Splicing Analysis Using Genomically Aligned ESTs. *Genome Res.* **11**(5): 889-900.
- Ladd and Cooper (2002) Finding signals that regulate alternative splicing in the postgenomic era *Genome Biology* **2002**, 3(11)Leipzig, J., Pevzner, P. and Heber, S. (2004) The Alternative Splicing Gallery (ASG): bridging the gap between genome and transcriptome. *Nucleic Acids Res.* **32**(13): 3977-83.
- Maniatis, T. and Tasic, B. (2002) Alternative pre-mRNA splicing and proteome expansion in metazoans. *Nature* **418**(6894): 236-43.
- Matlin, A.J., Clark, F. and Smith, C.W. (2005) Understanding Alternative Splicing: Towards a Cellular Code. *Nat. Rev. Mol. Cell. Biol.* **6**(5): 386-98.
- Modrek, B. and Lee, C. (2002) A Genomic view of Alternative Splicing. *Nature Genetics* **30**(1): 13-9.
- Moore and Silver (2008) Global analysis of mRNA splicing *RNA*. **14**: 197-203
- Pan, et al. (2008) Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nature Genetics* **40**.12.1413-
- Sorek, R., Shamir, R. and Ast, G. (2004) How prevalent is functional alternative splicing in the human genome? *Trends Genet.* **20**(2):68-71.
- Xing, T.Yu, Y.N.Wu, M.Roy, J.Kim & C.Lee (2006) An expectation-maximization algorithm for probabilistic reconstructions of full-length isoforms from splice graphs. *Nucleic Acids Res.* **34**(10): 3150-60
- Xing, Y. and Lee, C. (2006) Alternative Splicing and RNA Selection pressure – evolutionary consequences for eukaryotic genomes. *Nat. Rev. Genet.* **7**(7): 499-509.
- Wang and Cooper (2007) Splicing in disease: disruption of the splicing code and the decoding machinery *Nature Review Genetics* **8**.10.749-