Genomics & Medical Bioinformatics



Jacques S. Beckmann, Chair

A New Molecular Approach for Haplotyping in Large Population-Based Association Studies Benjamin Yakir, Dept. of Statistics, Hebrew University of Jerusalem

Authors: Ester Inbar, Benjamin Yakir and Ariel Darvasi, Hebrew University of Jerusalem Determination of haplotype frequencies (the joint distribution of genetic markers) in large population samples is a powerful tool for association studies.

Population haplotype frequencies evaluate linkage disequilibrium between markers. Haplotypes are of great value for association studies due to their greater extent of variability. Therefore, a single haplotype may capture any given functional polymorphism with higher statistical power than its SNP components. The statistical estimation of haplotype frequencies, usually employed in LD studies, requires the individual genotyping for each SNP in the haplotype, thus making it an expensive process.

In this talk, we describe a new method for direct measurement of haplotype frequencies in DNA pools, by allele-specific, long-range, amplification of the pool.

The proposed method allows high throughput genotyping of haplotypes composed of two SNPs in close vicinity (up to 20Kb).

We will discuss some of the statistical implications of applying this approach in large population based association studies.

Expression Profiling and the Quest for BRCAx

Zohar Yakhini, Agilent Technologies and Technion

Recent studies on molecular level classification of cancer cells produced results that strongly indicate the potential of gene expression profiling assays as diagnostic and segmentation tools and as a basis to the discovery of putative disease subtypes. At Agilent labs we are developing measurement and data analysis techniques that enable the acceleration of accurate and statistically sound expression based studies.

I will discuss several aspects of designing expression profiling assays and of analyzing resulting data. In particular I shall describe methods for un-supervised class discovery. In classified gene expression data an overabundance of genes that sharply separate the classes is typically observed. Reversing the relationship between overabundance and biological meaning we developed an approach to unsupervised class discovery. The output of the latter computational task is a partition of the set of samples which is supported by a statistically meaningful set of genes. Overabundance analysis is used as a figure of merit in our class discovery approach.

The genetic determinants of familial breast cancer that is not attributable to BRCA1/2 mutations are generally designated as BRCAx. I will discuss the application of class discovery and other data analysis tools in a BRCAx study recently completed with NHGRI collaborators.

Understanding the Emergence of Specific Information in Living Systems

Irun Cohen, Dept. of Immunology, Weizmann Institute of Science

Living systems *mine* information from their environments and *create* information within themselves; that's how they manage to live. Our task is to understand how *they* and *we* do it.