
Bioinformatics of Alternative Splicing in the Nervous System

(EIA-0218506-FY02)
Christopher B. Burge
Massachusetts Institute of Technology (MIT)

Almost every human cell contains a huge instruction manual called the genome with many thousands of pages (the genes), each of which tells the cell how to make a particular building block (protein) that it needs to live or grow or to perform its assigned function in the body. The cell uses this manual in a complicated way, first copying (transcribing) each page that it needs to a piece of scratch paper (the pre-mRNA), and then cutting and pasting (splicing) pieces of the scratch paper (the exons) together to form the final recipe (mRNA) for the protein product. Interestingly, this cutting and pasting is often carried out in different ways in different types of cells or under different conditions in a process called alternative splicing (AS), generating many different varieties of a protein under different conditions. Alternative splicing is particularly common in neurons, helping to generate protein variants whose properties are optimized to the local environment of the neuron. For example, AS is used to tune the electrical properties of ion channels which help different sensory neurons in the inner ear respond to different frequencies of sound. In addition, mutations that affect AS are associated with a number of neurodegenerative diseases. In order to paint a global picture of AS in the nervous system, we focus on events that are conserved since the divergence of human and mouse, as they are likely of primary biological importance, but relatively few such events are known. Here we describe sequence features that distinguish exons subject to evolutionarily conserved AS, which we call 'alternative-conserved exons' (ACEs), from other orthologous human/mouse exons and integrate these features into an exon classification algorithm, ACEScan. Genome-wide analysis of annotated orthologous human-mouse exon pairs identified ~2,000 predicted ACEs. Alternative splicing was verified in both human and mouse tissues using an RT-PCR-sequencing protocol for 21 of 30 (70%) predicted ACEs tested, supporting the validity of a majority of ACEScan predictions. By contrast, AS was observed in mouse tissues for only 2 out of 15 (13%) tested exons that had EST or cDNA evidence of AS in human but were not predicted ACEs, and was never observed for eleven negative control exons in human or mouse tissues. Such large-scale identification of ACEs now enables us to develop splicing-specific arrays to interrogate the cell and tissue-specific regulation of such events with regards to the nervous system.

PI Website

http://genes.mit.edu/burgelab

Publications

Yeo, G., Van Nostrand, Holste, D., E., Poggio, T. and Burge, C. B. (2005). "Identification and analysis of alternative splicing events conserved between human and mouse". *Proc. Natl. Acad. Sci USA* **102,** 2850-2855.

Pfarr, N., Prawitt, D., Kirschfink, M., Schroff, C., Knuf, M., Habermehl, P., Mannhardt, W., Zepp, F., Fairbrother, W., Loos, M., Burge, C. B. and Pohlenz, J. (2005). "Linking C5 deficiency to an exonic splicing enhancer mutation." *J. Immunol.* **174,** 4172-4177.
Wang, Z., Rolish, M., Yeo, G., Tung, V., Mawson, M. and Burge, C. B. (2004). "Systematic identification and analysis of exonic splicing silencers." *Cell* **119,** 831-845.

Yeo, G., Holste, D., Kreiman, G. and Burge, C. B. (2004). "Variation in alternative splicing across human tissues" *Genome Biol.* **5,** R74.

Yeo, G., Hoon, S., Venkatesh, B. and Burge, C. B. (2004). "Variation in sequence and organization of splicing regulatory elements in vertebrate genes." *Proc. Natl. Acad. Sci USA* **101,** 15700-15705.

Fairbrother, W. G., Holste, D., Burge, C. B. and Sharp, P. A. (2004). "Single nucleotide polymorphism-based validation of exonic splicing enhancers." *PLoS Biol.* **2,** e268.

Eng, L., Coutinho, G., Nahas, S., Yeo, G., Tanouye, R., Babaei, M., Dork, T., Burge, C., and Gatti, R. A. (2004). "Nonclassical splicing mutations in the coding and noncoding regions of the ATM gene: maximum entropy estimates of splice junction strengths." *Human Mutat.* **23,** 67-76.

Yeo, G. and Burge, C. B. (2004). "Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals." *J. Comp. Biol.* **11,** 377-394.