Estimation of Global And Local Genomic Sharing Using Genome-Wide SNP Data

Pak-Chung Sham
Genome Research Centre, University of Hong Kong and SGDP Centre, Institute of Psychiatry, King's College London

Purcell S
Center for Human Genetic Research, Massachusetts General Hospital, Boston

Neale BN
SGDP Centre, Institute of Psychiatry, King's College London

Daly M
The Broad Institute of Harvard & MIT, Cambridge

Recent technological developments have made possible the efficient genotyping of hundreds of thousands of SNPs in individual DNA samples. Genetic data of this quantity provide detailed information on the ancestry of individuals and on the relationship between pairs of individuals. We have developed a simple identity-by-state method to estimate the global genomic sharing between two individuals from the same population. This method was used to detect previously unknown relationships among individuals in the HapMap population samples. We are now developing a Hidden Markov Model (HMM) to obtain estimates of local genomic sharing between pairs of individuals, given the genotype data and the estimated global genomic sharing. The methodology will take some account of complications such as genotyping errors and linkage disequilibrium between SNPs, and result estimated probabilities of sharing none, one or both haplotypes by descent, at any section of the genome. The method will be implemented in the PLINK software (http://pngu.mgh.harvard.edu/purcell/plink/) and applied to the HapMap Phase 2 data. Finally, we will discuss the potential use of this methodology to gene-mapping studies of complex diseases and quantitative traits.

[Pak-Chung Sham, Genome Research Centre, Level 6, Lab Block, Li Ka Shing Faculty of Medicine, 21 Sassoon Road, Pokfulam, Hong Kong; pcsham@hku.hk]