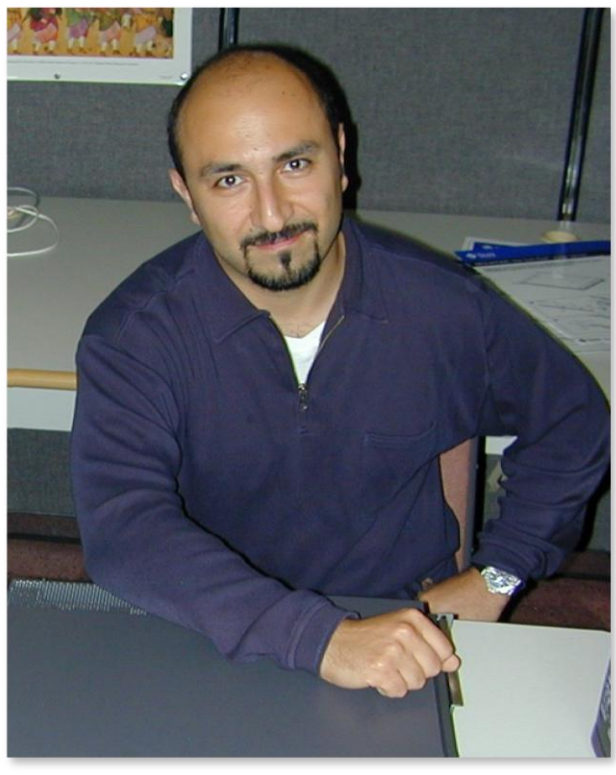
The School of Informatics and Computing Colloquium Series



Cenk Sahinalp

Simon Fraser University

Wednesday,
February 15, 2012
3:00-4:00 p.m.
Dogwood Rm, IMU

Discovery of Complex Genomic Rearrangements in Cancer using High-Throughput Sequencing

Abstract: High-throughput sequencing now allows us to identify patterns of rearrangement in tumor genomes, and has led to the discovery of Complex Genomic Rearrangements (CGRs) as a new cytogenetic feature of some cancers. Closed chain breakage and rejoining (CCBR), a generalization of reciprocal translocation, has been identifi ed in prostate cancer and acute lymphoblastic leukemia (ALL). CCBRs involve the balanced rearrangement of some n loci, and thus have the potential to fuse or interrupt up to n genes. In addition, dispersed throughout some breast cancer genomes are 'genomic shards', small fragments of DNA originating from elsewhere in the genome and inserted at the breakpoints of larger scale rearrangements.

Both CCBRs and genomic shards have implications for the discovery and characterization of gene fusions. A single CCBR can produce multiple gene fusions, and association of these related gene fusions may provide important information about a tumor's biology. Genomic shards have the potential to confound discovery of gene fusions. Discovery of a gene A to gene B rearrangement may falsely nominate an A-B gene fusion, when in fact the complete rearrangement is A-B-C: an A-C gene fusion with a shard of B inserted at the A-C breakpoint.

We have developed nFuse (http://n-fuse.sf.net), an algorithmic method for identifying CGRs in WGSS data based on shortest alternating paths in break- point graphs. We target our search for CGRs to those that underlie fusion transcripts predicted from matched high-throughput cDNA sequencing (RNA-Seq).

One of the most interesting results of nFuse on a xenograft of prostate tumor LTL-313H is a PCR-validated n-fusion composed of 9 genomic shards scaff olded by 8 fusion transcripts.

In a primary prostate tumor 963, we identi fied 5 CCBRs, 2 involving three loci, 2 involving four loci, and 1 involving 5 loci.

Biography: I was born in 1969, in Ankara, Turkey. I studied Electrical Engineering at Bilkent University for my B.Sc., and got my Ph.D. in Computer Science from University of Maryland, College Park. After my Ph.D. I first worked at Bell Labs, Murray Hill as a staff member and then at University of Warwick as a faculty. During my time at Warwick I was also affiliated with the Center for BioInformatics, University of Pennsylvania and frequently visited DIMACS, Bell Labs, and AT&T Research. Getting bored with being on the move all the time, I moved to Cleveland in 1999 and joined Case Western Reserve University as an Assistant Professor of computer science. There I became a founding member of Center for Computational Genomics and an adjunct faculty at the Department of Genetics, Case School of Medicine. I am now a Professor and Canada Research Chair in Computational Genomics (learn more about my research in French!) at Simon Fraser University where I direct the Lab for Computational Biology. I am also a visiting faculty at Department of Genome Sciences, University of Washington, a member of the MSFHR-CIHR funded Bioinformatics Training Program and an associate faculty at the Department of Molecular Biology and Biochemistry, SFU. I spent the 2007-2008 academic year at the University of Washington and Sabanci and Bilkent Universities in Turkey. My research focuses on biomolecular sequence analysis, RNA structure and interaction prediction, topological properties of biomolecular networks and more recently QSAR analysis.

SCHOOL OF INFORMATICS AND COMPUTING
INDIANA UNIVERSITY
Bloomington