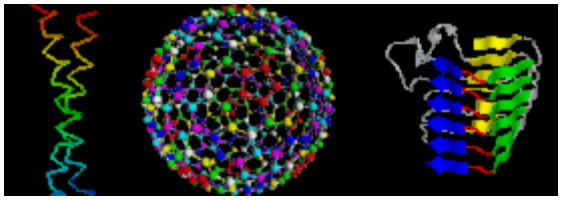
MIT
Department of Mathematics
& The Theory of
Computation Group
At CSAIL



Bioinformatics Seminar

Speaker: Gary Benson, Boston University
Title: Searching for Inverted Repeats in Genomic Sequences
Date: Monday, 3 May 2004
Time & Location:
Refreshments: 11:00 am in the Applied Mathematics Common Room at MIT's Building 2, Room 349
Talk: 11:30 am The Applied Mathematics Conference Room Building 2, Room 338
URL: http://www-math.mit.edu/compbiosem/

Abstract:

Inverted repeats (IRs) consist of two arms of homologous DNA – with one inverted and complemented relative to the other – around a central spacer region, and are in theory capable of forming stem-loop structures through pairing of the arms. Short IR pairing forms secondary structure in RNA sequences. Longer IR pairing in DNA can in principle form cruciform structures, which would be facilitated by negative DNA supercoiling. IRs have been associated with DNA replication, genomic instability, SMC protein binding and microRNA expression. Until recently, there has been no efficient program for detecting IRs in long genomic sequences.

In this talk, I discuss recent work on a program, Inverted Repeats Finder (IRF), for finding approximate inverted repeats. IRF works, in many ways, like an earlier program, Tandem Repeats Finder. Candidate IRs are detected by finding short, exact, reverse-complement matches of 4-7 nucleotides (k-tuples) between non-overlapping fragments of a sequence. Unlike the case for tandem repeats, the distance between matching tuples is not related to the size of the repeat. This makes for some interesting differences in the approach to candidate recognition.

The first application of IRF has been analysis of the recently completed human genome sequence (hg16). After masking of know repetitive elements, IRF detected 22,429 human IRs in 6 hours on a desktop PC. 159 IRs had arm lengths >8kb with a significant overabundance of these large IRs on the X and Y chromosomes. The presence on the Y chromosome of large, highly homologous IRs, which often contain genes expressed predominantly in testes, was recently described. Our analysis shows that there are similar large, homologous IRs on the X chromosome and that many of these also contain testes genes.

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