Cross-sample and cross-platform analysis of DNA copy number

by
Nancy R. Zhang
Stanford University
390 Serra Mall
Stanford, CA 94305-4065
nzhang@stanford.edu

Abstract

DNA copy number analysis involves the detection of chromosomal gains and losses using high-density microarray or genotyping platforms. Change-point methods have been applied successfully to detecting signals in single data sequences derived from one biological sample. However, it is common in association studies to have data from hundreds to thousands of biological samples. How should information be combined across samples to detect population level polymorphisms? It is also now common to have the same biological sample assayed using multiple experimental platforms. For example, in the Cancer Genome Atlas project, each biological sample is processed using Illumina, Affymetrix and Agilent chips. How should data be integrated across platforms to achieve higher accuracy? I will discuss the statistical issues underlying this problem and formulate the models and methods for cross-sample and cross-platform data integration.