Alternative Splicing: Session Introduction

H. Wang, U. Yang, C. Lee, and J. Blume

Pacific Symposium on Biocomputing 9:3-4(2004)

ALTERNATIVE SPLICING

H. WANG
Affymetrix Inc., Bioinformatics, 3380 Central Expressway, Santa Clara, CA 95051
hui_wang@affymetrix.com

C. LEE
Molecular Biology Institute, University of California, Los Angeles, CA 90095-1570
leec@mbi.ucla.edu

Human studies have estimated that approximately 30-60% of genes undergo alternative splicing and it has been shown that alternative splicing is an important regulatory mechanism often controlled by developmental or tissue-specific factors. Additionally, alternative splicing of a single gene sometimes produces functionally distinct proteins within the same tissue and, in some cases, gene isoforms have been associated with human diseases.

Alternative splicing is an important topic to cover in a conference concerning computational biology because, 1) it plays a significant role in human disease, 2) a wealth of data has become available and 3) recent technological advances make the problem of alternative splicing detectable by assays.

Alternative splicing plays a significant role in physiology and disease. Alternative splicing is a mechanism for generating a versatile repertoire of different proteins, perhaps with distinct functions, within individual cells. Its significance is clearly evident in highly specialized cells such as neurons. While the mechanism of splicing is becoming well-understood, our understanding of alternative splicing is just starting to emerge. Several new cis and trans splicing factors are suggested to be related to alternative splicing. Abnormal splicing can cause severe diseases. Many researchers are trying to determine the cause of aberrant splicing and to understand its disease association. Recent literature has implicated alternative splicing in a vast array of neurological diseases. Some studies describe a clear link between a mis-spliced gene and a disease state, while others imply an increased risk for a range of problems which include little understood diseases and general physiological changes.

A wealth of public data has become available. Recent bioinformatics efforts on splice variant analysis have enabled the discovery of splice variants.

Database mining techniques, such as EST clustering correlating with EST library information, offer enormous value in new variant as well as tissue and/or disease specific variant discovery. The availability of human, mouse, and rat genomes has expedited the discovery process. Most of the current studies are conducted either by aligning EST and mRNA sequences to genomes and/or by statistical pattern search approaches. Many alternative splicing related databases and tools have been developed.

Recent measurement technology advances have enabled analysis of splice variants. While microarray technology has become a standard method for gene expression profiling, most microarray design and analysis is limited to detecting and measuring changes of expression on a per gene basis. Being able to measure variant-level expression is important for accurate expression profiling, and consequently for obtaining a better understanding of the biological processes. Recently, several studies have applied microarray technology to this issue. Genomic tiling arrays and exon arrays can be used to identify co-regulated exons, which allows the inference of variant mixtures. Expression arrays with multiple probes have been retrospectively analyzed to identify exons that are differentially included or skipped in a tissue-specific manner. RNA-mediated ligation combined with arrays presents a novel method for detecting exon-exon junction information of known splice variants. Recently splice junction spanning oligonucleotides representing nearly all yeast splicing events have been used to monitor the genome-wide effects of splicing factor mutations in yeast, suggesting exon joining information can be accessed using oligonucleotide arrays. More recently, a human splice variant oligonucleotide microarray has been designed and an algorithm has been developed to deconvolute the absolute concentrations of each splice variant in the variant mixture. All of these efforts have shed lights on the future of variant specific expression analysis.

In this first PSB alternative splicing session we highlight work which studies several aspects of this interesting phenomenon. First we see how Sugnet et al and Kan et al develop comparative genomic methods using human and mouse genomes for the detection of alternative splicing. Then we see how Modrek et al applies a partial ordered alignment program for genome-wide splice variant detection, in seeking of minimizing both false positive and false negative errors when using EST data. Buhler et al present an alternative assay for the high throughput detection and quantification of splice variants using polymerase colonies. Sakai at el has tried to identify splicing pattern specific regulatory sequences. Zhang et al has developed a manually annotated alternatively spliced event database and an annotation system. Finally, Cline et al assesses the functional consequences of alternative splicing of transmembrance proteins.