

MURPA Seminar Friday 5th August 2011 at 9am: Rapid detection of gene-gene interactions in genome-wide association studies

Speaker: Vineet Bafna, Professor, Computer Science and Engineering UCSD

Venue: Seminar Room 135, Building 26 Monash Clayton

Abstract: In complex disorders, independently evolving locus pairs might interact to confer disease susceptibility, with only a modest effect at each locus. With genome-wide association studies on large cohorts, testing all pairs for interaction confers a heavy computational burden, and a loss of power due to large Bonferroni like corrections. Correspondingly, limiting the tests to pairs that show marginal effect at either locus, also has reduced power. Here, we describe an algorithm that discovers interacting locus pairs without explicitly testing all pairs, or requiring a marginal effect at each locus. The central idea of our algorithm is a mathematical transformation that maps 'statistical correlation between locus pairs' to 'distance between two points in a Euclidean space'. This enables the use of geometric properties to identify proximal points (correlated locus pairs), without testing each pair explicitly. For large data sets ($\sim 10^6$ SNPs), this reduces the number of tests from $\sim 10^{12}$ to $\sim 10^6$, significantly reducing the computational burden, without loss of power. The speed of the test allows for correction using permutation based tests. The algorithm is encoded in a tool called Rapid (RAPid Pair IDentification) for identifying paired interactions in case-control GWAS. We validated Rapid with extensive tests on simulated, and real data sets. On simulated models of interaction, Rapid easily identified pairs with small marginal effects. On the benchmark disease data sets from WTCCC, Rapid ran in ~ 30 CPU-minutes per data set, and identified many significant interactions. In many cases, the interacting loci were known to be important for the disease, but were not individually associated in the genome wide scan.

Bio:

Vineet Bafna joined the UCSD faculty on July 1, 2003, after seven years in the biosciences industry. He received his Ph.D. in Computer Science from Pennsylvania State University in 1994, and was a post-doctoral researcher at its Center for Discrete Mathematics and Theoretical Computer Science for two years. From 1996-99, Bafna was a senior investigator at SmithKline Beecham, conducting research on DNA signaling, target discovery and EST assembly. From 1999 to 2002, he worked at Celera Genomics, ultimately as Director of Informatics Research, at a time when Celera was decoding the human genome. He arrived at the Jacobs School from the Center for Advancement in Genomics, recently set up by Celera founder Craig Venter. Bafna has been on the Program Committees of the past two annual International Conferences on Research in Computational Biology (RECOMB), and is on the editorial board of the Journal of Bioinformatics and Computational Biology (JBCB). He received his B.Tech from the Indian Institute of Technology in 1989, and has published two dozen articles in major journals and conference proceedings.