

Hidden Markov Models for the Assessment of Chromosomal Alterations using High-throughput SNP Arrays

Ingo Ruczinski

Department of Biostatistics,
Johns Hopkins Bloomberg
School of Public Health,

Abstract

Many chromosomal alterations such as amplifications and deletions have been associated with disease. For the genome-wide assessment of such alterations, high-throughput single nucleotide polymorphism (SNP) arrays are often used, estimating DNA copy number and genotypes at up to 1 million loci. Hidden Markov Models (HMMs) are particularly useful for inferring chromosomal alterations from SNP array data, modeling the spatial dependence between neighboring SNPs. We propose a novel HMM for the assessment of the underlying chromosomal states, simultaneously integrating copy number estimates, genotype calls, and the corresponding measures of uncertainty when available. We also show how parent-of-origin effects can be assessed in family data, and discuss the publicly available software packages we have developed