7-8-2008

Identifying regulatory elements and gene boundaries with comparative genomic sequence analysis

Jim Lund
University of Kentucky

Poster presentation

Identifying regulatory elements and gene boundaries with comparative genomic sequence analysis

Jim Lund

Address: Department of Biology, University of Kentucky, Lexington, KY 40506, USA

Email: Jim Lund - jiml@uky.edu

For several groups of organisms multiple genomes have been sequenced either completely or to the unfinished shotgun sequencing stage. Comparative sequence analysis can identify functional sequence elements for pairs of organisms that diverged far enough in the past to allow mutational drift of non-conserved sequences. While analysis of a pair of genomes identifies many of these functional elements adding additional genomes allows additional information to be elicited.

Additional conserved sequence elements are identified as additional genomes are analyzed. These conserved sequence elements are often regulatory elements although they are difficult to classify with *in silico* analysis. For an individual gene, the set of associated conserved sequence elements and the organisms they are found in provides insight into the evolutionary history of the regulation of the gene.

The eight complete and unfinished shotgun sequenced nematode genomes and the dozen informative insect genome sequences were used to analyze conserved non-coding sequences in these groups of organisms. I have developed web-based software to allow researchers to explore and visualize sequence conservation that differs from previous work in analyzing conserved sequences in each pair of organisms rather than with respect to a single reference genome.

We have used this analysis to identify gene regulatory boundaries in the nematode *C. elegans*. The genomes of *C.* *elegans* and other nematodes have diverged enough that synteny regions are typically one or two genes long. I was able to associate conserved sequence elements to particular genes identifying the natural boundaries between genes and the extent of worm promoters.