Abstract of Presentation

Presentation Title:

Powerful Statistical Genetics Strategies Reinforce Genome-Wide Association Studies

Abstract :

Clarifying the genetic and environmental causes of common diseases requires efficient detection strategies and a sophisticated approach. Because each factor's effect should be moderate, the common disease-common variant hypothesis predicts the association study to be the most powerful study design. Therefore, we have developed an original high-throughput genotyping/association study system based on the JSNP database, from which our analysis of SNP genotype data from more than 13,000 genes led to a number of common disease findings, including a pioneering paper on myocardial infarction; the first report of a successful genome-wide association study in the world. Our groundbreaking construction and analysis of a genome-wide linkage disequilibrium map uncovered more complicated linkage disequilibrium patterns than those expected previously, with distant exonic SNP pairs exhibiting greater linkage disequilibrium than other intragenic SNPs. In addition, we identified tagging SNPs to reduce genotyping costs and then extended this strategy to the whole genome by participating in the International HapMap project, through which we constructed the foundation for genome-wide association studies by identifying a half-million genome-wide tagging SNPs. Detailed analyses of multiple genetic and environmental factors require advanced genetic models because their interactions likely increase individuals' disease risk in a complicated fashion. Integration of methods incorporating environmental factors and haplotypes with a step-wise procedure for multiple SNPs could contribute to identifying more plausible SNPs and combinations of SNPs/haplotypes as targets for further study. In addition, we recently proposed algorithms for phasing haplotypes of complicated patterns of copy-number-variations (CNVs) and SNPs, which could open a new era of analyzing structural variations' influence on common diseases. In the future, we plan to combine and analyze various types of data: genome-wide genotypes, environmental factors, and clinical information from many patients with various disease/drug-response phenotypes to identify background mechanisms and build mathematical theories that can facilitate our final goal, the development of personalized medicine.