Thesis Defense

On the Importance of Phase in Improving Detection of Shared Genomic Segments

Giulio Genovese

Tuesday May 25, 2010 ï 10:30 am ï Kemeny 004

Abstract

This thesis investigates techniques for the discovery and use of phase information in genetic studies. ``Phase'' refers to the precise assignment of the members of a different base pairs at the same genomic locus to two possible candidate homologous chromosomes. This ``missing'' information is usually not available in most genetic studies. We investigate the importance of phase information in linkage studies and we propose two methodologies to take advantage of inaccurate or incomplete but available phase information for the purpose of detection of identical by descent regions in the human genome, and to infer missing phase information starting from identical by descent regions in a fashion suitable for a large ñbiobankò of genotyped samples. We evaluate commonly used phase algorithms and show that for large pools of unrelated individuals, while good at discerning the relative phase among loci within the same linkage disequilibrium blocks, they perform poorly in the task of retrieving relative phase in between distinct linkage disequilibrium blocks, with the likely result of adding limited value to the problem of distinguishing identical by descent regions. Our approach should prove useful for the improved detection of shared genomic segments in future genetic studies.