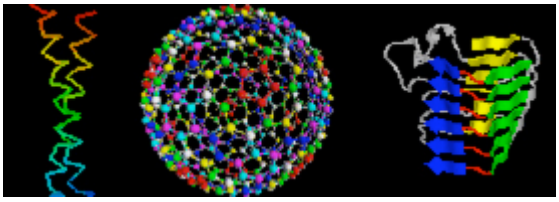
MIT
Department of Mathematics
& The Theory of
Computation Group
At CSAIL



Bioinformatics Seminar

Speaker: Elizabeth Thomas, Bauer Center for Genomics Research
Harvard University
Title: Doublets: a new player in genome evolution
Date: Monday, 14 March 2005
Time & Location:
Refreshments: 11 am in the Theory of Computation Lab at MIT's Building 32, Stata Center Room G-575
Talk: 11:30 am the Theory of Computation Lab at MIT's Building 32, Stata Center, Room G-575
URL: http://www-math.mit.edu/compbiosem/

Abstract:

Mammalian genomes are densely populated with long duplicated sequences. In this work, we demonstrate the existence of doublets, short duplications between 25 and 100 bp, distinct from previously described repeats. Each doublet is a pair of exact matches, separated by some distance. The distribution of these inter-match distances is strikingly nonrandom. An unexpectedly high number of doublets have matches either within 100 bp (adjacent) or at distances tightly concentrated ~1,000 bp apart (nearby). We focus our study on these proximate doublets. First, they tend to have both matches on the same strand. By comparing nearby doublets shared in human and chimpanzee, we can also see that these doublets seem to arise by an insertion event that produces a copy without markedly affecting the surrounding sequence. Most doublets in humans are shared with chimpanzee, but many new pairs arose after the divergence of the species. Doublets found in human but not chimpanzee are most often composed of almost tandem matches, whereas older doublets (found in both species) are more likely to have matches spaced by {approx}1 kb, indicating that the nearly tandem doublets may be more dynamic. The spacing of doublets is highly conserved.
So far, we have found clearly recognizable doublets in the following genomes: Homo sapiens, Mus musculus, Arabidopsis thaliana, and Caenorhabditis elegans, indicating that the mechanism generating these doublets is widespread. A mechanism that generates short local duplications while conserving polarity could have a profound impact on the evolution of regulatory and protein-coding sequences.

The seminar is co-hosted by Professor Peter Clote of Boston College's Biology and Computer Science Departments and MIT Professor of Applied Math Bonnie Berger. Professor Berger is also affiliated with CSAIL & HST.

Massachusetts Institute
of Technology
77 Massachusetts Avenue
Cambridge, MA 02139

For General Questions, please contact kvdickey@mit.edu