

Sequence and Structural Patterns in RNA

Prof. Michael Gribskov
Purdue University

Asst. Prof. T. Murlidharan Nair,
Departments of Biology and Computer Science/Informatics
Indiana University South Bend,
1700 Mishawaka Ave,
South Bend, IN 46634-7111
574-520-5068
mnair@iusb.edu

NSF award: \$620,091

Recent estimates suggest that more than 60% of human genes are alternatively spliced. Alternative splicing is considered as one of the primary mechanisms contributing to the complexity of the eukaryotic proteome. While the basic mechanism of RNA splicing via the spliceosome complex is understood in outline, many details remain to be elucidated. One of the most important questions is how the splicing process is regulated so as to produce alternatively spliced forms in specific tissues and/or developmental states. The signals involved are relatively subtle this is clear from the difficulty in computationally predicting gene models. Recent studies on the mechanism of alternative splicing suggest that alternatively spliced (AS) junctions are weaker, i.e. that the conserved donor and acceptor sequences are more weakly conserved, than constitutively spliced (CS) junctions and that AS exons are shorter than in CS exons. Unexpectedly, AS exons and their flanking introns have been found to be more highly conserved than CS exons, and although AS junctions have been found to be weaker, these junctions have a greater number of nearby exonic splicing silencer (ESS) elements than CS exons. We are interested in using machine learning and other statistical methods to identify signals (structural or sequence based) that explain the control of alternative splicing.