

# **“A Multilevel Hidden Markov Model of DNA Copy Number Variation”**

**Stephen Erickson, PhD**

University of Alabama at Birmingham  
Biostatistics Department

## **ABSTRACT**

DNA copy number variation (CNV) has received increasing attention over the past few years, following findings published in 2004 showing significantly greater CNV content in humans than previously thought (Sebat et al, Iafrate et al). Comparative genomic hybridization microarrays ("CGH arrays") and high-density genotyping microarrays ("SNP arrays") can assay CNV on a genome-wide scale. Several studies have revealed both common, inherited Variants, and unique, de novo variants in healthy, as well as diseased, subjects.

Most existing methodologies for inferring copy number state from microarray data have focused on analyzing arrays one at a time. A mixture, or multilevel, hidden Markov model (mHMM) of CNV, on the other hand, accounts both for the correlation of CNV state at different loci within an individual (i.e. along chromosomes) as well as between individuals. In this talk, I will introduce a general inferential framework, discuss computational and inferential challenges, and present some promising initial results on simulated data.